

Pectin Methylesterase Genes Influence Solid Wood Properties of *Eucalyptus pilularis*^{1[C][W][OA]}

Timothy R. Sexton^{2*}, Robert J. Henry, Chris E. Harwood, Dane S. Thomas, Luke J. McManus, Carolyn Raymond, Michael Henson, and Mervyn Shepherd

Cooperative Research Centre for Forestry, Hobart, Tasmania 7001, Australia (T.R.S., R.J.H., C.E.H., L.J.M., M.S.); Southern Cross Plant Science, Southern Cross University, Lismore, New South Wales 2480, Australia (T.R.S., R.J.H., C.R., M.S.); Queensland Alliance for Agriculture and Food Innovation, University of Queensland, Brisbane, Queensland 4072, Australia (R.J.H.); CSIRO Ecosystem Sciences, Hobart, Tasmania 7001, Australia (C.E.H.); Forests New South Wales, Land Management and Forestry Services, Coffs Harbour, New South Wales 2450, Australia (D.S.T., M.H.); South Australian Research and Development Institute, Waite Research Precinct, Urrbrae, South Australia 5064, Australia (D.S.T.); Forest Molecular Biology and Genetics Group, University of Melbourne, Creswick, Victoria 3363, Australia (L.J.M.); and Eucalyptus.biz Limited, Rosemead, Ripley, Surrey GU236DX, United Kingdom (M.H.)

This association study of *Eucalyptus pilularis* populations provides empirical evidence for the role of *Pectin Methylesterase* (*PME*) in influencing solid wood characteristics of *Eucalyptus*. *PME6* was primarily associated with the shrinkage and collapse of drying timber, which are phenotypic traits consistent with the role of pectin as a hydrophilic polysaccharide. *PME7* was primarily associated with cellulose and pulp yield traits and had an inverse correlation with lignin content. Selection of specific alleles in these genes may be important for improving trees as sources of high-quality wood products. A heterozygote advantage was postulated for the *PME7* loci and, in combination with haplotype blocks, may explain the absence of a homozygous class at all single-nucleotide polymorphisms investigated in this gene.

Wood is primarily composed of lignin, cellulose, and noncellulosic polysaccharides (including neutral and pectic polysaccharides), all of which interact to form a complex structure that provides the foundation for the massive stature unique to the woody life form known as “trees.” Wood properties are highly variable within and between tree species, a necessary attribute for adaptation to different environments (Carlquist, 1977). Quantitative genetic studies have revealed that wood properties are under strong additive genetic control in representative species across the *Eucalyptus* genus (Raymond et al., 2008; Stackpole et al., 2010a, 2010b). Because wood phenotypes are complex traits, it is expected that they will be controlled by many

genes, each with a small effect on the phenotype (Neale and Savolainen, 2004). Despite a basic knowledge of the biochemical pathways involved in lignin and carbohydrate (cellulose and noncellulosic polysaccharide) biosynthesis (Hertzberg et al., 2001), the relationship between physical wood properties and the activity of genes remains poorly understood.

Association studies are a means to bridge the understanding between phenotypic traits and underlying variation in the DNA sequence at specific candidate genes or dispersed genome wide. Differing from quantitative trait locus mapping, association studies use diverse individuals from natural populations, where generations of recombination have resulted in short stretches of linkage disequilibrium (LD; Butcher and Southerton, 2007). The rapid decay of LD in natural populations of forest trees means that genotype-phenotype associations are identified with the causal quantitative trait nucleotide (QTN) or sites in close proximity to the QTN. Therefore, markers developed by association studies should be informative for selection across the full diversity represented in tree breeding populations (Butcher and Southerton, 2007; Grattapaglia, 2007). The rapid decay of LD in forest trees also means that a high marker density would be needed for genome-wide association testing; for this reason, a candidate gene approach has been preferred (Neale and Kremer, 2011). Candidate gene-based association studies restrict analysis to regions of the

¹ This work was supported by the Cooperative Research Centre for Forestry (funding for the molecular and association studies).

² Present address: Department of Forest Sciences, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z4.

* Corresponding author; e-mail sexton.timothy@gmail.com.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Timothy R. Sexton (sexton.timothy@gmail.com).

^[C] Some figures in this article are displayed in color online but in black and white in the print edition.

^[W] The online version of this article contains Web-only data.

^[OA] Open Access articles can be viewed online without a subscription.

www.plantphysiol.org/cgi/doi/10.1104/pp.111.181602

genome most likely to influence target phenotypes. Good candidates for wood properties include genes involved in lignin and carbohydrate biosynthesis (Hertzberg et al., 2001). The preferred DNA markers for association studies are single-nucleotide polymorphisms (SNPs), which are the most abundant source of variation in the genome and underlie much of the morphological variation observed between individuals (Rafalski, 2002).

Candidate gene-based association studies in forest trees have demonstrated a number of correlations

between SNPs and phenotypic traits, including physical wood properties (Thumma et al., 2005, 2009; González-Martínez et al., 2007; Dillon et al., 2010; Sexton et al., 2010b; Beaulieu et al., 2011), resistance genes (Neale, 2007; Külheim et al., 2011), drought tolerance (González-Martínez et al., 2008), cold hardiness (Eckert et al., 2009), and timing of bud set (Ingvarsson et al., 2008; Holliday et al., 2010).

This experiment targets two genes of the *Pectin Methylesterase* (*PME*) gene family (EC 3.1.1.11) that catalyze the demethylesterification of the most abun-

Table 1. Principal components analysis illustrating the high degree of correlation among the 40 wood properties used in the association testing

The top three principal components (represented) explained 46% of the total variance in all wood properties analyzed. Loading values greater than 0.6 indicate strong correlations within components and are presented in boldface underlined numbers to assist with interpretation; strong negative correlations within component 3 are presented in italic underlined numbers.

Full Name	Abbreviated Name	Component		
		1	2	3
Lignin percentage	LIG	-0.49	-0.14	<u>0.64</u>
Lignin syringyl and guaiacyl ratio	S/G ratio	0.16	-0.03	-0.42
Cellulose percentage	CEL	<u>0.60</u>	0.12	<i>-0.72</i>
Screened pulp yield	SPY	0.41	0.10	<i>-0.77</i>
Total pulp yield	TPY	0.40	0.10	<i>-0.72</i>
Glucose percentage	GLC	<u>0.74</u>	0.05	<i>-0.52</i>
Arabinose percentage	ARA	0.34	-0.24	0.55
Galactose percentage	GAL	0.34	0.02	-0.14
Mannose percentage	MAN	-0.59	0.08	0.42
Xylose percentage	XYL	-0.35	0.11	0.34
Average air-dry density	ADDenkgm ³	<u>0.71</u>	-0.15	0.50
Core basic density	CBD	0.34	-0.24	0.42
Disc bulk density, outer wood (kg m ⁻³)	DiscBDa	0.20	-0.25	0.23
Disc bulk density, intermediate wood (kg m ⁻³)	DiscBDb	0.39	-0.43	0.43
Disc bulk density, inner wood (kg m ⁻³)	DiscBDc	0.59	-0.41	0.50
Director velocity squared (km ² s ⁻²)	DirVel ²	<u>0.62</u>	-0.25	-0.21
FAKOPP velocity squared (km ² s ⁻²)	FAKOPP ²	<u>0.71</u>	-0.30	-0.06
Modulus of elasticity (MPa)	MOE	<u>0.81</u>	-0.23	0.12
Modulus of rupture (MPa)	MOR	<u>0.69</u>	-0.32	0.31
Average hardness (kN)	AvgHard	0.52	-0.35	0.46
Block average radial shrinkage at 17% moisture	BR17	0.26	0.59	0.29
Block average radial shrinkage at 12% moisture	BR12	0.37	0.46	0.29
Block average radial shrinkage at 5% moisture	BR5	0.32	0.46	0.16
Core radial scans at 12% moisture	CR12scans	0.21	0.34	0.01
Block average tangential shrinkage at 17% moisture	BT17	0.13	<u>0.82</u>	0.21
Block average tangential shrinkage at 12% moisture	BT12	0.16	<u>0.79</u>	0.19
Block average tangential shrinkage at 5% moisture	BT5	0.15	<u>0.81</u>	0.19
Core average tangential shrinkage at 17% moisture	CT17	0.29	<u>0.77</u>	0.15
Core average tangential shrinkage at 12% moisture	CT12	0.33	<u>0.76</u>	0.13
Core average tangential shrinkage at 5% moisture	CT5	0.37	<u>0.76</u>	0.14
Block tangential-radial shrinkage ratio	T/R block	-0.27	0.12	-0.13
Core tangential-radial shrinkage ratio	T/R core	-0.14	-0.04	0.08
Average collapse grade	AvColGr	-0.08	0.58	0.24
Average maximum collapse	AvMaxcol	-0.01	0.46	0.24
Diameter at breast height at 104 months	DBH104	0.19	-0.20	0.43
Height at 104 months	HT104	0.37	-0.08	0.25
Spiral grain	Spiral	-0.02	-0.03	0.03
Interlocking degree	Interlock	-0.05	-0.08	0.09
Average pith	AvPith	-0.15	-0.01	-0.07
Average heartwood	AvHw	0.02	-0.19	0.30
Percentage of variance		17%	-16%	13%
Cumulative percentage		17%	32%	46%

