

P.A. Butcher · G.F. Moran

Genetic linkage mapping in *Acacia mangium*. 2. Development of an integrated map from two outbred pedigrees using RFLP and microsatellite loci

Received: 25 March 2000 / Accepted: 30 April 2000

Abstract An integrated genetic linkage map, comprised of 219 RFLP and 33 microsatellite loci in 13 linkage groups, was constructed using two outbred pedigrees of *Acacia mangium* Willd. The linkage groups ranged in size from 23 to 103 cM and the total map length was 966 cM. Individual maps were made for each pedigree and the ordering of loci was consistent with the integrated map. The use of two independent pedigrees allowed a comparison of recombination rates between linked loci in male and female meioses as well as between parents. Differences were confined to specific regions and were not uniform across the male and female genomes or between genotypes. The heterogeneity in recombination frequencies did not result in major differences in the ordering of loci between pedigrees; hence, the integrated map provides a sound basis for QTL detection, leading to marker-assisted selection in *A. mangium*. It also provides a reference map for comparative genome analysis in acacias. The co-dominant markers used for mapping provide a useful resource in population studies and for quality control in acacia breeding programs. Detection of a relatively high proportion of selfs in pods derived from flowers which were not emasculated (30%), compared with emasculated flowers (0.01%), indicates that emasculation is desirable for efficient delivery of control-crossed seed in acacia breeding programs.

Key words Genomic map · *Acacia mangium* · Recombination rate · Microsatellites · RFLP · Legumes · Polyads

Introduction

Acacia mangium Willd is a diploid forest tree ($2n = 2x = 26$) which, because of its rapid growth rates, adaptability to a wide range of sites and high pulp quality and yield (reviewed in Doran and Turnbull 1997), has become one of the most important plantation species in southeast Asia. Breeding programs for *A. mangium*, as for most other forest trees, have been based on recurrent selection from open-pollinated families. There is, however, a trend towards clonal forestry which opens opportunities for the use of marker-assisted selection. Genetic linkage maps provide an important tool for the location of genes linked to traits of economic importance which could be used in marker-assisted selection.

Mapping in forest trees generally relies on outbred pedigrees where genetic segregation is the result of meiotic recombination from both parents. While there have been no direct comparisons of the performance of inbred and outcrossed progeny of *A. mangium*, the observed decline in performance of individuals derived from a single mother tree in successive generations (Sim 1984) provides some evidence of inbreeding depression associated with high genetic load. This typically precludes the use of inbred lines in tree breeding programs (Griffin and Cotterill 1988; Williams and Savolainen 1996). In *A. mangium*, mapping pedigrees were therefore developed using outcrossed full-sib families from natural populations in New Guinea. Populations in this region had previously been identified as being the most heterozygous (Butcher et al. 1998). They have shown superior growth rates on a range of sites (Harwood and Williams 1992) and are the preferred source of breeding material in current tree improvement programs throughout southeast Asia.

A mapping strategy was developed based on two aspects of the reproductive biology of *A. mangium*. Firstly, the difficulty of producing large progeny arrays from controlled crosses due, in part, to the small size of flowers, and compound inflorescences with variable proportions of male flowers and low seed set (Sedgley et al.

Communicated by P.M.A. Tigerstedt

P.A. Butcher (✉) · 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

1992). Secondly, acacias produce compound pollen grains (polyads) so that all seeds within a pod may not be produced from independent meiotic events. The polyad is considered to be a mechanism for ensuring full pod-set following a single pollination event (Kenrick and Knox 1982). Cytological studies in acacias have shown that ovules in an ovary are derived from independent meiotic events (Buttrose et al. 1981). All ovules within a flower can be fertilised by a single polyad as the number of pollen grains in the polyad (16 in *A. mangium*) is the same as the maximum number of ovules within a flower (Sedgley et al. 1992). The polyads are derived from two rounds of mitosis preceding meiosis in the sporogenous cell (Newman 1933; Kenrick and Knox 1979). This implies that up to four seeds in a pod can be produced from a single meiosis. As tests for linkage are based on the assumption that all individuals in a mapping pedigree are derived from independent meioses, ideally only one individual per pod would be used for mapping. Where the number of pods produced is limited, the need for independence must be balanced with that of maximising the precision with which genetic linkage is measured. Precision is proportional to the number of individuals studied (Paterson 1996). A maximum of two seeds were therefore used from any one pod from each of two crosses.

Mapping with multiple full-sib pedigrees allows mapping of a larger number of loci than with a single pedigree, thereby increasing genomic coverage and/or marker density in specific genomic regions. Increased marker density increases the probability of identifying closely linked polymorphic loci for marker-assisted selection or for map-based cloning (see for example the *Phaseolus vulgaris* integrated map, Freyre et al. 1998). The use of two pedigrees provides a means of evaluating the repeatability of map construction and may provide an efficient means of validating quantitative trait loci (QTLs). Simulation studies indicate that the use of more than one full-sib pedigree increases the power to detect QTLs, especially where the QTLs explains more than 10% of the phenotypic variance (Muranty 1996). In other genera, mapping with multiple populations has provided evidence for chromosomal rearrangements (Beavis and Grant 1991; Kianian and Quiros 1992), gene duplication (Gentzmittel et al. 1995) and differences in recombination rates between sexes and genotypes (Fatmi et al. 1993; Sewell et al. 1999). A potential problem with pooling information from multiple pedigrees is that if heterogeneity of recombination is common, then assembling a consensus map for a species will be difficult (Beavis and Grant 1991; Plomion and O'Malley 1996).

In the present study, linkage data from two independent outcrossed pedigrees are combined to provide an integrated map for *A. mangium*. This is the first linkage map produced for an acacia and provides the basis for marker-assisted selection and comparative studies of genome organisation in other members of the genus.

Materials and methods

Development and testing of the mapping pedigrees

The mapping pedigrees consisted of two outbred two-generation crosses of *A. mangium*. The crosses were made between selected trees from four populations of *A. mangium*, as described in Butcher et al. (2000a). Pollination procedures followed Sedgley et al. (1992) with the exception that one-third of the flowers were not emasculated prior to pollination. This allowed investigation of the effects of emasculation on rates of germination, survival, selfing and contamination from foreign pollen. One hundred and forty progeny from cross A and 175 progeny from cross B were screened using ten RFLP loci to detect selfs and contaminants. Selfed progeny were identified using probes which hybridised to loci for which the male and female parent were fixed for alternate alleles (Fig. 1). Fully informative probes were used to maximise the probability of detecting individuals derived from foreign pollen.

