

## Genomics of *Eucalyptus* wood traits

Gavin F. Moran<sup>a\*</sup>, Karen A. Thamarus<sup>b</sup>, Carolyn A. Raymond<sup>c</sup>, Deyou Qiu<sup>a</sup>,  
Tom Uren<sup>a</sup> and Simon G. Southerton<sup>a</sup>

<sup>a</sup>CSIRO Forestry and Forest Products, PO Box E4008, Kingston, ACT 2604, Australia

<sup>b</sup>CSIRO Forestry and Forest Products and CRC Sustainable Production Forestry, PO Box E4008, Kingston, ACT 2604, Australia

<sup>c</sup>CSIRO Forestry and Forest Products and CRC Sustainable Production Forestry, GPO Box 252-12, Hobart, TAS 7001, Australia

(Received 5 July 2001; accepted 29 March 2002)

**Abstract** – A major focus of our research has been on using molecular technologies to guide breeding for high value wood and fibre traits in eucalypts. One approach has been to use genomic maps to locate and characterise QTL that control wood and fibre traits. A generic map for eucalypts has been developed that consists of codominant markers that can be assayed across the major commercial species. Also on the map are candidate genes for traits such as flowering and wood and fibre traits. QTL have been characterised for wood density, fibre length, pulp yield and microfibril angle. A number of these have been validated in a related pedigree. A subset of QTL colocalises with candidate genes. Research is now focussed on DNA sequence variation in different parts of these candidate genes and determining if this variation is associated with variation in wood properties. Of particular interest are genes involved in fibre development and cell wall formation. A third approach is to profile the expression of genes in the tissue of interest and determine if this relates to variation in trait performance. Microarrays are being used to relate expression of several thousand genes in xylem tissue to variation in traits such as microfibril angle and density.

**eucalypt / linkage maps / QTL / candidate genes / xylem / microarrays**

**Résumé** – **Génomique des critères de qualité du bois d'eucalyptus.** Notre recherche a été concentrée sur l'utilisation de technologies moléculaires pour guider l'amélioration génétique de critères importants de la qualité du bois et des fibres. On a développé une carte générique pour les eucalyptus à partir de marqueurs co-dominants intéressant les principales espèces d'intérêt commercial. Figurent également sur cette carte des gènes candidats concernant des caractères tels que la floraison et la qualité du bois et des fibres. On a mis en évidence des QTL pour la densité du bois, la longueur des fibres, le rendement en pâte et l'angle des microfibrilles. Un certain nombre ont été validés sur des familles apparentées à généalogie connue. Un sous ensemble de QTL sont co-localisés avec des gènes candidats. On concentre maintenant les recherches sur la variabilité de séquences ADN de ces gènes candidats pour savoir si elle est associée à celle des critères de qualité du bois. Sont d'un intérêt particulier les gènes impliqués dans le développement des fibres et la formation des parois cellulaires. Une troisième approche consiste à profiler l'expression des gènes dans les tissus en cause et à rechercher s'il existe d'éventuelles relations avec la variabilité des caractères. On utilise des microréseaux pour relier l'expression de plusieurs milliers de gènes du xylème avec la variabilité de caractères tels que l'angle des microfibrilles et la densité du bois.

**eucalyptus / carte de liaison / QTL / gènes candidats / xylème / microréseaux**

### 1. INTRODUCTION

Until recently the selection and breeding for wood and fibre traits has been significantly limited in eucalypts by the ability to measure rapidly large numbers of specific traits at reasonable cost [6, 15]. In Australia the traditional breeding

programs have focussed on yield, form etc. and, with automated coring, average wood density. In the last ten years non-destructive, cost effective methods have been developed for measuring a suite of wood and fibre traits such as microfibril angle, pulp yield, cellulose levels, fibre characteristics, etc. [6]. With the accurate determination of wood and fibre phenotypes for large numbers of individuals, the

\* Correspondence and reprints  
Tel.: 61 2 62818208; fax: 61 2 62818233; e-mail: gavin.moran@csiro.au

molecular genetic dissection of these traits became feasible. In the late 1980's the use of DNA markers became routine in forest trees and the construction of genetic maps a reality. The second generation of these maps in eucalypts are proving to be strong tools in locating and mapping genetic components (QTL) of wood and fibre traits [18]. As part of this process, and given the availability of the complete DNA sequence of *Arabidopsis* and soon rice, candidate genes are now being placed on these maps. Here a candidate gene is defined as any gene putatively involved in trait variation, based on its biological function and/or its map position [14]. When such genes collocate with QTL positions on a map there is the incentive to look at variation in the gene itself and see if this relates to variation in traits. The arrival of microarray technology has enabled the monitoring of the expression of many genes simultaneously in specific tissues. It is being used to determine what genes are expressed in xylem tissue and how that expression relates to variation in wood and fibre traits. In this paper the current state of Australian research in these areas on the temperate eucalypts, *E. globulus* and *E. nitens*, is summarised with emphasis on the considerable advantages of the integration of these approaches.