dant class of pectinaceous polysaccharides, homogalacturonan (HG; Wolf et al., 2009). The degree and pattern of methylesterification of HG have strong influences on the water-holding capacity of pectin gels (Willats et al., 2001).

In plant genomes, *PMEs* are a large gene family where paralogs have various biological roles, including wood formation, vegetative and reproductive processes, as well as plant-pathogen interactions (Pelloux et al., 2007). In wood formation, *PME* genes play roles in cell wall formation, stiffening, and remodeling (Moustakas et al., 1991; Al-Qsous et al., 2004; Wolf et al., 2009). Up- and down-regulation of *PME* in transgenic hybrid *Populus tremula* × *Populus tremuloides* (poplar) suggest an important role in wood formation, where it is localized in expanding wood cells and is involved in mechanisms determining cell/fiber width and length (Siedlecka et al., 2008).

The candidate genes investigated (*PME6* and *PME7*) were originally identified in southern blue gum (*Eucalyptus globulus*) where both genes are expressed in the intermediate and basal stem tissues, as well as the phloem and xylem (Goulao et al., 2011). *PME6* and *PME7* were also noted as good candidates for association studies because they are influenced by balancing selection with a high number of transsubgeneric SNPs shared between the *Monocalyptus* and *Symphyomyrtus* subgenera as well as excessively high SNP densities. It was proposed that balancing selection in these genes would be reflected by observed differences in the quantitative wood properties between *Eucalyptus* trees (T.R. Sexton, unpublished data).

The model species for this association study is *Eucalyptus pilularis* (blackbutt), which is an important hardwood species endemic to Australia. Timber from natural forests and plantations of *E. pilularis* is used for building construction, flooring, plywood, poles, railway sleepers, veneer, and pulp (Boland et al., 2006). Association testing demonstrated that SNPs in *PME6* and *PME7* were associated with variation in economically important wood properties of *E. pilularis*.

RESULTS

Relationships among wood properties were examined using principal components analysis (Table I). As expected, all traits were highly correlated within three components; for example, cellulose, Glc, and density had high loading values (greater than 0.6) in component 1. In this analysis of 40 wood phenotypes, three components accounted for 46% of the cumulative variation and were termed (1) cellulose and pulp yield traits, (2) shrinkage and collapse traits, and (3) lignin traits. The correlations within components 1 to 3 were consistent with published accounts for other *Eucalyptus* species; for example, acoustic wave velocity and pulp yield are positively correlated in both *Eucalyptus dunnii* and *E. pilularis* (Raymond et al., 2008, 2010).

Considering the full population data set ($n = 561$) and correction for kinship, a total of 29 associations were identified to be significant at the individual test level (Supplemental Table S1). Of these, 10 were significant when P values were adjusted for multiple comparisons (Table II). Of the eight SNPs examined, five were significantly associated with at least one trait, and two of these, *PME7* position 391 and *PME6* position 204, had significant associations with multiple wood properties correlated in principal component 1, 2, or 3.

PME7 Is Primarily Associated with Lignin and Cellulose Traits

The SNP at position 391 of *PME7* was significantly associated with lignin, cellulose, and pulp yield (Table II; Fig. 1; Supplemental Table S1), traits correlated in principal components 1 and 2. Oddly, only two genotypes were observed (G/G and C/G). It seems unusual that the homozygous C/C genotype was not observed, despite large numbers of individuals with the alternate genotypes; consequently, there was a significant departure from Hardy-Weinberg equilibrium (HWE) at this locus (Table III). The C/G heterozygotes had a significantly higher percentage of

Table II. Significant associations reported in the full population of 561 *E. pilularis* trees between wood traits and SNPs in *PME6* or *PME7*

Test values reported were derived from a general linear model using 1,000 permutations and were significant following correction for multiple comparisons (P -adj). The P -perm value indicates significance at the permutation normalized SNP-wise level, while the P -adj value indicates significance at the experiment-wise level (corrected for multiple comparisons). NA, Not applicable.

Trait	SNP	P -perm	P -adj	r^2	Population Mean	SE	Homozygous A Mean	SE	Heterozygous Mean	SE	Homozygous B Mean	SE
BT17	<i>PME6_204</i>	0.004	0.006	0.0227	7.92	±0.14	7.78 ^a	±0.14	9.35	±0.52	9.34	±0.81
CT17	<i>PME6_204</i>	0.005	0.015	0.0166	7.95	±0.10	7.86 ^a	±0.10	8.99	±0.36	10.21	±0.76
CT5	<i>PME6_204</i>	0.002	0.02	0.0141	11.15	±0.09	11.07 ^a	±0.09	12.01	±0.31	13.21	±0.39
T/Rcore	<i>PME6_268</i>	0.006	9.99E-04	0.0259	3.00	±0.13	2.92 ^a	±0.11	5.65	±2.34	NA	NA
AvMaxcol	<i>PME6_456</i>	0.01	0.05	0.0173	9.52	±0.16	9.42 ^a	±0.16	11.42	±0.74	NA	NA
CEL1	<i>PME7_391</i>	0.004	0.013	0.0162	42.22	±0.04	42.08	±0.06	42.33 ^b	±0.06	NA	NA
LIG1	<i>PME7_391</i>	0.002	0.004	0.0178	28.36	±0.04	28.50 ^a	±0.06	28.25 ^b	±0.05	NA	NA
SPY1	<i>PME7_391</i>	9.99E-04	0.002	0.0194	47.93	±0.08	47.66	±0.11	48.15 ^b	±0.10	NA	NA
TPY1	<i>PME7_391</i>	0.002	0.003	0.0181	49.47	±0.09	49.16	±0.13	49.72 ^b	±0.12	NA	NA
MOR	<i>PME7_441</i>	0.007	0.04	0.0186	120.15	±0.78	119.61	±0.80	129.15 ^a	±2.61	NA	NA

^aSelect genotype for improved solid wood properties.

^bSelect genotype for improved pulp properties.

