# 4 Eucalypts

Alexander A. Myburg[1], Brad M. Potts[2], Cristina M. Marques[3], Matias Kirst[4], Jean-Marc Gion[5], Dario Grattapaglia[6], and Jacqueline Grima-Pettenatti[7]

[1] Department of Genetics, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria 0002, South Africa, *e-mail*: zander.myburg@fabi.up.ac.za
[2] School of Plant Science, and CRC for Forestry, University of Tasmania, Private Bag 55, Hobart, Tasmania 7001, Australia
[3] RAIZ-Direcçao de Investigacao Florestal ITQB II Av. Republica, Apartado 127, 2781-901 Oeiras, Portugal
[4] School of Forest Resources and Conservation, 363 Newins-Ziegler Hall, University of Florida, Gainesville, FL 32611-0410, USA
[5] Programme Arbres et Plantations, CIRAD Forêt TA 10/C, Campus de Baillarguet 34398, Montpellier Cedex 5, France
[6] EMBRAPA Genetic Resources and Biotechnology SAIN Parque Rural CP 0272 and Graduate Program in Genomic Sciences and Biotechnology, Universidade Catolica de Brasília-SGAN, 916 modulo B, 70790-160 DF Brasilia, Brazil
[7] UMR UPS/CNRS 5546, Pôle de Biotechnologies Végétales, 24 chemin de Borde Rouge, BP42617, Auzeville, Tolosane 31326 Castanet Tolosan, France

## 4.1 Introduction

### 4.1.1 History of the Crop

#### Origin and Evolution

*Eucalyptus* tree species, commonly referred to as eucalypts, are among the most planted hardwoods in the world (Doughty 2000). They are generally long-lived, evergreen species belonging to the predominantly southern-hemisphere, angiosperm family Myrtaceae (Ladiges et al. 2003). They are native to Australia and islands to its north (Potts and Pederick 2000; Ladiges et al. 2003), where they occur naturally from sea level to the alpine tree line, from high rainfall to semiarid zones, and from the tropics to latitudes as high as 43° south (Williams and Woinarski 1997). Eucalypts are the dominant or codominant species of virtually all vegetation types in Australia except rainforest, the vegetation of the central arid zone, and higher montane regions (Wiltshire 2004). They are generally sclerophyllous and adapted to low nutrient soils (Eldridge et al. 1993; Florence 1996; Specht 1996) and fire (Pryor 1976; Ashton 2000; Burrows 2002).

The eucalypt lineage is old, possibly extending back to the Late Cretaceous – ca. 70 million years ago (Hill et al. 1999; Ladiges et al. 2003; Crisp et al. 2004). Their ancestors were likely to have been widely dispersed on the supercontinent of Gondwana, as there are macrofossils ascribed to eucalypts of Eocene (55 to 34 Mya) age from northeastern Australia (Rozefelds 1996) and possibly Patagonia (Hill et al. 1999) and of Miocene (27 to 10 Mya) age from New Zealand (Pole et al. 1993) and Australia (Hill et al. 1999). The tectonic isolation of Australia (ca. 32 Mya) led to cooler, drier, and more seasonal climates and consequently a transition from a rainforest-dominated flora to Australia's unique sclerophyll flora (Hill et al. 1999; Ladiges et al. 2003; Crisp et al. 2004; Hill 2004). There is little doubt that the current dominance of the Australian continent by eucalypts is relatively recent and linked with the onset of severe aridity during the Late Miocene (10 to 7 Mya) and the present climatic system of extreme wet-dry glacial cycles that commenced around 2.9 Mya (Crisp et al. 2004). The increasing prevalence of fire played a significant role in the transformation of the Australian biota over this drying period (Kershaw et al. 1994), with eucalypts believed to have expanded from drier, central regions of the continent into more coastal environments climatically suitable for fire-sensitive rainforest taxa (Hill et al. 1999). The arrival of Aborigines on the Australian continent at least 55,000 years ago and the instigation of "fire-stick farming" would have continued this shift (Kershaw et al. 1994; Bowman 1998).

The latest molecular dating (Crisp et al. 2004) argues that divergence of the eucalypt genera and subgenera predated the final development of an ocean between Australia and Antarctica ca. 32 Mya. While molecular dating is contentious (Ladiges and Udovicic 2005), it is suggested that diversification proceeded steadily for at least 30 millions years before Australia was isolated and continued thereafter. However, de-

Genome Mapping and Molecular Breeding in Plants, Volume 7
Forest Trees
C. Kole (Ed.)
© Springer-Verlag Berlin Heidelberg 2007

spite over 100 eucalypt species having been sequenced for ITS (Steane et al. 2002), there is still insufficient sampling to test for more recent rapid radiation (Crisp et al. 2004).

### Domestication of the Eucalypts

Following their discovery by Europeans in the late 18th century, eucalypts were spread rapidly around the world (Zacharin 1978; Jacobs 1981; Eldridge et al. 1993; Doughty 2000). They were introduced early on into countries such as India (c. 1790), France (c. 1804), Chile (1823), Brazil (1825), South Africa (1828), and Portugal (1829) (Doughty 2000; Potts et al. 2004) and rapidly spread as the fast growth and good adaptability of eucalypt species were recognized. Initially the botanical gardens of southern Europe played a major role in their introduction to other parts of the world, including Africa and South America (Zacharin 1978). Later in the 19th century, large quantities of seed were distributed directly from Australia. In other cases, seed was collected from local exotic plantings and, where multispecies plantings occurred, this often contained $F_1$ hybrids (Potts and Dungey 2004). While these $F_1$ hybrids may have performed well, subsequent seed collection from them often resulted in plantations that performed poorly in subsequent generations and were extremely variable, e.g., the Río Claro hybrid in Brazil (Campinhos and Ikemori 1977; Brune and Zobel 1981) and the Mysore hybrid in India (Varghese et al. 2000). In many countries where eucalypts have been introduced for a long time and continually reproduced from local seed sources, they have formed landraces adapted to the specific environment of the country (Eldridge et al. 1993).

The history of eucalypt breeding is detailed in Eldridge et al. (1993) and Potts (2004). Some of the earliest breeding was undertaken by French foresters in Morocco in 1954-1955 (Eldridge et al. 1993). The rise of industrial eucalypt plantations in the 1960s saw a more formal approach to genetic improvement with, for example, the commencement of the Florida *E. grandis* breeding program in 1961 (Franklin 1986), *E. globulus* breeding in Portugal in 1965-1966 (Dillner et al. 1971; Potts et al. 2004), and establishment of large provenance tests of *E. camaldulensis* in many countries (Eldridge et al. 1993). Major advances occurred in the 1970s with, for example, the first commercial plantings of selected clones derived from hardwood cuttings in the Congo (Martin and Quillet 1974) followed by Aracruz in Brazil and the establishment in many countries of the first large base population trials of species such as *E. urophylla* (Eldridge et al. 1993) and *E. globulus* (Potts et al. 2004). These trials were established from open-pollinated seed lots collected from rangewide provenance collections and formed the base for deployment and breeding populations in many countries (Eldridge et al. 1993). Many other major international base population trials were established through the 1980s for species such as *E. grandis*, *E. tereticornis*, and *E. viminalis* and more intensive collections of elite provenances identified in earlier collections (Eldridge et al. 1993).

Eucalypts are still at the early stages of domestication compared to crop species, with most breeding programs only one or two generations removed from the wild. However, eucalypts are fast becoming among the most advanced genetic material in forestry, with stock originating from the *E. grandis* program in Florida already in its sixth generation (Potts 2004). Domestication of eucalypts has proceeded faster in countries like Brazil and South Africa that rely on plantations for their eucalypt wood than in Australia, where up until the 1990s wood products of eucalypts were derived almost entirely from native forests. As with other forest tree genera with long generation times, there is the potential for eucalypt domestication programs to benefit tremendously from genomic-era molecular technologies that could significantly speed the process of genetic improvement.

## 4.1.2 Botany

### Taxonomy

In the broad sense, eucalypts encompass species of the genera *Eucalyptus* L'Hérit., *Corymbia* (Hill and Johnson 1995), and *Angophora* Cav. (Ladiges 1997, Table 1). A key feature of the majority of eucalypts is the fusion of either the petals and/or sepals to form an operculum from which the eucalypts derive their name (from the Greek *eu*, "well," and *calyptos*, "covered," Eldridge et al. 1993; Ladiges 1997). The operculum appears to have evolved independently in different eucalypt lineages and has not evolved in *Angophora* (Ladiges 1997). There is some debate as to whether the *Corymbia* and *Angophora* genera (bloodwood taxa) warrant separation from the genus *Eucalyptus* (non-bloodwood taxa) in the strict sense. This split is supported by several independent molecular studies (e.g., Sale et al. 1996; Ladiges and Udovicic 2000, Udovicic and Ladiges 2000; Steane et al. 2002; Whittock et al.

2003) and, following Ladiges et al. (2003), adopted herein. The latest taxonomic revision (Brooker 2000) of the eucalypts recognizes just over 700 species that belong to 13 main evolutionary lineages (Table 1) but still treats the bloodwood eucalypts as subgenera of *Eucalyptus*. Most species belong to the subgenus *Symphyomyrtus*, and it is mainly species from three sections of this subgenus that are used in plantation forestry (Table 1).

Eucalypts encompass an exceptional level of genetic diversity. They range in habit from shrubs and multistemmed mallees to giant trees and include the tallest flowering plants on earth (*Eucalyptus regnans* – 96 m; Boland et al. 1985; Wardell-Johnson et al. 1997; Potts and Pederick 2000; Potts et al. 2003). The major subgenera exhibit different ecological and reproductive characteristics (Pryor 1976; Florence 1996; Ladiges 1997), and closely related species are usually ecologically differentiated (Florence 1996; Williams and Woinarski 1997). Within species, marked genetic differentiation between populations is the norm rather than the exception (Pryor and Johnson 1971, 1981; Potts and Wiltshire 1997; Dutkowski and Potts 1999). Genetic variation between populations in quantitative traits is often continuous and clinal, paralleling environmental gradients associated with changes in, for example, latitude, continentality, or altitude (Pryor 1976; Potts and Wiltshire 1997). Indeed, many recognized species intergrade along such gradients, resulting in complexes of closely related species where no clear morphological discontinuity is apparent (Pryor 1976; Jordan et al. 1993; Holman et al. 2003).

## Breeding System

Eucalypt flowers are occasionally solitary (e.g., *E. globulus*) but often occur in clusters of three or

**Table 1.** Major evolutionary lineages within the eucalypts. The alignment of Pryor and Johnson's (1971) genera and subgenera with Brooker's (2000) subgenera. Pryor and Johnson's classification was informal, but widely used for 30 years. The number of species in each of Brooker's subgenera is indicated and examples of well-known forestry species are given. Most species used in plantation forestry, particularly outside Australia, are from Brooker's sections *Maidenaria* (e.g., *E. globulus*, *E. nitens*, *E. viminalis*), *Exsertaria* (e.g., *E. camaldulensis*, *E. tereticornis*), and *Latoangulatae* (e.g., *E. grandis*, *E. saligna*, *E. urophylla*) in the subgenus *Symphyomyrtus* (from Potts 2004)

| Pryor & Johnson's subgenera/genera | Brooker's subgenera | No. of species | Examples of well-known forestry species |
|---|---|---|---|
| *Angophora* (genus) | *Angophora*[a] | 7 | |
| *Blakella* | *Blakella*[a] | 15 | |
| *Corymbia* | *Corymbia*[a] | 67 | *C. torelliana, C. citridora, C. variegata C. maculata* |
| *Eudesmia* | *Eudesmia* | 19 | |
| *Gaubaea* | *Acerosa* | 1 | |
| *Gaubaea* | *Cuboidea* | 1 | |
| *Idiogenes* | *Idiogenes* | 1 | *E. cloeziana* |
| *Monocalyptus* | *Primitiva* | 1 | |
| *Monocalyptus* | *Eucalyptus* | 110 | *E. regnans, E. delegatensis, E. obliqua, E. marginata, E. fastigata* |
| *Symphyomyrtus* | *Cruciformes* | 1 | *E. guilfoylei* |
| *Symphyomyrtus* | *Alveolata* | 1 | *E. microcorys* |
| *Symphyomyrtus* | *Symphyomyrtus* | 474 | *E. camaldulensis, E. exserta, E. globulus, E. grandis, E. nitens, E. paniculata, E. robusta, E. saligna, E. tereticornis, E. urophylla, E. viminalis* |
| *Telocalyptus* | *Minutifructus*[b] | 4 | *E. deglupta* |

[a] The subgenera *Blakella* and *Corymbia* had previously been treated as a separate genus *Corymbia* (Hill and Johnson 1995) and the subgenus *Angophora* treated as a genus (Hill and Johnson 1995), and this approach has been adopted in the text
[b] A recent molecular study suggests that these species belong to subgenus *Symphyomyrtus* (Whittock 2003)

more in umbels or terminal inflorescences. The eucalypt flower is normally bisexual, with numerous stamens that expand outward after operculum shed to form the conspicuous floral display (Pryor 1976). Eucalypts are predominantly animal pollinated, with vectors encompassing a wide variety of insects, birds, marsupials, and a few bat species (House 1997; Hingston et al. 2004; Southerton et al. 2004; Barbour et al. 2005). They have a mixed mating system but are generally preferential outcrossers, with high levels of outcrossing maintained by protandry and various incomplete pre- and postzygotic barriers to self-fertilization (Potts and Wiltshire 1997; Pound et al. 2002a, b, 2003). The postzygotic barriers include intense selection against the products of inbreeding (Hardner and Potts 1995, 1997; Potts and Wiltshire 1997). Outcrossing rates may vary between (Butcher and Williams 2002; McDonald et al. 2003) and within trees even of the same species. For example, outcrossing is often higher in denser stands (Hardner et al. 1996) and at the top of the tree canopy (Patterson et al. 2004b; Hingston and Potts 2005). Biparental inbreeding may also occur in open-pollinated progenies from native stands due to related individuals growing in close spatial proximity (Eldridge et al. 1993; Hardner et al. 1998). The seed of most species has no specialized mechanisms for dispersal and is normally deposited within a distance of twice the tree height (Potts and Wiltshire 1997).

The major eucalypt subgenera do not hybridize (Griffin et al. 1988; c.f. Stokoe et al. 2001), and, while there are significant barriers to hybridization between species within subgenera, these are often weak (Potts et al. 2003; Potts and Dungey 2004). Natural hybridization and intergradation between recognized taxa is common in nature (Griffin et al. 1988; Potts and Wiltshire 1997; Byrne and Macdonald 2000; Butcher et al. 2002), often making delineation of species difficult (Pryor and Johnson 1971). Natural introgression may be cryptic, and only detectable at the molecular level (McKinnon et al. 2001a, 2004), indicating ancient hybridization and gene flow between lineages. At the other extreme, first-generation hybridization is actively occurring between exotic plantation eucalypts and native populations (Barbour et al. 2002, 2003, 2005). Artificial hybrid combinations have been produced, and in general hybrid inviability tends to increase with increasing taxonomic distance between the parents, but there are exceptions (Griffin et al. 1988; Potts and Dungey 2004).

## Genome Size and Structure

The size and structure of the genome of eucalypts is reviewed in Poke et al. (2005). Consistent with most myrtaceous genera, eucalypts are diploids with a haploid chromosome number of 11 (Eldridge et al. 1993). Reports of higher numbers of chromosomes have not been verified, e.g., *E. cladocalyx* is $2n = 22$ (R. Wiltshire unpublished data) and not $2n = 24$ (Eldridge et al. 1993). While polypoidy has been artificially induced (Janki-Ammal and Khosla 1969), there are no reports of polyploidy in nature. Grattapaglia and Bradshaw's (1994) estimates of the haploid genome size of several eucalypt species and hybrids based on flow cytometry range from 370 to 700 million base pairs (Mbp). They estimated the average haploid genome size for *Symphyomyrtus* species to be 650 Mbp (Grattapaglia and Bradshaw 1994). While no *Angophora* species were included in their study, the two *Corymbia* species examined had a haploid genome size (370 and 390 Mbp) substantially smaller than the *Eucalyptus* species studied. Hybrids had intermediate DNA content, and there was no evidence of polyploidy. Estimates will vary with laboratory technique, and a new estimate for the size of the *E. globulus* genome, for example, is 644 Mbp (Pinto 2005), which is 20% larger than Grattapaglia and Bradshaw's (1994) estimate of 530 Mbp. Nevertheless, the haploid genome size of *Eucalyptus* species would generally appear to be slightly larger than those of the current plants whose genomes have been completely sequenced (125 Mbp for *Arabidopsis thaliana*, 420 to 466 Mbp for two rice varieties, Fukuoka et al. 1998, and 473 Mbp for *Populus tricarpa*, Poke et al. 2005), and is clearly many-fold smaller than the genome size estimated for *Pinus* species, which have large regions of repetitive DNA (Ahuja 2001). There were no studies of the structure of the eucalypt genome before the advent of accessible DNA marker technology (see below).

As in most angiosperms, the inheritance of both the mitochondrial (Vaillancourt et al. 2004) and chloroplast (Byrne et al. 1993; McKinnon et al. 2001b) genomes appears to be fully maternal. While the structure of the mitochondrial genome has not been studied, the *E. globulus* chloroplast has been completely sequenced and shows very high homology of coding regions, with chloroplast sequences from species such as *Nicotiana tobacum* and *Oenothera elata* but considerable divergence in the noncoding regions (Steane 2005). Broadly transferable microsatellite regions have been identified from

the *E. globulus* chloroplast sequence (Steaneet al. 2005). A hypervariable region, $J_{la}$, also occurs in the intergenic spacer on either side of the junction between the large single-copy region and the inverted repeat A in *Symphyomyrtus* but appears absent from *Monocalyptus* species (Vaillancourt and Jackson 2000).

**Molecular Diversity**

Historically, allozymes were the main molecular markers used for genetic diversity studies in eucalypts and revealed their mixed mating system, high levels of genetic diversity and heterozygosity, and greater population differentiation in regionally distributed species than in widespread or local species (reviewed in Moran 1992; Potts and Wiltshire 1997). However, over the last decade these markers have been replaced in diversity studies by more informative DNA marker technologies such as random amplified polymorphic DNA (RAPD), restriction fragment length polymorphism (RFLP), amplified fragment length polymorphism (AFLP), and microsatellite analysis (Shepherd and Jones 2005), and more recently, by the analysis of direct sequence variation in functional genes (Poke et al. 2003; McKinnon et al. 2005; Thumma et al. 2005). These markers have revealed that eucalypt gene pools have a hierarchy of spatial genetic structure from clonal patches (e.g., Rossetto et al. 1999; Smith et al. 2003; Jones et al. 2005) and family groups (Skabo et al. 1998) that extend over tens of meters, to localized population differention over hundereds of meters (McGowen et al. 2001), and broad-scale genetic differentiation over many hundreds of kilometers (Byrne et al. 1998; Butcher et al. 2002; Jones et al. 2002; Holman et al. 2003; Astorga et al. 2004; Potts et al. 2004). As expected, maternally inherited, haploid chloroplast markers, which are dispersed only through the seed, exhibit greater population differentiation than nuclear markers (Byrne and Macdonald 2000; Jones et al. 2005).

## 4.1.3 Economic Importance

Eucalypts are renowned as species with fast growth, straight form, valuable wood properties, wide adaptability to soils and climates, and ease of management through coppicing (Eldridge et al. 1993; Potts 2004). They are now found in more than 90 countries where the various species are grown for products as diverse as sawn timber, mine props, poles, firewood, pulp, charcoal, essential oils, honey, and tannin, as well as for shade, shelter, and soil reclamation (Doughty 2000). They are an important source of fuel and building material in rural communities in countries such as India, China, Ethiopia, Peru, and Vietnam. However, it is the great global demand for short-fiber pulp that has driven the massive expansion of eucalypt plantations throughout the world during the 20th century (Turnbull 1999). Their high fiber count relative to other wood components, coupled with the uniformity of fibers relative to other angiosperm species, has caused high demand for eucalypt pulp for coated and uncoated free-sheet paper, bleach board, sanitary products (fluff pulp), and, secondarily, for top liner on cardboard boxes, as a corrugating medium, and as a filler in long-fiber conifer products such as newsprint and containerboard (Kellison 2001). New technologies are also increasing interest in the use of plantation eucalypts for sawn wood, veneer, and medium-density fiberboard and as extenders in plastic and molded timber (Kellison 2001; Bermúdez Alvite et al. 2002; Shield 2004; Waugh 2004).

While precise global figures are difficult to obtain, a conservative estimate is 9.5 million ha of industrial eucalypt plantations in 1999, with the vast majority of these established since the 1950s, and predicted to reach 11.6 million ha in 2010 (Raga 2001). The majority of plantations consist of only a few eucalypt species and hybrids. The most important are *E. grandis*, *E. globulus*, and *E. camaldulensis*, which together with their hybrids account for about 80% of the plantation area; they are followed by *E. nitens*, *E. saligna*, *E. deglupta*, *E. urophylla*, *E. pilularis*, *Corymbia citriodora*, and *E. teriticornis* (Eldridge et al. 1993; Waugh 2004). Market favorites for pulpwood are *E. grandis and E. urophylla* and their hybrids in tropical and subtropical regions and *E. globulus* in temperate regions. There is considerable interest in introgressing superior *E. globulus* pulp traits into the tropical and subtropical genetic backgrounds.

## 4.1.4 Classical Breeding Objectives

The main breeding programs in eucalypts worldwide are focused on improving profit from industrial pulpwood plantations (Borralho 2001; Kanowski and Bor-

ralho 2004; Raymond and Apiolaza 2004). In such systems, the key objective traits traditionally considered to drive profits are volume production per hectare, wood density, and pulp yield (Borralho et al. 1993; Greaves and Borralho 1996; Greaves et al. 1997; Wei and Borralho 1999), although their economic weights will vary depending upon whether plantations are part of a vertically integrated production system (Borralho et al. 1993) or only for wood-chip export (Apiolaza et al. 2005). Traits such as pest and disease resistance (e.g., Coutinho et al. 1998; Soria and Borralho 1998; Milgate et al. 2005), adaptability (e.g., frost resistance; Cauvin and Potts 1991; Tibbits et al. 1991), drought resistance (Toroet al. 1998), and survival (Chambers and Borralho 1997) are only important as far as they impact on one or more of these objective traits. Secondary wood property traits of interest to pulp producers include the quantity or quality of extractives or lignin in the wood that affect the economic and/or environmental cost of pulping (Poke et al. 2004; Raymond and Apiolaza 2004). Breeding programs directly linked with paper mills are also exploring other wood properties that impact on paper quality such as fiber length, courseness, and wall dimensions (Raymond and Apiolaza 2004), as well as studying genetic effects on paper quality (Cotterill et al. 1999).