To balance the competing aims of maximising the number of progeny in the mapping pedigrees and the probability of sampling independent recombination events, a maximum of two seeds were used from any one pod. Linkage analysis was based on marker genotypes of 108 plants from cross A and 123 plants from cross B.

DNA isolation and marker development

DNA isolation followed Butcher et al. (1998). RFLP and microsatellite markers were developed from *A. mangium* genomic libraries following procedures described in Butcher et al. (2000a) for RFLPs and Butcher et al. (2000b) for microsatellites. RFLP loci are denoted by the prefix g and SSR loci by the prefix Am.

Data analysis

Loci were scored concurrently for progeny of the two pedigrees, and fragment length similarities were used to infer orthologous loci and alleles. All loci were co-dominant. Alleles were numbered consecutively, starting with the largest fragments labelled 1. Loci were scored independently by two people to minimise scoring and the interpretation errors. Where there was doubt about whether loci were orthologous they were given different subscripts in the two pedigrees. If both mapped to the same position, the subscripts were removed. Segregation of the loci was tested for deviation ($P < 0.5$) from expected Mendelian segregation by χ^2 analysis (see Butcher et al. 2000a,b).

Individuals within the mapping pedigrees should all have the same degree of relatedness. This was examined using datasets containing the male haplotypes. The similarity between pairs of individuals was estimated using group-average clustering analysis in GENSTAT 5 Release 4.1 (Payne et al. 1987).

To determine whether the use of two seeds from a pod was likely to bias recombination estimates, the male haplotypes were examined along each linkage group. This was done separately for

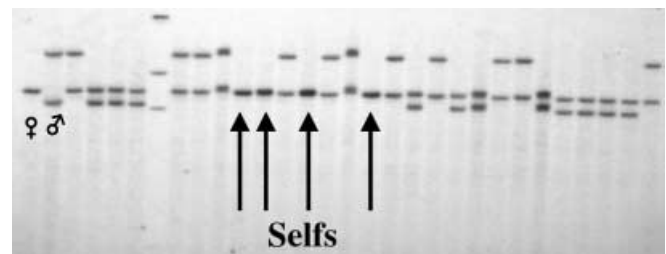


Fig. 1 Identification of selfs amongst progeny from cross A at RFLP locus g592 (λ HindIII size marker in lane 7)

crosses A and B. If a crossover occurred at a given point in one individual from the pair, but not in the other, it could be assumed that different recombination events were sampled. The number of pairs where different crossovers were detected between individuals from the same pod (observed) were summed for each linkage group. This was compared with equivalent data for the same number of pairs of randomly selected individuals from different pods (expected) using a χ^2 test.

Linkage analysis

Analysis of linkage between loci was carried out using JOINMAP version 2.0 (Stam and van Ooijen 1995). Pairwise recombination frequencies were estimated by maximum likelihood. Linked markers were initially placed into linkage groups with $\text{LOD} \geq 5$ for cross B and $\text{LOD} \geq 4$ for cross A. These LOD scores were selected to avoid significant jumps in χ^2 values within linkage groups. Map distances were calculated using Kosambi's mapping function (Kosambi 1944). Following construction of preliminary maps, the banding patterns of probes which produced co-segregating loci were compared. If the same pattern was observed, the results for one of the loci was deleted from the data set.

To integrate the two maps, pairwise recombination frequencies for the two pedigrees were combined into a single data set and re-analysed for linkage using JOINMAP. Differences in recombination rates between linked loci were tested using a χ^2 test of heterogeneity in JOINMAP. Where there were differences in ordering between the integrated map and the single-pedigree maps, loci were entered as fixed orders reflecting those in the majority of the haploid maps. The order was accepted if it did not result in a significant jump in the χ^2 value.

Results were compared with those from the gene-ordering algorithm of MULTIMAP version 1.1 (Matise et al. 1994) (<http://linkage.rockefeller.edu/multimap>). For two-generation outbred pedigrees the linkage phase of alleles is not known a priori and must be determined from the progeny segregation data (Maliepaard et al. 1997). In JOINMAP, phase is determined internally, based on the recombination frequency (Stam and van Ooijen 1995). The inclusion of phase in the input for MULTIMAP, inferred from progeny segregation data, increased the number of loci which were mapped.

Heterogeneity of meiotic recombination

To investigate sex-related differences in recombination and to determine whether the ordering of loci was consistent between sexes, loci for crosses A and B were divided into maternal and paternal haploid datasets. For loci showing segregation of two alleles where both parents were heterozygous, only the homozygotes were included in the analysis as the origin of the alleles in the heterozygotes could not be determined. Recombination rates between linked loci were estimated independently for the male and female parents of each cross. These were then merged and differences between recombination rates determined using χ^2 tests for heterogeneity. Male and female maps were constructed using procedures described for the diploid data to determine whether the ordering of loci was consistent between sexes.

Table 1 The number of pods produced following cross-pollination of *A. mangium*, with and without emasculation, showing the number of progeny derived from selfing and foreign pollen

Item	Cross A		Cross B	
	Emasculated	Not emasculated	Emasculated	Not emasculated
Florets pollinated	774 (48 spikes)	173	230 (18 spikes)	197
Pods produced	54 (0.07%)	13 (0.08%)	59 (25.7%)	11 (5.6%)
Seeds produced	317	75	470	72
Progeny tested	118	22	158	18
Selfs	0	4 (3 pods)	1	3 (2 pods)
Foreign pollen	0	1	0	1

Estimation of genome length and map coverage

Estimates of genome length, $E(G)$ (Hulbert et al. 1988; method 3 in Chakravarti et al. 1991), and expected genome coverage, $E(C_n)\%$ (Bishop et al. 1983), were calculated from pairwise segregation data for marker pairs above a threshold LOD of 4, excluding loci which were heterozygous for the same alleles in the male and female parents. Confidence intervals for genome length were calculated following the method of Gerber and Rodolphe (1994). The observed genome coverage was calculated from the results of map construction as described in Nelson et al. (1994). These procedures are detailed in Echt and Nelson (1997).

Results

Error detection in the mapping pedigree

The number of pods produced from florets that were pollinated with and without emasculation is summarized in Table 1. Twenty percent of pods in cross A and 16% of pods in cross B were produced from florets which were not emasculated. Pod set in cross B was 20% higher from emasculated flowers but no different in cross A to that from flowers which were not emasculated. Germination and survival were 8% and 12% lower respectively amongst seeds produced without emasculation (data not shown). Screening of over 300 progeny with ten RFLP probes revealed that 30% of pods produced from flowers which were not emasculated contained selfs or seed derived from foreign pollen compared with less than 1% of pods from emasculated flowers (Table 1). The proportion of contaminants among the progeny was lower (23%) reflecting the lower survival rate from flowers which were not emasculated. Where a self was detected in a pod from a flower which was not emasculated and a second progeny was tested from that pod, it was also found to be a self.