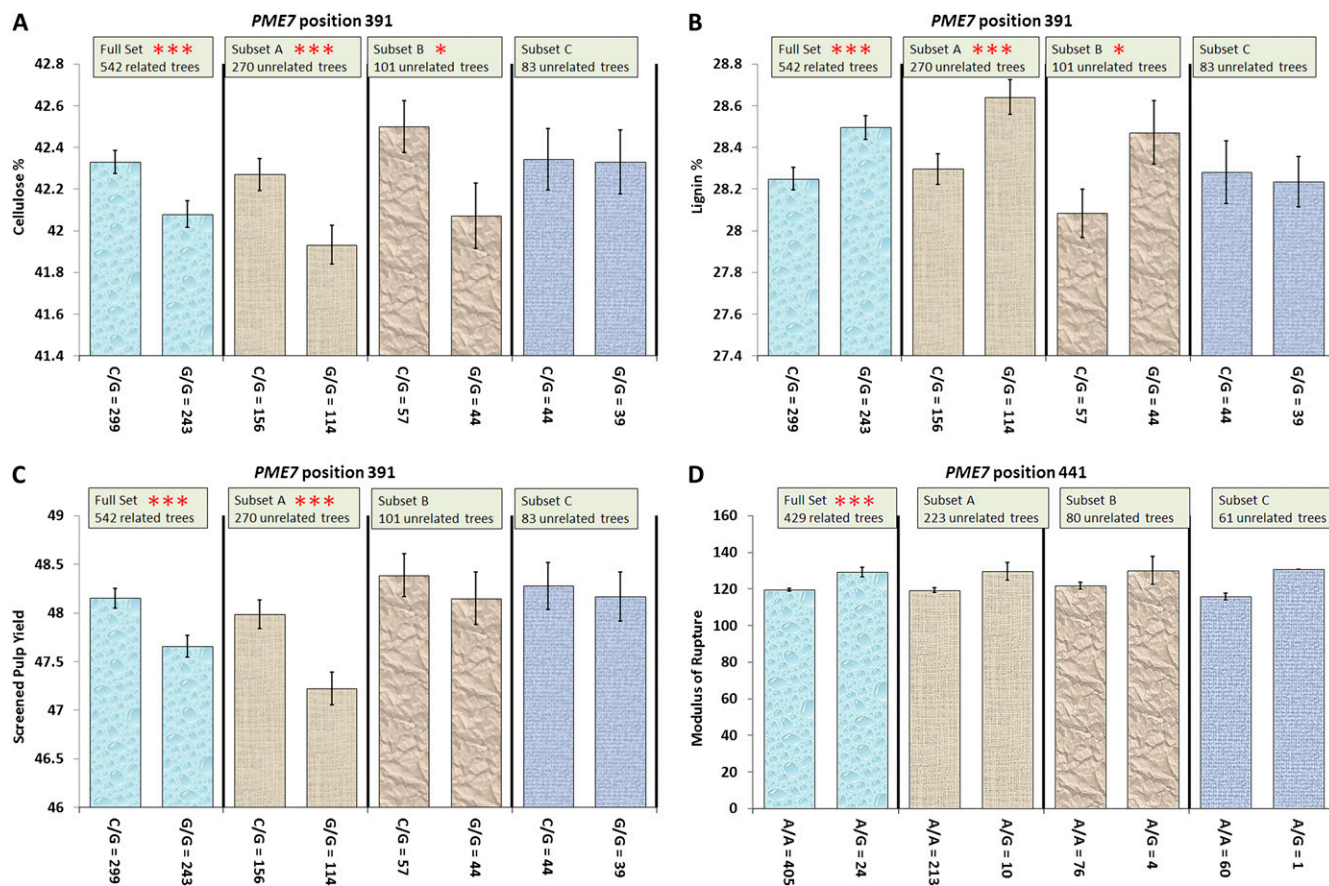


Figure 1. A to C, Associations between *PME7* position 391 and cellulose (A), lignin (B), and screened pulp yield (C). D, Association between *PME7* position 441 and MOR. Error bars represent SE , using $\alpha = 0.05$; three asterisks indicate significance following correction for multiple tests, and one asterisk indicates significance at the single-marker test level. Trees were segregated into genotype categories, and corresponding phenotypes were plotted. [See online article for color version of this figure.]

cellulose and pulp yield (Fig. 1; Table II). In contrast, the G/G homozygotes had significantly more lignin.

Another significant association was identified between a SNP in *PME7* 50 bp downstream at position 441 and the modulus of rupture (MOR; Table II; Fig. 1; Supplemental Table S1). MOR is the point at which wood fibers fail/break due to bending and is a trait positively correlated with cellulose, density, and acoustic velocity in principal component 1 (Table I). The alternate homozygote (G/G) was also missing, but, as indicated by the lower minor allele frequency (MAF), the G/G homozygote at position 441 was rare and not sampled. This was supported by no departure from HWE (Table III). Relative to the A/A homozygote, the heterozygous A/G genotype had a significantly higher MOR (Fig. 1; Table II).

PME6 Is Primarily Associated with Cellulose, Pulp Yield, and Shrinkage Properties

The SNP at position 204 of *PME6* was significantly associated with wood shrinkage during drying, par-

ticularly with tangential shrinkage in blocks and cores at 17% moisture content (Table II; Fig. 2; Supplemental Table S1). Shrinkage is a trait influencing wood defect and is correlated with collapse in principal component 2. Only two individuals were observed with a homozygous T/T genotype. These two individuals were excluded from the association testing because of the small sample size. However, their phenotypes were plotted with the other genotypes for visual comparison (Fig. 2). All genotypes were observed at ratios expected under HWE (Table III). Relative to the trees with the homozygous C/C genotype, trees with the heterozygous C/T genotype had significantly higher levels of shrinkage (Fig. 2).

Another significant association was identified between the SNP in *PME6* at position 268 and the ratio of tangential to radial shrinkage at 12% moisture content in cores (Table II; Fig. 2; Supplemental Table S1). Compared with trees with the homozygous T/T genotype, the heterozygous A/T trees had significantly higher levels of shrinkage (Fig. 2). The A/A homozygote was not sampled because the A allele was rare,

Table III. Genotypes, counts, and population parameters for 11 SNPs from *PME6* or *PME7* in *E. pilularis* ($n = 561$)

Significant departures from HWE ($P \leq 0.001$) are designated by asterisks. Columns are as follows: gene and SNP position; AA, amino acid change; genotypes and respective counts; ObsHET and PredHET, observed and expected heterozygosity; HWpval, P value for HWE test; %Geno, percentage data set completeness; MAF; and whether the *E. pilularis* (P) SNP is shared with *E. globulus* (G) or *E. pyrocarpa* (Y). SNPs from *PME7* at positions 30, 149, and 153 were excluded from the association testing.

SNP	AA	Genotypes	Counts	ObsHET	PredHET	HWpval	%Geno	MAF	Shared
<i>PME6</i> _204	–	TT, CT, CC	2, 41, 503	0.075	0.079	0.457	97.3	0.041	PY
<i>PME6</i> _268	S/T	TT, AT	545, 15	0.027	0.026	1	99.8	0.013	PG
<i>PME6</i> _456	–	CT, CC	26, 535	0.046	0.045	1	100	0.023	PY
<i>PME7</i> _166	K/E	CA, AA	505, 43	0.922	0.497	<0.001*	97.7	0.461	PY
<i>PME7</i> _186	–	CT, CC	504, 41	0.925	0.497	<0.001*	97.1	0.462	PY
<i>PME7</i> _373	–	AG, AA	511, 7	0.986	0.5	<0.001*	92.3	0.493	PYG
<i>PME7</i> _391	–	GG, GC	249, 304	0.55	0.399	<0.001*	98.6	0.275	PY
<i>PME7</i> _441	–	AG, AA	27, 421	0.06	0.058	1	79.9	0.03	PY
<i>PME7</i> _30	–	GG, GA	545, 3	0.005	0.005	1	97.7	0.003	PYG
<i>PME7</i> _149	T/K	CA	514	1	0.5	<0.001*	91.6	0.5	P
<i>PME7</i> _153	–	TT, CT	3, 513	0.994	0.5	<0.001*	92	0.497	PYG

with a MAF of 1.3%, and the genotypes showed no departure from HWE (Table III).