## 2. CANDIDATE GENES AND GENETIC LINKAGE MAPS

In forest trees, the Holy Grail has been early selection using molecular markers as tags to QTL for high value traits such as pulp yield and wood density. Primarily anonymous markers have been used, either dominant markers such as RAPDs and AFLPs or codominant RFLPs and microsatellite markers. In eucalypts RFLPs were developed first [3] and were the principal markers used to construct a genetic linkage map in *E. nitens* [4]. In contrast the first generation maps made from interspecific pedigrees were based initially on RAPDs [9, 19] and subsequently on AFLPs [12].

In the meantime microsatellite loci were developed in a number of eucalypts [2, 5, 8, 17] and have been shown to be highly informative and generally assayable across the important commercial species. Subsequent mapping has included candidate genes and has led to more powerful second generation maps [7, 18] and moves toward a generic map for eucalypts.

As stated previously candidate genes may be identified empirically as genes that collocate to the map position of a QTL or they could be candidates because they belong to biochemical pathways known to be involved in trait expression. For wood and fibre traits, examples would be genes in the cellulose and lignin biosynthetic pathways.

Of fifty putative genes on the genetic linkage map of *E. globulus*, forty-one genes correspond to known proteins or have homology to genes of known function (*figure 1*) [18]. The loci seemed to be randomly distributed in the genome, with little evidence that genes encoding enzymes in the same

**Table I.** Types of known potential candidate genes for wood traits mapped in *E. globulus*.

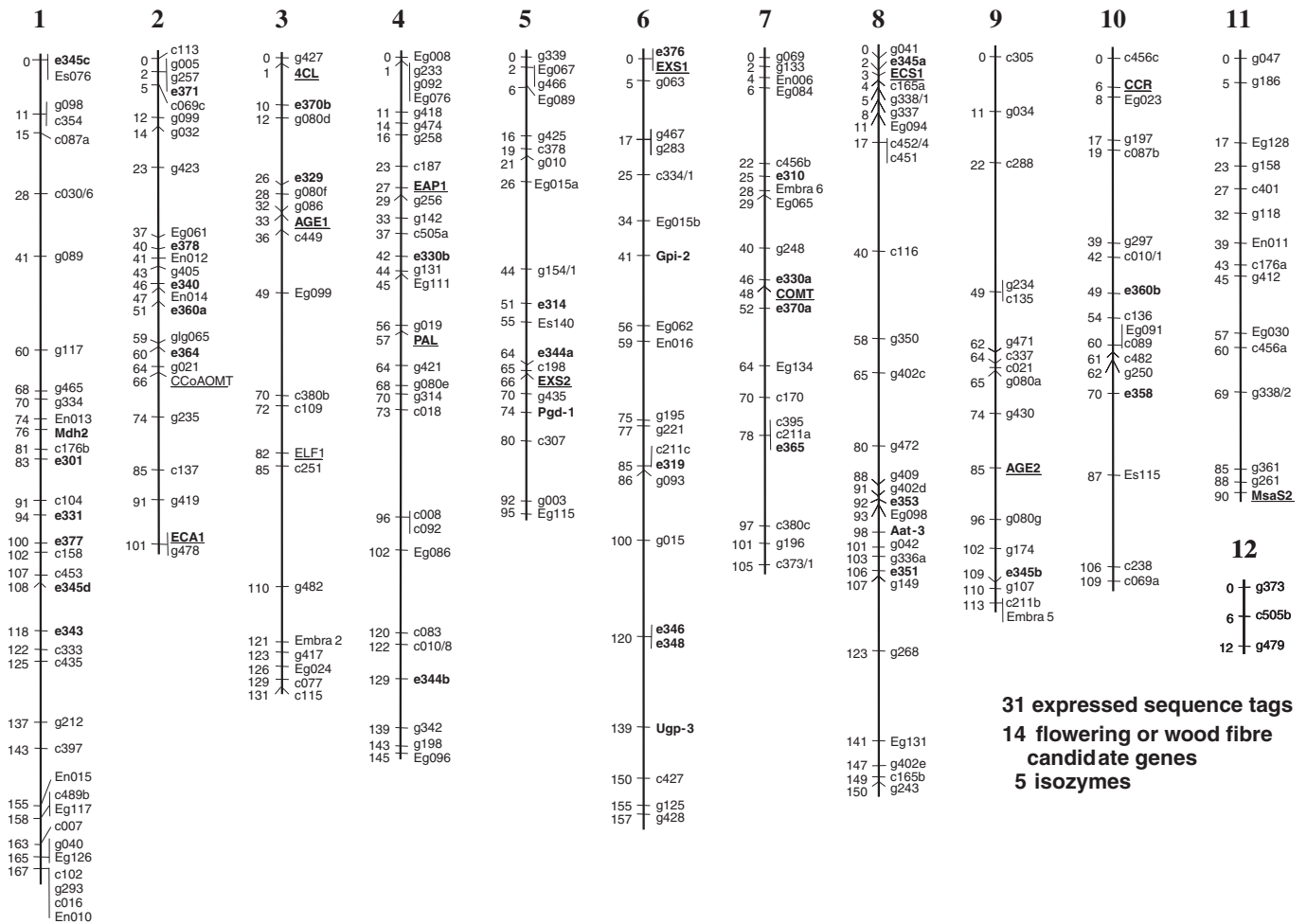
Type of candidate gene	Numbers	Sources
Cambial-specific ESTs	22	Bossinger [1]
Lignin biosynthesis	5	Thamarus [18]
Cell wall polysaccharide biosynthesis	5	Southerton CSIRO unpublished