There is increasing interest in breeding eucalypts for sawn timber, veneers, and reconstituted wood products worldwide (Kube and Raymond 2005; Raymond 2002; Raymond and Apiolaza 2004). Key wood properties believed to affect sawn timber and composites are given in Table 2 (see also Shield 2004). There is increasing work on identifying breeding objectives, economic weights, and cost-effective selection traits for these production systems (Greaves et al. 2004a, b). However, this is complicated by the range of products, silvicultural regimes and production systems, changing technologies over the longer rotations, and requirements of plantations for multiple products through thinning or changes in product pricing (Shield 2004).

While wood and cellulose production is the key focus of most breeding programs, there are small eucalypt breeding programs aiming to improve the yield of alternative products such as leaf volatile oils (Boland et al. 1991; Doran and Bell 1995; Byrne 1999) and for enhanced ecosystem services by selection of salt-resistant genotypes (Meddings et al. 2001; Dale 2002). Decreasing plant propagation costs by, for example, selection for vegetative propagability (de Assis 2001; Cañas et al. 2004) or increased seed production (McGowen et al. 2004) also occurs as a secondary objective in many breeding or deployment programs (Raymond and Apiolaza 2004).

### 4.1.5 Classical Breeding Achievements

As with most forest trees, large and cost-effective genetic gains have been achieved in the early stages of eucalypt domestication, simply through species (Jacobs 1981) and provenance selection (Eldridge et al. 1993), followed by individual (family) selection and establishment of clonal or seedling seed orchards or clonal propagation of elite selections for direct deployment (Eldridge et al. 1993; Kanowski and Borralho 2004; Potts 2004). Subsequent population improvement has also demonstrated significant genetic gain through recurrent selection in an open-pollinated breeding population (Reddy and Rockwood 1989), or

**Table 2.** Key wood properties for a range of product classes (modified from Raymond 2002)

| Pulp and paper | Sawn timber | Composites |
|---|---|---|
| Basic density | Basic density and gradient | Basic density |
| Pulp yield/cellulose content | Microfibril angle | Lignin content |
| Fiber length | Strength and stiffness | Extractives content |
| Lignin content and composition | Dimensional stability | Cellulose content |
| | Shrinkage and collapse | |
| | Tension wood | |
| | Knot size | |
| | Incidence of decay, spiral grain, and end splits | |

sublines (Griffin 2001; Sanhueza and Griffin 2001), possibly coupled with open- or controlled-pollinated nucleus populations of the most elite selections or specialized breeds (Potts 2004). For species that are easily propagated vegetatively, such as *E. grandis*, clonally propagated breeding populations have further enhanced gains (Snedden and Verryn 2004). Overlapping generation breeding using a "rolling front" strategy has also been adopted by programs in Australia and Portugal (Borralho and Dutkowski 1998; McRae et al. 2004b). Major advances were made in the 1990s, with the addition of wood density and pulp yield with growth into the breeding objectives as the two traits that account for over 70% of the benefits from breeding for pulp production (Borralho 2001). With such an objective, two generations of selection of *E. globulus* for a vertically integrated eucalypt pulp production system were expected to increase income by 1.5% and decrease production costs by 16%, saving US$ 7.2 million per annum for a 250,000-ton pulp mill (Kanowski and Borralho 2004).

Eucalypt hybrids, either $F_1$s or composites, have long been deployed in eucalypt forest through vegetative propagation and are a significant component of eucalypt plantation forestry, particularly in the tropics and subtropics (de Assis 2000; Vigneron and Bouvet 2000; Potts and Dungey 2004). Such hybrids have been classically selected based on large-scale hybrid production and clonal testing, through either (1) initial screening of seedlings for vegetative propagability or (2) remobilizing mature selections through coppicing, sequential grafting, or in vitro techniques and then introducing them into clonal tests. The main hybrids used in industrial plantations are *E. grandis* × *urophylla*, *E. grandis* × *camaldulensis* and varieties including at least one of *E. saligna*, *E. pellita*, *E. exserta*, and *E. tereticornis*. Such hybrids are planted on a relatively large scale in Brazil and the Congo, although sizeable plantations also occur in China, Indonesia, and South Africa. The deployment of selected clones of *E. urophylla* × *grandis* in Brazil and the Congo has been a major success in overcoming canker and disease susceptibity of *E. grandis* (Campinhos and Ikemori 1989; Vigneron and Bouvet 2000). Indeed in Brazil, a combination of improvements in genetics, silviculture, and propagation has increased the mean annual growth of eucalypt plantations from $10\,m^3\ ha^{-1}\ yr^{-1}$ in the 1960s to more than $40\,m^3\ ha^{-1}\ yr^{-1}$ at present (Binkley and Stape 2004).

### 4.1.6 Future Perspective: Challenges for Molecular Breeding of *Eucalyptus*

The next decade will see major advances in our understanding of the eucalypt genome and molecular breeding (Grattapaglia 2004; Poke et al. 2005; Thumma et al. 2005). The genome of a eucalypt tree (that of an *E. camaldulensis* clone) is being sequenced at Kazusa DNA Research Institute in Japan (T. Hibino and S. Tabata personal communication; Myburg 2004), and the *Eucalyptus* Genome Network (EUCAGEN, www.eucagen.org) has been initiated to coordinate the generation of further genomic resources for *Eucalyptus*. Furthermore, the *Eucalyptus* research community has been invited to submit a proposal for the sequencing of the *E. grandis* genome by the US Department of Energy (DOE), a proposal that will, if successful, lead to the public release of the *E. grandis* genome sequence (6–8 coverage) before the end of 2008 (J. Tuskan personal communication). These milestones are bound to lead to rapid advances in the development of molecular breeding tools and molecular genetic improvement of eucalypts. The development of genetically modified eucalypts has been slow compared with, for example, *Populus* species (Potts et al. 2003). The development of fully tested genetically modified clones to the stage of large-scale planting is a slow process (Griffin 1996), and regulatory and certification requirements associated with the use of genetically modified trees are likely to limit their operational use in the foreseeable future (Burley and Kanowski 2005). Nevertheless, other molecular technologies offer great potential to contribute to eucalypt breeding through many avenues including quantifying genetic diversity and relationships, breeding systems, gene flow and fingerprinting for quantifying contamination and clone identification, QTL detection, and molecular breeding through marker- or gene-assisted selection. Major advances are still to be made through better definition of breeding and deployment objectives and economic weights (Raymond and Apiolaza 2004), enhanced accuracy of genetic evaluation through more sophisticated trial design (Williams et al. 2002) and analysis (Dutkowski et al. 2002; Costa e Silva et al. 2005), and exploitation of the major advances in quantitative genetic analysis that have occurred over the last decade (e.g., Soria et al. 1998; Fernandez and Toro 2001; Costa e Silva et al. 2004; de Resende and Thompson 2004; McRae et al. 2004a). Genetic gain is a function of selection

intensity, and there are now clear opportunities for quantum increases in the size of pedigreed breeding populations through the recent advances that have occurred in controlled pollination (Harbard et al. 1999; Williams et al. 1999; Patterson et al. 2004a; de Assis et al. 2005) and vegetative propagation techniques (e.g., mini- and microcuttings; de Assis 2001), increasingly cheaper techniques for nondestructive assessment of key wood properties (e.g., near-infrared reflectance and raman spectroscopy; Schimleck et al. 1996; Downes et al. 1997; Raymond and Apiolaza 2004), and industrial-scale systems for data management and genetic evaluation (McRae et al. 2004a). The challenge facing molecular breeding is its integration into existing eucalypt breeding programs in order to deliver gains to industrial forests above that which can be achieved by equivalent investment in other already well-integrated technologies and strategies.

This chapter provides an overview of the status of genome mapping and molecular breeding of *Eucalyptus* tree species. Emphasis is placed on genetic and physical mapping of eucalypt genomes and on new approaches to locating genes and regulatory regions that control quantitative traits in *Eucalyptus* plantations. The last section of the chapter provides a brief future perspective on genome research in *Eucalyptus* and its impact on the molecular domestication of this important fiber crop. Readers are also referred to other recent reviews of genome research and molecular breeding in *Eucalyptus* (Moran et al. 2002; Potts 2004; Shepherd and Jones 2005; Grattapaglia 2004; Poke et al. 2005).

## 4.2 Genetic Linkage Mapping of Eucalypt Genomes

Genetic markers are DNA phenotypes that reflect differentiation among individuals, populations, and species. The availability of large numbers of highly polymorphic and neutral genetic markers whose inheritance and segregation can be followed through generations has allowed the construction of genetic linkage maps for many plant species, including several eucalypt species. Most genetic mapping studies in plants have relied on the availability of inbred lines, near-isogenic lines, or backcross progeny of inbred parents. However, available pedigrees for the majority of outbred angiosperm tree species, such as the eucalypts, generally involve only two parents and their full-sib progeny or one parent and its maternal half-sib progeny. Such outbred mapping pedigrees suffer from several limitations (discussed below) that complicate genetic linkage mapping. Novel mapping strategies, such as the "two-way pseudotestcross design" discussed later in this section, have been developed to overcome these limitations, allowing the generation of single-tree genetic linkage maps for selected individuals of tree species. The high level of genetic diversity, the ability to generate large progeny sets, the ability to clone segregating progeny, the relatively small genome size, and the low proportion of repetitive DNA of *Eucalyptus* facilitated early interest in genetic linkage mapping in this genus (Grattapaglia and Bradshaw 1994). Most eucalypt linkage maps have been constructed from segregating half-sib or full-sib families, with family sizes of up to 200 individuals. To maximize the detection of genetic polymorphism at the DNA level, highly heterozygous parents have been selected, most often from different species. Wide interspecific crosses in *Eucalyptus* have revealed genetic barriers to crossing (Myburg et al. 2004) and have resulted in distorted segregation ratios in mapping progeny, which can lead to biased estimates of recombination and false linkage (Lorieux et al. 1995). The development of codominant markers for *Eucalyptus* (Brondani et al. 1998, 2002) stimulated the first comparative mapping studies, which may lead to more efficient use of genetic mapping information for molecular breeding and evolutionary studies in *Eucalyptus*. In this section, we briefly review practical considerations for the construction of genetic linkage maps in *Eucalyptus* and we discuss potential applications of genetic mapping in molecular breeding of these tree species.

### 4.2.1 DNA Isolation for Genetic Mapping

The isolation of relatively pure genomic DNA from a large number of individuals is the first step in any genetic mapping project. *Eucalyptus* total genomic DNA has been isolated from leaf tissue mostly using modified versions of the published protocols of Doyle and Doyle (1990), Saghai-Maroof et al. (1984), or Wagner et al. (1987). The presence of secondary metabolic compounds and carbohydrates in eucalypt leaves can interfere with DNA quality, especially if leaf samples have to spend a considerable time in transit

to the laboratory. Tissue condition before extraction is therefore usually a critical factor (Ferreira and Grattapaglia 1995). Myburg et al. (2001) proposed a high-throughput 96-well DNA extraction protocol based on the QIAGEN 96-well mouse tail DNA extraction kit, modified for plant DNA extraction chemistry, as well as tissue homogenization using the FastPrep instrument (QBiogene). The convenience and throughput of this method is further improved by collecting *Eucalyptus* leaf discs directly into 2-ml tubes that already contain a small amount of a desiccant such as silica gel crystals. Such samples can be stored at room temperature until DNA isolation and are stable for long periods of time.

### 4.2.2 Marker Availability

DNA-based marker techniques vary in DNA requirements, cost of development and assay, technical expertise, genetic information, and transferability across taxa. Genetic marker analysis and mapping in eucalypts have progressed at the pace of the development of new marker technologies. Isozyme markers were first used in 1983 to study mating systems in natural and exotic populations of *Eucalyptus* (Moran and Bell 1983). These markers were, however, limited in genome coverage and in the amount of genetic polymorphism that could be assayed. Restriction fragment length polymorphism (RFLP) markers were used first in phylogenetic analyses of eucalypt chloroplast DNA (Steane et al. 1991) and later to construct *Eucalyptus* genetic linkage maps (Byrne et al. 1995; Thamarus et al. 2002). RFLP analysis has not found widespread use in *Eucalyptus* as it is technically demanding, labor intensive, and time consuming and requires the previous development and labeling of probes.

The advent of the polymerase chain reaction (PCR) (Mullis and Faloona 1987) allowed the development of many new markers based on DNA polymorphisms. Random amplified polymorphic DNA (RAPD) markers (Williams et al. 1990) greatly facilitated the rapid construction of genetic linkage maps for several eucalypt species (Grattapaglia and Sederoff 1994; Vaillancourt et al. 1994; Verhaegen and Plomion 1996; Bundock et al. 2000; Gan et al. 2003). As in other plant species, the main limiting feature of RAPD markers in *Eucalyptus* is their dominant behavior, relatively low multiplex ratio, and low repeatability across research laboratories (Jones et al. 1997). Amplified fragment length polymorphism (AFLP) analysis (Voset al. 1995) has proven to be more robust, and the technique samples much larger numbers of loci in high-resolution sequencing gels. This has allowed the construction of high-density, AFLP-based genetic maps of individual *Eucalyptus* trees using only small number of oligonucleotide primers and minute amounts of DNA (Marques et al. 1998). Multiplexed AFLP genotyping using fluorescently labeled primers further increased the throughput and convenience of AFLP analysis in *Eucalyptus* (Myburg et al. 2003). As with RAPD markers, the main disadvantage of AFLP markers has been their dominant behavior, but also the fact that there are commercial restrictions on the use of the technology (Keygene, Wageningen, The Netherlands). Furthermore, AFLP markers are generally not shared among different outbred eucalypt parents, which has made the integration of different AFLP maps difficult, except where shared parents are used in mapping pedigrees (Myburg et al. 2004). The development of transportable, PCR-based and codominant simple sequence repeat (SSR), or microsatellite markers for *Eucalyptus* (Byrne et al. 1996; Brondani et al. 1998; Glaubitz et al. 2001; Steane et al. 2001; Brondani et al. 2002) revolutionized the genetic mapping of eucalypt trees. SSR markers have been added to previously constructed genetic linkage maps (Brondani et al. 1998, 2002; Marques et al. 2002) and mapped in combination with other markers (Bundock et al. 2000; Thamarus et al. 2002). Limitations of the use of SSRs for mapping include the effort required to develop markers and the fact that only one locus is sampled with each primer pair.

More recently, gene-based molecular markers have been integrated into existing maps (Bundock et al. 2000; Gion et al. 2000; Thamarus et al. 2002). These are no doubt the "ultimate" markers for genetic linkage mapping, as potential candidate loci may be identified empirically from known genes that collocate with quantitative trait loci (QTLs). However, the multiplex ratio of these markers remains low, as most genes are mapped one at a time. In contrast, microarray technology holds great potential for genotyping large numbers of anonymous markers in *Eucalyptus* (Lezar et al. 2004) and may form the basis of high-density genetic mapping in the future. The use of short oligonucleotide probes (25 mer) on microarrays may in the future allow direct mapping of genes through the detection of single nucleotide polymorphisms (SNPs) in genes (Kirst 2004).

### 4.2.3
### Map Construction

Relative to most crop plant species, eucalypts are still essentially undomesticated and outbred, and as such highly heterozygous, plant species. Linkage analysis is somewhat complicated by the varying numbers of marker alleles (up to four) that may be present at each locus. Moreover, linkage phases of markers are generally unknown. Despite these difficulties, many pollen and seed parent genetic linkage maps have been constructed in *Eucalyptus* in the last decade (Tables 3 and 4) using informative genetic markers that segregate in appropriate pedigrees. Grattapaglia and Sederoff (1994) implemented the "two-way pseudotestcross model" to construct individual-tree genetic linkage maps of the two parents of an interspecific full-sib cross of *E. grandis* and *E. urophylla*. This model allowed the use of dominant markers and inbred line mapping approaches in this oubred pedigree and resulted in three types of segregating markers: (1) 1:1 testcross markers inherited from the pollen parent, (2) 1:1 testcross markers inherited from the seed parent, and (3) 3:1 intercross markers inherited from both parents. Dominant intercross markers (such as AFLPs) and codominant markers (such as SSRs) have been used to establish homology and investigate synteny between parental maps (Verhaegen and Plomion 1996; Brondani et al. 1998, 2002; Marques et al. 1998, 2002; Bundock et al. 2000; Myburg et al. 2003). Higher-resolution comparative mapping was achieved through a "double pseudobackcross" mapping strategy in $F_2$ interspecific backcross families of *Eucalyptus* (Myburg et al. 2003). This approach allowed the use of shared testcross and intercross markers to align the maps of the $F_1$ hybrid with those of the two backcross parents.

Genetic linkage mapping is based on recombination rates that may differ between male and female meiosis. Myburg et al. (2003) did not observe differences in whole-genome AFLP recombination rates of seed and pollen parents in an *E. grandis* × *E. globulus* hybrid pseudobackcross with *E. grandis* as the male and *E. globulus* as the female parent. Thamarus et al. (2002) also did not find significant differences in the female and male recombination rates among consecutive pairs of fully informative loci. This has also been verified with SSR markers by Brondani et al. (2002).

Maps constructed with outbred progenies and a reasonable number of fully informative

**Table 3.** *Eucalyptus* genetic linkage maps constructed with dominant markers

| Species | Population | Number of markers | Number of linkage groups | Observed map length (cM) | Reference |
|---|---|---|---|---|---|
| *E. grandis* | 62 $F_1$ | 240 RAPD | 14 | 1552 | Grattapaglia and Sederoff 1994 |
| *E. urophylla* | 62 $F_1$ | 251 RAPD | 11 | 1101 | Grattapaglia and Sederoff 1994 |
| *E. gunnii* × *E. globulus* | 72 $F_2$ and 10 BC | 75 RAPD | 9 | 255 | Vaillancourt et al. 1994 |
| *E. urophylla* | 93 $F_1$ | 269 RAPD | 11 | 1331 | Verhaegen and Plomion 1996 |
| *E. grandis* | 93 $F_1$ | 236 RAPD | 11 | 1415 | Verhaegen and Plomion 1996 |
| *E. globulus* | 91 $F_1$ | 268 AFLP | 16 | 967 | Marques et al. 1998 |
| *E. tereticornis* | 91 $F_1$ | 285 AFLP | 14 | 919 | Marques et al. 1998 |
| *E. urophylla* | 82 $F_1$ | 245 RAPD | 23 | 1505 | Gan et al. 2003 |
| *E. tereticornis* | 82 $F_1$ | 264 RAPD | 23 | 1036 | Gan et al. 2003 |
| *E. grandis* | 156 BC | 438 AFLP | 11 (comparative) | 1335 | Myburg et al. 2003 |
| *E. globulus* | 177 BC | 367 AFLP | 11 (comparative) | 1405 | Myburg et al. 2003 |
| *E. grandis* × *E. globulus* $F_1$ paternal | 156 grandis BC | 518 AFLP | 11 (comparative) | 1448 | Myburg et al. 2003 |
| *E. grandis* × *E. globulus* $F_1$ maternal | 177 globulus BC | 577 AFLP | 11 (comparative) | 1318 | Myburg et al. 2003 |

codominant markers (such as RFLP or SSR) have been integrated in species maps for *E. nitens* (Byrne at al. 1995) and *E. globulus* (Thamarus et al. 2002). In most cases, linkage analysis was performed using MAPMAKER (Lander et al. 1987), GMENDEL (Liu and Knapp 1990), or JoinMap (Stam 1993). Often, subsets of markers were selected for their intensity, ease of scoring, reduced missing data, size, and spacing in order to construct framework maps with increased marker order reliability.

### 4.2.4 Physical Genome Size vs. Genetic Map Size

As discussed in the previous section, the first estimates of nuclear DNA content for *Eucalyptus* (Grattapaglia and Bradshaw 1994) ranged from 0.77 pg/2C for *E. citriodora* to 1.47 pg/2C for *E. saligna*, corresponding to a haploid genome size range of 370 to 700 Mbp. At the time it was not clear how physical map size (in bp) would relate to genetic map sizes (in cM), but several mapping experiments have since shed light

**Table 4.** *Eucalyptus* genetic linkage maps constructed with codominant markers

| Species | Population | Number of markers | Number of linkage groups | Observed map length | Reference |
|---|---|---|---|---|---|
| *E. nitens* | 4 grandparents, 2 parents, 118 $F_2$ outbred progeny | 210 RFLP, 125 RAPD, 4 isozyme | 12 (integrated) | 1462 cM | Byrne et al. 1995 |
| *E. grandis* | 94 $F_1$ | 19 SSR | 9 | Added to Grattapaglia and Sederoff 1994 | Brondani et al. 1998 |
| *E. urophylla* | 94 $F_1$ | 17 SSR | 6 | Added to Grattapaglia and Sederoff 1994 | Brondani et al. 1998 |
| *E. urophylla* | 201 $F_1$ | 8 genes | 5 | Added to Verhaegen and Plomion 1996 | Gion et al. 1996 |
| *E. grandis* | 201 $F_1$ | 4 genes | 3 | Added to Verhaegen and Plomion 1996 | Gion et al. 1996 |
| *E. globulus* (KI2) | 165 $F_1$ | 153 RAPD, 16 SSR | 11 | 701 cM | Bundock et al. 2000 |
| *E. globulus* (G164) | 94, 71, 165 $F_1$ | 173 RAPD, 21 SSR, GPI-2 | 13 | 1013 cM | Bundock et al. 2000 |
| *E. globulus* | 73 $F_1$ | 34 SSR | 8 | Added to Marques et al. 1998 | Marques et al. 2002 |
| *E. tereticornis* | 73 $F_1$ | 34 SSR | 8 | Added to Marques et al. 1998 | Marques et al. 2002 |
| *E. grandis* | 92 $F_1$ | 63 SSR | 11 | Added to Grattapaglia and Sederoff 1994 maps | Brondani et al. 2002 |
| *E. urophylla* | 92 $F_1$ | 53 SSR | 10 | Added to Grattapaglia and Sederoff 1994 | Brondani et al. 2002 |
| *E. globulus* | 148 outbred $F_1$ | 204 RFLP, (31 EST and 14 genes) 40 SSR, 5 isozyme | 12 (integrated) | 1375 cM | Thamarus et al. 2002 |

on this question. Observed map lengths for *E. grandis* average 1,434 cM (Grattapaglia and Sederoff 1994; Verhaegen and Plomion 1996; Myburg et al. 2003), 1,312 cM for *E. urophylla* (Grattapaglia and Sederoff 1994; Verhaegen and Plomion 1996; Gan et al. 2003), 1,092 cM for *E. globulus* (Marques et al. 1998; Bundock et al. 2000; Thamarus et al. 2002; Myburg et al. 2003), 978 cM for *E. tereticornis* (Marques et al. 1998; Gan et al. 2003) and 1,462 cM for *E. nitens* (Byrne et al. 1995) (Tables 3 and 4). Differences among reported map lengths for the same species might result from differences in the number of framework markers and the stringency level adopted for assigning locus order and grouping. Based on Grattapaglia and Bradshaw's (1994) genome estimates of 640 Mbp/1C for *E. grandis*, 650 Mbp/1C for *E. urophylla*, 530 Mbp/1C for *E. globulus*, and 580 Mbp/1C for *E. tereticornis*, 1 cM map distance would be equivalent to 446 kbp in *E. grandis*, 495 kbp in *E. urophylla*, 485 kbp in *E. globulus*, and 593 kbp in *E. tereticornis*. In *Arabidopsis*, 1 cM is equivalent on average to 230 kbp, only about half the genetic/physical ratio for eucalypts (Verhaegen and Plomion 1996). Myburg et al. (2003) suggested that the 20% genome size difference between *E. grandis* and *E. globulus* could be the result of many small and dispersed chromosomal duplications or deletions, reflecting a pattern of dispersed genome expansion in *Eucalyptus*.