The similarity analysis showed that no individuals were unusually similar. The highest level of similarity was 81% and 83% for two individuals (from different pods) in crosses A and B respectively. All individuals merged into a single group by 48% similarity, indicating

Fig. 2 Genetic linkage map for *A. mangium*. Linkage groups for cross A are shown on the left and those for cross B on the right of an integrated map (AB). Loci are listed on the right and recombination distances (cM, Kosambi) on the left of each linkage group. RFLP markers are denoted by the prefix *g* and microsatellites by the prefix *Am*. Loci with distorted segregation ratios ($P \leq 0.05$) are marked with an *asterix*

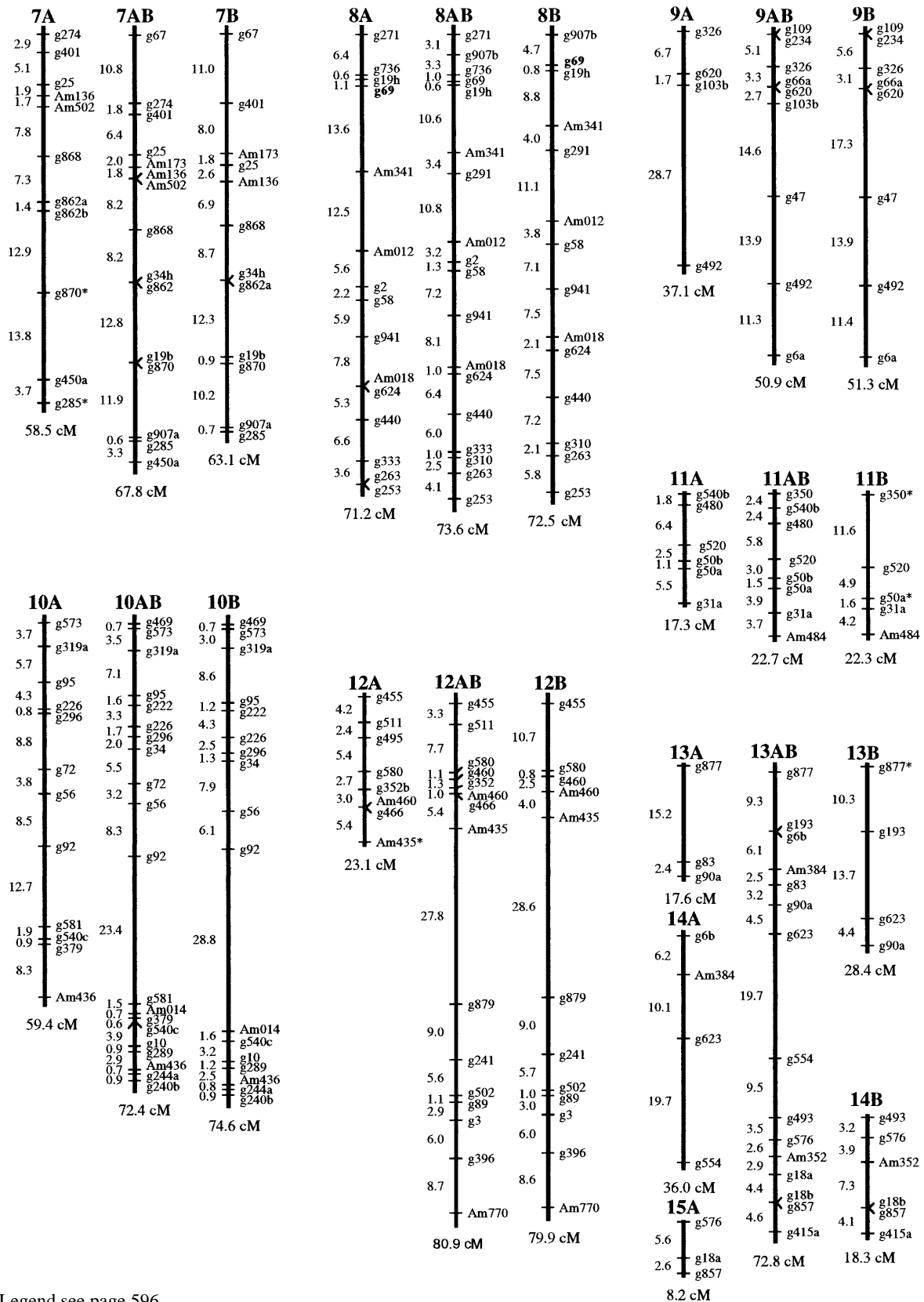


Fig. 2 Legend see page 596

that no individuals were unusually distinct from the rest of the mapping population.

Less recombination was detected between pairs of individuals from the same pod than between those from different pods (cross A 407 vs 424; cross B 566 vs 579). However, the differences were not statistically significant ($\chi^2 P \leq 0.5$), indicating the recombination estimates in this study were not significantly biased as a result of using a maximum of two seeds per pod.

Segregation of markers

Segregation data from 263 loci were initially used for map construction. Twelve RFLP loci were removed as they produced identical genotypes to at least one other probe. One hundred and fifty one RFLP and 24 microsatellite loci in cross A and 170 RFLP and 30 microsatellite loci in cross B were assigned to linkage groups. Nineteen percent of loci in cross A and 24% of loci in cross B were fully informative. A total of 102 loci were polymorphic in both crosses (Table 2). One hundred and forty eight RFLP probes (87%) gave a single locus. Of these, 62 RFLP loci, together with 21 microsatellite loci which were polymorphic in both pedigrees, were considered orthologous by default. An additional 19 loci from probes which revealed more than one locus were orthologous as they mapped to the same position.

Loci with distorted segregation ratios (Butcher et al. 2000a,b), marked by an asterisk in Fig. 2, were randomly distributed throughout the genome [cross A ten loci (5.7% of RFLPs); cross B nine loci (4.5% of RFLPs)] as expected using a 5% probability criterion. No SSR loci and only one RFLP locus deviated significantly from the expected segregation ratios in both crosses. The RFLP locus caused a significant jump in χ^2 in cross B and was removed from this data set. The absence of any clustering of loci with distorted segregation ratios suggests there is no biological basis for the observed distortion.