The last significant association was between the SNP at position 456 of *PME6* and average maximum collapse (Table II; Fig. 2; Supplemental Table S1). Collapse is also a trait causing wood defect and is also correlated with shrinkage in principal component 2. Relative to the homozygous C/C genotype, the rarer heterozygous C/T genotype was associated with significantly higher levels of collapse, a trend consistent in the full population and all three subsets (Fig. 2). The T/T homozygous genotype was not sampled because the T allele was rare, with a MAF of 2.3%, and this locus showed no departure from HWE (Table III).

DISCUSSION

Associations Demonstrate Expected Genotype-Phenotype Links

In this study, associations were identified between two *PME* genes and a range of solid wood properties in *E. pilularis*. Associations in *PME7* were primarily with cellulose, pulp yield, and lignin, traits correlated in principal components 1 and 3. In contrast, the associations with *PME6* were primarily with shrinkage and collapse, traits correlated in principal component 2. The associations between *PME6* and *PME7* with multiple wood traits are consistent with the current understanding of wood formation and the biochemical activity of *PME* (Hertzberg et al., 2001; Willats et al., 2001; Kirst et al., 2004; Goulao et al., 2011).

In transgenic poplar, up- and down-regulation of *PME* localized in expanding wood cells has an effect on cell/fiber width and length (Siedlecka et al., 2008). Longer and thicker fibers increase the proportion of cellulose. Consequently, pulp yield is also increased, along with acoustic wave velocity (Raymond et al., 2010). The association between *PME7* and cellulose might be explained due to competition for common precursor molecules, as the biosynthesis of pectin,

cellulose, and hemicellulose all occur in the carbohydrate pathway (Hertzberg et al., 2001). The inverse association between *PME7* and lignin is not as straightforward, given that lignin and carbohydrate biosynthesis occur in independent biochemical pathways; however, this inverse correlation may be explained by competition for carbon between the lignin and carbohydrate pathways (Kirst et al., 2004).

The associations between *PME6* and shrinkage may be attributed to the water-holding (binding) capacity of pectic polysaccharides. Pectins are highly hydrophilic polysaccharides and, when demethylated, form gels that bind large volumes of water. The most abundant class of pectinaceous polysaccharides is HG, which is demethylated by *PME* (Wolf et al., 2009). Differing degrees and patterns of methylation of HG are known to influence the water-holding capacity of pectic gels (Willats et al., 2001). Trees with more water bound in their cell walls would be expected to show higher levels of shrinkage, as their timber is dried.

In addition, pit membranes facilitate the movement of free water between cells and contain high concentrations of pectin. The relative abundance of acidic versus methylated pectins of pit membranes is known to differ between species with low versus high ionic effect (Gortan et al., 2011). It has been proposed that weakening, distortion, and pit aspiration as timber is dried can explain differences in the drying behavior (shrinkage and collapse) of sapwood in Norway spruce (*Picea abies*; Rosner et al., 2010).

Although six SNPs seemed to interact with wood properties in some way, the two most intriguing associations were with *PME7* position 391 and *PME6* position 204. Neither of these SNPs resulted in clear alterations to the theoretically translated protein sequence of these enzymes. Because of the rapid decay of LD in *Eucalyptus* (Thumma et al., 2005; Thavamanikumar et al., 2011), the QTN-causative polymorphisms are likely to be located in close proximity to the SNP positions where associations have been demonstrated. Previously, these genes were sequenced, and multiple nonsynonymous (amino

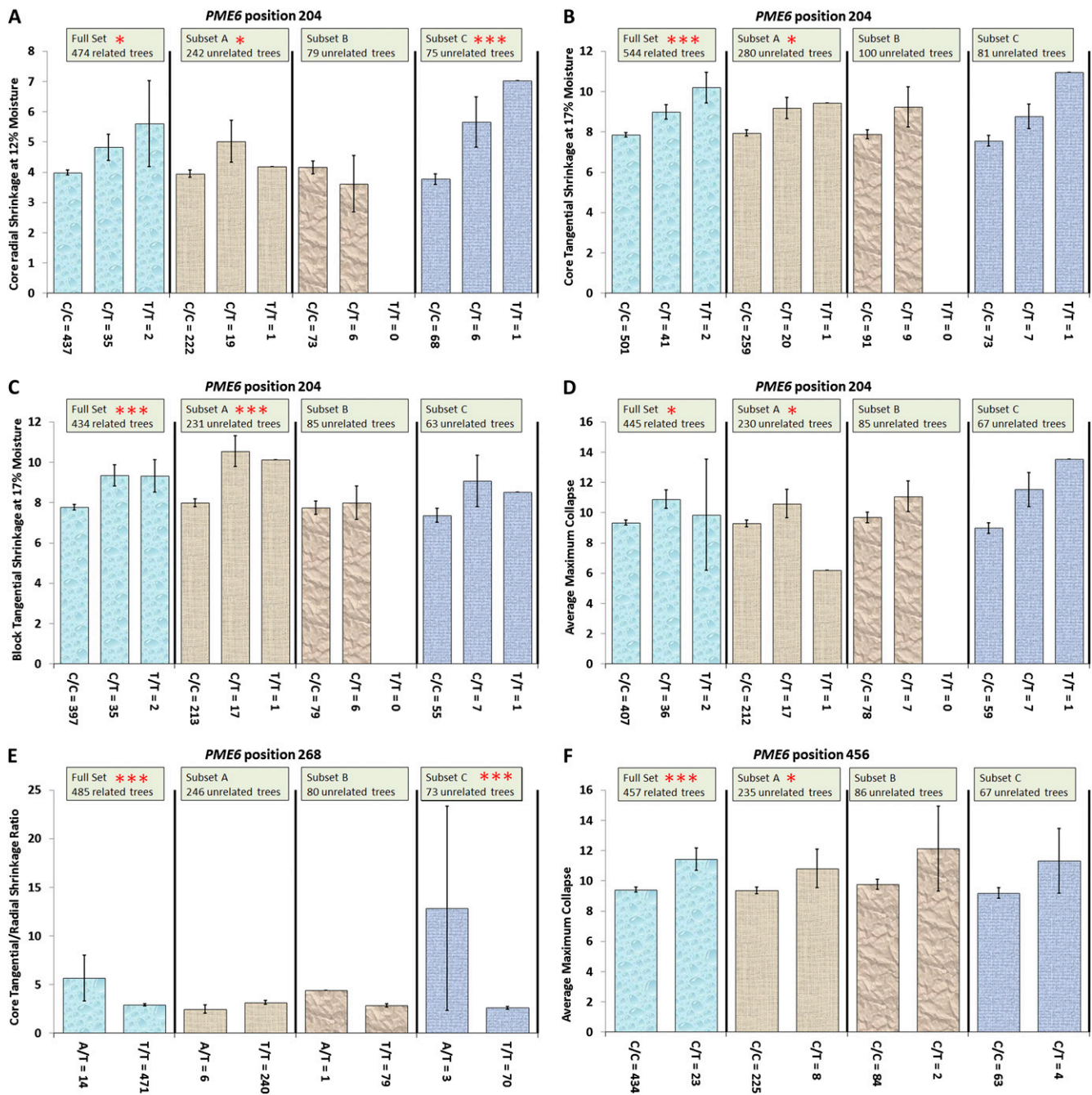


Figure 2. A to D, Associations between *PME6* position 204 and core radial scans at 12% moisture content (A), core tangential shrinkage at 17% moisture content (B), block tangential shrinkage at 17% moisture content (C), and average maximum collapse (D). The T/T genotypes were excluded from the association testing but included here for visual comparison. E, Association with *PME6* position 268 and core tangential-radial shrinkage ratio. F, Association with SNP *PME6* 456 and average maximum collapse. Error bars represent SE, using $\alpha = 0.05$; three asterisks indicate significance following correction for multiple tests, and one asterisk indicates significance at the single-marker test level. Trees were segregated into genotype categories, and corresponding phenotypes were plotted. [See online article for color version of this figure.]