### 4.2.5 Segregation Distortion

The distorted segregation of markers in genetic maps is generally indicative of the segregation of genomic incompatibilities or genetic effects that result in differential fitness of the gametes or progeny through which marker alleles are transmitted. In general, mapping studies in intraspecific crosses of *Eucalyptus* species (Byrne et al. 1995; Thamarus et al. 2002) have reported lower proportions of distorted markers (6% and 0.8%, respectively) than studies in interspecific crosses. However, Grattapaglia and Sederoff (1994) did not find significant numbers of markers with distorted segregation in interspecific $F_1$ progeny of *E. grandis* and *E. urophylla*. Bundock et al. (2000) reported significant segregation distortion in intraspecific maps of an *E. globulus* female parent, but no distortion in the *E. globulus* male parent. Verhaegen and Plomion (1996) found that 8% of RAPD markers exhibited segregation distortion in an interspecific cross of *E. urophylla* and *E. grandis*. They also observed the clustering of distorted loci, as previously reported in *E. nitens* (Byrne et al. 1995), and subsequently in *E. globulus* (Bundock et al. 2000), *E. urophylla*, and *E. tereticornis* (Gan et al. 2003). Such clustering of loci may indicate a biological basis for the observed distortion (reviewed in Bundock et al. 2000). Marques et al. (1998) reported that 15% of AFLP bands displayed skewed segregation ratios in a cross of *E. tereticornis* and *E. grandis*. Myburg et al. (2004) observed an average of 28% transmission ratio distortion of AFLP markers in the parental maps of a wide interspecific backcross pedigree of *E. grandis* and *E. globulus*. The distorted markers were located in distinct regions of the parental maps. Relatively high levels of distortion were observed in the backcross parents, which was unexpected and suggested that there may be genetic variability within *E. grandis* and *E. globulus* for genetic factors that affect hybrid fitness (Myburg et al. 2003). No overall suppression of recombination was detected in the $F_1$ hybrid, relative to the *E. grandis* and *E. globulus* backcross parents, suggesting that the observed segregation distortion was not related to genomic incompatibilities that may reduce recombination between *E. grandis* and *E. globulus* homologs. Gan et al. (2003) also reported a high proportion of distorted loci (33%) for *E. tereticornis* in an interspecific cross with *E. urophylla*, which was higher than what had been identified in previous studies (Marques et al. 1998). It is evident from these studies that there is great potential to use high-resolution genome mapping to elucidate the genetic basis of intraspecific and interspecific crossing barriers in *Eucalyptus*. The increased use of codominant markers in genetic mapping studies will further allow the differentiation of gametic and zygotic mechanisms of distortion (Myburg et al. 2003).

### 4.2.6 Comparative Mapping

Comparative mapping studies rely on a set of common transferable markers that segregate in multiple species or pedigrees of interest. Comparative genetic mapping is a powerful approach to study genetic differentiation, genomic structure, genome evolution, and reproductive isolation in diverging species. In *Eucalyptus*, the verification of synteny and collinear-

ity will add enormous value to the identification and isolation of genes and the validation of gene and QTL positions in multiple pedigrees. One of the immediate objectives would be to align current genetic linkage maps of species in the subgenus *Symphyomyrtus* and particularly those in the sections *Maidenaria*, *Exsertaria*, and *Latoangulatae*, which contain most of the commercially planted eucalypt species (Table 1).

Verhaegen and Plomion (1996) used 25 RAPD intercross markers to identify homeology between LGs (LGs) in *E. urophylla* and *E. grandis*, both members of the section *Latoangulatae*. Despite the limitations of RAPD markers regarding locus specificity, the authors successfully used some of the RAPD markers that had been previously mapped by Grattapaglia and Sederoff (1994) to identify four homologous LGs in *E. urophylla* and three in *E. grandis*. Marques et al. (1998) used 19 AFLP intercross markers to suggest homeology between LGs in *E. globulus* (section *Maidenaria*) and *E. tereticornis* (section *Exsertaria*). Brondani et al. (1998) developed 20 SSR markers for *Eucalyptus* and verified synteny in six *E. grandis* and *E. urophylla* LGs. With the same set of SSR, homeology was extended between these species and two *E. globulus* maps (Bundock et al. 2000) and, later, an *E. tereticornis* and another *E. globulus* map (Marques et al. 2002). Marques et al. (2002) reported synteny of SSR loci and QTLs for vegetative propagation in four eucalypt species. Brondani et al. (2002) contributed an additional 50 SSR markers for comparative mapping in *Eucalyptus*. These markers were used to establish collinearity and synteny among 10 LGs of *E. grandis* and *E. urophylla*. In their double-pseudotestcross mapping pedigree, Myburg et al. (2003) were able to align maps of all 11 LGs of *E. grandis*, *E. globulus*, and their $F_1$ hybrid, and the authors reported very high collinearity between the genomes of these two species, which represent one of the widest crosses among commercially planted eucalypt species.

Although it is possible that smaller genome rearrangements exist, all of the mapping studies above signify high genome collinearity within the subgenus *Symphyomyrtus*. The morphological similarity and genetic compatibility in sexual crosses among eucalypts from the same subgenus may also suggest high levels of DNA sequence conservation. This is supported by results of SSR primer transferability studies (Byrne et al. 1996; Brondani et al. 2002; Marques et al. 2002), which suggest that prospects for transferring microsatellite mapping information across eucalypt species are excellent. The prospects seem good for obtaining a subgenuswide reference map with SSR and gene-based markers in the foreseeable future.

### 4.2.7 Future Perspective: Integration and Application of Genetic Linkage Maps

Integrated genetic linkage maps (defined below) will provide a framework for the application of molecular breeding, as well as for addressing questions of population and conservation genetics of eucalypt species. The construction of integrated linkage maps in outbred eucalypt pedigrees relies on the use of codominant markers that are informative (polymorphic) in both parental maps of the pedigree, allowing complete integration of the two parental maps. Ideally, such codominant markers will also be informative in other related and unrelated crosses, allowing integration, or at least alignment of multiple genetic linkage maps. If successful, *Eucalyptus* researchers will in the future be able to use genetic information from multiple mapping pedigrees to guide population (association) genetic studies and, ultimately, perform molecular breeding of eucalypt species and hybrids. The construction of a genuswide reference map for *Eucalyptus*, or a series of connected linkage maps that represent the genus, is now both a commercial and an academic research target. A reasonable number (>400) of codominant markers such as microsatellites, gene and EST markers are currently available, allowing the alignment of existing linkage maps in the three commercially important sections of the subgenus *Symphyomyrtus*. The *Genolyptus* project in Brazil is developing and mapping more than 1,200 new SSR markers identified from different sources, including a large EST database and a growing number of BAC-end sequences (Grattapaglia et al. 2004). Some of these markers will allow the anchoring of genomic clone (BAC) contigs to existing genetic linkage maps, supporting the construction of localized physical maps focused on genomic regions in which a significant concentration of QTLs for important traits have been detected. The identification of genes in these regions will assist functional genomics studies, providing insights on the mechanisms of gene expression and regulation. Besides helping to characterize QTLs, the integration of known genes into linkage

maps will also be useful for gross comparative mapping of *Eucalyptus* with model plant species (e.g., *Arabidopsis*) and different tree genera such as *Populus* and *Pinus*. Information exchange and comparison among laboratories and independent experiments will speed candidate gene mapping, QTL verification, and the implementation of marker-assisted selection (MAS) in eucalypt breeding.

## 4.3 QTL Mapping in *Eucalyptus*

### 4.3.1 Historical Perspective on QTL Analysis

The foundation for quantitative trait locus (QTL) analysis was established almost a century ago. Thomas H. Morgan (1910) discovered the principles of linkage mapping and demonstrated that genes were linked and could be placed into groups that were equal in number to that of the haploid chromosomes in *Drosophila*. The genetic distance between genes was later defined by Haldane (1919) based on recombination frequencies. Finally, the connection between segregating genetic elements and quantitative traits was proposed by Karl Sax in 1923, who observed that "inherent size differences are apparently dependent on Mendelizing factors," but that in "most cases many factors are involved and simple ratios are not obtained." More sophisticated approaches were later introduced that used the genotypic information from two adjacent loci to more accurately estimate the effect and position of QTLs in marker intervals (Paterson et al. 1988; Lander and Botstein 1989) and control for the effect of other markers associated with the trait (Jansen and Stam 1994; Zeng 1994). Another limitation of the early days of QTL analysis was the lack of broad marker coverage of mapped genomes. Isozymes and restriction fragment length polymorphisms, or RFLPs (Botstein et al. 1980), partially resolved this limitation. High-density maps were later developed using several PCR-based methods (Mullis and Faloona 1987; Williams et al. 1990; Welsh and McClelland 1990; Vos et al. 1995) and, more recently, high-throughout single nucleotide polymorphism (SNP) genotyping methods (Chen and Sullivan 2003). This section provides a brief overview of the application of QTL mapping in *Eucalyptus*, a summary of the traits that have been targeted for QTL mapping, and a discussion of new integrative genomics approaches for the molecular dissection of quantitative traits in *Eucalyptus*.

### 4.3.2 QTL Mapping in *Eucalyptus*, Limitations and Advantages

Success in QTL identification relies on crossing two individuals with alternative alleles at loci affecting a trait of interest and following the segregation of the parental genomes in the progeny. Multiple genetic loci should be genotyped, with ample genome coverage, so that all of the genomic regions that may affect the trait are evaluated. The first genomewide QTL analysis carried out in plants identified few QTLs with large effect (Paterson et al. 1988, 1991; Stuber et al. 1992), suggesting that loci linked to traits of interest could be rapidly incorporated into traditional breeding programs, through MAS. The early success in crop plants was, however, not expected to be repeated in forest tree species, and the predictions were that QTL identification was going to be limited (Strauss et al. 1992).

The primary limitation of QTL analysis in forest tree genera such as *Eucalyptus* is related to the types of crosses that are the basis for QTL identification methods in agricultural crops (e.g., backcrosses and $F_2$intercrosses). Traditional backcrosses and $F_2$s rely on the creation of inbred lines, which cannot be readily generated for forest tree species due to high genetic load and inbreeding depression. Consequently, a typical full-sib tree family may have up to four alleles segregating at any given locus, often in unknown linkage phase. In conifers, the problem can be partially overcome by the analysis of maternally inherited haploid tissue from the seed megagametophyte (Adams and Joly 1980). This alternative is not possible in *Eucalyptus* and other woody angiosperms. Nonetheless, novel mapping designs, such as the two-way pseudotestcross strategy (Grattapaglia et al. 1994), have allowed genetic mapping and QTL analysis to be pursued in highly heterozygous forest tree species using models and software developed for inbred lines.

Another limitation was associated with the long-lived nature of forest tree species. Perennial plants undergo significant morphological changes throughout development, and different sets of genes may contribute to phenotypic variation from the juvenile to mature phase. Developmental changes in wood quality and growth are very significant in conifers (Zobel

and Sprague 1998). These changes are also observed in most *Eucalyptus* species, although they are less pronounced. Changes in the set of genes that control a quantitative trait may also be due to environmental variation, as perennial plants have to adapt to changing conditions throughout multiple growing seasons. Seasonal and year-to-year variation in sources of biotic and abiotic stress imply that different physiological mechanisms may need to be activated for plant survival. Some genetic loci may be vital for survival and have significant pleiotropic effects on growth in one year while being essentially "unnecessary" the next, when the stress is no longer existent. Changes in the major genetic loci regulating growth and wood density were documented by Verhaegen et al. (1997), over a period of 18 months, in an $F_1$ cross of *E. grandis* and *E. urophylla*. None of the QTLs detected at one time (18, 24, and 36 months of age) could be repeated throughout the entire experiment, suggesting that different loci are contributing to phenotypic variation during different stages of development. Weng et al. (2002) showed in *Pinus* that the variance explained by major QTLs decreases over time, suggesting the increased complexity of quantitative traits with aging of the tree. The contribution of different sets of genes to quantitative variation during development may lead to lower power of detecting QTLs, as a phenotype measured at rotation age essentially represents the cumulative effect of many distinct genes.

QTL mapping trials in forest trees generally sample more environmental variation per site than equivalent experiments in crop plants that can often be grown in high density, or sometimes even in greenhouses. Where replicated trails are possible (clonal propagation of mapping progeny) QTLs may be detected in one site while being absent in other sites. Numerous studies in crop plants have identified common QTLs in several environments and QTLs that are specific to certain sites, suggesting genotype by environment interactions (Paterson et al. 1991; Teulat et al. 2001; Hittalmani et al. 2003; Leon et al. 2003). Studies that evaluate the conservation of QTL in different environments have not been published in forest tree species, to our knowledge. A partnership of industry and governmental research institutions in Brazil, the Genolyptus project, has initiated the planting of several clonally propagated populations in a broad range of locations and environments. This study will shed light on the stability of QTLs in different environments in *Eucalyptus* (D. Grattapaglia personal communication). Because QTL experiments in forest tree species typically require large field plantations that are carried out in field sites that are highly heterogeneous, it is likely that confounding effect from the environment (spatial variation) can diminish the power of QTL detection. Environmental variation can be accounted for by analysing spatial variation in field sites (e.g., Dutkowski et al. 2002) and by assessing phenotypes in multiple sites.

These and other limitations suggested early on that QTL identification of traits of interest for commercial forestry would be largely unsuccessful and possibly would be limited to the analysis of traits of chemical or morphological nature (Strauss et al. 1992). However, early studies, carried out in *Eucalyptus*, *Populus*, and *Pinus*, showed that QTLs for growth and wood-quality traits could be readily identified (Groover et al. 1994; Bradshaw and Stettler 1995; Grattapaglia et al. 1995). Nevertheless, QTL studies carried out in *Eucalyptus* and other forest tree species have typically identified fewer QTLs and explained a smaller proportion of the phenotypic variation, compared to agronomic crops. QTL scans for a variety of traits in crop species have been able to identify an average of four major QTLs jointly explaining ca. 46% of the phenotypic variance (Kearsey and Farquhar 1998). The limited power to detect QTLs in forest tree mapping experiments relative to agricultural species may be due to high environmental variation in tree plantations, smaller populations tested, and developmental variation.

Part of the success in the identification of QTLs is forest tree species is likely due to the wide genetic variation in the material that is commonly used in tree breeding programs. Also, the essentially undomesticated character of forest tree species and large population sizes suggest that the majority of commercially important alleles have not been lost or fixed yet by selection. Although there have been major reductions in the gene pool during seed collections destined for breeding programs in other parts of the world, the gene pool remains essentially intact in natural populations in Australia and adjacent islands. The large population sizes and wide pollen dispersion that are characteristic of outcrossing, undomesticated tree species ensure that new alleles are maintained in the population and not lost by genetic drift. Therefore, QTL studies in eucalypt species typically detect QTLs simply because there is large variation. Efforts based on interspecific hybrids of eucalypts, in particular $F_2$ interspecific progeny (e.g., Myburg 2001), have the added advantage of exploring variation in alleles that

underlie differentiation among species, frequently because of the unique characteristics and environmental adaptation of eucalypt species.

### 4.3.3 QTLs Identified in *Eucalyptus*

QTL mapping in forest trees has been applied to the identification of genetic loci associated with variation in biomass productivity (height, diameter, volume), stem form, wood properties (wood density and composition, fiber traits, bark composition), vegetative propagation, biotic/abiotic stress response, development, foliar chemistry, and inbreeding depression. QTL analysis of transcript levels, measured by cDNA microarray analysis of thousands of genes, has also been reported recently (Kirst et al. 2005). A brief description of the most relevant QTL studies carried out in *Eucalyptus* follows.

#### Biomass Productivity

Traits associated with biomass productivity (growth) have been the most studied by QTL analysis in *Eucalyptus* (Grattapaglia et al. 1996; Byrne et al. 1997a; Verhaegen et al. 1997; Thamarus et al. 2004; Kirst et al. 2004), and tree species in general, because of their commercial importance and ease of phenotyping. A direct comparison among studies is complicated by the fact that they were generally carried out using different pedigrees, markers, cross designs (e.g.,half-sib, full-sib $F_1$, and backcrosses), species, biomass productivity estimators (i.e., height, diameter, volume, and growth rate), ages and QTL detection methods (e.g., single-marker vs. composite interval mapping). Grattapaglia and colleagues (1996) first identified QTLs for circumference at breast height in a 6.5-year-old half-sib population of *E. grandis*. A set of 300 individuals was genotyped with 77 RAPD markers, and 3 QTLs explaining 13.7% of the phenotypic variation were identified. Byrne et al. (1997a) identified three QTLs for seedling height (explaining between 10.3 and 14.7% of phenotypic variation) in two three-generation pedigrees of *E. nitens*. A first indication of the level of conservation of QTLs during the lifetime of *Eucalyptus* was established by Verhaegen et al. (1997) in the analysis of an $F_1$ hybrid of *E. grandis* and *E. urophylla*. A progeny set of 200 $F_1$ individuals was measured at 18, 26, and 38 months for "vigor" (a combination of height, circumference at breast height, and volume), and QTLs were identified for each age. None of the QTLs were detected consistently throughout the entire experiment, but the majority (68%) were detected in two consecutive ages. The phenotypic variation explained ranged from 5 to 14%. More recently, three QTLs for height and diameter growth were identified in a pseudobackcross of *E. grandis* and *E. globulus* (Kirst et al. 2004). The study was part of an effort to integrate gene expression data and QTL analysis to identify genes controlling quantitative variation (described below).

#### Wood Quality

Significant effort has been dedicated to the phenotyping and mapping of QTLs for wood physical and chemical composition traits in *Eucalyptus*. The first wood-quality-trait QTLs were identified by Grattapaglia and colleagues (1996) in a half-sib population of *E. grandis*. Five QTLs that controlled almost half of the genetic variation for wood-specific gravity were identified. These initial studies were limited by the technology for phenotyping wood physical and chemical property traits, which carried high cost and labor requirements. Recently, novel methods for wood-quality phenotyping have been developed and applied in forestry. These include near-infrared spectrometry (Schimleck et al. 1996), SilviScan (x-ray densitometry combined with automated scanning x-ray diffraction and image analysis), mass spectrometry (Evans and Ilic 2001), computer tomography X-ray densitometry (CT scan), and pyrolysis molecular beam mass spectrometry (pyMBMS; Tuskan et al. 1999). Some methods are based on predictions based on measurements of associated traits and can be less precise compared to direct measurement methods. Lack of precision can, however, be compensated in part by the size of the populations that can be surveyed by these methods.

Myburg (2001) demonstrated the application of indirect, high-throughput phenotyping of *Eucalyptus* wood-quality traits by NIR for QTL mapping in a pseudobackcross of *E. grandis* and *E. globulus*. Approximately 300 individuals that had been previously genotyped with AFLP markers were analyzed by NIR, and predictions were made for pulp yield, alkali consumption, basic density, fiber length and coarseness, and several wood chemical properties (lignin, cellulose, and extractives). Basic density was also measured directly. A comparison between QTLs identified by direct and indirect measurements yielded a few common and several distinct QTLs between methods. More recently, Thamarus and colleagues (2004) used novel high-throughput and

traditional methods to quantify wood density, fiber length, pulp yield, and microfiber angle (MFA) in two full-sib families of *E. globulus* that shared a common parent. Pulp yield and cellulose content were determined by NIR, and MFA was quantified by SilviScan. QTLs for all traits could be detected in both populations (with the exception of fiber length), including three QTLs in common genetic regions on both crosses for wood density, one for pulp yield, and one for microfibril angle. The proportion of phenotypic variation explained by the QTLs identified in both crosses ranged from 3.2 to 15.8%.

### Abiotic Stress

*Eucalyptus* species have adapted to a broad range of environmental conditions. *Eucalyptus* species are naturally found in tropical and temperate regions, in high altitude and coastal plains, and in the central Australian desert. The different species exhibit considerable variation for different sources of abiotic stress.

*Eucalyptus* species display considerable variation for salt tolerance, and QTLs for this trait were identified by Dale et al. (2000). Six $F_1$ hybrid populations from three salt-tolerant *E. camaldulensis* and two highly productive, salt-sensitive *E. grandis* genotypes were created and established in field plantations. In addition, five ramets from each of 192 genotypes from one of the populations were established in a greenhouse, and growth and leaf chloride content were phenotyped after a 2-week treatment. QTL analysis of the mapping population identified three QTLs in a RAPD-derived genetic map using multiple-interval mapping. Three QTLs were contributed by the salt-tolerant parent (*E. camaldulensis*) and three by the less tolerant parent (*E. grandis*). Individual QTL effects ranged from 3 to 5% relative to the population mean. The other populations were planted in field trials, but no significant correlation was found between the level of chloride in leaves in the greenhouse population and the same family grown in the field.