pathway are not randomly distributed across the genome. For instance the five lignin genes are on different linkage groups, the two xylan synthase genes [16] on linkage groups 5 and 6. The 32 candidate genes for wood and fibre traits include 10 known genes and 22 cambial-specific expressed sequence tags (ESTs). The source of these genes is given in *table 1*.

## 3. QTL FOR WOOD AND FIBRE TRAITS

Studies have been undertaken to characterise QTL for wood and fibre traits in *E. globulus*, the major hardwood plantation species in Australia. The aim is to use this information to better understand the genetic components of these important traits and to develop more efficient and cost effective selection procedures. Two fullsib pedigrees were chosen that had a common male parent and were planted in field trials in 1990 at seven sites across southern Australia (Thamarus, pers. comm.). Each family was represented by four or more replicates per site and a 5 tree row plot per replicate. One of these was an interprovenance cross with 148 progeny and was used as the test pedigree. It was also used to construct the detailed genetic linkage map shown in *figure 1* [18]. The second cross was an intraprovenance cross with 135 progeny and was used to validate QTL.

At the age of 7 years, three wood cores were taken from all progeny and used to measure basic wood density, fibre length, percent cellulose content, predicted pulp yield and microfibril angle. Details of methods of assessment of traits are given by Raymond [15]. ANOVAs were performed for progeny of the test pedigree to test for significant marker-trait associations at the 1% level of probability. The data set consisted of genotypes at 249 codominant loci and quantitative data for the 5 traits. The locations of putative QTL for these traits are shown in *figure 2*. As shown on the map, there were five putative QTL for basic wood density, four for microfibril angle, three for pulp yield, and three for both cellulose content and fibre length. There were QTL locations on all linkage groups except 3 and 6 and a minimum of 13 separate QTL positions. Clearly there could be pleiotropic effects and some QTL positions may affect two traits. Some of these are not unexpected such as the QTL positions on linkage groups 4 and 10, which are significant for both pulp yield and cellulose. However, it is interesting to note that there is a QTL for cellulose on linkage 11 that does not occur for pulp yield.



**Figure 1.** A genetic linkage map of *E. globulus* with ESTs, candidate genes and isozyms in bold. Anonymous RFLPs and microsatellite loci are in ordinary type. See [18] for details of map construction.

Similarly, two QTL for microfibril angle on linkage groups 7 and 8 collocate with QTL for pulp yield and fibre length respectively.