Linkage analysis for crosses A and B

One hundred and sixty three loci in cross A were assigned to 16 linkage groups using JOINMAP's JMGRP module and a LOD threshold of 4. One group contained only two RFLP loci which were linked to each other but not strongly linked to other loci, and one locus was unlinked. One hundred and eighty seven loci in cross B were initially assigned to 15 linkage groups using a LOD threshold of 5. Two groups were combined to form linkage group 10 (LOD 3.5) based on evidence from cross A. No significant jump in χ^2 resulted from combining the groups. There were no unlinked loci. Loci segregating from the male and female parents, or both, were spread across the linkage groups. The ordering of loci in the two maps was consistent with only minor differences in the order of closely linked loci, marked in bold on Fig. 2.

Table 2 Number of markers which were polymorphic in more than one parent and/or mapping pedigree

Mapping population	Segregating markers		
	Intra-pedigree	Inter-pedigree	Total ^a
Cross A maternal	47	71	84
Cross A paternal	47	66	79
Cross B maternal	65	76	95
Cross B paternal	65	72	91
Integrated	–	102	136

^a Total number of markers represented in intra- and/or inter-pedigree categories

Comparison of meiotic recombination and ordering of loci among parents

To investigate heterogeneity in recombination frequencies, the data for each cross were divided into maternal and paternal haploid data sets. The linkage data from the four independent mapping parents (251 loci) contained 136 markers which were common to at least two of the haploid data sets (Table 2). There were significant differences in recombination frequencies between the pairs of markers in cross A listed in Table 3. The majority of differences were on linkage group 2. This linkage group also contained the highest number of orthologous markers (ten) raising the question of whether the number of differences which were detected simply reflect sampling error. No significant differences in the rate of recombination were detected between male and female meioses in cross B. Recombination frequencies are averaged in JOINMAP so, where there are differences between male and female meioses, or between crosses, the distances on the integrated map are approximate.

The ordering of loci on the maternal and paternal maps of cross A was, however, consistent (data not shown) with only one reversal in order of Am164 and g474 at the end of linkage group 1. The distance between the two markers was 1 cM in the female map and 3.5 cM in the male. In cross B there was minor rearrangement in the order of the closely linked markers, g630, g441 and g374, on linkage group 2. These markers were separated by a distance of 0.7 and 2.3 cM in the maternal and paternal maps respectively. Am136 was placed after Am173 in linkage group 7 on the maternal map and before g25 on the paternal map, and the order of Am770 and g396 on linkage group 12 was reversed. Given the small distances involved, these differences could simply reflect statistical inaccuracies in the estimated recombination frequencies (Maliepaard et al. 1997). They did not result in changes in marker order on the diploid maps.

Integrated map

The pairwise recombination frequencies for each cross were combined and linkage groups for the merged data

Table 3 Comparisons of recombination frequencies between linked loci for male and female meioses in cross A. Only those significantly different between sexes ($P \leq 0.05$) are shown

Linkage group	Interval	Male	Female	Male/female	χ^2	P
2	g593–g474	0.140	0.035	4.00	4.30	0.038
2	g403–Am429	0.389	0.167	2.33	6.99	0.008
2	g403–g374	0.185	0.056	3.33	4.46	0.035
	g403–g441					
	g403–g630					
2	Am429–g374	0.478	0.315	1.51	5.65	0.018
2	Am429–g441	0.469	0.324	1.44	4.55	0.033
	Am429–g630					
2	Am429–g869	0.092	0.284	0.32	13.72	0.000
8	g2hp–g941	0.055	0.138	0.40	4.39	0.036
10	g319a–g95	0.136	0.015	8.97	7.95	0.005

Table 4 Comparisons of recombination frequencies between linked loci for cross A and cross B. Only those significantly different between crosses ($P \leq 0.05$) are shown

Linkage group	Interval	Cross A	Cross B	Cross A/B	χ^2	P
1	g334–g618	0.099	0.024	4.10	6.13	0.013
1	g19a–g618	0.346	0.213	1.62	4.81	0.028
1	g342–g90b	0.115	0.252	0.46	4.82	0.028
2	g441–Am368	0.261	0.396	0.65	4.47	0.035
2	g926–g308	0.036	0.133	0.27	5.60	0.018
2	g308–g403	0.074	0.267	0.28	8.06	0.005
2	g926–g403	0.056	0.168	0.33	4.73	0.030
4	g436 and g649	0.202	0.360	0.56	4.94	0.026
10	g56 and g92a	0.155	0.059	2.63	3.99	0.046
11	g31a and g50a	0.113	0.061	1.86	6.92	0.009

set formed at $\text{LOD} \geq 4$ using the JMPWG program of JOINMAP; 249 loci were mapped in 14 linkage groups (13 linkage groups at $\text{LOD} \geq 3.5$). The two markers which were linked to each other, but not to other markers in cross A (g499 and g569), were initially grouped on linkage group 3 but caused significant jumps in χ^2 values and were therefore omitted from the map. All other markers were mapped. Combining the data from the two pedigrees provided additional linkage information which enabled several of the smaller linkage groups to be combined. The previously unmapped marker from cross A was also mapped (g396 on linkage group 12).

The 13 linkage groups of the integrated map span 966 cM with an average distance between markers of 4.6 cM. Three regions of the map had distances between adjacent markers greater than 20 cM, namely g92a–g581 on group 10 (23.4 cM), Am435–g879 on group 12 (27.8 cM) and g623–g554 on group 13 (27.8 cM). These linkages were all supported by linkage between more than one pair of markers and LOD scores higher than 3. For example, g92a on linkage group 10 is linked to both g56 (lod=13.5) and Am436 (lod 3.6); g879 on group 12 is linked to g502 (lod 15.7) and g580 (lod 3.8); and g623 on linkage 13 is linked to Am352 (lod 3.9) and g90a (lod 54.2).

The allocation of loci into linkage groups from JOINMAP was verified using the MULTIMAP program. Fewer loci were ordered within linkage groups using MULTIMAP, however the ordering of the mapped loci ($\text{LOD} \geq 3$) was consistent with JOINMAP. The number of mapped loci was increased by inputting phase data, de-

rived from progeny segregation, and reducing the LOD score; however, the latter resulted in inconsistencies in marker order in several linkage groups. Map lengths were approximately 10% shorter in JOINMAP. A similar trend has been reported in several other studies (van Ooijen et al. 1994; Devey et al. 1996; Sewell et al. 1999) and has been attributed to the different mapping algorithms used. MULTIMAP uses adjacent marker pairs only to calculate distances whereas JOINMAP uses all pairwise estimates above a pre-defined LOD threshold. In addition, the likelihood method in MULTIMAP assumes an absence of interference, so where there is interference JOINMAP will produce shorter maps (Qi et al. 1996).