Frost tolerance has been another major target for QTL identification (Byrne et al. 1997b; Fullard and Moran 2003), mostly for the purpose of introgressing frost-tolerance alleles from species adapted to temperate regions and high altitude into the fast-growing, commercial species from topical and subtropical areas. Byrne and colleagues (1997b) identified two QTLs explaining 8 and 11% of the variation for frost tolerance in a cross of *E. nitens* involving parents selected from regions of high and low frost incidence. Fullard and Moran (2003) identified QTLs for frost tolerance in a cross involving *E. globulus* (tolerant) and a frost-susceptible hybrid of *E. urophylla* × *E. grandis*.

### Biotic Stress

Inheritance of disease resistance in forest species is in many cases explained by Mendelian factors. Major genes that control tolerance to pathogens in forest tree species were primarily identified in loblolly and sugar pines (Devey et al. 1995; Wilcox et al. 1996; Harkins et al. 1998) and in poplar hybrids (Newcombe et al. 1996; Stirling et al. 2001). Examples of quantitative inheritance have also been identified (Newcombe and Bradshaw 1996; Kubisiak et al. 1997), but the proportion of the phenotypic variation explained by the QTL is typically much larger than for traits such as growth and wood quality.

Very few studies have attempted to identify QTLs for disease resistance in *Eucalyptus* (Shepherd et al. 1995; Freeman et al. 2003; Junghans et al. 2003). However, very compelling evidence for the major gene model observed in pines and poplar has also been demonstrated in *Eucalyptus* (Junghans et al. 2003). Analysis of the progeny from ten full-sib families of *E. grandis* by Junghans and colleagues (2003) suggested the presence of a major locus for rust resistance. Bulked segregant analysis (BSA) of one large full-sib population identified a RAPD marker that cosegregates with the resistance locus and may be useful for MAS or introgression of the resistance locus into other genetic backgrounds.

### Vegetative Propagation

The ease of propagating certain *Eucalyptus* species/genotypes by rooted cuttings is a major advantage for commercial clonal forestry. With vegetative propagation, the genetic components that contribute to a superior genotype, including dominance and epistatic effects, are captured and potentially multiplied into field plantations. There is substantial variation in rooting ability within and among *Eucalyptus* species. Vegetative propagation traits also have reasonable heritabilities, typically higher than growth traits. Grattapaglia et al. (1995) first described the identification of QTLs for fresh weight of micropropagated shoot, stump sprouting, and rooting ability of cuttings in an $F_1$ hybrid population of *E. grandis* and *E. urophylla*. For all traits that were quantified, a large proportion (>60%) of the genetic variation could be explained by the detected QTLs. Marques et al. (1999) car-

ried out a similar analysis in a population of 315 genotypes from a pseudotestcross of *E. tereticornis* and *E. globulus*. *E. tereticornis* typically has higher rooting ability, but its pulping quality is inferior to that of *E. globulus*. A broad variety of traits were evaluated, and QTLs could be identified for mortality of cuttings, adventitious rooting, sprouting ability, and other propagation-related traits. The results from these two studies were compared by regenotyping the two populations with a set of common SSR markers to establish synteny between LGs and align them for comparison of QTL location (Marques et al. 2002). Surprisingly, some QTLs could be detected in homeologous LGs, which could be due to variation in orthologous genes. An ancient allele for vegetative propagation ability could have been fixed in a common ancestor to one of the species in each cross, which would explain why it segregated in the two $F_1$ populations. However, the phylogenetic classification of *E. grandis* and *E. urophylla* into a different section of *Symphyomyrtus* (*Latoangulatae*) than *E. tereticornis* and *E. globulus* (*Exsertaria* and *Maidenaria*, respectively) does not support this hypothesis. Alternatively, considering the broad support interval of QTLs, there is a high probability that two QTLs overlap by chance.

### Developmental Traits

QTLs have been reported for a number of developmental traits in *Eucalyptus*, including flowering precocity (Missiaggia and Grattapaglia 2005) and leaf physical and chemical characteristics (Byrne et al. 1997a; Shepherd et al. 1999).

### QTL Analysis of Gene Expression Levels

QTL analysis in *Eucalyptus* has focused mostly on traits of importance to the forestry industry. However, molecular phenotypes such as transcript (mRNA) and protein expression levels may also be analyzed using traditional QTL detection methods. Transcript abundance variation has been shown to be genetically controlled and heritable (Dumas et al. 2000; Karp et al. 2000; Brem et al. 2002; Wayne and McIntyre 2002; Schadt et al. 2003; Yvert et al. 2003). A genomic description of transcriptional regulation has recently emerged in our work in *Eucalyptus* (Kirst et al. 2004, 2005) following similar demonstrations in yeast and mice (Brem et al. 2002; Schadt et al. 2003; Yvert et al. 2003). Genetic mapping and transcript level information were integrated to define genomic regions involved in the regulation of gene expression by expression QTL (eQTL) analysis. Transcript level variation was measured using a cDNA microarray platform for 2,605 genes in a segregating population of 91 individuals from a pseudobackcross of an $F_1$ hybrid of *E. grandis* and *E. globulus* (Myburg et al 2003), which was backcrossed to a different *E. grandis* parent (Kirst et al. 2005). QTL analysis of gene expression variation in the *E. grandis* backcross family identified genomic regions that harbor regulatory sequences controlling, in *cis*- or *trans*-, the expression of 1,067 genes, or 41% of the genes represented on the microarray. Of the 1,067 genes for which eQTLs were detected, 811 were located in the paternal ($F_1$ hybrid parent) map and 451 in the maternal (*E. grandis* backcross parent) map. The eQTLs for 195 genes mapped to both parental maps, the majority to nonhomologous LGs, suggesting *trans*-regulation by different loci in the two genetic backgrounds. For 821 genes a single eQTL was identified that explained up to 70% of the transcript level variation (Kirst et al. 2005).

The description of the genetic architecture of transcript variation in *Eucalyptus* allowed for inference about the relationship among expressed genes. In several instances, the transcript abundance of genes that were part of the same biochemical pathway (such as the lignin biosynthetic pathway) was shown to be regulated by a single genetic locus (i.e., a shared eQTL, Kirst et al. 2004). Some genes were regulated by multiple genetic loci and shared several eQTLs. Many eQTL hotspots have been shown, in our studies and others (Brem et al. 2002; Yvert et al. 2003; Kirst et al. 2004), to include genes associated with the same metabolic and regulatory pathways, suggesting coordinated regulation of pathway genes by specific regulatory loci. The correlation of gene expression patterns in segregating progeny can also extend our knowledge about genes involved in these pathways. The cDNAs representing previously uncharacterized or hypothetical genes, whose transcript levels are strongly correlated with those of genes with known function, may be associated with the same pathway or biological process. Similarly, new functions can tentatively be assigned to previously characterized genes that had not been described in the context revealing pleiotropic action of these genes. A major limitation in this type of study in *Eucalyptus* is the lack of a completed genome sequence, because without it the relative locations of large numbers of genes and their eQTLs cannot be determined. This information is required to assess whether the genetic control of gene expression variation is in *cis*- or *trans*-, for each gene. The completion

of the current *E. camaldulensis* genome sequencing project will provide the first opportunity to dissect the genetic control of gene expression levels in such a fashion. In the meantime, master eQTLs that control gene expression of important biochemical pathways may have great value for molecular breeding in *Eucalyptus*.

### 4.3.4 Future Perspective: from QTL to Gene

QTL analysis identifies broad genomic regions linked to one or more polymorphisms that cause variation in a phenotype. These broad regions, or QTL support intervals, typically span over 10 to 20 cM or more in most studies carried out in *Eucalyptus* and other forest tree species. Previous estimates of the length of *Eucalyptus* genetic maps have ranged between 1300 and 1,500 cM (Grattapaglia and Sederoff 1994; Byrne et al. 1995; Verhaegen and Plomion 1996; Myburg et al. 2003), with some exceptions (Marques et al. 1998). Therefore, for the haploid genome size of different *Eucalyptus* species (370 to 700 megabase pairs, Grattapaglia and Bradshaw 1994), the average physical distance of 1 cM is expected to range between 300 and 600 kbp. An average QTL would span 3 to 12 Mbp. There are currently no accurate estimates of the number of genes in the *Eucalyptus* genome. Based on the number of genes identified in the sequenced genome of other angiosperms (*Arabidopsis thaliana* ~ 26,000 genes, *Populus trichocarpa* ~ 40,000), an estimate of 30,000 to 50,000 genes seems reasonable. Therefore, a typical QTL interval of 10 to 20 cM may contain anywhere from 200 to 1,000 genes, depending on the precision of the QTL estimate and the gene density in the QTL interval.

Identifying the gene that underlies a QTL remains a major challenge in any organism. The standard strategy has been to fine-scale map the QTL by saturating the region of interest with a large number of makers and evaluating recombination in large progeny sets, to define the location of the QTL at a very high resolution (<1 cM). Chromosome walking can then be undertaken to identify genes in the region. If the resolution is sufficiently high, it can precisely identify the gene that controls the quantitative variation. The large populations that are required and the labor-intensive effort have limited the application of positional cloning to a few examples in crop plants (reviewed in Morgante and Salamini 2003). Positional cloning has not been attempted for any of the QTLs identified in *Eucalyptus*. However, efforts have been made to clone a gene associated with disease-resistance loci in poplar, but the effort has been hindered by the presence of a low-recombination region in the vicinity of the resistance locus. The cost of identifying markers closely linked to quantitative loci and identifying recombinants represents a substantial challenge in forest genetics.

As discussed above, an integrative genomics approach that combines QTL mapping with gene expression analysis may be an alternative to traditional methods of identifying genes underlying QTLs. Genetic analysis of complex traits has normally been carried out by correlating genotypic and phenotypic variation in segregating populations and identifying molecular markers associated with a quantitative trait through QTL analysis and association studies. Complex traits could also be analyzed from the perspective of transcript variation, as it represents an intermediate stage between genotype and phenotype. In contrast to anonymous markers, the transcript levels of thousands of specific genes can be monitored in a segregating population using genomic tools such as microarray analysis that assess genomewide variation in gene expression.

Characterizing the transcriptome of a segregating family can provide valuable information for the dissection of complex traits. Several approaches that integrate genotypic and transcript level variation for the identification of candidate genes have been suggested, but the most powerful one involves the characterization of large progeny sets. Schadt et al. (2003) identified candidate genes by detecting eQTLs that colocalized with QTLs for obesity in mice. Using a similar strategy we have identified candidate genes underlying QTLs for growth and wood density by contrasting the transcript levels for individuals that inherited alternative genotypes for significant phenotypic QTLs in *Eucalyptus* (Kirst et al. 2004). We have also been able to identify specific genetic loci that regulate metabolic pathways and confirmed that variation regulating the expression of metabolically related genes correlates with synthesis of the pathway products (Kirst et al. 2004). For lignin biosynthesis, our previous studies identified genes encoding enzymes of the phenylpropanoid, shikimate, and methionine pathways (all involved in lignin biosynthesis) that have a common eQTL, which overlaps with QTLs for growth in two LGs. Genetic mapping of these genes indicates *trans*- regulation of the pathway genes. These preliminary results indicate that the integra-

tion of different genetic information streams (genetic maps, transcript levels, and traditional phenotypic traits) collected from a segregating population can be highly effective in the dissection of the genetic networks that control variation in complex traits such as lignin content and growth. Since most QTL mapping studies in *Eucalyptus* have been performed in interspecific pedigrees, this integrative genomics approach holds the added promise of elucidating how eucalypt species have become differentiated in the genetic control of gene expression patterns and how this affects observed differences in commercially important traits.

## 4.4 Gene Mapping in *Eucalyptus*

### 4.4.1 From Anonymous Markers to Candidate Genes

The mapping and the characterization of genes involved in the control of complex traits constitute a new stage in our understanding of genome organization and function in forest trees. New strategies exist to go from the phenotype of a complex trait to the gene(s) contributing all, or part of, its variability. One such strategy would be to screen mutants for a given phenotype (morphological, biochemical, or molecular), followed by fine-scale genetic mapping of the mutant and identification of the mutated gene. Forward genetics approaches such as these have, unfortunately, not been feasible in forest trees, mainly due to the long generation times and outbred nature of forest tree species and the difficulty of screening for the adult phenotypes that we are interested in. Alternatively, complex traits can be broken down into elementary components, each with Mendelian inheritance (i.e., quantitative trait loci, or QTLs), particularly at the physiological or biochemical levels (Paterson et al. 1988; Damerval et al. 1994), as discussed in the previous section. These studies can be followed up with the colocalization of QTLs and candidate genes on genetic maps in order to identify positional candidate genes that can be targeted for further confirmatory studies such as association genetic analysis in tree populations (e.g., Thumma et al. 2005), or transgenic studies.

QTL-candidate gene colocalization studies are presently limited by the low resolution of QTL mapping studies performed so far in *Eucalyptus*. As discussed in the previous section, the confidence interval of a QTL is often between 10 and 20 cM (Mangin et al. 1994). In *Eucalyptus*, such an interval could span up to 12 Mbp (Grattapaglia and Bradshaw 1994), i.e., potentially hundreds of genes. The low resolution of QTL characterization is one of the major factors limiting their usefulness in MAS and gene cloning. Nevertheless, of the possible strategies to identify genes underlying QTLs, the candidate gene approach is certainly the simplest. As QTL information is gathered in increasing numbers of unrelated pedigrees, the association of some candidate genes and QTLs may gain enough support to warrant their classification as positional candidates for the traits involved.

Functional candidate genes affecting the expression of a trait of interest can be selected a priori on the basis of the known biochemical and metabolic pathways affecting this trait. The effect of a candidate gene can be estimated using a traditional QTL analysis in which a polymorphism inside the candidate gene is used as an additional genetic marker. Such analyses, of course, cannot exclude the possibility of another closely linked candidate gene in the same QTL interval and has to be followed up by populationwide association studies. The colocalization of QTLs and genes of known function has already been reported in several plant species (Goldman et al. 1994; Causse et al. 1995; Byrne et al. 1997a; Prioul et al. 1997; de Vienne et al. 1999; Pelleschi et al. 1999). In forest trees like *Eucalyptus* species, such strategies are now being developed to more efficiently identify positional candidate genes. These approaches are generally based on the mapping of expressed sequences, i.e., known or unknown genes,with an a priori knowledge of metabolic pathways involved in the trait of interest. The main interest in direct candidate gene mapping is to get around the requirement of linkage disequilibrium between the gene and flanking markers, directly targeting its functional variability. In this section, we briefly summarize the genes and traits of interest for candidate-gene mapping in *Eucalyptus* and we give an overview of the methods used to detect polymorphism and map candidate genes in this genus.

### 4.4.2 Traits and Genes of Interest in *Eucalyptus*

In *Eucalyptus*, growth, wood quality, disease resistance, and vegetative-propagation-related traits are considered the most important commercial traits for

breeding programs. Several genetic factors with major effect on the variation of these traits have been detected by classical QTL analysis (see previous section). However, the identification of candidate genes underlying these QTLs is usually not reported. Indeed, the mapping of known candidate genes is more or less advanced according to the trait of interest.

**Growth**

Several studies have reported the mapping of major QTLs for growth in controlled *Eucalyptus* crosses (Grattapaglia et al. 1996; Byrne et al. 1997a; Verhaegen et al. 1997). Even if the QTL detection indicated functional variability of some genes for growth variation, the identity of these genes remains undetermined. Growth-related traits are a good example of how a candidate gene approach is still difficult. The complexity of growth phenotypes (due to the many metabolic pathways and the relatively large environmental effects that may be involved) makes it very difficult to a priori select putative candidate genes to map. Furthermore, major growth QTLs detected in interspecific pedigrees are often related to hybrid abnormality or viability effects, which are generally not useful for breeding purposes.

**Vegetative Propagation**

Vegetative propagation traits are important for the development and deployment of clonal varieties of *Eucalyptus*. Major QTLs for propagation traits have been detected on *Eucalyptus* genetic maps (Marques et al. 1999, 2002). Gion et al. (2000) mapped two candidate genes regulated by auxin, which are thought to be involved in vegetative propagation traits. These genes were cloned and sequenced from *E. globulus* roots during the symbiosis between *E. globulus* and *Pisolythus* (Carnero Diaz et al. 1996; Nehls et al. 1998). However, the study of colocalization between these genes and propagation-related traits did not reveal any effect of gene variability on phenotypic variation in an interspecific cross of *E. urophylla* and *E. grandis* (Gion 2001).

**Wood Quality**

Wood quality is today the most important trait for which gene discovery is in progress. Many genomic studies have reported the analysis of genes expressed during wood formation and, more particularly, during xylogenesis (Hertzberg et al. 2001; Israelsson et al. 2003; Yang et al. 2003; Egertsdotter et al. 2004; Yang et al. 2004a, b; Paux et al. 2004, 2005; Foucart et al. 2006). Some important metabolic pathways producing the chemical components of wood, like lignin or cellulose, are well known. Indeed, the first example of a wood property for which genomic data were available for association studies was that of lignification genes. Several structural and regulatory genes involved in lignin biosynthesis are known in *Eucalyptus*, including those encoding components of the common phenylpropanoid pathway (PAL, C3H, C4H, COMT, CCoAOMT, and 4CL) and those of the monolignol-specific pathway like CCR and CAD as well lignin regulatory factors such as MYB transcription factors (Goicoechea et al. 2005). All of these genes seem to be good candidates for QTL colocalization studies with wood-quality and lignin-content QTLs.

**Other Traits and Genes**

Some other genes that are not directly linked to the targeted traits of breeding programs, like architectural and floral development traits, have also been localized in *Eucalyptus* (Thamarus et al. 2002). Resistance to rust has been studied in *E. grandis* from Brazil (Junghans et al. 2003), foliar oil composition in *E. grandis* (Shepherd et al. 1999), and frost tolerance in *E. nitens* (Byrne et al. 1997b). More than the knowledge of metabolic pathway involved, the choice of mapped genes depends to a great extent on the availability of gene sequences in the public databases.

### 4.4.3 *Eucalyptus* Species and Populations Used for Gene Mapping

Currently, the genomes of only five of the major *Eucalyptus* species used for plantation forestry have been mapped with relatively high map coverage: *E. globulus*, *E. grandis*, *E. urophylla*, *E. tereticornis*, and *E. nitens* (Tables 3 and 4). Different works have reported a high level of genetic variability in these species for phenotypic data (Potts and Jordan 1994; Chambers et al. 1997; MacDonald et al. 1997) and for molecular data (Byrne et al. 1994, 1996; Martins-Corder and Lopez 1997; Brondani et al. 1998). In spite of the high levels of reported intraspecific variability, other results suggest high conservation at the genome and sequence level between different *Symphyomyrtus* species (Byrne et al. 1996; Gion et al. 2000, 2005), allowing the use of gene sequences from one species to target and map the homologous gene in other species. The relative ease of making interspe-

cific crosses between *Symphyomyrtus* species also suggests high levels of genome conservation, despite differences in genome size of up to 20% (Grattapaglia and Bradshaw 1994). Interspecific hybridization has been exploited in *Eucalyptus* breeding programs to combine desirable traits from different species. This explains why most mapping populations in *Eucalyptus* have been based on interspecific crosses (Tables 3 and 4).

The species generally used for gene sequencing correspond globally to the same *Eucalyptus* species used for genetic mapping. However, the genomes of some species that have been used for gene sequencing like *E. gunnii* and *E. camaldulensis* have not been mapped yet. Fortunately, these species also belong to the *Symphyomyrtus* subgenus (*Maidenaria* and *Exsertaria* sections) and they are closely related to the mapping species. The high degree of marker transferability among *Eucalyptus* species constitutes a unique opportunity for comparative gene and genome mapping in these tree species.

### 4.4.4 Genomic Resources for Gene Mapping in *Eucalyptus*

High-throughput functional genomics efforts such as EST projects are an important source of sequence data to develop new molecular markers for gene mapping (Gupta and Rustgi 2004). The interest in these sequences for genetic mapping depends on their nature and variability.

#### *Eucalyptus* Gene Sequences Available for Marker Development

During the last decade, the number of *Eucalyptus* gene sequences in public databases has increased dramatically (Fig. 1) but trails far behind that available for other trees such as pines and poplar. The first *Eucalyptus* gene sequence deposited into a public database was an mRNA sequence for a lignin gene (CAD) in *E. gunnii* (Feuillet et al. 1993). For the next decade, the number of eucalypt sequences in the database was fewer than one thousand, including several different types of sequences (nuclear, chloroplast, and mitochondrial sequences). The most prevalent sequences were ribosomal sequences used for phylogenetic studies, which represented 39% of available sequences compared to 25% for EST sequences. After 2003, the sequence number increased 40-fold, principally due to EST development from four *Eucalyptus* species: *E. grandis*, *E. tereticornis*, *E. globulus*, and *E. gunnii* (Kirst et al. 2004; Paux et al. 2004; Foucart et al. 2006, P. Sivadon personnal communication). These sequences, which currently represent 92% of the total number of *Eucalyptus* entries, constitute an interesting library for candidate-gene mapping. Around 15,000 entries corresponding to ESTs of *Eucalyptus* are freely available in the EMBL nucleotide database. Private *Eucalyptus* functional genomics projects are also in progress in countries such as Japan (Sato et al. 2005) and Brazil (Grattapaglia 2004), and very large EST databases have been generated by private consortia such as Arborgen. The registration of these sequences in the international databases would constitute a major asset for the international scientific community.

The 15,000 *Eucalyptus* EST and mRNA sequences in EMBL were developed from different types of tissues: flower and carpel (8% of the sequences), seedling (8%), leaf (6%), differentiating xylem (77%), root (<1%), and ectomycorrhiza (<1%). Although xylem ESTs are clearly overrepresented, the multiple origins of around 3,000 ESTs constitute an interesting collection of expressed gene sequences for the mapping of different types of genes affecting traits of interest. The majority of these sequences will be useful for gene mapping because most (99%) are longer than 150 bp, allowing the design of specific primers to amplify genomic DNA in other genotypes and/or species.