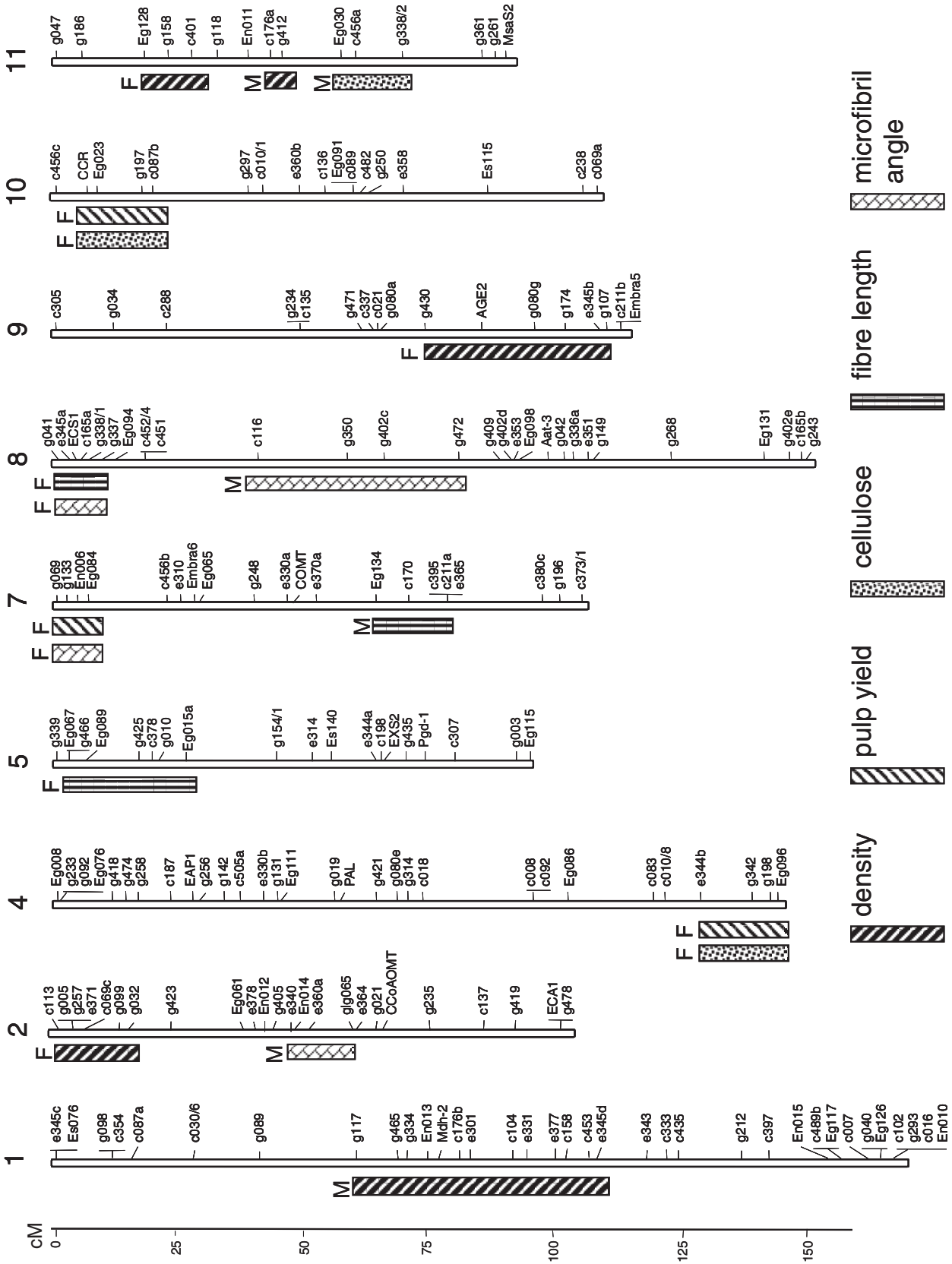
It is also noticeable that for most of the QTL, the effect is primarily coming from one side of the cross. Since the two pedigrees have a common male parent, but different female parents, the use of the second pedigree for validation can only strictly check QTL effects from the male parent. Thus *table II*

shows which QTL/marker associations were significant from ANOVA analyses in the second cross and also those not strictly tested as yet (i.e. those sourced from the female parent in the test pedigree). A total of at least seven QTL appear to be validated with this data. Some clearly were not validated despite accurate assessments of traits. It is noteworthy that for density and pulp yield, which are known to be fairly highly heritable [15], no QTL have been invalidated. In the QTL analyses some false positives were found, probably in part because total progeny sample sizes were limited and progeny were located across a number of field trials. Generally, individual QTL explained or accounted for only a small proportion of the total phenotypic variation, similar to that found in loblolly pine [13]. For instance the different wood density QTL explained from 5.4% to 6.5% of the variation. Work is now in progress to test the transference and occurrence of these QTL in *E. nitens*. The checking of QTL across studies from other laboratories [10] will require integration of maps, which is now possible with the mapping of microsatellites and known genes.