The ordering of loci was consistent between the two crosses despite significant differences in recombination rates between closely linked loci in the two pedigrees (Table 4). These differences were mainly in linkage groups 1 and 2 and the latter were due to lower frequencies in female meioses in cross A. The remaining differences were attributable to variation in recombination rates between crosses. No differences were detected for linkage groups 3, 5–9, 12 and 13. For the integrated map this does not appear to be due to there being fewer informative markers (ie segregating in both crosses) in which to investigate homogeneity in these linkage groups. While the proportion of orthologous loci ranged from 23% in linkage group 6 to 52% in linkage group 2, there was no relationship between the number of differences in recombination detected and the number of informative markers per linkage group.

Table 5 Number of markers used for map construction, estimated genome length, and expected and observed genome coverage for maternal and paternal maps

Mapping population	Number of mapped markers	Map coverage ^a	Estimated genome length ^a	95% Confidence interval ^a	Expected coverage (%)	Observed coverage (%)
A maternal	106	703	1468	1289,1704	86	55
A paternal	107	763	1490	1468,1960	85	52
B maternal	125	834	1564	1319,1919	90	54
B paternal	126	798	1673	1412,2051	89	58
Integrated	249	966	Mean 1548			62

^a Map units in centimorgans using the Kosambi mapping function

Estimated map length and genome coverage

The expected total length of the genome estimated using method 3 of Chakravarti et al. (1991) and maternal and paternal linkage data for locus pairs with a LOD score ≥ 4 for each cross, ranged from 1470 to 1670 cM (Table 5). This is within the range of other forest trees (Table 6). The estimated genome coverage was 85–90%; however, the observed genome length was considerably less (62%) indicating the need for more markers to be mapped.

Discussion

Quality control

RFLPs provided an effective means for detecting contamination and selfing in breeding populations. The relatively high proportion of selfs in pods produced from flowers which were not emasculated indicates that selection for outcross pollen is not strong in *A. mangium*. This is in contrast to other Australian acacias in which estimates of outcrossing rates at the viable seed stage, in natural populations, are very high (Moran et al. 1989; Muona et al. 1991; McGranahan et al. 1997). In *A. mangium* flowers have a short female phase which reduces the effectiveness of the protogynous outcrossing mechanism (Sedgley et al. 1992) and no difference has been found in ovule penetration between self and outcross pollen (Sedgley et al. 1992). The bagging of flowers presumably overcomes any temporal barrier which may otherwise prevent selfing in this species. Emasculatation reduced the number of pods containing progeny derived from selfing or foreign pollen from 30% to less than 1%. The higher rate of germination and survival among progeny produced from emasculated flowers may also reflect the lower proportion of selfs. Despite the additional effort required, emasculatation is therefore desirable in any controlled crossing program for *A. mangium*.

Integrated linkage map

The integrated map represents the most-likely order of markers given the available data. Differences in recom-

bination frequencies in the two crosses may cause some distortion of map distances between the relevant pairs of markers on the integrated map. However, the close correspondence between marker distances on the A and B maps (Fig. 2) shows that any distortion caused by combining these data was not extreme.

The high proportion of common markers, together with the consistency of gene order between the two individual maps, provide a measure of reliability for the integrated map. The integrated map was established on the basis of 102 common markers out of a total of 249 (41%). The proportion of common markers for linkage groups 1–13 was 47, 52, 36, 39, 46, 23, 44, 61, 33, 45, 38, 24 and 33% respectively. The order of unique markers in regions of the genome containing a low density of common markers will, however, be less reliable than in regions with a high density of common markers.

The distribution of markers in the 13 linkage groups of the integrated map was reasonably uniform with only three regions where the distance between two adjacent markers was greater than 20 cM. Similar gaps have been reported on most plant RFLP maps and may represent either regions of high recombination or genomic regions which were not sampled with probe isolation techniques. Based on estimates that the power to detect QTLs does not significantly increase for marker densities greater than one every 10–20 cM (Darvasi et al. 1993), the map should provide a firm basis for the location of QTLs. This, however, assumes 100% coverage of the genome.

The *A. mangium* map is shorter than most maps reported for forest trees (Table 6), reflecting either incomplete genome coverage or a lower rate of recombination. *A. mangium* has relatively low restriction fragment length polymorphism (Butcher et al. 1998). Positive correlations between the level of RFLP and rates of recombination have been observed within genomes (Begun and Aquadro 1992; Aguade and Langley 1994; Dvorak et al. 1998; Kraft et al. 1998), and a relationship between the level of DNA diversity and the overall rate of recombination in different genomes would not be unexpected. Lower levels of recombination have also been reported in other Leguminosae (Trifoleae) (reviewed by Young et al. 1996) and make map-based cloning difficult.

Table 6 Comparison of the observed map lengths and estimated map lengths for linkage maps developed for forest trees using RFLP and/or SSR markers

Species	Number of markers	Number of linkage groups/haploid chromosome number	Length of map (cMK)	Average distance between markers (cMK)	Estimated genome length (cMK)	Observed genome coverage (%)	Reference
<i>Acacia mangium</i>	217 RFLP 32 SSR	13/13	966	4.6	1550	62	Present study
<i>Populus tremuloides</i>	54 RFLP 3 isozyme	14/19	664	15.4	ne ^b	26 ^d	Lui and Furnier 1993
<i>P. trichocarpa</i> × <i>P. deltoides</i>	111 RAPD 215 RFLP 17 STS	35/19	1261	9.5 ^a	2400–2800	49	Bradshaw et al. 1994
<i>Eucalyptus nitens</i>	207 RFLP 119 RAPD 4 isozyme	12/11	1462	4.4	ne	ne	Byrne et al. 1995
<i>E. grandis</i>	240 RAPD 19 SSR	14/11	1551(f)	12.2	1620(f)	96	Grattapaglia and Sederoff 1994; Brondani et al. 1998
<i>E. urophylla</i>	251 RAPD 17 SSR	11/11	1101(m)	10.2	1160(m)	95	As above
<i>Quercus robur</i>	271 RAPD 18 SSR 10 SCAR 6 isozyme	12/12	893(f) 921(m)	11.1	1192(f) 1235(m)	85	Barreneche et al. 1998
<i>Cryptomeria japonica</i>	77 RFLP 12 RAPD 1 isozyme	13/11	887	10.3	ne	ne	Mukai et al. 1995
<i>Pseudotsuga menziesii</i>	125 RFLP 15 RAPD	17/13	1062	7.5	ne	ne	Jermstad et al. 1998
<i>Picea abies</i>	38 SSR 246 AFLP	29/12	2198	9.3	2840	77	Paglia et al. 1998
<i>Pinus radiata</i>	157 RFLP 16 SSR	14/12	1223	7	1660 ^c	75 ^c	Devey et al. 1999
<i>Pinus strobus</i>	64 RAPD 5 SSR	17/12	1204	14	2080	58	Echt and Nelson 1997
<i>Pinus taeda</i>	211 RFLP 12 SSR	20/12	1281	4.3	1660 ^c	75 ^c	Devey et al. 1999
<i>Pinus taeda</i>	278 RFLP 67 RAPD 12 isozyme	14/12	1227	3.4	1340(f) 1980(m)	74	Sewell et al. 1999