Various methods have been used to generate EST databases for *Eucalyptus*. For example, Paux et al. (2004) generated a xylem subtractive library using the suppression subtractive hybridization (SSH) technique (Diatchenko et al. 1996) in order to obtain a cDNA library enriched in xylem-specific sequences. In total, 224 unique sequences (unigenes) were obtained with an average length of 382 bp. The functional classification of these EST sequences, according to the MIPS standard (Paux et al. 2004), revealed a high proportion of sequences classified either as "no hits" (44%) or as "proteins of unknown function" (17%). For those sequences where putative function could be assigned (39%), several functional categories were represented, reflecting the complexity of secondary xylem in woody angiosperms. Some of these ESTs developed from *E. gunnii* were tested for their transferability across the *Eucalyptus* genus (Gion et al. 2005). Primers were designed according to the *E. gunnii* sequences and PCR amplification performed on genomic DNA from 30 different species (3 genotypes/species). The transferability decreases with in-

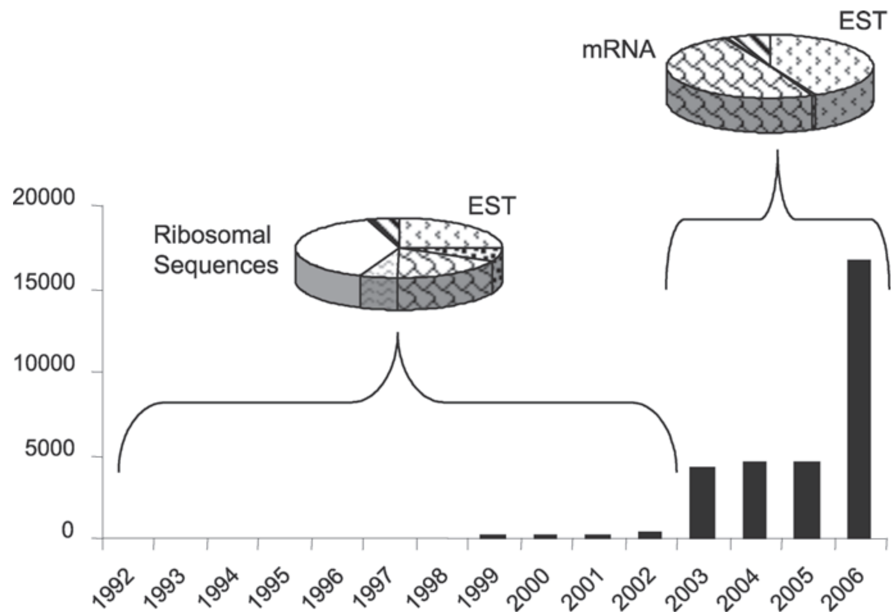


**Fig. 1.** Number of *Eucalyptus* sequences in the EMBL database from 1992 to 2006 (May). The relative proportion of sequence types is indicated from 1992 to 2002 and from 2003 to 2006; EST (◻), chloroplast or mitochondrial genes (◻), full-length mRNA and complete gene (◻), SSR (◻), ribosomal gene (☐), cleaved amplified polymorphic sequence (■), and unassigned DNA sequences (◻)

creasing phylogenetic distance from *E. gunnii*, from 94% for the subgenus *Symphyomyrtus* sections *Maidenaria*, *Latoangulatae*, and *Exsertaria*, to 68% for the subgenus *Monocalyptus* and only 48% for the genus *Corymbia*. This work revealed the high potential of *Eucalyptus* ESTs for comparative mapping and association studies between candidate genes and quantitative traits at the subgenus or genus level.

In the near future, we can expect that the number of available *Eucalyptus* sequences in public databases will continue to increase. The idea of a cooperative effort for the development of *Eucalyptus* genomic resources coordinated by the *Eucalyptus* Genome Network (EUCAGEN) could be a major asset for gene-mapping efforts in *Eucalyptus*. It would then be necessary to develop new technologies that would allow the high-throughput mapping of thousands of ESTs in *Eucalyptus*.

## 4.4.5 Gene Mapping in *Eucalyptus*: Technologies and Perspectives

The first generation of *Eucalyptus* genetic maps was mostly based on dominant, anonymous markers that allowed rapid saturation of LGs in order to achieve high map coverage (Sect. 4.2, Table 3). Several genetic maps were also constructed for *Symphyomyrtus* species using different types of codominant molecular markers (Table 4). These genetic maps have been used to detect the presence of major effect QTLs for complex traits in *Eucalyptus* (discussed in Sect. 4.3.3). More recently, with the availability of *Eucalyptus* genomic resources, it has become feasible to add gene loci to these maps, and several studies have reported the location of important target genes on these maps.

### Genes and Marker Types Used for Gene Mapping

Byrne et al. (1995) reported the first gene mapping results corresponding to four isozyme loci mapped on *E. nitens* genetic maps. The three enzyme systems used were malate dehydrogenase (MDH1 and MDH2), shikimic acid dehydrogenase (SDH), and phosphogluconate dehydrogenase (PDH). The polymorphism revealed in *E. nitens* intraspecific $F_1$ progeny allowed the mapping of the *MDH1*, *MDH2*, *SDH*, and *PDH* gene loci on four different LGs, 6, 1, 12, and 5, respectively. More recently, these same and additional isozyme loci were mapped in the *E. globulus* genome (Thamarus et al. 2002). However, the most commonly used markers for gene mapping in *Eucalyptus* have been PCR-based markers. The recent availability of *Eucalyptus* sequences has allowed the design of specific or degenerate primers for the amplification of the orthologous gene or fragment in the species of interest. Even where the design of *Eucalyptus*-specific primers has not been feasible, degenerate primers based on sequences from other species have been used with success. As an example, the PAL gene was mapped in *E. urophylla* genetic maps by Gion et al. (2000) using degenerated primers developed from ten different accession numbers.

Several genes involved in the lignin biosynthesis pathway have been mapped in *E. grandis*, *E. urophylla*, and *E. globulus* (Gion et al. 2000; Myburg 2001; Thamarus et al. 2002). Some of these genes, such as phenylalanine ammonia-lyase (PAL), Caffeate O-methyltransferase (COMT), 4 coumarate Coenzyme A ligase (4CL), and Caffeoyl coenzyme A O-methyltransferase (CCoAOMT), belong to the common phenylpropanoid pathway. Two others, cinnamyl alcohol dehydrogenase (CAD) and cynnamoyl CoA reductase (CCR), belong to the monolignol-specific pathway. In *E. grandis* and *E. urophylla*, all of these genes were significantly linked (LOD > 11) to previously mapped RAPD markers (Gion et al. 2000). The mapping of lignin biosynthesis genes gives us for the first time an overview of the genome organization in *Eucalyptus* of a set of genes involved in the same biosynthetic pathway. The fact that these genes are located in five distinct genomic regions is a favorable situation for QTL-candidate gene colocalization efforts. It also increases the probability of finding transgressive genotypes with favorable alleles for multiple lignin genes.

In *E. globulus*, 31 cambium-specific ESTs were mapped by Thamarus et al. (2002) and found to be located on 10 different LGs. With an average of 3 ESTs per LG, this EST set will be useful for comparative mapping in other *Eucalyptus* species. Carocha et al. (2004) have started the mapping of 224 expressed candidate genes and 83 additional functional candidate genes in *E. globulus*. A success rate of 90% was achieved with the amplification of targeted gene fragments in *E. globulus* using primers designed on *E. gunnii* sequences. The mapping of the same set of expressional candidate genes is in progress in *E. urophylla* and *E. grandis* in the context of a European collaboration.

Four flowering genes have been mapped in *E. globulus* genetic maps (Moran et al. 2002). *ELF1*, *EAP*, *AGE1*, and *AGE2* were mapped on LG 3, 4, 3, and 9, respectively. *ELF1* was recently also mapped on LG 4 of *E. urophylla* (C. Boudet and J.-M. Gion unpublished data). Junghans et al. (2003) reported the mapping of a monogenic resistance locus for the *Puccinia psidii* resistance gene 1 (*Ppr1*) relative to previously mapped RAPD markers.

#### Methods Used to Detect Polymorphism in *Eucalyptus*

The genomic resources obtained by EST sequencing, or complete cDNA characterization, constitutes an important source of nonanonymous genetic markers in *Eucalyptus*, provided that polymorphism can be detected in these sequences. For the genetic mapping of such sequence tagged sites (STS), several techniques can be used. The simplest methods include the generation of PCR-RFLP, or cleaved amplified polymorphic sequence (CAPS) markers, which is based on the digestion of amplified gene fragments with specific restriction enzymes. The digestion products are then observed on agarose gels after ethidium bromide staining (Tragoonrung et al. 1992). Other techniques, like thermal gradient gel electrophoresis (TGGE, Riesner et al. 1992) or denaturing gradient gel electrophoresis (DGGE), are based on the comparison of the stability of amplified DNA fragments under specific thermal or denaturing conditions. These techniques seem a priori simple, but they require highly controlled conditions. Finally, the most commonly used method in *Eucalyptus* has been the single-strand conformation polymorphism (SSCP) technique, which is based on the specific secondary structure of single-strand DNA under non-denaturing conditions (Orita et al. 1989).

### 4.4.6 Future Perspective: Comparative Gene Mapping and Candidate-Gene Analysis in *Eucalyptus*

In addition to its use for QTL characterization, gene mapping is a useful approach to obtain a set of codominant markers for comparative mapping studies in *Eucalyptus*. For the moment, the number of common markers used for genetic mapping of *Eucalyptus* species is not enough to realize detailed comparative mapping of different eucalypt species. Fewer than ten gene-based markers (PGD, MDH, ELF1, 4CL, COMT, CCR, PAL, CAD, and CCoAOMT) have been mapped in more than one genetic linkage map, allowing the identification of homologies between the maps of different species. To date, only a small number of LGs of *E. globulus*, *E. urophylla*, *E. grandis*, and *E. nitens* could be identified as homologs (Fig. 2). The addition of more ESTs on *Eucalyptus* genetic maps will allow higher-resolution comparative mapping and investigation of genome evolution in this genus.

A major goal of gene mapping in *Eucalyptus* has been to determine the proportion of the variation in quantitative traits that can be explained by the segregation of allelic forms of candidate genes in view of using these genes in molecular breeding. This has been achieved by using candidate genes as molecular

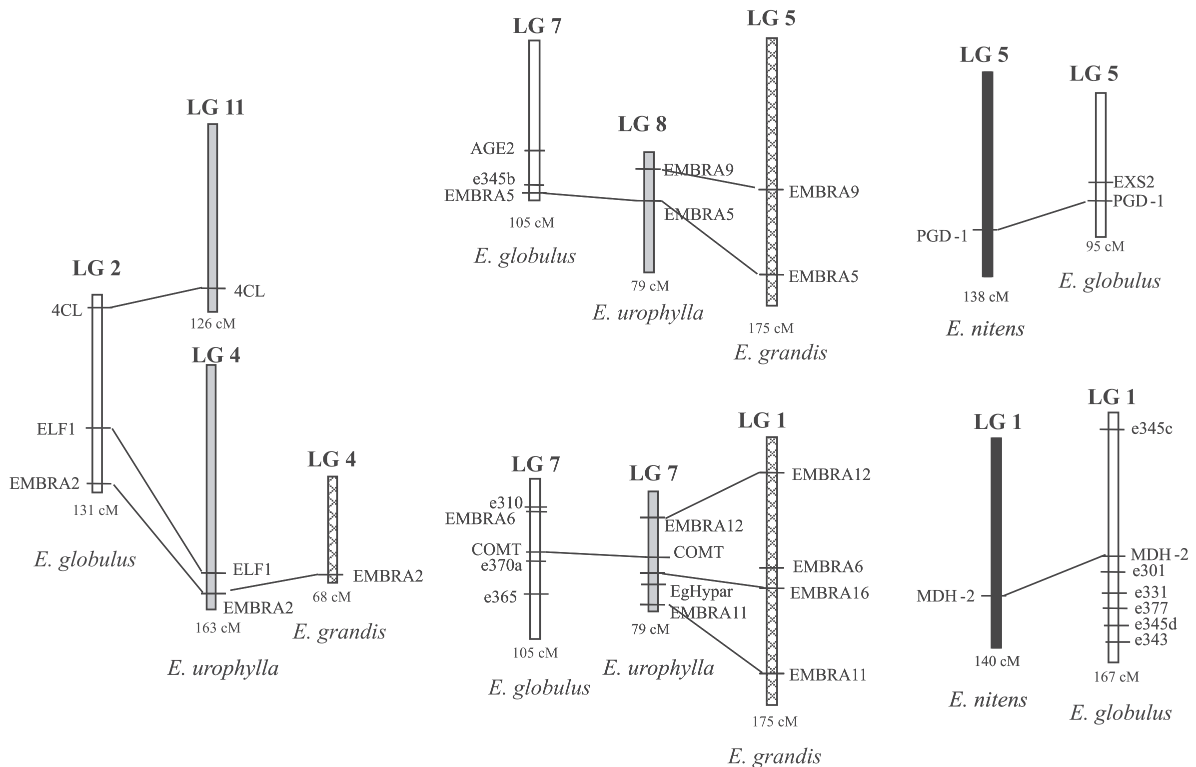


**Fig. 2.** Comparative mapping of LGs of *E. globulus* ( ), *E. urophylla* ( ), *E. grandis* ( ), and *E. nitens* ( ) based on Thamarus et al. (2002), Gion et al. (2000), Brondani et al. (1998), and Byrne et al. (1995), respectively. Only codominant markers are represented on the LGs. A total of five LGs are identified as homologs

markers in a normal QTL analysis and so determine the colocalization of genes and QTLs. Some results on gene and QTL colocalization have been published in *Eucalyptus*. In *E. globulus*, the CCR gene locus was mapped in a region of colocating with pulp yield and cellulose QTLs on LG 10 (Thamarus et al. 2004). Another colocalization was found between the *ECS1* gene, a *Eucalyptus* homolog of cellulose synthase, and a fiber length QTL on LG 8. This gene encodes a key enzyme in the deposition of cellulose microfibrils in plant cell walls. A hydroxymethyltransferase gene was also found to colocate with a wood density QTL on LG 1, while a p-glycoprotein was colocated with a pulp yield QTL on LG 4 (Thamarus et al. 2004). The association of candidate-gene loci with QTLs is suggestive, but not indicative, of functional variability. Even tight marker-QTL linkage may represent large genomic regions that may harbor the actual gene or regulatory sequence underlying the QTL. All of these observed colocalizations need to be validated by association studies in more complex populations or by independent QTL studies in other pedigrees. For example, the colocalization between the CCR gene and a QTL for lignin content in *E. globulus* has also been observed in LG 6 of *E. urophylla* (Gion 2001). In this case, the CCR gene explained 15% of the variation in lignin content observed in an interspecific *E. urophylla* × *E. grandis* family. These results in two different species suggest that the CCR gene may be a good candidate gene for molecular breeding for wood quality in *Eucalyptus*.

## 4.5 Physical Mapping and Map-Based Cloning in *Eucalyptus*

The construction of resources for physical mapping of whole genomes has been a key component of map-based cloning and genome-sequencing efforts in a number of species. Currently, the preferred approach for physical mapping of complex genomes involves the construction of a library of large-insert bacterial artificial chromosome (BAC) clones (Shizuya et al. 1992), followed by their assembly into a structured set using restriction enzyme fingerprinting

technologies (Marra et al. 1997; Schein et al. 2004). The rationale behind physical map construction for any genome is the reliance on random breakage of the genome structure and the subsequent ordering of these genome pieces minimizing gaps and aiming for a most likely order. A physical map can therefore be defined as an ordered set of clones comprising a near-contiguous path across the target genome. Such clone-based maps are used to validate sequence assembly order, supply long-range linking information for assembled sequences, anchor sequences to genetic maps, supply contigs for positional cloning efforts, and provide templates for closing gaps in full genome sequencing. Fingerprint maps are also a critical resource for subsequent functional genomics studies because they provide a redundant and ordered sampling of the genome with clones. In other words, once assembled, physical maps are a key resource in the dissemination of clones by known location for use in several disciplines of biology. This section provides a brief overview of physical mapping methodologies and their application in genome mapping in *Eucalyptus*. The potential use of these technologies for map-based cloning and future molecular breeding in *Eucalyptus* is discussed.

### 4.5.1 Physical Mapping Methodologies and Their Use in *Eucalyptus*

With the technological advances in the ability to clone and fingerprint larger DNA fragments, physical maps have been constructed for a growing number of species. Large-insert BAC fingerprint maps have been developed for humans (McPherson et al. 2001), some animal species including model organisms *Drosophila* (Hoskins et al. 2000), mouse (Gregory et al. 2002), rat (Krzywinski et al. 2004), and economically important species such as chicken (Wallis et al. 2004) and cow (http://www.bcgsc.ca/lab/mapping/bovine). Physical maps have also been published for some plant species including rice (Tao et al. 2001; Chen et al. 2002), sorghum (Klein et al. 2000), *Arabidopsis* (Chang et al. 2001), and soybean (Wu et al. 2004), while work is well advanced for other species such as poplar (http://www.bcgsc.ca/gc/poplar), maize (Coe et al. 2002), wheat (http://wheat.pw.usda.gov), and grape (M. Morgante personal communication).

An excellent and thorough review of the different fingerprinting techniques, detailing the pros and cons of each and the science and statistical issues behind physical map assembly, was recently published (Meyers et al. 2004). Several factors play an important role in the assembly of contigs or a full physical map, including the genome size, the degree of repetitiveness of the genome, the composition and age of repetitive elements, and the fingerprinting method used. The fingerprinting method for building physical maps involves the restriction enzyme digestion of each BAC clone into fragments, which are then separated by electrophoresis and detected. Overlapping clones derived from the same genomic region produce patterns of shared restriction fragments, seen as shared bands on a gel. The proportion of shared bands is indicative of the degree of overlap. This is typically evaluated by the "Sulston cutoff score," i.e., the probability that the number of bands matched between any two clones is coincidental (Sulston et al. 1988). This statistical evaluation and assembly is typically carried out using the only physical map assembly software available today, FPC - Fingerprinted Contigs (Soderlund et al. 2000). The overlap across numerous clones is then used to order the clones into contigs.

Different methods exist to generate clone fingerprints that vary by the number of restriction enzymes used and the detection method. Most published physical maps to date have been constructed using single enzyme digestion, agarose gel electrophoresis, and conventional DNA stains. Variations of this basic method have been published, taking advantage of multiple enzyme digestion, multicolor fluorescent detection technology, and high-resolution capillary electrophoresis (Ding et al. 1999, 2001; Luo et al. 2003). These different methods provide more or less complex fingerprints, i.e., variable numbers of bands that, in turn, have a significant impact on the extent of the overlap among clones and eventually on the quality of the physical map and number of gaps left. Meyers et al. (2004) showed that multiple-enzyme multicolor fluorescent methods that result in more complex fingerprints are superior to single-enzyme, agarose resolution methods. This is intuitive as the observed bands serve as anchors along the BAC clone. The larger the number of anchors available, the greater the confidence in declaring an overlap between clones. Conversely, with larger numbers of bands, a highly significant cutoff score is attained even at lower overlap percentages between clones. This is true, however, only up to a certain number of bands per clone, after which false overlaps will likely occur. Recently, the first physical map constructed using fluorescent fin-

gerprinting technology and capillary electrophoresis was published for *Pennicillium* (Xu et al. 2005). In that work it was shown that the selection of fingerprinting enzymes was crucial to quality map construction and that the fingerprints labeled with one or two colors, resulting in 40 to 70 bands per clone, were assembled into much better quality maps than those labeled with three or four colors.

Physical genome maps are generally built with libraries that offer between 10- and 15-fold redundant coverage, i.e., on average, each region of the map is represented by 10 to 15 clones. Coverage redundancy is critical to achieving map contiguity and is used to provide evidence that individual clones are not cloning artifacts but high-fidelity representations of the underlying genome. Physical maps assembled with higher information content fingerprinting methods will require a lower coverage redundancy as they allow detection of smaller overlaps with high confidence. Furthermore, the way that the BAC clones were produced also determines the final quality of the physical map in terms of gaps. As pointed out by Meyers et al. (2004), two kinds of gaps can occur: (1) assembly gaps resulting from the false-negative rate that is determined with the choice of the cutoff and correspond to our inability to detect existing overlaps between clones; and (2) physical gaps result from regions that are not covered in the clone collection. BAC clones produced by restriction enzyme digestion will have larger gaps and deeper coverage in some regions of the genome than a set of clones produced in a completely random way, such as through mechanical shearing. The use of different libraries produced by digesting genomic DNA with different enzymes will reduce physical gaps that might result from biased restriction site distribution.

In a similar way to genetic map construction, physical map assembly involves initially an automatic step carried out at higher statistical stringency to avoid false positives. Following this automatic assembly, a manual analysis at lower stringency is carried out when contigs are joined to build longer scaffolds. At this stage, linkage and ordering information of mapped microsatellites together with knowledge of the BAC clones to which they were mapped is very valuable for ordering contigs relative to one another. After physical maps are constructed, redundancy is unnecessary for a complete representation of the genome, and most map-driven sequencing efforts use the map to select a minimal tiling path, i.e., a minimal set of overlapping clones that together provides complete coverage across a genomic region. For the human genome, for example, after assembling a physical map with 415,000 BAC clones at a deep coverage of 15×, a set of 32,855 validated clones were selected to cover 99% of the whole genome (Krzywinski et al. 2004).

If on one side genetic mapping of *Eucalyptus* has evolved quite rapidly in the last years, particularly so with the advent of RAPD, AFLP, and, more recently, microsatellite markers, no published report has come out yet on the construction of BAC library resources or physical maps for species of the genus. A complete physical map for *Eucalyptus* will certainly represent a great experimental resource for years to come. With this concept in mind this task was started in the context of the *Genolyptus* project in Brazil (Grattapaglia 2004). A *Eucalyptus grandis* BAC library with over 70% of the inserts averaging 150 kbp has been built (S. Brommonschenkel personal communication). Genomic DNA for library contruction was isolated from a pure *E. grandis* tree of good growth and adaptability in Brazil, and originally from Australian Coffs Harbor provenance. A first set of 20,160 partially digested *Hind*III clones was arrayed, covering an estimated 4× of the 630 Mbp of the *E. grandis* genome. Verification of genome coverage was carried out by PCR using a set of previously mapped single-locus microsatellite markers and STS markers for single-copy genes using a simple bidimensional pooling strategy. Such single-locus markers identified on average two to five BAC clones and provided a satisfactory initial experimental confirmation of the genome coverage of this library. An analysis of cpDNA contamination using chloroplast-specific microsatellites showed it to be less than 1%.