**Table II.** QTL for wood and fibre traits in *E. globulus*.

Trait	No. Putative QTL <sup>a</sup>	Validated QTL <sup>b</sup>	QTL not tested <sup>c</sup>
Density	5	2	2
Pulp yield	3	2	1
Cellulose levels	3	0	1
Fibre length	3	1	1
Microfibril angle	4	2	0

<sup>a</sup>As shown in *figure 2*, <sup>b</sup>QTL that were significant in the second pedigree at the 1% level of significance from ANOVA analyses, <sup>c</sup>QTL segregating of female side in test pedigree.



**Figure 2.** Location of putative QTL for wood and fibre traits on the *E. globulus* linkage map. Only linkage groups with QTL locations are shown. F and M indicate positive QTL effect comes from female and male parents respectively.

#### 4. QTL AND CANDIDATE GENE VARIATION

An outcome of the QTL studies has been a greater focus on the candidate genes and anonymous cDNA RFLP probe loci associated with QTL map positions. Work is underway to sequence cDNA probes mapped within a few centiMorgans of wood density and other QTL and see whether they are homologous to known genes. A second approach is to look at DNA sequence variation in the few candidate genes collocating to the characterised QTL. Here the approach being adopted is to focus on open-pollinated trees from across the range of *E. nitens* and *E. globulus*, which exist in field trials and have for which trait data can readily be obtained. The aim is to test for associations between allelic variation at the sequence level and variation in traits such as wood density. However the environmental variation within such trials will have to be kept to a minimum and substantial sample sizes may be required.

#### 5. XYLEM GENE EXPRESSION AND VARIATION IN WOOD TRAITS

A cDNA library from xylem tissue of *E. grandis* was constructed and used to characterise the eucalypt homologues of several known genes involved in cell wall formation (Southerton unpublished). The library has also been used to create xylem microarrays containing about 4500 cDNA genes. It is envisaged that this will allow changes in expression (messenger RNA abundance) of all these genes to be measured simultaneously in relation to changes in wood and fibre traits. This parallel hybridisation of high density arrays of nucleic acids on glass slides (DNA microarrays) promises to define much more effectively both known and unknown genes that are expressed in particular biological tissues [11].

The first application of these xylem microarrays has been to compare patterns of xylem gene expression in tension wood and non-tension (relaxed) wood in eucalypts. Fifteen genes were found to be up- or down- regulated more than twofold compared to normal wood (Qiu, Uren and Southerton unpublished). Two of these shared strong homology with hypothetical *Arabidopsis* genes of unknown function identified during the sequencing of the *Arabidopsis* genome. It was noteworthy that four genes, some of which were highly expressed, belong to a family of arabinogalactan proteins. This result seems analogous to that found using a different approach in compression wood in softwoods [20]. Tension wood is a significant problem in young plantations of *E. globulus* in Australia, and these results provide a new starting point to understand the molecular processes controlling this trait.

The strategy is to identify candidate genes for wood fibre traits by analysing gene expression in xylem from 7-year-old progeny in a controlled cross pedigree of *E. nitens*. RNA is isolated from xylem tissue for each progeny to assess gene

expression and cores taken from each progeny will allow quantifying of phenotypes for wood traits. At the same time marker assays will allow location of existing and perhaps new QTL for the wood fibre traits. RNA will be pooled from trees having phenotypic extremes of traits and will be used to synthesise fluorescent cDNA probes, for differentially probing of the xylem microarrays. RNA pools could also be made from trees possessing different alleles at loci associated with QTL. Candidate genes will be those that are up-regulated or down-regulated and will be sequenced and mapped. Those genes collocating with QTL will be targeted for further analysis. It can be seen that the integration of a number of technologies not available 10 years ago, promise an unravelling of the molecular basis of wood formation and may ultimately lead to more efficient selection methods for desirable wood fibre traits.

#### 6. THE WAY FORWARD – INTEGRATING TECHNOLOGIES

In future our mapping of additional loci will be selective and focussed on candidate genes targeted because of their detection in association with desirable traits through gene expression and other molecular methods. In addition our research approaches will integrate, as much as possible, the new and emerging molecular and wood assessment technologies. The ultimate goal is to understand which genes are critical in controlling trait formation so that propagules with higher value wood and fibre traits can be produced through marker-assisted breeding, selection at the actual genes, or manipulation of the genes through genetic engineering.