acid-altering) SNPs were identified in close proximity to each locus where significant associations were identified (T.R. Sexton, unpublished data). Unfortunately, it was not possible to genotype these nonsynonymous SNPs with current technology, because excessively high nu-

cleotide diversity in these genes prevented the design of most SNP assays. This is a common problem encountered when genotyping in *Eucalyptus*, a genus noted for high heterozygosity and high nucleotide diversity (Sexton et al., 2010a; Grattapaglia et al., 2011). Sequenc-

ing haplotypes is an alternative approach capable of genotyping in highly heterozygous and polymorphic genotypes. Sequencing is an attractive alternative to traditional genotyping by primer extension, because SNP discovery and genotyping can be achieved in a single step, and this approach is now cost effective in light of rapid developments in second and third generation sequencing.

Strong Association between the SNP at Position 391 of *PME7* and Wood Quality

The SNP in *PME7* at position 391 was significantly associated with a number of traits that were highly correlated in the principal components: (1) cellulose and pulp yield traits, and (3) lignin traits (Tables I and II; Fig. 1). The homozygous G/G genotype was significantly associated with high lignin but also inversely correlated with lower cellulose and pulp yield (Fig. 1). The effect size for each of these associations was low ($r^2 < 0.019$), with phenotypic gains from the population mean to the G/G genotype mean being greater than 2 SE (Table II). The effect size was larger in the

smaller unrelated subsets, for example, associations with Glc and director (acoustic wave) velocity² in subset B ($n = 101$), and both had high r^2 values of 0.06 (Supplemental Table S1). This might be attributed to the Beavis effect, where phenotypic variances of small populations/samples tend to be overestimated (Beavis, 1998). Given that *E. pilularis* is a species primarily grown for solid wood products, increasing the frequency or fixing the G allele in the breeding population is desirable to increase lignin and density.

It is intriguing that the alternate homozygous genotype of *PME7* position 391 was not observed, despite the large sample size and high numbers of the alternate genotypes. *PME* orthologs do have pleiotropic functions, such as breaking seed dormancy (Ren and Kermod, 2000). Therefore, one explanation for the absence of the C/C genotype in our sample may be delayed seed germination, where slower germinating seeds would be inadvertently removed from the population in the nursery. Alternatively, the absence of the homozygous C/C genotype could be explained by a lethal recessive allele.

LD and the resulting haplotype blocks might explain why the SNPs sampled in *PME7* only have two

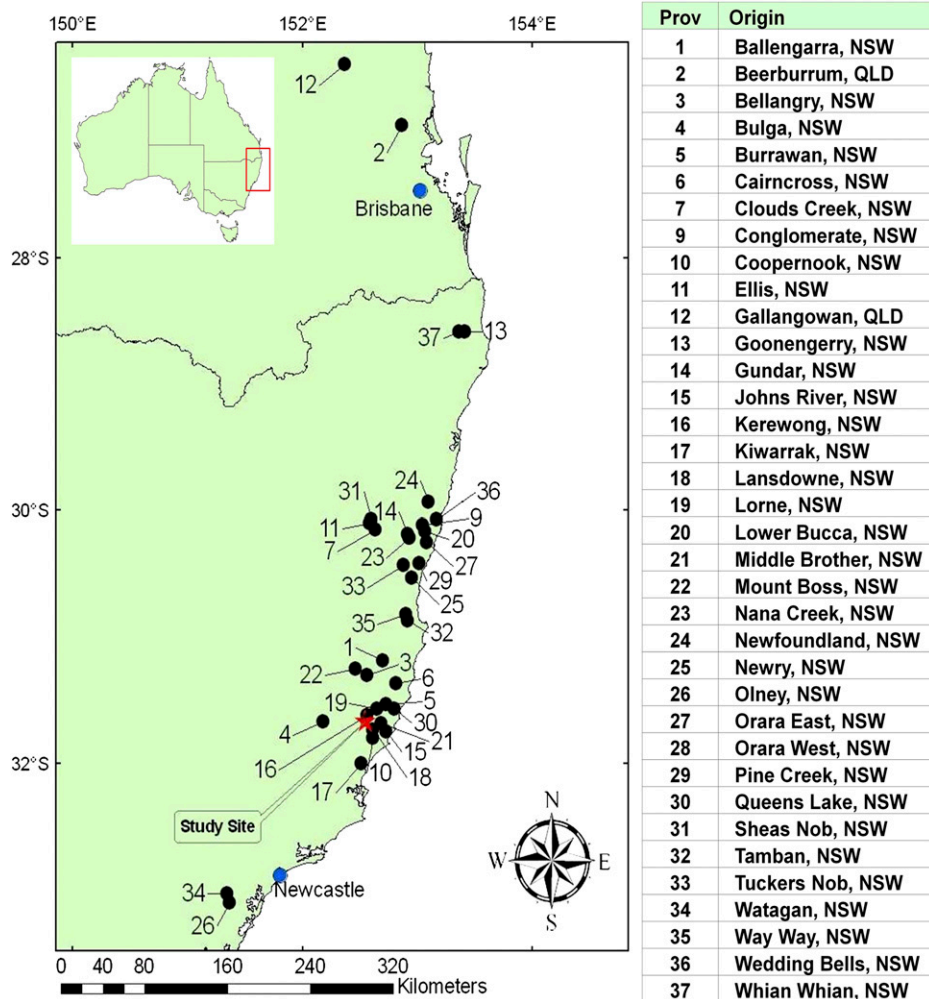


Figure 3. Geographical distribution of provenances represented in the association population planted within the study site at Hannam Vale, New South Wales (NSW), Australia. QLD, Queensland. [See online article for color version of this figure.]

genotype classes. Significant departures from HWE meant that haplotyping and evaluation of LD in *PME7* would be inappropriate (Table III). Forcing the analysis with all five SNPs in this gene did indicate strong LD (data not shown).

In *PME6*, detection of rare alleles better explained the lack of a homozygous genotype class, where each SNP in this gene had a low MAF and showed no significant departures from HWE (Table III). This sampling of three rare alleles in *PME6* was not unusual; many populations of forest trees have been found to have a high abundance of rare SNPs (with low MAF), and it is postulated that this is attributed to large population sizes and an out-crossing mating system (Eldridge et al., 1993; Novaes et al., 2008; Grattapaglia et al., 2011; T.R. Sexton, unpublished data).