Fingerprinting of the 20,160 *E. grandis* BAC clones is currently under way using the method of Luo et al. (2003). The clones are digested with four rare cutting enzymes and a frequent cutter. Each of the rare cutters leaves a different single-stranded overhang that is then filled in with a distinct, labeled ddNTP using the SNPshot labeling kit (Applied Biosystems). A fifth fluorochrome is used for the internal size standard, and the digested products are resolved on a 5-dye ABI 3100 automated sequencer. To obtain reproducible fingerprints, the DNA minipreps have to yield sufficient and consistent amounts of BAC DNA. As no PCR is involved in the process, the quality of the profile of restriction fragments visualized as fluorescent peaks, i.e., the fingerprint, is fully dependent on the quantity and quality of BAC DNA.

Using the BAC resource, a directed search for the full genomic sequences of a number of candidate genes involved in wood chemical composition has been undertaken (Grattapaglia et al. 2004). Primers were obtained from the literature and from consensus sequences in ORFs derived from *Eucalyptus* EST clusters. A hierarchical pooling strategy was employed to rapidly arrive at the BAC clone containing the gene of interest. BAC DNA midipreps were performed by pooling the 96 clones from a single plate. The 210 pooled DNA samples were, in turn, pooled in 35 superpools of six 96-well plates, i.e., six DNA midiprep samples. Step one of the BAC library screening was carried out with 35 PCR reactions plus the genomic DNA control. The six pools were then screened for a single positive superpool followed by colony-PCR screening of the positive 96-well plate to arrive at the target BAC, thus totaling 140 PCRs to get to a gene. This strategy has allowed us to land on BACs for genes that code for the following proteins: bXylan, UDP-Glucose, Tubulin, Xyloglucan endotransdycosilase, MYB, Pectate lyase, CCR, CCoAOMT, CAD, PAL, COMT, F5H, C4H, and 4CL. Although more than one BAC clone was found for all the genes but 4CL, supporting the 4× coverage, a single BAC was selected for future shotgun sequencing or primer walking. New primer sets pointing out of the amplified segment were designed to provide the full genomic sequence by subsequent walks. Alternatively, shotgun sequencing and assembly of a single BAC clone is being considered by selecting the shortest BAC clone identified in the library that contains the gene of interest. With the full genomic sequence in hand it should be possible to identify regulatory regions and carry out a detailed analysis of polymorphism in a set of individuals by resequencing specific upstream regions in an association mapping approach.

Parallel to clone fingerprinting, both ends of the 20,160 clones have been sequenced. These BAC end sequences could be aligned to the 4× shotgun sequencing now in progress for *E. camaldulensis* at Kazusa DNA Research Institute in Japan (and to the *E. grandis* genome sequence possibly produced by DOE) to provide long-range linking information for assembled sequences and templates for closing gaps in the draft genome sequence. Even though the BAC library is from *E. grandis* and the shotgun sequence from *E. camaldulensis*, the extensive genome homology found across species of the same subgenus *Symphyomyrtus* should allow using the BAC end sequences as useful connectors for closing gaps in interesting regions of the genome. The BAC ends will also be a prolific source of microsatellite and STS markers to help anchor the BAC clones to a linkage map. In a sample sequencing analysis of the *E. grandis* genome based on 7,395 random shotgun clones, Lourenço (2004) used the software TROLL (Castelo et al. 2002) to identify a total of 319 microsatellites for which primers could be designed. Based on this study, it is expected that the almost 40,000 phred-20 BAC end sequences should provide at least 1000 nonredundant microsatellites (R. Brondani personal communication).

In a reverse approach, a set of ∼120 framework mapped microsatellite markers has been used to screen the 4× BAC library with the objective of anchoring specific BACs to a reference linkage map. Microsatellites were selected based on (1) robustness of amplification and allele interpretation; (2) absence of locus duplication sometimes seen for microsatellites in *Eucalyptus*; (3) high information content that allowed the marker to be positioned on three independently constructed linkage maps; (4) their relative map position and ordering, attempting to cover the largest extension of the genome from the recombination standpoint, at regular intervals; (5) fluorochrome labeling of the marker so as to allow multiplexed analysis. The same superpool-pool screening strategy used to fish out gene-containing BAC clones is being used to map microsatellites to BACs with the added advantage of simultaneous screening of several microsatellites labeled with different colors in the same PCR. Due to the 4× redundancy, screening of the BAC clones results in the discovery of more than one positive BAC clone. As not all BAC clones will necessarily fall into contigs during the future physical map assembly, all the BAC clones that contain anchoring microsatellites are sought.

Because most *Eucalyptus* maps have been constructed using nontransferable markers (RAPD or AFLP) or with increasingly less used RFLP probes, current numbering of *Eucalyptus* LGs is not consistent across maps. With the increased use of microsatellites this problem could be mitigated. However, even in the existing maps that include microsatellites (Brondani et al. 2002; Thamarus et al. 2002), LG numbering is attributed arbitrarily. BAC clones containing mapped microsatellite markers, once mapped to chromosomes using FISH, will ultimately allow assigning of the correct number to the LGs of *Eucalyptus* maps according to the chromosome number defined by its pachytene length or centromere position.

*E. grandis* was the obvious choice for a first BAC library and physical map resource as it constitutes the genetic base of most tropically planted elite germplasm. However, due to the special wood properties of the species and the great interest of breeders, a second BAC library from *E. globulus* has been constructed (S. Brommonschenkel personal communication). Accumulated evidence indicates that genomic homology and locus ordering between *E. grandis* and *E. globulus* are very high (Marques et al. 2002). It is therefore not in our immediate plans to build a physical map of *E. globulus* but rather to use the physical/genetic mapping information derived from *E. grandis* to identify and explore specific genomic regions in *E. globulus* by using its BAC library. Based on the map information derived from *E. grandis*, it will be possible, for example, to clone the full homolog gene from *E. globulus* and thus compare in detail both the coding and regulatory regions that could be responsible for differential patterns of gene expression and resulting phenotypic variation.

### 4.5.2 Map-Based Cloning in *Eucalyptus*

Map-based or positional cloning relies on the identification of closely linked markers to the target trait and then uses chromosome walking or landing to identify, isolate, and characterize the gene(s) responsible for the trait (Wicking and Williamson 1991). Map-based cloning requires genetic segregation for the trait of interest and genetic mapping of the trait close to discrete molecular markers. During chromosome walking, BAC clones are screened with a marker (PCR-based or RFLP) that functions as a starting point. New markers are then developed complementary to sequences of the same BAC clone that are adjacent to the starting point, and these are then used to identify additional BAC clones overlapping the one selected as the starting point. The procedure is used repetitively, working away from the starting point. This is a time-consuming approach hindered by the large amounts of DNA that often have to be crossed and by the high frequency of repetitive DNA. Based on the advances of genetic mapping and high-throughput molecular marker technologies, Tanksley et al. (1995) proposed and demonstrated a new paradigm for positional cloning of genes in complex plant genomes. Chromosome landing involves the isolation of one or more DNA markers at a physical distance from the targeted gene that is less than the average insert size of the genomic library being used for clone isolation. The DNA markers are then used to screen the library and isolate (or "land" on) the clone containing the gene, without any need for chromosome walking and its associated problems. Chromosome landing has been the main strategy used for map-based cloning of genes in plant genomes.

Finding and isolating a gene by map-based cloning in a complex genome requires an integrated and complex set of powerful genetic and genomic tools. Genetic maps based on recombination frequencies among markers are the starting point. However, they provide only megabase-level resolution. In *Eucalyptus*, 1% recombination between two markers or between a marker and a gene should correspond on average to 500,000 base pairs of DNA, which can contain several tens of genes (Grattapaglia and Bradshaw 1994). Furthermore, most QTL-mapping experiments in *Eucalyptus* to date have only achieved resolutions of between 10 and 30 cM, still very far from what is needed for efficient positional cloning. The availability of a physical map that is aligned and anchored to a reference genetic map constructed with transferable microsatellites would greatly facilitate positional cloning in *Eucalyptus*. However, even without a complete physical map, by simply screening the BAC library, in a chromosome landing approach, localized physical maps can be constructed representing specific regions of the genome delimited by molecular markers flanking the QTL of interest. This would require much higher resolution genetic mapping than is currently performed in *Eucalyptus* (i.e., mapping populations of more than 1,000 individuals).

Map-based cloning strategies are particularly appropriate when the biochemical basis of the target trait is unclear, thus precluding the use of a candidate-gene approach. Over the last few years a number of genes responsible for QTLs have been cloned in plants, although still restricted to rice, maize, tomato, and *Arabidopsis* (see Morgante and Salamini 2003 for a review). With the exception of the maize *tb1* gene cloned by transposon tagging, the others were all identified by positional cloning. These genes code for transcription factors, proteins involved in metabolism or active in signal transduction pathways. Two of the three genes identified as responsible for QTLs for flowering times in rice were homologs of *Arabidopsis* genes affecting flowering time. Comparative genome analysis of a *Eucalyptus* genome draft to the *Arabidopsis* genome could significantly accelerate the prediction

of the molecular basis of common traits between these two species.

Limited Mendelian genetic analysis has been performed in *Eucalyptus* so far, mostly due to the fact that very few simply inherited traits have been identified in eucalypt species that could be considered as interesting and potential targets for positional cloning. Exceptions could be the recently mapped major QTL for *Puccinia psidii* rust resistance in *E. grandis* (Junghans et al. 2003) and the early-flowering QTL also mapped in *E. grandis* (Missiaggia and Grattapaglia 2005). Most economically relevant traits mapped in *Eucalyptus* are, in fact, multifactorial and controlled by unknown genes that can be genetically mapped as QTLs but not easily identified. Nevertheless, traits related to wood quality typically display high heritability. Major-effect QTLs found and validated for traits such as wood basic density could be targeted for positional cloning. In human genetics the availability of microsatellite maps was a key step toward the subsequent positional cloning of numerous human disease genes (Kong et al. 2002) as it allowed comparing and consolidating segregation data from different affected families. In *Eucalyptus* the situation is similar. With the availability and widespread use of a reference microsatellite map that allows detailed comparative mapping and validation of major QTLs across multiple pedigrees, the perspectives of map-based cloning are tangible. An interesting example could be a QTL for wood density mapped in *E. globulus* by Thamarus et al. (2004). As pointed out in that report, this QTL could be the same as that reported in *E. urophylla/E. grandis* (Verhaegen et al 1997; Gion 2001) since the LGs appear homologous on the basis of common microsatellite markers.

To consider the possibility of positional cloning in *Eucalyptus*, high-resolution mapping experiments are required first, to attempt narrowing down the most probable location of a target QTL to a specific BAC contig. High-resolution mapping involves (1) choosing appropriate segregating pedigrees, (2) generating extended progeny sets of thousands of plants, (3) carrying out accurate phenotyping, and (4) genotyping with high-throughput markers that allow saturation of the target genomic regions followed by the conversion of cosegregating markers to sequence tagged markers. At this level of resolution, crop plants benefit from the availability of near-isogenic lines (NILs) differing specifically for the QTL region and on the analysis of thousands of progenies from their cross. This has been the common way by which genes have been map-based cloned in crop plants and *Arabidopsis*. A more generally applicable alternative has been bulk segregant analysis (BSA), which involves screening phenotypically contrasting bulks for regional mapping (Michelmore et al. 1991).

In genetically heterogeneous *Eucalyptus*, the development of NILs does not seem to be a feasible goal in the near future. BSA, on the other hand, has been successfully used for high-resolution mapping of the *Ppr1* locus that confers resistance to *P. psiidi* rust (Junghans et al. 2003). For complex wood and growth traits, well-planned and large (>2,000) interspecific outbred $F_2$ populations from species with contrasting phenotypes have to be generated. Phenotypic differences between species may be due to fixed alleles within the species so that populations typically segregate abundantly both genetically and phenotypically. $F_2$ backcross populations involving *E. grandis* and *E. globulus* have been successfully used for QTL (Myburg 2001) and expression QTL mapping (Kirst et al. 2004, 2005). Within the genetic limitations of *Eucalyptus*, these would be the best populations for map-based cloning of genes responsible for QTLs. Given the improved methods for controlled pollination (Harbard et al. 1999) ,generating large progeny sizes can be routinely done in *Eucalyptus*. Backcross and intercross $F_2$ families with over 2,000 trees were generated and planted in the experimental network of the *Genolyptus* project in Brazil to be potentially used to clone validated QTLs.

Precise phenotypic assessment of large segregating progeny sets is an essential aspect in the high-resolution mapping step. This should be achievable in *Eucalyptus* for simply inherited binary traits such as disease resistance or flowering mutations. For example, high-resolution mapping allowed landing on single BAC clones containing target genes in heterozygous fruit trees with small genomes such as *Citrus* (Deng et al. 2001) and plum (Claverie et al. 2004). Both reports, however, aimed at discrete disease-resistance phenotypes. Continuously inherited traits such as wood properties represent a challenge. Cloning each one of the segregating progeny individuals would maximize heritability for the traits and provide a way to phenotype plants at a much higher level of accuracy. Cloning would also allow detailed phenotyping by destructive sampling when whole tree measurements are needed. Improved statistical mapping methods will also play a crucial role when attempting to narrow down the most probable location of the QTL and identify a BAC contig comprising the desired gene(s).

High-resolution mapping requires complementary genotyping technologies to microsatellites. Microsatellites are key to arrive at and carry out interpedigree validation of genomic segments that contain QTLs. However, the number of microsatellites available in a particular genomic region is still insufficient to allow fine mapping at below centiMorgan resolution. Maybe in the near future, with the availability of a shotgun draft of *Eucalyptus*, it will be possible to develop thousands of microsatellites or at least have the possibility of mining new microsatellites on demand to target specific regions of interest. At this time, markers such as AFLP or variations thereof that intensively scan the genome for single-nucleotide and indel polymorphism would be the choice (Myburg et al. 2001; Lezar et al. 2004). Alternatively, end sequences of the BAC clones that constitute a contig can be used to derive both microsatellite and STS markers. The success of high-resolution mapping, however, is also very much influenced by the rate of recombination in the target region. An instructive example was reported in poplar where high-resolution mapping was carried out around a gene that confers resistance to poplar leaf rust, a clearly defined phenotype with single-gene inheritance. A high-resolution map was constructed with 19 AFLP markers spanning 2.73 cM. However, a severely reduced recombination rate in the region failed to delimit the *MXC3* locus within a 300-kbp interval defined by the overlapping BAC clones (Stirling et al. 2001). In this case, the most straightforward solution would be to shotgun sequence and assemble the whole BAC contig, which should not be a problem with current costs and throughput of sequencing.

### 4.5.3 Future Perspective: Overcoming Challenges to Map-Based Cloning in *Eucalyptus*

Given the challenges faced in positionally cloning a gene in *Eucalyptus*, efforts should be spent to enrich a reference microsatellite genetic map with hundreds, or even thousands, of candidate genes. Genes mapped and colocalized with QTLs or simply inherited phenotypes would make appealing positional candidate genes. This is true especially for those genes whose deduced function suggests that it could be the source of genetic variation in the trait in question. With the current availability of EST collections and the forthcoming genomic shotgun of *E. camaldulensis*, large numbers of gene sequences will be available to allow the design of PCR-based strategies or oligoarray genotyping (Borevitz et al. 2003) to carry out a massive gene-mapping effort (M. Kirst personal communication).

Finally, once a BAC clone carrying the target gene or a positional candidate gene is identified, an efficient transformation system of *Eucalyptus* will be required to carry out formal complementation tests. Although some transformation protocols have recently been developed for *Eucalyptus* (Kawazu et al. 2003; Tournier et al. 2003), all of these are still genotype dependent. While this is a limitation when planning to transform a specific elite clone, it should not represent a problem in the final stage of a map-based cloning work if the phenotype altered by the cloned gene can be accurately measured in the easily transformable genotype.

In concluding this section, it is clear that there are still several genomic and genetic tools that need to be developed or improved before one can seriously consider map-based cloning in *Eucalyptus*. While the genomic tools should come online in the next few years, the biggest challenge that deserves thinking about and planning still lies in the clear definition of a target gene or QTL and the development and phenotyping of the appropriate populations. QTL mapping and validation together with association mapping experiments will, in effect, be the fundamental information resources for any future map-based cloning effort in *Eucalyptus* and thus should receive continued and improved attention by the eucalypt genome community. The challenge thereafter will be to use this information to develop molecular breeding tools that can be successfully integrated into current eucalypt breeding programs.

## 4.6 Conclusions: Opportunities for Genome Research and Molecular Breeding in *Eucalyptus*

*Eucalyptus* tree species and their hybrids form the basis of the largest hardwood plantation crop in the world. These plantations produce the raw materials for multibillion-dollar processing and manufacturing industries based on wood fiber and timber. In some of these industries, the emphasis is shifting from traditional products such as pulp and paper to novel, lignocellulose-based polymers for use in future processing technologies. With the recent increase in oil-

based energy prices, there is also renewed interest in more efficient use of the renewable biomass (e.g., lignin and hemicellulose) traditionally viewed as low-value byproducts of pulping. The idea of using the tree as a biorefinery for a range of secondary products (e.g., ethanol) will be made feasible as much by the new energy economy as by our ability to breed or genetically engineer trees that are much more suitable for such processing. Fast-growing eucalypt hybrids are already able to capture very large amounts of biomass per unit land. Further genetic improvement and domestication of eucalypt trees will provide a truly renewable source of timber, fiber, energy, and bioproducts for future generations. As discussed throughout this chapter, the completion of the *Eucalyptus* genome sequence and development of associated genomic mapping resources will be important endeavors in the molecular domestication of this important fiber crop. To conclude this chapter, we provide a brief summary of opportunities that exist for genome research and molecular breeding in *Eucalyptus*.

### 4.6.1 A Community Linkage Map

Research communities of model and crop plant species have benefited much from the availability of shared mapping resources such as recombinant inbred line (RIL) populations (Lister and Dean 1993) or doubled-haploid populations (Lu et al. 1996). Such populations have provided a stable source of DNA so that many researchers could contribute to the mapping of markers and genes. The opportunity now exists to generate a shared mapping pedigree for *Eucalyptus*, preferably using the *E. camaldulensis* or *E. grandis* clones donated for genome sequencing in interspecific crosses. The $F_1$ progeny (or later $F_2$ backcross progeny) can be immortalized by maintaining the plants as pruned hedges for future leaf sampling and DNA isolation. Such a resource will allow *Eucalyptus* researchers worldwide to contribute to the mapping of large numbers of genes and markers onto a shared community linkage map, which will provide a useful link between genetic linkage data and the genome sequence (see below).

### 4.6.2 An Integrated Physical and Genetic Linkage Map

The *E. camaldulensis* genome sequencing effort in Japan is proceeding without a genetic or physical map of the particular genotype (T. Hibino personal communication). This means that the genome sequence will eventually be made available as a large number of unordered contigs. An opportunity exists to generate a high-density (and high-resolution) genetic linkage map of *E. camaldulensis* and to use the same markers to anchor sequence (BAC) contigs onto the genetic linkage map. High-throughput marker systems such as AFLP could be used to genotype mapping progeny and BAC pools in order to obtain an integrated genetic and physical mapping framework (Klein et al. 2000). However, microarray-based marker systems such as Diversity Array Technology (DArT, Jaccoud et al. 2001; Lezar et al. 2004) or oligoarray markers (e.g., Borevitz et al. 2003) now hold the promise to simultaneously map thousands of anonymous or gene-based markers onto genetic and physical maps (M. Kirst personal communication). Such a high-resolution, integrated physical and genetic linkage map will facilitate map-based cloning of genes underlying QTLs. The *E. grandis* BAC library and BAC end sequences constructed by the *Genolyptus* project in Brazil (see previous section) will provide further opportunities for cross-linking of genetic and physical (sequence) data.

### 4.6.3 Comparative Genome Mapping

A generic genotyping chip (e.g., DArT array) and a set of highly transferable SSR markers will be useful for comparative genome mapping in *Eucalyptus*. As discussed earlier in the chapter, the first priority would be to construct comparative genetic maps of species in the sections *Maidenaria*, *Exsertaria*, and *Latoangulatae* in the subgenus *Symphyomyrtus*. The availability of comparative maps of the commercially important eucalypts will allow the integration of QTL information and the dissection of QTL multiallelism. Moreover, high-density genotyping (e.g., using microarrays) will provide the opportunity to study detailed, genomewide patterns of genome evolution and reproductive isolation in *Eucalyptus* (Myburg et al. 2004), which will greatly assist hybrid breeding in this genus. A genotyping chip with thousands of markers will also facilitate genomewide selection in intra- and interspecific pedigrees and will be particularly useful for advanced-generation hybrid breeding. It may even make backcross introgression feasible if combined with accelerated breeding techniques.

### 4.6.4 Association Genetics

The successful application of molecular breeding technologies in *Eucalyptus* will depend heavily on our ability to first demonstrate the breeding value of molecular polymorphisms (such as SNPs or SNP haplotypes) in well-designed association genetic studies. The first generation of association genetic studies in forest trees was based on the analysis of allelic diversity in typically fewer than 20 selected candidate genes. These studies have revealed that, although some candidate gene associations can be detected (e.g., Thumma et al. 2005), it is very difficult to predict an appropriate set of candidate genes for any trait of interest. We simply do not understand the distribution of molecular genetic variation in tree genomes well enough to predict the location of trait-altering polymorphisms. Instead, a much larger (and unbiased) set of genes and regulatory sequences have to be interrogated for allelic diversity and trait association. The most valuable associations may be present in genes (or their promoters) that would not necessarily be predicted to be candidates for a trait of interest. It will, however, be far too expensive for average *Eucalyptus* research groups to survey such large numbers of genes using current SNP discovery protocols based on allele sequencing.

Several alternatives to sequencing have been proposed to discover and type allelic polymorphism in a large set of individuals. One very promising technique is Ecotilling (Comai et al. 2004), an adaptation of the mutation detection technology used in Targeting Induced Local Lesions in Genomes (TILLING, Colbert et al. 2001). SNPs, small insertions and deletions, and microsatellites can be efficiently detected and typed in gene regions of up to 800 bp using this technique. It is also a low-cost technology that may provide an efficient way to survey natural variation in *Eucalyptus*. However, array-based genotyping (e.g., Borevitz et al. 2003; Hazen et al. 2005) may be the only approach that will allow sufficient throughput for association genetic analysis of thousands of genes at a time. Ultradense arrays of short (25 nt) oligonucleotides tiling gene and promoter regions may be useful to simultaneously discover and type SNPs and other polymorphisms in tree populations (Kirst 2004). Such arrays may even be useful for molecular breeding once marker-trait associations have been firmly established.