#### REFERENCES

- [1] Bossinger G., Leitch M.A., Isolation of cambium specific genes from *Eucalyptus globulus* Labill, in: Savidge R., Barnett J., Napier R. (Eds.), Cell and Molecular Biology of Wood Formation, BIOS Scientific, Oxford, 2000, pp. 203–207.
- [2] Brondani R.P.V., Brondani C., Tarchini R., Grattapaglia D., Development, characterization and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla*, Theor. Appl. Genet. 97 (1998) 816–827.
- [3] Byrne M., Moran G.F., Murrell J.C., Tibbits W.N., Detection and inheritance of RFLPs in *Eucalyptus nitens*, Theor. Appl. Genet. 89 (1994) 397–402.
- [4] Byrne M., Murrell J.C., Allen B., Moran G.F., An integrated genetic linkage map for eucalypts using RFLP, RAPD and isozyme markers, Theor. Appl. Genet. 91 (1995) 869–875.
- [5] Byrne M., Marques-Garcia M.I., Uren T., Smith D.S., Moran G.F., Conservation and genetic diversity of microsatellite loci in the genus *Eucalyptus*, Aust. J. Bot. 44 (1996) 331–341.
- [6] Downes G.M., Hudson I.L., Raymond C.A., Dean G.H., Michell A.J., Schimleck L.R., Evans R., Muneri A., Sampling Plantation Eucalypts for Wood and Fibre Properties, CSIRO Publishing, Melbourne, 1997.
- [7] Gion J.M., Rech P., Grima-Pettenati J., Verhaegen D., Plomion C., Mapping candidate genes in *Eucalyptus* with emphasis on lignification genes, Mol. Breed. 6 (2000) 441–449.

- [8] Glaubitz J.C., Emebiri L.C., Moran G.F., Dinucleotide microsatellites from *Eucalyptus sieberi*: inheritance, diversity and improved scoring of single-base differences, *Genome* 44 (2001) 1041–1045.
- [9] Grattapaglia D., Sederoff R.R., Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: Mapping strategy and RAPD markers, *Genetics* 137 (1994) 1121–1137.
- [10] Grattapaglia D., Bertolucci F.L.G., Penchel R., Sederoff R.R., Genetic mapping of quantitative trait loci controlling growth and wood quality traits in *Eucalyptus grandis* using a maternal half-sib family and RAPD markers, *Genetics* 144 (1996) 1205–1214.
- [11] Lockhart D.J., Winzeler E.A., Genomics, gene expression and DNA arrays, *Nature* 405 (2000) 827–836.
- [12] Marques C.M., Araújo J.A., Ferreira J.G., Whetten R., O'Malley D.M., Liu B.-H., Sederoff R., AFLP genetic maps of *Eucalyptus globulus* and *E. tereticornis*, *Theor. Appl. Genet.* 96 (1998) 727–737.
- [13] Neale D.B., Sewell M.M., Brown G.R., Molecular dissection of the quantitative inheritance of wood property traits in loblolly pine, *Ann. For. Sci.* 59 (2002) 595–605.
- [14] Pflieger S., Lefebvre V., Causse M., The candidate gene approach in plant genetics: a review, *Mol. Breed.* 7 (2001) 275–291.
- [15] Raymond C.A., Genetics of *Eucalyptus* wood properties, *Ann. For. Sci.* 59 (2002) 525–531.
- [16] Somerville C., Cutler S., Use of genes encoding xylan synthase to modify plant cell wall composition, International Patent Number WO 98/55596, 1998.
- [17] Steane D.A., Vaillancourt R.E., Russell R., Powell W., Marshall D., Potts B.M., Development and characterisation of microsatellite loci in *E. globulus* (Myrtaceae), *Silvae Genet.* 50 (2001) 89–91.
- [18] Thamarus K.A., Groom K., Murrell J., Byrne M., Moran G.F., A genetic linkage map for *Eucalyptus globulus* with candidate loci for wood, fibre and floral traits, *Theor. Appl. Genet.* 104 (2002) 379–387.
- [19] Verhaegen D., Plomion C., Genetic mapping in *Eucalyptus urophylla* and *Eucalyptus grandis* using RAPD markers, *Genome* 39 (1996) 1051–1061.
- [20] Zhang Y., Sederoff R.R., Allona I., Differential expression of genes encoding cell wall proteins in vascular tissues from vertical and bent pine trees, *Tree Physiol.* 20 (2000) 457–466.