Strong Association between the SNP at Position 204 of *PME6* and Shrinkage

The most highly represented associations in *PME6* were with a synonymous SNP at position 204 and a number of traits that were highly correlated into principal components: (1) cellulose and pulp yield traits as well as (2) shrinkage and collapse traits (Tables I and II; Fig. 2). The homozygous T/T genotype was only sampled in two trees and therefore excluded from the association tests. When the phenotypes were segregated into genotypic categories for visual analysis, the alternate C/C homozygote class showed the lowest levels of shrinkage (Fig. 2). The heterozygous genotype displayed an intermediate phenotype, suggesting that there was additive gene action at this locus. The effect size of each association at *PME6* position 204 was low ($r^2 < 0.023$), with predicted gains less than 1 SE above the population mean for the C/C genotype for most traits (Table II). Like associations with other SNPs, the effect size was larger in the smaller unrelated subsets; for example, associations with core radial shrinkage at 12% moisture content in subset C ($n = 84$) had a high r^2 value of 0.11 (Supplemental Table S1), although again, this may indicate overestimation due to the Beavis effect (Beavis, 1998). Given that *E. pilularis* is a species primarily grown for sawn wood products, fixing the C allele in the breeding population would be desirable to reduce the shrinkage of drying timber.

Balancing Selection in *PME* and Tree Breeding

Previous studies have suggested that both *PME6* and *PME7* are under balancing selection; therefore, these genes are likely to have adaptive functions that help the tree to cope with biotic and abiotic stresses (T.R. Sexton, unpublished data). The significant departures from HWE and the corresponding high abundance of heterozygous genotypes in *PME7* suggested that a heterozygous advantage may be the mechanism balancing selection at this locus. The mechanism of balancing selection in *PME6* is not as clear because the

SNPs genotyped are rare and showed no departures from HWE.

This study suggests that the high nucleotide diversity maintained by balancing selection in these *PME* genes (T.R. Sexton, unpublished data) is also associated with the structural composition of wood. Tree breeders should be mindful that these *PME* genes may be under balancing selection; therefore, reductions of diversity at these loci may leave a population more vulnerable to disease or abiotic stresses. This concept needs further investigation, but this paper has identified genes and specific alleles that could be used as models to initiate further studies.

Implications for Breeding and Deployment Systems

We have shown how SNPs in two *PME* genes influence many wood properties of *E. pilularis*. The variance (r^2) explained by these associations is low but typical for complex traits and all associations reported in other forest trees (Thumma et al., 2005, 2009; González-Martínez et al., 2007, 2008; Neale, 2007; Ingvarsson et al., 2008; Dillon et al., 2010; Sexton et al., 2010b). Selections based on individual SNPs are expected to provide only small gains (Grattapaglia and Resende, 2010). Nevertheless, small gains may be economically advantageous even if only representing a very minor shift from the population mean. As demonstrated in the livestock industry (Goddard and Hayes, 2009), investments in marker development can be rapidly recouped by phenotypic gains among the massive numbers of individuals descended from genotyped parents.

MATERIALS AND METHODS

SNP Discovery

SNPs were identified in a pooled sample of 30 trees representing the natural geographical range of *Eucalyptus pilularis*. Both *PME6* and *PME7* were amplified from this pooled DNA and sequenced on an Illumina GAIIX (T.R. Sexton, unpublished data).

Population and Phenotyping

We investigated 561 destructively sampled trees, which were descended from 284 open pollinated families and grown in a field trial at Hannam Vale, New South Wales, Australia (31°40' S, 152°33' E). These families were collected from well-separated mother trees from 37 provenances of *E. pilularis* (Fig. 3) and represented a large part of the natural geographical distribution of this species (Johnson and Stanton, 1993; Johnson and Nikles, 1997; Boland et al., 2006). Each of the 561 trees was destructively sampled at 112 months and had undergone extensive phenotypic assessment of quantitative traits, including basic density, wood stiffness, wood chemistry (cellulose, lignin, and hemicellulose), wood shrinkage, and collapse during drying (Pelletier et al., 2008; Raymond et al., 2008, 2009), as well as hardness, modulus of elasticity, and MOR (Henson and Smith, 2007).

Principal Components Analysis

Many of the wood traits were highly correlated, so principal components analysis was performed (using PASW statistics 18) in order to establish relationships between traits (Table I).

Sequenom Genotyping

SNPs previously characterized in *PME6* and *PME7* (T.R. Sexton, unpublished data) were genotyped across these same 561 individuals for which phenotypic data had been collected. The availability of high-resolution SNP discovery data (T.R. Sexton, unpublished data) permitted use of the standard Sequenom iPLEX Gold Assay for genotyping. Most SNPs identified in these genes by sequencing could not be genotyped by primer extension because of the high SNP density; ignoring this information would lead to mispriming and subsequent assay failure (Sexton et al., 2010a; Grattapaglia et al., 2011). Assays were designed based on previously ascertained *E. pilularis* sequences for *PME6* and *PME7* (T.R. Sexton, unpublished data). All SNP positions in these genes were coded in the International Union of Pure and Applied Chemistry ambiguity code, and candidate SNPs desired for genotyping were converted to the format required for the Sequenom assay design (3.1) software.

A total of 11 SNPs were successfully genotyped with more than 75% dataset completeness. Only eight SNPs were used for association testing, as three SNPs from *PME7* at positions 30, 149, and 153 were excluded because limited numbers of individuals were represented in the alternate genotype class or classes (Table III).

Association Testing

Five SNPs from *PME7* and three from *PME6* (for a total of eight SNPs) were used for association testing with 40 wood-quality traits. The sample size was nominally 561 related trees, but numbers in each test may be lower because of missing data for either genotypic or phenotypic data sets.

Population structure is known to lead to false positives in association studies (Pritchard and Rosenberg, 1999). However, in this study of *E. pilularis*, correction for population structure was unnecessary and not applied because the association population represents provenances previously classified as one structure group, north of Sydney (Shepherd et al., 2010).

Because *Eucalyptus* trees are diploid, each SNP will have only two nucleotide alleles in an individual tree (the genotype), and within the population three possible genotypes may exist: the heterozygote and two alternate homozygote genotypes. These molecular data are often unbalanced where a population contains unequal numbers of individuals corresponding to each genotype; therefore, a general linear model of TASSEL 2.1 was used for association testing. This model uses a Σ -restricted model that is a good test of marker effects, even when the data are unbalanced (Bradbury et al., 2007; Buckler et al., 2009).

As the full sample ($n = 561$) included trees that were known to be half-sibs, a kinship coefficient was used as covariate data in the model, partitioning out covariance due to relationship. The degree of genetic covariance caused by polygenic effects was defined as positive for a pair of related individuals and zero for a pair of unrelated individuals (Yu et al., 2006). The kinship matrix was produced with SPAGeDi (Hardy and Vekemans, 2002), using genotypes from 132 SNPs, following an approach described previously (Loiselet et al., 1995; Yu et al., 2006; Bradbury et al., 2007).

Association testing using eight SNPs and 40 traits meant that 320 comparisons were made, so permutation testing was used to correct for multiple comparisons. Following the guidelines of Churchill and Doerge (1994), 1,000 permutations were applied for a 0.05 significance level. Two P values are reported. The first is a SNP-wise value (P -perm), which is a normalized value following the permutations (Supplemental Table S1). The second P value reported is an experiment-wise permutation test value (P -adj), which is corrected for multiple comparisons and is derived from a step-down MinP procedure (Ge et al., 2003; Buckler et al., 2009).

Validation of Association Testing

To ensure that the results of this study were not biased by relationships among trees, associations identified in a large population of 561 related trees were validated in at least one of three smaller subsets composed of unrelated family members: subset A ($n = 284$), subset B ($n = 101$), and subset C ($n = 83$). All subsets are derived from the larger population; subsets B and C are open-pollinated siblings of subset A. Tree phenotypes were segregated into genotypic categories based on the allele combination, and graphs were plotted for each subset, including the mean and se. All test values from the full population and the subsets are presented in Supplemental Table S1.