### 4.6.5 Integrative Genomics

The ability to measure the transcript levels of thousands of genes in mapping pedigrees has created the opportunity to study the genetic regulation of global gene expression patterns in *Eucalyptus* (Kirst et al. 2005). As discussed earlier in this chapter, transcript abundance data can be integrated with genetic linkage data to map expression QTLs (eQTLs) underlying transcript-level variation of individual genes. An eQTL represents a polymorphism in or near a gene (i.e., cis-acting), or in a transcription factor of the gene (i.e., trans-acting), that affects the expression level of the gene. Trans-acting polymorphisms in key regulatory genes are visualized by the colocalization (clustering) of eQTLs identified for the individual target genes of the putative transcription factor. When performed in $F_2$ interspecific pedigrees, eQTL mapping becomes a powerful tool to study differentiation in the regulation of gene expression among eucalypt species. Similarly, eQTL studies in intraspecific crosses will expose polymorphisms that underlie gene expression variation within species. Such polymorphisms may explain much of the phenotypic differentiation (within and between species) that is of interest to eucalypt breeders. Furthermore, eQTL-trait associations can be confirmed by the colocation of eQTLs and trait QTLs (Kirst et al. 2004), much like QTL-candidate gene colocation is used to identify positional candidate genes. Markers flanking eQTLs may therefore be valuable tools for marker-assisted breeding, once the breeding value of the eQTLs has been demonstrated.

eQTL studies in *Eucalyptus* are currently constrained by the high cost of microarray analysis and the availability of a whole-transcriptome arrays. Large numbers of arrays have to be performed in order to achieve adequate statistical confidence and power for eQTL detection. At US $300 to US $1,000 per hybridization (cDNA vs. oligoarray technologies) and at least one hybridization per progeny member, the cost of an eQTL mapping experiment will typically be more than US $100,000. Furthermore, eQTL experiments only provide data on gene expression variation in a specific tissue at a specific developmental stage. Whole-transcriptome (>30,000 gene) oligonucleotide arrays will become available for eQTL mapping as the *E. camaldulensis* genome-sequencing effort progresses (T. Hibino personal communincation). In the meantime, several private consortia will be able to

synthesize large (>20,000 gene) arrays based on large EST collections.

Other forms of integrative genomics may also provide opportunities for genome research in *Eucalyptus*. For example, the increasing array of technologies for large-scale analysis of protein abundance may allow the integration of proteome and genome mapping data to identify polymorphisms that underlie variation in protein abundance, i.e., protein expression QTLs or pQTLs (also called protein quantity loci or PQLs, Amiour et al. 2003). Similarly, large-scale analysis of metabolite levels (metabolomics) can be integrated with genome mapping data to detect loci (mQTLs) underlying variation in metabolite levels. Variation in the proteome may be especially informative since transcript abundance is not sufficient to predict the structure, function, amount, and activity of the proteins in the cell. The complexity of the proteome is also much higher than that of the transcriptome. For instance, the human genome, which consists of ca. 30,000 genes, is expected to encode between 200,000 and 2 million proteins, mostly due to alternative splicing and posttranslational modification. Very little is known about the situation in plants, but it is assumed that plant proteome complexity will be of a similar order of magnitude (Rose et al. 2004).

The dissection of the genetic basis of the variation in individual protein amounts may prove to be a very powerful approach to select "candidate proteins," as illustrated first by studies on the effect of water stress in maize (Riccardi et al. 2004; Vincent et al. 2005). More recently, the interest and predictive potential of the PQL/candidate protein approach in forest trees was demonstrated by Plomion et al. (2004). The authors demonstrated the colocalization of the gene encoding glutamine synthetase, the corresponding PQL and a QTL for early height growth providing a strong indication that the enzyme is involved in the control of juvenile growth variation in maritime pine. The integration of transcriptome, proteome, metabolome, and genome mapping data clearly holds great promise for the identification of trait-linked markers for molecular breeding in *Eucalyptus*.

## 4.6.6 Comparative Genomics

The completion of the genome sequences of *Arabidopsis thaliana* (The Arabidopsis Genome Initiative 2000) and *Oryza sativa* (Yu et al. 2001; Goff et al. 2002), the recent public release of the complete genome sequence of *Populus trichocarpa* (http://genome.jgi-psf.org/poptr1/poptr1.home.html), the ongoing *E. camaldulensis* genome sequencing effort, and a possible DOE-funded *E. grandis* genome-sequencing effort will provide many excellent opportunities for comparative genomics research. Fundamental questions can be addressed regarding the evolution of woody and herbaceous growth forms. Comparative analysis of the *Eucalyptus* and poplar genomes will allow researchers to determine whether these two woody genera have evolved different solutions for challenges such as being large, long-lived organisms facing a multitude of biotic and abiotic stresses throughout their lifetimes.

Poplar was chosen to be sequenced as a model tree genome because it fulfilled a number of criteria that were required for a model system: (1) a relatively small genome size (500 Mbp), only 4 times larger than the genome of *Arabidopsis*, but 50 times smaller than the genome of pine; (2) diploid inheritance ($n = 19$); (3) facile clonal propagation; (4) rapid growth; and (5) the availability of an efficient transformation procedure via *Agrobacterium*. Although most eucalypt species have very similar characteristics such as relatively small genome sizes (ranging from 370 to 700 Mbp), diploid inheritance ($n = 11$), facile clonal propagation, and fast growth, the lack of efficient genetic transformation methods has hampered *Eucalyptus* from becoming the preferred model tree. Moreover, the huge commercial potential of eucalypts has fostered a situation in which access to genomic resources is restricted to a small number of private research consortia. These limitations may be overcome with the eventual public release of the *Eucalyptus* genome sequence and the development of public resources for genome research by the *Eucalyptus* Genome Network (EUCAGEN). Nevertheless, *Eucalyptus* research will benefit tremendously from model studies in *Arabidopsis* and poplar, which will help to focus hypothesis-driven research in *Eucalyptus* and expedite the development of molecular breeding tools for this important fiber crop.

## References

Adams WT, Joly RJ (1980) Genetics of allozyme variants in loblolly pine. J Hered 71:33–40

Ahuja MR (2001) Recent advances in molecular genetics of forest trees. Euphytica 121:173–195

Amiour N, Merlino M, Leroy P, Branlard G (2003) Chromosome mapping and identification of amphiphilic proteins of hexaploid wheat kernels. Theor Appl Genet 108:62–72

Apiolaza LA, Raymond CA, Yeo BJ (2005) Genetic variation of physical and chemical wood properties of *Eucalyptus globulus*. Silvae Genet 54:160–166

Ashton DH (2000) Ecology of eucalypt regneration. In: Brown BN (ed) Diseases and Pathogens of Eucalypts. CSIRO, Collingwood, Australia, pp 35-46

Astorga R, Soria F, Basurco F, Toval G (2004) Diversity analysis and genetic structure of *Eucalyptus globulus* Labill. In: Borralho NMG, Pereira JS, Marques C, Coutinho J, Madeira M, Tomé M (eds) '*Eucalyptus* in a Changing World'. RAIZ, Instituto Investigação de Floresta e Papel, Aveiro, Portugal, pp 351–358

Barbour R, Potts BM, Vaillancourt RE, Tibbits WN, Wiltshire RJE (2002) Gene flow between introduced and native *Eucalyptus* species. New For 23:177–191

Barbour RC, Potts BM, Vaillancourt RE (2003) Gene flow between introduced and native *Eucalyptus*: exotic hybrids are establishing in the wild. Aust J Bot 51:429–439

Barbour RC, Potts BM, Vaillancourt RE (2005) Pollen dispersal from exotic eucalypt plantations. Conserv Genet 6:253–257

Bermúdez Alvite JD, Touza Vázquezm MC, Infante FS (2002) Manual da Madeira de Eucalipto Comum. CIS-Madera, Xunta de Galicia

Binkley D, Stape JL (2004) Sustainable management of eucalypt plantations in a changing world. In: Tomé M (ed) IUFRO Conference: *Eucalyptus* in a Changing World, 11-15 Oct 2004, RAIZ, Instituto Investigação de Floresta e Papel, Aveiro, Portugal

Boland DJ, Brooker MIH, Chippendale GM, Hall N, Hyland BPM, Johnston RD, Kleinig DA, Turner JD (1985) Forest Trees of Australia. Australian Government Publishing Service, Melbourne, Australia

Boland DJ, Brophy JJ, House APN (1991) *Eucalyptus* leaf oils. Inkata, Melbourne/Sydney, Australia

Borevitz JO, Liang D, Plouffe D, Chang HS, Zhu T, Weigel D, Berry CC, Winzeler E, Chory J (2003) Large-scale identification of single-feature polymorphisms in complex genomes. Genome Res 13:513–523

Borralho NMG (2001) The purpose of breeding is breeding for a purpose. In: Barros S (ed) Developing the Eucalypt of the Future. IUFRO International Symposium, 10-15 Sept 2001, INFOR, Chile

Borralho NMG, Dutkowski GW (1998) Comparison of rolling front and discrete generation breeding strategies for trees. Can J For Res 28:987–993

Borralho NMG, Cotterill PP, Kanowski PJ (1993) Breeding objectives for pulp production of *Eucalyptus globulus* under different industrial cost structures. Can J For Res 23:648–656

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32:314–331

Bowman D (1998) Tansley Review No. 101 - The impact of Aboriginal landscape burning on the Australian biota. New Phytol 140:385–410

Bradshaw HD, Stettler RF (1995) Molecular genetics of growth and development in Populus. 4. Mapping QTLs with large effects on growth, form, and phenology traits in a forest tree. Genetics 139:963–973

Brem RB, Yvert G, Clinton R, Kruglyak L (2002) Genetic dissection of transcriptional regulation in budding yeast. Science 296:752–755

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

Brondani RPV, Brondani C, Grattapaglia D (2002) Towards a genus-wide reference linkage map for *Eucalyptus* based exclusively on highly informative microsatellite markers. Mol Genet Genom 267:338–347

Brooker MIH (2000) A new classification of the genus *Eucalyptus* L'Her. (Myrtaceae). Aust Syst Bot 13:79–148

Brune A, Zobel BJ (1981) Genetic base populations, gene pools and breeding populations for *Eucalyptus* in Brazil. Silvae Genet 30:146–149

Bundock PC, Hayden M, Vaillancourt RE (2000) Linkage maps of *Eucalyptus globulus* using RAPD and microsatellite markers. Silvae Genet 49:223–232

Burley J, Kanowski PJ (2005) Breeding strategies for temperate hardwoods. Forestry 78:199–208

Burrows GE (2002) Epicormic strand structure in *Angophora*, *Eucalyptus* and *Lophostemon* (Myrtaceae) - implications for fire resistance and recovery. New Phytol 153:111–131

Butcher PA, Otero A, McDonald MW, Moran GF (2002) Nuclear RFLP variation in *Eucalyptus camaldulensis* Dehnh. from northern Australia. Heredity 88:402–412

Butcher PA, Williams ER (2002) Variation in outcrossing rates and growth in *Eucalyptus camaldulensis* from the Petford region, Queensland; evidence of outbreeding depression. Silvae Genet 51:6–12

Byrne M (1999) High genetic identities between three oil mallee taxa, *Eucalyptus kochii* ssp. *kochii* ssp. *plenissima* and *E. horistes*, based on nuclear RFLP analysis. Heredity 82:205–211

Byrne M, Macdonald B (2000) Phylogeography and conservation of three oil mallee taxa, *Eucalyptus kochii* ssp *kochii*, ssp *plenissima* and *E. horistes*. Aust J Bot 48:305–312

Byrne M, Moran GF, Tibbits WN (1993) Restriction map and maternal inheritance of chloroplast DNA in *Eucalyptus nitens*. J Hered 84:218–220

Byrne M, Moran GF, Murrell JC, Tibbits WN (1994) Detection and inheritance of RFLPs in *Eucalyptus nitens*. Theor Appl Genet 89:397–402

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

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

Byrne M, Murrell JC, Owen JV, Kriedemann P, Williams ER, Moran GF (1997a) Identification and mode of action of quantitative trait loci affecting seedling height and leaf area in *Eucalyptus nitens*. Theor Appl Genet 94:674–681

Byrne M, Murrell JC, Owen JV, Williams ER, Moran GF (1997b) Mapping of quantitative trait loci influencing frost tolerance in *Eucalyptus nitens*. Theor Appl Genet 95:975–979

Byrne M, Parrish TL, Moran GF (1998) Nuclear RFLP diversity in *Eucalyptus nitens*. Heredity 81:225–233

Campinhos E, Ikemori YK (1977) Tree improvement program of *Eucalyptus* spp.: preliminary results. In: Third World Consultation on Forest Tree Breeding. CSIRO, Canberra, Australia, pp 717-738

Campinhos E, Ikemori YK (1989) Selection and management of the basic population *Eucalyptus grandis* and *E. urophylla* established at Aracruz for the long term breeding programme. In: Matheson AC (ed) IUFRO Conference, Breeding Tropical Trees: Population Structure and Genetic Improvement Strategies in Clonal and Seedling Forestry. Oxford Forestry Institute, Oxford, UK and Winrock International, Arlington, VA, Pattaya, Thailand, pp 169–175

Cañas I, Soria F, López G, Astorga R, Toval G (2004) Genetic parameters for rooting trait in *Eucalyptus globulus* (Labill.). In: Tomé M (ed) IUFRO Conference: *Eucalyptus* in a Changing World, 11-15 Oct 2004, RAIZ, Instituto Investigação de Floresta e Papel, Aveiro, Portugal, pp 159–160

Carnero Diaz M, Martin F, Tagu D (1996) Eucalypt alpha-tubulin: cDNA cloning and increased level of transcripts in ectomycorrhizal root system. Plant Mol Biol 31:905–910

Carocha VJ, Melo M, Araújo JA, Borralho N, Marques CMP (2004) Candidate gene mapping in *E. globulus*. In: Tomé M (ed) IUFRO Conference: *Eucalyptus* in a Changing World, 11-15 Oct 2004, RAIZ, Instituto Investigação de Floresta e Papel, Aveiro, Portugal, pp 359-372

Castelo AT, Martins W, Gao GR (2002) TROLL - tandem repeat occurrence locator. Bioinformatics 18:634–636

Causse M, Rocher JP, Pelleschi S, Barriere Y, De Vienne D, Prioul JL (1995) Sucrose phosphate synthase: an enzyme with heterotic activity correlated with maize growth. Crop Sci 35:995–1001

Cauvin B, Potts BM (1991) Selection for extreme frost resistance in *Eucalyptus*. In: Schönau APG (ed) Intensive Forestry: The Role of Eucalypts. Proceedings of the IUFRO Symposium, P202-01 Productivity of Eucalypts, 2-6 Sept 1991. Southern African Institute of Forestry, Durban, South Africa, vol 1, pp 209–220

Chambers PGS, Borralho NMG (1997) Importance of survival in short-rotation tree breeding programs. Can J For Res 27:911–917

Chambers PGS, Potts BM, Tilyard PG (1997) The genetic control of flowering precocity in *Eucalyptus globulus* ssp. *globulus*. Silvae Genet 46:207–214

Chang YL, Tao Q, Scheuring C, Ding K, Meksem K, Zhang HB (2001) An integrated map of *Arabidopsis thaliana* for functional analysis of its genome sequence. Genetics 159:1231–1242

Chen M, Presting G, Barbazuk WB, Goicoechea JL, Blackmon B, Fang G, Kim H, Frisch D, Yu Y, Sun S, Higingbottom S, Phimphilai J, Phimphilai D, Thurmond S, Gaudette B, Li P, Liu J, Hatfield J, Main D, Farrar K, Henderson C, Barnett L, Costa R, Williams B, Walser S, Atkins M, Hall C, Budiman MA, Tomkins JP, Luo M, Bancroft I, Salse J, Regad F, Mohapatra T, Singh NK, Tyagi AK, Soderlund C, Dean RA, Wing RA (2002) An integrated physical and genetic map of the rice genome. Plant Cell 14:537–545

Chen X, Sullivan PF (2003) Single nucleotide polymorphism genotyping: biochemistry, protocol, cost and throughput. Pharmacogenomics J 3:77–96

Claverie M, Dirlewanger E, Cosson P, Bosselut N, Lecouls AC, Voisin R, Kleinhentz M, Lafargue B, Caboche M, Chalhoub B, Esmenjaud D (2004) High-resolution mapping and chromosome landing at the root-know nematode resistance locus Ma from Myrobalan plum using a large-insert BAC DNA library. Theor Appl Genet 109:1318–1327

Coe E, Cone K, McMullen M, Chen SS, Davis G, Gardiner J, Liscum E, Polacco M, Paterson A, Sanchez-Villeda H, Soderlund C, Wing R (2002) Access to the maize genome: an integrated physical and genetic map. Plant Physiol 128:9–12

Colbert T, Till BJ, Tompa R, Reynolds S, Steine MN, Yeung AT, McCallum CM, Comai L, Henikoff S (2001) High-throughput screening for induced point mutations. Plant Physiol 126:480–484

Comai L, Young K, Till BJ, Reynolds SH, Greene EA, Codomo CA, Enns LC, Johnson JE, Burtner C, Odden AR, Henikoff S (2004) Efficient discovery of DNA polymorphisms in natural populations by Ecotilling. Plant J 37:778–786

Costa e Silva J, Borralho NMG, Potts BM (2004) Additive and non-additive genetic parameters from clonally replicated and seedling progenies of *Eucalyptus globulus*. Theor Appl Genet 108:1113–1119

Costa e Silva J, Dutkowski GW, Borralho NMG (2005) Across-site heterogeneity of genetic and environmental variances in the genetic evaluation of *Eucalyptus globulus* trials for height growth. Ann For Sci 62:183–191

Cotterill P, Macrae S, Brolin A (1999) Growing eucalypt for high-quality papermaking fibres. Appita J 52:79–83

Coutinho TA, Wingfield MJ, Alfenas AC, Crous PW (1998) *Eucalyptus* rust: a disease with the potential for serious international implications. Plant Dis 82:819–825

Crisp M, Cook L, Steane D (2004) Radiation of the Australian flora: what can comparisons of molecular phylogenies across multiple taxa tell us about evolution of diversity in present-day communities? Philos Trans R Soc Lond B Biol Sci 359:1551–1571

Dale GT (2002) Salt tolerant eucalypts for commercial forestry: Progress and promise. In: Private Forestry - Sustainable, Accountable and Profitable, Session Paper 351. Proc of the

Australian Forest Growers, 2002 National Conference, 13-16 Oct 2002, Albany, Western Australia

Dale GT, Aitken KS, Sasse JM (2000) Development of salt-tolerant *E. camaldulensis* × *E. grandis* hybrid clones using phenotypic selection and genetic mapping. In: Proc QFRI/CRC-SPF Symp - Hybrid breeding and genetics of forest trees, Noosa, Australia, pp 227–233

Damerval C, Maurice A, Josse JM, de Vienne D (1994) QTL underlying gene product variation: A novel perspective for analysing regulation of genome expression. Genetics 137:289–301

de Assis TF (2000) Production and use of *Eucalyptus* hybrids for industrial purposes. In: Nikles DG (ed) Hybrid Breeding and Genetics of Forest Trees. Proc QFRI/CRC-SPF Symp, 9-14 April, Noosa, Dept of Primary Industries, Brisbane, Australia, pp 63–74

de Assis TF (2001) The evolution of technology for cloning *Eucalyptus* in a large scale. In: IUFRO Conf: Developing the Eucalypt of the Future, 10-15 Sept 2001, INFOR, Valdivia, Chile

de Assis TFd, Warburton P, Harwood C (2005) Artificially induced protogyny: an advance in the controlled pollination of *Eucalyptus*. Aust For 68:27–33

de Resende MDV, Thompson R (2004) Factor analytic multiplicative mixed models in the analysis of multi-environment forest field trials. In: Tomé M (ed) IUFRO Conference: *Eucalyptus* in a Changing World, 11-15 Oct 2004, RAIZ, Instituto Investigação de Floresta e Papel, Aveiro, Portugal, pp 123–129

de Vienne D, Leonardi A, Damerval C, Zivy M (1999) Genetics of proteome variation for QTL characterization: application to drought-stress responses in maize. J Exp Bot 50:291–302

Deng Z, Huang S, Ling P, Yu C, Tao Q, Chen C, Wendell MK, Zhang HB, Gmitter FG Jr (2001) Fine genetic mapping and BAC contig development for the citrus tristeza virus resistance gene locus in *Poncirus trifoliata* (Raf.). Mol Genet Genom 265:739–747

Devey ME, Delfino-Mix A, Kinloch, BB Jr, Neale DB (1995) Random amplified polymorphic DNA markers tightly linked to a gene for resistance to white pine blister rust in sugar pine. Proc Natl Acad Sci USA 92:2066–2070

Diatchenko L, Lau YFC, Campbell AP, Chenchik A, Moqadam F, Huang B, Lukyanov S, Lukyanov K, Gurskaya N, Sverdlov ED, Siebert PD (1996) Suppression subtractive hybridisation: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. Proc Natl Acad Sci USA 93:6025–6030

Dillner B, Ljunger Å, Herud OA, Thune-Larsen E (1971) The breeding of *Eucalyptus globulus* on the basis of wood density, chemical composition and growth rate. Timber Bull Eur 23:120–151

Ding Y, Johnson MD, Colayco R, Chen YJ, Melnyk J, Schmitt H, Shizuya H (1999) Contig assembly of bacterial artificial chromosome clones through multiplexed fluorescence-labeled fingerprinting. Genomics 56:237–246

Ding Y, Johnson MD, Chen WQ, Wong D, Chen YJ, Benson SC, Lam JY, Kim YM, Shizuya H (2001) Five-color-based high-information-content fingerprinting of bacterial artificial chromosome clones using type IIS restriction endonucleases. Genomics 74:142–154

Doran JC, Bell RE (1995) Influence of genetic factors on yield of monoterpenes in leaf oils of *Eucalyptus camaldulensis* and implications for tree breeding. New For 8:363–373