^a Based on 19 largest linkage groups^b Not estimated^c Based on estimated genome length for *P. taeda* from Sewell et al. (1999)^d Percentage of an estimated genome length of 2400–2800 cM for *Populus* (Bradshaw et al. 1994)

Heterogeneity in recombination

The use of two independent pedigrees, together with co-dominant markers, a proportion of which segregated in both sexes (33%) and/or both pedigrees (41%), allowed the comparison of recombination rates in male and female meiosis and between genotypes in *A. mangium*. Genome-wide trends were not evident. Significant differences in meiotic recombination rates between the sexes were concentrated in a single linkage group in one pedigree (Table 3). Sex-related differences are not uncommon in linkage studies of plants, including those of forest trees. In gymnosperms, evidence indicates that recombination occurs more frequently in male meiosis, for example *Pinus radiata* (Moran et al. 1983), *Pinus taeda* (Groover et al. 1995; Sewell et al. 1999) and *Pinus pinaster* (Plomion and O'Malley 1996), while in the angiosperms *Lycopersicon* (de Vicente and Tanksley 1991; van Ooijen et al. 1994) and *Lilium* (Burt et al. 1991) recombination is greater in female meiosis. Environmental factors have been suggested as a possible cause of differences in recombination between the sexes in gymnosperms where male and female meioses occur in different seasons (Beavis and Grant 1991). The *A. mangium* mapping pedigrees were produced in a seed orchard in the same flowering season and the differences in recombination are therefore more likely to be genetically rather than environmentally based.

The male and female *A. mangium* maps were based on recombination in four different parents and the observed differences in recombination frequency may therefore be due to a genotype effect rather than a sex effect. This could be tested using reciprocal crosses. Genotypic variation in the frequency of recombination has been observed in other legumes, for example *Glycine max* (Pfeiffer and Vogt 1990).

The differences in recombination between crosses in *A. mangium* were largely confined to two regions of the genome (linkage groups 1 and 2). They were between closely linked loci and did not result in major changes in the ordering of loci. As a result they did not cause major problems in the integration of maps.

The consistency of marker order in the four haploid maps provides evidence that there has been no intraspecific chromosomal rearrangements among the four *A. mangium* trees. While this is not unexpected, given that the trees come from the same geographic region (New Guinea), it could be speculated that genome rearrangements are more likely to occur where there has been a high level of genetic differentiation among populations. This has been documented for the Australian and New Guinea populations of *A. mangium* (Butcher et al. 1998). The limitation may reside in the power of mapping algorithms to statistically detect linkage reorganisation on a broad scale.

Estimated genome coverage

The expected total length of the genome estimated for the maternal and paternal data sets of both crosses ranged from 1470 to 1670 cM and the expected coverage of the genome from 85 to 90% (Table 6). Using the average estimated genome length of 1550 cM, a DNA content of 2.277 pg/2 C Mbp (Mukherjee and Sharma 1995), and given that 1 pg of DNA = 0.965×10^9 bp (Bennett and Smith 1976), the average physical equivalent per unit of genetic distance is 710 kbp/cM. This value is within the range reported for other legumes, for example 758 kb/cM for *Phaseolus vulgaris* (Nodari et al. 1993); 517 kbp/cM for *Glycine max* (Diers et al. 1992); and 1500 kbp/cM in *Medicago sativa* (Kiss et al. 1993) but is at the upper end of the range reported for angiosperm trees; 350 kbp/cM for *Eucalyptus nitens* (Byrne et al. 1995) and 400 to 600 kbp/cM for *Eucalyptus grandis* and *Eucalyptus urophylla* (Grattapaglia and Sederoff 1994; Verhaegen and Plomion 1996). Higher average physical distance estimates reduce the possibility of the map-based cloning of genes.

One source of error in these calculations is the reported variability in the nuclear DNA concentrations in acacias. Not only is there significant variation in the amount of nuclear DNA in different acacia species, associated with their evolution and divergence (Mukherjee and Sharma 1995), but values also differ among authors. 4 C values reported for diploid Australian acacias range from 4.4 pg in *Acacia auriculiformis* to 8.4 pg in *Acacia simsii* (Mukherjee and Sharma (1995)). The values often differed from those reported by Bennett et al. (1998) which ranged from 4 C=2.2 pg for *Acacia modesta*, 4 C=3.3 pg for *A. auriculiformis* to 8.2 pg in *Acacia victoriae*.

Applications

The integrated map is based on four parents from the New Guinea region which is the preferred source of breeding material in current tree improvement programs for *A. mangium*. It will therefore provide a sound basis on which to carry out molecular breeding of this species. It also provides a useful reference map for comparative studies of genome organisation in other acacias.

The integrated map consolidated the linkage groups from two unrelated pedigrees. The advantages of mapping multiple pedigrees include: a larger number of loci are mapped, gene order and map distances are estimated more accurately, and alterations in these values, possibly due to chromosomal rearrangements affecting one of the parents, are easily detected (Kianian and Quiros 1992). In addition, polymorphic loci common to more than one pedigree, which may have a higher chance of segregating in other crosses, are identified. Use of multiple pedigrees increases the power to detect QTLs (Muranty 1996) and the map can be used to compare QTLs identified in different genetic backgrounds (Beavis and Grant 1991).

Knowledge of linkage relationships between loci provides a baseline for selecting markers spanning the genome for use in population studies. Mapped loci at different levels of linkage are being used to develop new statistical approaches to assess the demographic history of a species. For example, the pattern of variation across unlinked microsatellite loci has been used to test whether population size has been constant or increasing (Goldstein et al. 1999). Variation among closely linked microsatellite loci has been employed to measure gene flow from introduced populations to native populations (Estoup et al. 1999).

The linkage relationships in this map will be used to locate genes affecting quantitative traits of economic importance in *A. mangium*. Replicated clonal trials have been established from the mapping populations and the quantitative data will be employed to locate markers linked to genes controlling resistance to the phyllode rust *Atelocaula digitata*. This disease is emerging as a potentially serious problem in *A. mangium* plantations and nurseries in South-east Asia (Old et al. 1997).