Calculation of HWE

All 561 trees were included for tests of HWE using the software package GEVALT (Davidovich et al., 2007). Significant departures from HWE in *PME7* suggested that these SNPs were inappropriate for the calculation of LD (Table III). Only three SNPs were genotyped in *PME6*, but they showed little evidence of LD between them (data not shown).

Avoiding Paralogs

Based on studies in other plant species, *PMEs* are coded by a large gene family (Pelloux et al., 2007; Wolf et al., 2009). Therefore, precautions were necessary to minimize the potential to inadvertently confound paralogous gene sequences. During SNP discovery, primer specificity was carefully cross-validated against the *Eucalyptus* genome version 1.0, 8 \times assembly (<http://eucalyptusdb.bi.up.ac.za/>; T.R. Sexton, unpublished data). However, this problem could also exist within the genotyping data, where SNP assays utilized different primer pairs that amplify much shorter products. The unusual observation of 100% heterozygosity in *PME7* at position 149 could be attributed to the sampling of gene paralogs, so this locus was removed from the analysis (Table III). Despite this potential problem, SNPs used in the association testing were not attributed to paralogs, as demonstrated by the observation of homozygous individuals in the population as well as by evidence from a SNP in *PME7* at position 391, where a controlled back-cross between parents (G/G \times G/C) resulted in a 1:1 Mendelian inheritance among six progeny.

Genotyping primers were designed based on *E. pilularis*-specific sequences *PME6* (AB591320) and *PME7* (AB591324), and SNP positions mentioned correspond to these references. Initial gene discovery was based on orthologous genes from *E. globulus* *PME6* (DQ376136) and *PME7* (DQ376137).

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Table S1. Results from association testing, including the population, traits, corresponding SNP, P value, corrected P value, and R^2 value for each test.

ACKNOWLEDGMENTS

Forests NSW and Forest & Wood Products Australia provided phenotypic data. The Illumina sequencing and Sequenom genotyping were performed at Southern Cross Plant Genomics. We thank Mark Edwards, Stirling Bowen, Frances Elliott, Gary Ablett, Saravanan Thavamani Kumar, Bill Joe, and Shawn Mansfield for helpful discussions and Bala Reddy Thumma for reviewing a draft of the manuscript.

Received June 8, 2011; accepted October 25, 2011; published November 3, 2011.

LITERATURE CITED

- Al-Qsous S, Carpentier E, Klein-Eude D, Burel C, Mareck A, Dauchel H, Gomord V, Balangé AP (2004) Identification and isolation of a pectin methyltransferase isoform that could be involved in flax cell wall stiffening. *Planta* **219**: 369–378
- Beaulieu J, Doerksen T, Boyle B, Clément S, Deslauriers M, Beauseigle S, Blais S, Poulin P-L, Lenz P, Caron S, et al (2011) Association genetics of wood physical traits in the conifer white spruce and relationships with gene expression. *Genetics* **188**: 197–214
- Beavis WD (1998) QTL analyses: power, precision, and accuracy. In AH Paterson, ed, *Molecular Dissection of Complex Traits*. CRC Press, New York, pp 145–162
- Boland DJ, Brooker MIH, Chippendale GM, Hall N, Hyland BPM, Johnson RD, Kleinig D, McDonald MW, Turner JD (2006) *Forest Trees of Australia*, Ed 5. CSIRO Publishing, Melbourne, Australia
- Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES

- (2007) TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**: 2633–2635
- Buckler E, Casstevens T, Bradbury PJ, Zhang Z** (2009) Trait Analysis by aSSociation, Evolution and Linkage (TASSEL), Version 2.1. User Manual. Cornell University, Ithaca, NY
- Butcher P, Southerton S** (2007) Marker-assisted selection in forestry species. In E Guimarães, J Ruane, B Scherf, A Sonnino, J Dargie, eds, Marker-Assisted Selection: Current Status and Future Perspectives in Crops, Livestock, Forestry and Fish. Food and Agriculture Organization of the United Nations, Rome, p 284
- Carlquist S** (1977) Ecological factors in wood evolution: a floristic approach. *Am J Bot* **64**: 887–896
- Churchill GA, Doerge RW** (1994) Empirical threshold values for quantitative trait mapping. *Genetics* **138**: 963–971
- Davidovich O, Kimmel G, Shamir R** (2007) GEVALT: an integrated software tool for genotype analysis. *BMC Bioinformatics* **8**: 36
- Dillon SK, Nolan M, Li W, Bell C, Wu HX, Southerton SG** (2010) Allelic variation in cell wall candidate genes affecting solid wood properties in natural populations and land races of *Pinus radiata*. *Genetics* **185**: 1477–1487
- Eckert AJ, Bower AD, Wegrzyn JL, Pande B, Jermstad KD, Krutovsky KV, St Clair JB, Neale DB** (2009) Association genetics of coastal Douglas fir (*Pseudotsuga menziesii* var. *menziesii*, Pinaceae). I. Cold-hardiness related traits. *Genetics* **182**: 1289–1302
- Eldridge K, Davidson J, Harwood C, Wyk G** (1993) Eucalypt Domestication and Breeding. Oxford University Press Inc, New York
- Ge Y, Dudoit S, Speed T** (2003) Resampling-based multiple testing for microarray data analysis. *TEST* **12**: 1–77
- Goddard ME, Hayes BJ** (2009) Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nat Rev Genet* **10**: 381–391
- González-Martínez SC, Huber D, Ersoz E, Davis JM, Neale DB** (2008) Association genetics in *Pinus taeda* L. II. Carbon isotope discrimination. *Heredity* **101**: 19–26
- González-Martínez SC, Wheeler NC, Ersoz E, Nelson CD, Neale DB** (2007) Association genetics in *Pinus taeda* L. I. Wood property traits. *Genetics* **175**: 399–409
- Gortan E, Nardini A, Salleo S, Jansen S** (2011) Pit membrane chemistry influences the magnitude of ion-mediated enhancement of xylem hydraulic conductance in four Lauraceae species. *Tree Physiol* **31**: 48–58
- Goulao LF, Vieira-Silva S, Jackson PA** (2011) Association of hemicellulose- and pectin-modifying gene expression with *Eucalyptus globulus* secondary growth. *Plant Physiol Biochem* **49**: 873–881
- Grattapaglia D** (2007) Marker-assisted selection in *Eucalyptus*. In E Guimarães, J Ruane, B Scherf, A Sonnino, J Dargie, eds, Marker-Assisted Selection: Current Status and Future Perspectives in Crops, Livestock, Forestry and Fish. Food and Agriculture Organization of the United Nations, Rome, p 252
- Grattapaglia D, Resende M** (2010) Genomic selection in forest tree breeding. *Tree Genet Genomes* **7**: 241–255
- Grattapaglia D, Silva-Junior OB, Kirst M, de Lima BM, Faria DA, Pappas GJ Jr** (2011) High-throughput SNP genotyping in the highly heterozygous genome of *Eucalyptus*: assay success, polymorphism and transferability across species. *BMC Plant Biol* **11**: 65
- Hardy OJ, Vekemans X** (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol Ecol Notes* **2**: 618–620
- Henson M, Smith HJ** (2007) Achievements in forest tree genetic improvement in Australia and New Zealand. 1. *Eucalyptus pilularis* Smith tree improvement in Australia. *Aust For* **70**: 4–10
- Hertzberg M, Aspeborg H, Schrader J, Andersson A, Erlandsson R, Blomqvist K, Bhalerao R, Uhlén M, Teeri TT, Lundeberg J, et al** (2001) A transcriptional roadmap to wood formation. *Proc Natl Acad Sci USA* **98**: 14732–14737
- Holliday JA, Ritland K, Aitken SN** (2010) Widespread, ecologically relevant genetic markers developed from association mapping of climate-related traits in Sitka spruce (*Picea sitchensis*). *New Phytol* **188**: 501–514
- Ingvarsson PK, Garcia MV, Luquez V, Hall D, Jansson S** (2008) Nucleotide polymorphism and phenotypic associations within and around the phytochrome B2 locus in European aspen (*Populus tremula*, Salicaceae). *Genetics* **178**: 2217–2226
- Johnson GI, Nikles GD** (1997) Plan for Developing and Deploying Genetically-Improved Varieties of Blackbutt (*Eucalyptus pilularis* Smith) in New South Wales. Forest Research and Development Division, State Forests of New South Wales, Sydney
- Johnson IG, Stanton RR** (1993) Thirty Years of Eucalypt Species and Provenance Trials in New South Wales: Survival and Growth Trials Established from 1961 to 1990. Research Paper Number 20. Research Division Forestry Commission of New South Wales, Sydney
- Kirst M, Myburg AA, De León JP, Kirst ME, Scott J, Sederoff R** (2004) Coordinated genetic regulation of growth and lignin revealed by quantitative trait locus analysis of cDNA microarray data in an interspecific backcross of *eucalyptus*. *Plant Physiol* **135**: 2368–2378
- Külheim C, Yeoh SH, Wallis IR, Laffan S, Moran GF, Foley WJ** (2011) The molecular basis of quantitative variation in foliar secondary metabolites in *Eucalyptus globulus*. *New Phytol* **191**: 1041–1053
- Loiselle BA, Sork VL, Nason J, Graham C** (1995) Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *Am J Bot* **82**: 1420–1425
- Moustakas AM, Nari J, Borel M, Noat G, Ricard J** (1991) Pectin methyl-esterase, metal ions and plant cell-wall extension: the role of metal ions in plant cell-wall extension. *Biochem J* **279**: 351–354
- Neale DB** (2007) Genomics to tree breeding and forest health. *Curr Opin Genet Dev* **17**: 539–544
- Neale DB, Kremer A** (2011) Forest tree genomics: growing resources and applications. *Nat Rev Genet* **12**: 111–122
- Neale DB, Savolainen O** (2004) Association genetics of complex traits in conifers. *Trends Plant Sci* **9**: 325–330
- Novaes E, Drost DR, Farmerie WG, Pappas GJ Jr, Grattapaglia D, Sederoff RR, Kirst M** (2008) High-throughput gene and SNP discovery in *Eucalyptus grandis*, an uncharacterized genome. *BMC Genomics* **9**: 312
- Pelletier MC, Henson M, Boyton S, Thomas D, Vanclay J** (2008) Genetic variation in dimensional stability of *Eucalyptus pilularis* assessed using increment cores and test blocks. *N Z J For Sci* **38**: 194–210
- Pelloux J, Rustérucci C, Mellerowicz EJ** (2007) New insights into pectin methyl-esterase structure and function. *Trends Plant Sci* **12**: 267–277
- Pritchard JK, Rosenberg NA** (1999) Use of unlinked genetic markers to detect population stratification in association studies. *Am J Hum Genet* **65**: 220–228
- Rafalski A** (2002) Applications of single nucleotide polymorphisms in crop genetics. *Curr Opin Plant Biol* **5**: 94–100
- Raymond C, Shepherd M, Sexton TR, Henson M** (2009) Quantitative and molecular genetic control of wood properties and chemistry in *Eucalyptus pilularis*. In L Apiolaza, S Chauhan, J Walker, eds, Revisiting Eucalypts 2009. Wood Technology Research Centre, University of Canterbury, Christchurch, New Zealand, pp 13–28
- Raymond CA, Henson M, Pelletier M-C, Boyton S, Joe B, Thomas D, Smith H, Vanclay J** (2008) Improving Dimensional Stability in Plantation-Grown *Eucalyptus pilularis* and *E. dunnii*. Forest and Wood Products Australia, Melbourne, Australia, pp 1–65
- Raymond CA, Thomas DS, Henson M** (2010) Predicting pulp yield and pulp productivity of *Eucalyptus dunnii* using acoustic techniques. *Aust For* **73**: 91–97
- Ren C, Kermodé AR** (2000) An increase in pectin methyl esterase activity accompanies dormancy breakage and germination of yellow cedar seeds. *Plant Physiol* **124**: 231–242
- Rosner S, Konnerth J, Plank B, Salaberger D, Hansmann C** (2010) Radial shrinkage and ultrasound acoustic emissions of fresh versus pre-dried Norway spruce sapwood. *Trees Struct Funct* **24**: 931–940
- Sexton T, Henry R, McManus L, Bowen S, Shepherd M** (2010a) Capture of assay template by multiplex PCR of long amplicons for genotyping SNPs and InDels with MALDI-TOF mass spectrometry. *Mol Breed* **25**: 471–480
- Sexton TR, Henry RJ, McManus LJ, Henson M, Thomas D, Shepherd M** (2010b) Genetic association studies in *Eucalyptus pilularis* Smith (blackbutt). *Aust For* **73**: 254–258
- Shepherd M, Sexton TR, Thomas D, Henson M, Henry RJ** (2010) Geographical and historical determinants of microsatellite variation in *Eucalyptus pilularis*. *Can J For Res* **40**: 1051–1063
- Siedlecka A, Wiklund S, Péronne M-A, Micheli F, Lesniewska J, Sethson I, Edlund U, Richard L, Sundberg B, Mellerowicz EJ** (2008) Pectin methyl esterase inhibits intrusive and symplastic cell growth in developing wood cells of *Populus*. *Plant Physiol* **146**: 554–565

- Stackpole D, Vaillancourt R, de Aguilar M, Potts B** (2010a) Age trends in genetic parameters for growth and wood density in *Eucalyptus globulus*. *Tree Genet Genomes* **6**: 179–193
- Stackpole DJ, Vaillancourt RE, Downes GM, Harwood CE, Potts BM** (2010b) Genetic control of kraft pulp yield in *Eucalyptus globulus*. *Can J For Res* **40**: 917–927
- Thavamanikumar T, McManus LJ, Tibbits JFG, Bossinger G** (2011) The significance of single nucleotide polymorphisms (SNPs) in *Eucalyptus globulus* breeding programs. *Aust For* **74**: 23–29
- Thumma BR, Matheson BA, Zhang D, Meeske C, Meder R, Downes GM, Southerton SG** (2009) Identification of a cis-acting regulatory polymorphism in a eucalypt COBRA-like gene affecting cellulose content. *Genetics* **183**: 1153–1164
- Thumma BR, Nolan ME, Evans R, Moran GF** (2005) Polymorphisms in cinnamoyl CoA reductase (CCR) are associated with variation in microfibril angle in *Eucalyptus* spp. *Genetics* **171**: 1257–1265
- Willats WGT, Orfila C, Limberg G, Buchholt HC, van Alebeek G-JWM, Voragen AGJ, Marcus SE, Christensen TMIE, Mikkelsen JD, Murray BS, et al** (2001) Modulation of the degree and pattern of methyl-esterification of pectic homogalacturonan in plant cell walls: implications for pectin methyl esterase action, matrix properties, and cell adhesion. *J Biol Chem* **276**: 19404–19413
- Wolf S, Mouille G, Pelloux J** (2009) Homogalacturonan methyl-esterification and plant development. *Mol Plant* **2**: 851–860
- Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, et al** (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet* **38**: 203–208