Acknowledgements The assistance of Juliet Ward with DNA extractions and raising plants in the glasshouse; Robyn Bell, Jane Harbard and Peter Pomroy (Qld Forest Research Institute) with cross-pollinations; Christine Hackett from the Scottish Crops Research Institute; Colin Matheson and Charlie Bell for advice on statistical analysis; Curt Brubaker and Mike Devey for reviewing the manuscript are all gratefully acknowledged. Financial support was provided by CSIRO's Australian Tree Seed Centre.

References

- Aguade M, Langley CH (1994) Polymorphism and divergence in regions of low recombination in *Drosophila*. In: Golding B (ed) Non-neutral evolution. Theories and molecular data. Chapman and Hall, New York, pp 67–77
- Barreneche T, Bodenec C, Lexter C, Trontin J-F, Fluch S, Streiff R, Plomion C, Roussel G, Steinkellner H, Burg K, Favre J-M, Glossl J, Kremer A (1998) A genetic linkage map of *Quercus robur* L. (pedunculate oak) based on RAPD, SCAR, microsatellite, minisatellite, isozyme and 5S rDNA markers. *Theor Appl Genet* 97:1090–1103
- Beavis WD, Grant D (1991) A linkage map based on information from four F₂ populations of maize (*Zea mays* L.). *Theor Appl Genet* 82:636–644
- Begun DJ, Aquadro CF (1992) Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. *Nature* 356:519–520
- Bennett MD, Smith JB (1976) Nuclear DNA amounts in angiosperms. *Phil Trans Roy S Lond B* 274:227–274
- Bennett MD, Cox AV, Leitch IJ (1998) Angiosperm DNA C-values database. <http://www.rbgekew.org.uk/cval/database1.html>
- Bishop DT, Canning C, Skolnick M, Williamson JA (1983) The number of polymorphic DNA clones required to map the human genome. In: Wier BS (ed) Statistical analysis of DNA sequence data. Marcel-Dekker, New York, pp 181–200
- Bradshaw HD, Villar M, Watson BD, Otto KG, Stewart S, Stettler RF (1994) Molecular genetics of growth and development in *Populus*. III. A genetic linkage map of a hybrid poplar composed of RFLP, STS and RAPD markers. *Theor Appl Genet* 89:167–178
- 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
- Burt A, Bell G, Harvey PH (1991) Sex differences in recombination. *J Evol Biol* 4:259–277
- 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 (2000a) 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
- Butcher PA, Decroocq S, Gray Y, Moran GF (2000b). Development, inheritance and cross-species amplification of microsatellite markers from *Acacia mangium*. *Theor Appl Genet* (in press)
- Buttrose MS, Grant WJR, Sedgley M (1981) Floral development in *Acacia pycnantha* Benth. in Hook. *Aust J Bot* 29:385–395
- Byrne M, Murrell JC, Allen B, Moran GF (1995) An integrated genetic linkage map for eucalypts using RFLP, RAPD and isozyme markers. *Theor Appl Genetics* 91:869–875
- Chakravarti A, Lasher LK, Reefer JE (1991) A maximum likelihood method for estimating genome length using genetic linkage data. *Genetics* 128:175–182
- Darvasi A, Weinreb A, Minke V, Weller JI, Soller M (1993) Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. *Genetics* 134:943–951
- 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
- Devey ME, Sewell MM, Uren TL, Neale DB (1999) Comparative mapping in loblolly and radiata pine using RFLP and microsatellite markers. *Theor Appl Genet* 99:656–662
- Diers BW, Keim P, Fehr WR, Shoemaker RC (1992) RFLP analysis of soybean seed protein and oil content. *Theor Appl Genet* 83:13–26
- Doran JC, Turnbull JW (eds) (1997) Australian trees and shrubs: species for land rehabilitation and farm planting in the tropics. ACIAR, Canberra, pp 178–181
- Dvorak J, Luo M-C, Yang Z-L (1998) Restriction fragment length polymorphism and divergence in the genomic regions of high and low recombination in self-fertilizing and cross-fertilizing *Aegilops* species. *Genetics* 148:423–434
- Echt CS, Nelson CD (1997) Linkage mapping and genome length in eastern white pine (*Pinus strobus* L.). *Theor Appl Genet* 94:1031–1037
- Estoup A, Cornuet J-M, Rousset F, Guyomard R (1999) Juxtaposed microsatellite systems as diagnostic markers for admixture: theoretical aspects. *Mol Biol Evol* 16:898–908
- Fatmi A, Poneleit CG, Pfeiffer TW (1993) Variability of recombination frequencies in the Iowa Stiff Stalk synthetic (*Zea mays* L.). *Theor Appl Genet* 86:859–866
- Freyre R, Skroch PW, Geffroy V, Adam-Blondon A-F, Shirmohamadali A, Johnson WC, Llaca V, Nodari RO, Pereira PA, Tsai S-M, Tohme J, Dron M, Nienhuis J, Vallejos CE, Gepts P (1998) Towards an integrated linkage map of common bean. 4. Development of a core linkage map and alignment of RFLP maps. *Theor Appl Genet* 97:847–856
- Gentzittel L, Vear F, Zhang Y-Z, Berville A, Nicolas P (1995) Development of a consensus linkage RFLP map of cultivated sunflower (*Helianthus annuus* L.). *Theor Appl Genet* 90:1079–1086
- Gerber S, Rodolphe F (1994) An estimation of the genome length of maritime pine (*Pinus pinaster* Ait.). *Theor Appl Genet* 88:289–292
- Goldstein DB, Roemer GW, Smith DA, Reich DE, Bergman A, Wayne RK (1999) The use of microsatellite variation to infer population structure and demographic history in a natural model system. *Genetics* 151:797–801
- Grattapaglia D, Sederoff R (1994) Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-test-cross: mapping strategy and RAPD markers. *Genetics* 137:1121–1137
- Griffin AR, Cotterill PP (1988) Genetic variation in growth of out-crossed, selfed and open-pollinated progenies of *Eucalyptus regnans* and some implications for breeding strategy. *Silvae Genet* 37:124–131

- Groover AT, Williams CG, Devey ME, Lee JM, Neale DB (1995) Sex-related differences in meiotic recombination frequency in *Pinus taeda*. *J Hered* 86:157–158
- 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
- Hulbert SH, Ilott TW, Legg EJ, Lincoln SE, Lander ES, Michelmore RW (1988) Genetic analysis of the fungus, *Bremia lactucae*, using restriction fragment length polymorphisms. *Genetics* 120:947–958
- Jermstad KD, Bassoni DL, Wheeler NC, Neale DB (1998) A sex-averaged genetic linkage map in coastal Douglas-fir [*Pseudotsuga menziesii* (Mirb.) Franco var 'menziesii'] based on RFLP and RAPD markers. *Theor Appl Genet* 97:762–770
- Kenrick J, Knox RB (1979) Pollen development and cytology in some Australian species of *Acacia*. *Aust J Bot* 27:413–427
- Kenrick J, Knox RB (1982) Function of the polyad in reproduction of *Acacia*. *Ann Bot* 50:721–727
- Kianian SF, Quiros CF (1992) Generation of a *Brassica oleracea* composite RFLP map: linkage arrangements among various populations and evolutionary implications. *Theor Appl Genet* 84:544–554
- Kiss GB, Csanádi G, Kálmán K, Kaló P, Ökrész L (1993) Construction of a basic genetic map for alfalfa using RFLP, RAPD, isozyme and morphological markers. *Mol Gen Genet* 238:129–137
- Kosambi DD (1944) The estimation of map distance from recombination values. *Ann Eugen* 12:172–175
- Kraft T, Säll T, Magnusson-Rading I, Nilsson N-O, Halldén C (1998) Positive correlation between recombination rates and levels of genetic variation in natural populations of sea beet (*Beta vulgaris* subsp. *maritima*). *Genetics* 150:1239–1244
- Lui Z, Furnier GR (1993) Inheritance and linkage of allozymes and restriction fragment length polymorphisms in Trembling Aspen. *J Hered* 84:419–424
- Maliepaard C, Jansen J, Van Ooijen JW (1997) Linkage analysis in a full-sib family of an outbreeding plant species: overview and consequences for applications. *Genet Res* 70:237–250
- Matise TC, Perlin M, Chakravarti A (1994) Automated construction of genetic linkage maps using an expert system (MultiMap): a human genome linkage map. *Nature Genet* 6:384–90
- McGranahan M, Bell JC, Moran GF, Slee M (1997) High genetic divergence between geographic regions in the highly outcrossing species *Acacia aulacocarpa* (Cunn. ex Benth.). *For Genet* 4:1–13
- Moran GF, Bell JC, Hilliker J (1983) Greater meiotic recombination in male vs female gametes in *Pinus radiata*. *J Hered* 74:62
- Moran GF, Muona O, Bell JC (1989) Breeding systems and genetic diversity in *Acacia auriculiformis* and *A. crassicaarpa*. *Biotropics* 21:250–256
- Muona O, Moran GF, Bell JC (1991) Hierarchical patterns of correlated mating in *Acacia melanoxylon*. *Genetics* 127:6–9–626
- Mukai Y, Suyama Y, Tsumura Y, Kawahara T, Yoshimaru H, Kondo T, Tomaru N, Kuramoto N, Murai M (1995) A linkage map for sugi (*Cryptomeria japonica*) based on RFLP RAPD and isozyme loci. *Theor Appl Genet* 90:835–840
- Mukherjee W, Sharma AK (1995) In situ nuclear DNA variation in Australian species of *Acacia*. *Cytobios* 83:59–64
- Muranty H (1996) Power of tests for quantitative trait loci detection using full-sib families in different schemes. *Heredity* 76:156–165
- Nelson CD, Kubisiak TL, Stine M, Nance WL (1994) A genetic linkage map of longleaf pine (*Pinus palustris* Mill.) based on random amplified polymorphic DNAs. *J Hered* 85:433–439
- Newman IV (1933) Studies in the Australian Acacias. II. The life history of *Acacia baileyana* (F.v.M.). Part I. Some ecological and vegetative features, spore production and chromosome number. *J Lin. Soc Lond Bot* 49:145–171
- Nodari RO, Tsai SM, Gilbertson RL, Gepts P (1993) Towards an integrated map of common bean. 2. Development of an RFLP-based linkage map. *Theor Appl Genet* 85:513–520
- Ooijen JW van, Sandbrink JM, Vrieling M, Verkerk R, Zabel P, Lindhout P (1994) An RFLP linkage map of *Lycopersicon peruvianum*. *Theor Appl Genet* 89:1007–1013
- Old KM, Hood IA, Yuan ZQ (1997) Diseases of tropical acacias in northern Queensland. In: Old KM, Lee SS, Sharma JK (eds) Diseases of tropical acacias. Proc Int Wkshp, Subanjeriji, South Sumatra, 28 April–3 May 1996. CIFOR Special Publication, pp 1–22
- Paglia GP, Olivieri AM, Morgante M (1998) Towards second-generation STS (sequence-tagged sites) linkage maps in conifers: a genetic map of Norway spruce (*Picea abies* K.). *Mol Gen Genet* 258:466–478
- Paterson AH (1996) Making genetic maps. In: Paterson AH (ed) Genome mapping in plants. Academic Press, San Diego, pp 23–40
- Payne RW, Lane PW, Ainsley AE, Bicknell KE, Digby PGN, Harding SA, Leech PK, Simpson HR, Todd AD, Verrier PJ, White RP (1987) Genstat 5 Reference Manual. Clarendon, Oxford
- Pfeiffer TW, Vogt (1990) Variability for recombination frequencies in the AP12 soybean population. *Crop Sci* 30:545–549
- Plomion C, O'Malley DM (1996) Recombination rate differences for pollen parents and seed parents in *Pinus pinaster*. *Heredity* 77:341–350
- Qi X, Stam P, Lindhout P (1996) Comparison and integration of four barley genetic maps. *Genome* 39:379–394
- Sedgley M, Harbard J, Smith R-MM, Wickneswari R, Griffin AR (1992) Reproductive biology and interspecific hybridisation of *Acacia mangium* and *Acacia auriculiformis* A. Cunn. ex Benth. (Leguminosae: Mimosoideae). *Aust J Bot* 40:37–48
- Sewell MM, Sherman BK, Neale DB (1999) A consensus map for loblolly pine (*Pinus taeda* L.). I. Construction and integration of individual linkage maps from two outbred three-generation pedigrees. *Genetics* 151:321–330
- 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
- Stam P, van Ooijen JW (1995) JOINMAP version 2.0: software for the calculation of genetic linkage maps. CPRO-DLO, Wageningen
- Verhaegen D, Plomion C (1996) Genetic mapping in *Eucalyptus urophylla* and *Eucalyptus grandis* using RAPD markers. *Genome* 39:1051–1061
- Vicente MC de, Tanksley SD (1991) Genome-wide reduction in recombination of backcross progeny derived from male versus female gametes in an interspecific cross of tomato. *Theor Appl Genet* 83:173–178
- Williams CG, Savolainen O (1996) Inbreeding depression in conifers: implications for breeding strategy. *For Sci* 42:102–117
- Young ND, Weeden NF, Kochert G (1996) Genome mapping in legumes (Fam. Fabaceae). In: Paterson AH (ed) Genome mapping in plants. Academic Press, San Diego, pp 211–228