Doughty RW (2000) The *Eucalyptus*. A natural and commercial history of the gum tree. The Johns Hopkins Univ Press, Baltimore, MD and London, UK

Downes GM, Hudson IL, Raymond CA, Dean GH, Michell AJ, Schimleck LR, Evans R, Muneri A (1997) Sampling plantation eucalypts for wood and fibre properties. CSIRO, Melbourne, Australia

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Dumas P, Sun Y, Corbeil G, Tremblay S, Pausova Z, Kren V, Krenova D, Pravenec M, Hamet P, Tremblay J (2000) Mapping of quantitative trait loci (QTL) of differential stress gene expression in rat recombinant inbred strains. J Hypertens 18:545–551

Dutkowski GW, Potts BM (1999) Geographical patterns of genetic variation in *Eucalyptus globulus* ssp. *globulus* and a revised racial classification. Aust J Bot 46:237–263

Dutkowski GW, Costa e Silva J, Gilmour AR, Lopez GA (2002) Spatial analysis methods for forest genetic trials. Can J For Res 32:2201–2214

Egertsdotter U, van Zyl LM, MacKay J, Peter G, Kirst M, Clark C, Whetten R, Sederoff R (2004) Gene expression during formation of earlywood and latewood in loblolly pine: expression profiles of 350 genes. Plant Biol 6:654–663

Eldridge K, Davidson J, Harwood C, van Wyk G (1993) Eucalypt domestication and breeding. Clarendon, Oxford, UK

Evans R, Ilic J (2001) Rapid prediction of wood stiffness from microfibril angle and density. For Prod J 51:53–57

Fernandez J, Toro MA (2001) Controlling genetic variability by mathematical programming in a selection scheme on an open-pollinated population in *Eucalyptus globulus*. Theor Appl Genet 102:1056–1064

Ferreira ME, Grattapaglia D (1995) Introdução ao Uso de Marcadores RAPD e RFLP em Análise Genética. EMBRAPA - Cenargen, Brazil

Feuillet C, Boudet AM, Grima-Pettenati J (1993) Nucleotide sequence of a cDNA encoding cinnamyl alcohol dehydrogenase from *Eucalyptus*. Plant Physiol 103(4):1447

Florence RG (1996) Ecology and Silviculture of Eucalypt Forests. CSIRO, Collingwood, Australia

Foucart C, Paux E, Ladouce N, San-Clemente H, Grima-Pettenati J, Sivadon P (2006) Transcript profiling of a xylem vs phloem cDNA subtractive library identifies new genes expressed during xylogenesis in *Eucalyptus*. New Phytol 170:739–752

Franklin EC (1986) Estimation of genetic parameters through four generations of selection in *Eucalyptus grandis*. In: IUFRO Joint Meeting of Working Parties on Breeding Theory, Progeny Testing and Seed Orchards, 12-17 Oct 1986, Williamsburg, VA

Freeman JS, Milgate AW, Vaillancourt RE, Mohammed CL, Potts BM (2003) Genetic control of resistance to *Mycosphaerella* leaf disease in *Eucalyptus globulus*. Abstr 19th Genet Congr, 6-11July 2003, Melbourne, Australia, p 155

Fukuoka S, Namai H, Okuno K (1998) RFLP mapping of the genes controlling hybrid breakdown in rice (*Oryza Sativa* L.). Theor Appl Genet 97:446–449

Fullard KJ, Moran G (2003) Identification of frost tolerance genes in a *Eucalyptus* hybrid cross. Tree Biotechnology 2003, Umeå, Sweden

Gan S, Shi J, Li M, Wu K, Wu J, Bai J (2003) Moderate-density molecular maps of *Eucalyptus urophylla* S.T. Blake and *E. tereticornis* Smith genomes based on RAPD markers. Genetica 118:59–67

Gion J-M (2001) Etude de l'architecture génétique des caractères complexes chez l'*Eucalyptus*: des marqueurs anonymes aux gènes candidats. PhD thesis, University of Rennes I, France

Gion J-M, Rech P, Grima-Pettenati J, Verhaegen D, Plomion C (2000) Mapping candidate genes in *Eucalyptus* with emphasis on lignification genes. Mol Breed 6:441–449

Gion J-M, Cabannes E, Poitel M, Vigneron Ph (2005) Conservation of xylem preferentially expressed sequence tags in several subgenera of *Eucalyptus* L'Hérit. In: Plant & Animal Genome XIII conf, San-Diego

Glaubitz JC, Emebiri LC, Moran GF (2001) Dinucleotide microsatellite from *Eucalyptus sieberi*: Inheritance, diversity, and improved scoring of single-base differences. Genome 44:1014–1045

Goff SA, Ricke D, Lan TH, Presting G, Wang RL, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D, Hutchinson D, Martin C, Katagiri F, Lange BM, Moughamer T, Xia Y, Budworth P, Zhong JP, Miguel T, Paszkowski U, Zhang SP, Colbert M, Sun WL, Chen LL, Cooper B, Park S, Wood TC, Mao L, Quail P, Wing R, Dean R, Yu YS, Zharkikh A, Shen R, Sahasrabudhe S, Thomas A, Cannings R, Gutin A, Pruss D, Reid J, Tavtigian S, Mitchell J, Eldredge G, Scholl T, Miller RM, Bhatnagar S, Adey N, Rubano T, Tusneem N, Robinson R, Feldhaus J, Macalma T, Oliphant A, Briggs S (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp *japonica*). Science 296:92–100

Goicoechea M, Lacombe E, Legay S, Milhaevic S, Rech P, Jauneau A, Lapierre C, Pollet B, Verhaegen D, Chaubet-Gigot N, Grima-Pettenati J (2005) *EgMYB2*, a new transcriptional activator from *Eucalyptus* xylem, regulates secondary cell wall formation and lignin biosynthesis. Plant J 43:553–567

Goldman IL, Rocheford TR, Dudley JW (1994) Molecular markers associated with maize kernel oil concentration in an Illinois high protein X Illinois low protein cross. Crop Sci 34:908–915

Grattapaglia D (2004) Integrating genomics into *Eucalyptus* breeding. Gen Mol Res 3:369–379

Grattapaglia D, Bradshaw HD (1994) Nuclear DNA content of commercially important *Eucalyptus* species and hybrids. Can J For Res 24:1074–1078

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

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

Grattapaglia D, Bertolucci FLG, Penchel R, Sederoff R (1996) 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 144:1205–1214

Grattapaglia D, Lourenço RT, Santos SN, Povoa A, Carocha V, Pappas GJ Jr, Brommonschenkel SH (2004) Fishing for complete lignification genes in a BAC library of *Eucalyptus grandis*. Abstr of the 50th Brazilian Genet Congr, p 1126. ISBN 85-89109-04-6

Greaves BL, Borralho NMG (1996) The influence of basic density and pulp yield on the cost of eucalypt kraft pulping: a theoretical model of tree breeding. Appita J 49:90–95

Greaves BL, Borralho NMG, Raymond CA (1997) Breeding objectives for plantation eucalypts grown for production of kraft pulp. For Sci 43:465–472

Greaves B, Hamilton M, Pilbeam D, Dutkowski G (2004a) Genetic variation in commercial properties of six and fifteen year-old *Eucalyptus globulus*. In: Tomé M (ed) IUFRO Conf: *Eucalyptus* in a Changing World, 11-15 Oct 2004, RAIZ, Instituto Investigação de Floresta e Papel, Aveiro, Portugal, pp 103-104

Greaves BL, Dutkowski GW, McRae TA (2004b) Breeding objectives for *Eucalyptus globulus* for products other than kraft pulp. In: Tomé M (ed) IUFRO Conf: *Eucalyptus* in a Changing World, 11-15 Oct 2004, RAIZ, Instituto Investigação de Floresta e Papel, Aveiro, Portugal, pp 175–180

Gregory SG, Sekhon M, Schein J, Zhao S, Osoegawa K, Scott CE, Evans RS, Burridge PW, Cox TV, Fox CA, Hutton RD, Mullenger IR, Phillips KJ, Smith J, Stalker J, Threadgold GJ, Birney E, Wylie K, Chinwalla A, Wallis J, Hillier L, Carter J, Gaige T, Jaeger S, Kremitzki C, Layman D, Maas J, McGrane R, Mead K, Walker R, Jones S, Smith M, Asano J, Bosdet I, Chan S, Chittaranjan S, Chiu R, Fjell C, Fuhrmann D, Girn N, Gray C, Guin R, Hsiao L, Krzywinski M, Kutsche R, Lee SS, Mathewson C, McLeavy C, Messervier S, Ness S, Pandoh P, Prabhu AL, Saeedi P, Smailus D, Spence L, Stott J, Taylor S, Terpstra W, Tsai M, Vardy J, Wye N, Yang G, Shatsman S, Ayodeji B, Geer K, Tsegaye G, Shvartsbeyn A, Gebregeorgis E, Krol M, Russell D, Overton L, Malek JA, Holmes M, Heaney M, Shetty J, Feldblyum T, Nierman WC,

Catanese JJ, Hubbard T, Waterston RH, Rogers J, de Jong PJ, Fraser CM, Marra M, McPherson JD, Bentley DR (2002) A physical map of the mouse genome. Nature 418:743–750

Griffin AR (1996) Genetically modified trees - the plantations of the future or an expensive distraction? Commonw For Rev For Rev 75:169–175

Griffin AR (2001) Deployment decisions - capturing the benefits of tree improvement with clones and seedlings. In: IUFRO Conf: Developing the Eucalypt of the Future, 10-15 Sept 2001, INFOR, Valdivia, Chile

Griffin AR, Burgess IP, Wolf L (1988) Patterns of natural and manipulated hybridisation in the genus *Eucalyptus* L'Herit. – a review. Aust J Bot 36:41–66

Groover A, Devey M, Fiddler T, Lee J, Megraw R, Mitchel-Olds T, Sherman B, Vujcic S, Williams C, Neale D (1994) Identification of quantitative trait loci influencing wood specific gravity in an outbred pedigree of loblolly pine. Genetics 138:1293–1300

Gupta PK, Rustgi S (2004) Molecular markers from the transcribed/expressed region of the genome in higher plants.Funct Integr Genom 4:139–62

Haldane JBS (1919) The combination of linkage values and the calculation of distances between loci of linked factors. J Genet 8:299–309

Harbard JL, Griffin AR, Espejo J (1999) Mass controlled pollination of *Eucalyptus globulus*: a practical reality. Can J For Res 29:1457–1463

Hardner CM, Potts BM (1995) Inbreeding depression and changes in variation after selfing *Eucalyptus globulus* subsp. *globulus*. Silvae Genet 44:46–54

Hardner CM, Potts BM (1997) Post-dispersal selection under mixed-mating in *Eucalyptus regnans*. Evolution 51:103–111

Hardner CM, Vaillancourt RE, Potts BM (1996) Stand density influences outcrossing rate and growth of open-pollinated families of *Eucalyptus globulus*. Silvae Genet 45:226–228

Hardner CM, Potts BM, Gore PL (1998) The relationship between cross success and spatial proximity of *Eucalyptus globulus* ssp. *globulus* parents. Evolution 52:614–618

Harkins DM, Johnson GN, Skaggs PA, Mix AD, Dupper GE, Devey ME, Kinloch BB, Neale DB (1998) Saturation mapping of a major gene for resistance to white pine blister rust in sugar pine. Theor Appl Genet 97:1355–1360

Hazen SP, Harmon FG, Pruneda-Paz FG, Schulz TF, Ecker JR, Borevitz JO, Kay SA (2005) Rapid linkage mapping in *Arabidopsis* by microarray genotyping: In: Proc Plant & Animal Genomes XIII Conf San Diego, W218

Hertzberg M, Aspeborg H, Schrader J, Andersson A, Erlandsson R, Blomqvist K, Bhalerao R, Uhlen M, Teeri TT, Lundeberg J, Sundberg B, Nilsson P, Sandberg G (2001) A transcriptional roadmap to wood formation. Proc Natl Acad Sci USA 98:14732–14737

Hill RS (2004) Origins of the southeastern Australian vegetation. Philos Trans R Soc B 359:1537–1549

Hill KD, Johnson LAS (1995) Systematic studies in the eucalypts 7. A revision of the bloodwoods, genus *Corymbia* (Myrtaceae). Telopea 6:185–504

Hill R, Truswell E, McLoughlin S, Dettmann M (1999) Evolution of the Australian flora: fossil evidence. In: George AS (ed) Flora of Australia, vol 1. Co-published by CSIRO and Australian Biological Resources Study, pp 251–320

Hingston AB, Potts BM (2005) Pollinator activity can explain variation in outcrossing rates within individual trees. Aust Ecol 30:319–324

Hingston AB, Potts BM, McQuillan PB (2004) The swift parrot, *Lathamus discolor* (Psittacidae), social bees (Apidae) and native insects as pollinators of *Eucalyptus globulus* ssp. *globulus* (Myrtaceae). Aust J Bot 52:371–379

Hittalmani S, Huang N, Courtois B, Venuprasad R, Shashidhar HE, Zhuang JY, Zheng KL, Liu, GF, Wang GC, Sidhu JS, Srivantaneeyakul S, Singh VP, Bagali PG, Prasanna HC, McLaren G, Khush GS (2003) Identification of QTL for growth- and grain yield-related traits in rice across nine locations of Asia. Theor Appl Genet 107:679–690

Holman JE, Hughes JM, Fensham RJ (2003) A morphological cline in *Eucalyptus*: a genetic perspective. Mol Ecol 12:3013–3025

Hoskins RA, Nelson CR, Berman BP, Laverty TR, George RA, Ciesiolka L, Naeemuddin M, Arenson AD, Durbin J, David RG, Tabor PE, Bailey MR, DeShazo DR, Catanese J, Mammoser A, Osoegawa K, de Jong PJ, Celniker SE, Gibbs RA, Rubin GM, Scherer SE (2000) A BAC-based physical map of the major autosomes of *Drosophila melanogaster*. Science 287:2271–2274

House SM (1997) Reproductive biology of eucalypts. In: Woinarski J (ed) Eucalypt Ecology: Individuals to Ecosystems. Cambridge University Press, Cambridge, UK, pp 30–56

Israelsson M, Eriksson ME, Hertzberg M, Aspeborg H, Nilsson P, Moritz T (2003) Changes in gene expression in the wood-forming tissue of transgenic hybrid aspen with increased secondary growth. Plant Mol Biol 52:893–903

Jaccoud D, Peng K, Feinstein D, Kilian A (2001) Diversity arrays: a solid state technology for sequence information independent genotyping. Nucleic Acids Res 29:e25

Jacobs MR (1981) Eucalypts for Planting. FAO, Rome

Janki-Ammal EK, Khosla SN (1969) Breaking the barrier to polyploidy in the genus *Eucalyptus*. Proc Ind Acad Sci (Bangalore) 20:248–249

Jansen RC, Stam P (1994) High resolution of quantitative traits into multiple loci via interval mapping. Genetics 136:1447–1455

Jones CJ, Edwards KJ, Castaglione S, Winfield MO, Sala F, Vandewiel C, Bredemeijer G, Vosman B, Matthes M, Daly A, Brettschneider R, Bettini P, Buiatti M, Maestri E, Malcevschi A, Marmiroli N, Aert R, Volckaert G, Rueda J, Linacero R, Vazquez A, Karp A (1997) Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. Mol Breed 3:381–390

Jones RC, McKinnon GE, Potts BM, Vaillancourt RE (2005) Genetic diversity and mating system of an endangered tree, *Eucalyptus morrisbyi*. Aust J Bot 53:366–377

Jones RC, Steane DA, Potts BM, Vaillancourt RE (2002) Microsatellite and morphological analysis of *Eucalyptus globulus* populations. Can J For Res 32:59–66

Jordan GJ, Potts BM, Kirkpatrick JB, Gardiner C (1993) Variation in the *Eucalyptus globulus* complex revisited. Aust J Bot 41:763–785

Junghans DT, Alfenas AC, Brommonschenkel SH, Oda S, Mello EJ, Grattapaglia D (2003) Resistance to rust (*Puccinia psidii* Winter) in *Eucalyptus*: mode of inheritance and mapping of a major gene with RAPD markers. Theor Appl Genet 108:175–180

Kanowski PJ, Borralho NMG (2004) Economics of tree improvement. In: Youngquist JA (ed) Encyclopedia of Forest Science. Elsevier, Oxford, pp 1561–1568

Karp CL, Grupe A, Schadt E, Ewart SL, Keane-Moore M, Cuomo PJ, Kohl J, Wahl L, Kuperman D, Germer S, Aud D, Peltz G, Wills-Karp M (2000) Identification of complement factor 5 as a susceptibility locus for experimental allergic asthma. Nat Immunol 1:221–226

Kawazu T, Doi K, Kondo K (2003) Process for transformation of mature trees of *Eucalyptus* plants. US Patent 6,563,024

Kearsey MJ, Farquhar AGL (1998) QTL analysis in plants: where are we now? Heredity 80:137–142

Kellison RC (2001) Present and future uses of eucalypts wood in the world. In: Barros S (ed) Developing the Eucalypt of the Future. IUFRO Int Symp, 10-15 Sept 2001, INFOR, Chile

Kershaw AP, Martin HA, McEwen MJRC (1994) The Neogene: a period of transition. In: Hill RS (ed) History of the Australian Vegetation: Cretaceous to Recent. Cambridge University Press, Cambridge, UK, pp 299-327

Kirst M (2004) Microarrays beyond gene expression: discovery of candidate genes for quantitative traits to large scale SNP genotyping. In: Li B, McKeand M (eds) Forest Genetics and Tree Breeding in the Age of Genomics: Progress and Future. IUFRO Joint Conf Div 2, 1-5 Nov 2005, Double Tree Guest Suites, Charleston, SC, p 341

Kirst M, Myburg AA, De Leon 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

Kirst M, Basten CJ, Myburg A, Zeng Z-B, Sederoff R (2005) Genetic architecture of transcript level variation in differentiating xylem of *Eucalyptus* hybrids. Genetics 169:2295–2303

Klein PE, Klein RR, Cartinhour SW, Ulanch PE, Dong JM, Obert JA, Morishige DT, Schlueter SD, Childs KL, Ale M, Mullet JE (2000) A high-throughput AFLP-based method for constructing integrated genetic and physical maps: progress toward a sorghum genome map. Genome Res 10:789–807

Kong A, Gudbjartsson DF, Sainz J, Jonsdottir GM, Gudjonsson SA, Richardsson B, Sigurdardottir S, Barnard J, Hallbeck B, Masson G, Shlien A, Palsson ST, Frigge ML, Thorgeirsson TE, Gulcher JR, Stefansson K (2002) A high-resolution recombination map of the human genome. Nat Genet 31:241–247

Krzywinski M, Wallis J, Gosele C, Bosdet I, Chiu R, Graves T, Hummel O, Layman D, Mathewson C, Wye N, Zhu B, Albracht D, Asano J, Barber S, Brown-John M, Chan S, Chand S, Cloutier A, Davito J, Fjell C, Gaige T, Ganten D, Girn N, Guggenheimer K, Himmelbauer H, Kreitler T, Leach S, Lee D, Lehrach H, Mayo M, Mead K, Olson T, Pandoh P, Prabhu AL, Shin H, Tanzer S, Thompson J, Tsai M, Walker J, Yang G, Sekhon M, Hillier L, Zimdahl H, Marziali A, Osoegawa K, Zhao S, Siddiqui A, de Jong PJ, Warren W, Mardis E, McPherson JD, Wilson R, Hubner N, Jones S, Marra M, Schein J (2004) Integrated and sequence-ordered BAC- and YAC-based physical maps for the rat genome. Genome Res 14:766–779

Kube PD, Raymond CA (2005) Breeding to minimise the effects of collapse in *Eucalyptus nitens*. For Genet 12:23–34

Kubisiak TL, Hebard FV, Nelson CD, Zhang JS, Bernatzky R, Huang H, Anagnostakis SL, Doudrick RL (1997) Mapping resistance to blight in an interspecific cross in the genus *Castanea* using morphological, isozyme, RFLP, and RAPD markers. Phytopathology 87:751–759

Ladiges PY (1997) Phylogenetic history and classification of eucalypts. In: Woinarski JCZ (ed) Eucalypt Ecology: Individuals to Ecosystems. Cambridge University Press, Cambridge, UK, pp 16–29

Ladiges PY, Udovicic F (2000) Comment on a new classification of the eucalypts. Aust Syst Bot 13:149–152

Ladiges PY, Udovicic F (2005) Comment on the molecular dating of the age of eucalypts. Aust Syst Bot 18:291–293

Ladiges PY, Udovicic F, Nelson G (2003) Australian biogeographical connections and the phylogeny of large genera in the plant family Myrtaceae. J Biogeogr 30:989–998

Lander ES, Botstein D (1989) Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185–199

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

Leon AJ, Andrade FH, Lee M (2003) Genetic analysis of seed-oil concentration across generations and environments in sunflower. Crop Sci 43:135–140

Lezar S, Myburg AA, Berger DK, Wingfield MJ, Wingfield BD (2004) Development and assessment of microarray-based DNA fingerprinting in *Eucalyptus grandis*. Theor Appl Genet 109:1329–1336

Lister C, Dean C (1993) Recombinant inbred lines for mapping RFLP and phenotypic markers in *Arabidopsis thaliana*. Plant J 4:745–750

Liu BH, Knapp SJ (1990) GMENDEL: a programme for Mendelian segregation and linkage analysis of individual or multiple progeny using log-likelihood ratios. J Herd 81:407

Lorieux M, Goffinet B, Perrier X, González de León D, Lanaud C (1995) Maximum-likelihood models for mapping genetic markers showing segregation distortion. 1. Backcross populations. Theor Appl Genet 90:73–80

Lourenço RT (2004) Genomic structure of three mega basespairs of genomic shotgun DNA of *Eucalyptus grandis*: nucleotide content, repetitive sequences and genes. MSc Thesis, Univ of Campinas, Brazil

Lu C, Shen L, Tan Z, Xu Y, He P, Chen Y, Zhu L (1996) Comparative mapping of QTLs for agronomic traits of rice across environments using a doubled haploid population. Theor Appl Genet 93:1211–1217

Luo MC, Thomas C, You FM, Hsiao J, Ouyang S, Buell CR, Malandro M, McGuire PE, Anderson OD, Dvorak J (2003) High-throughput fingerprinting of bacterial artificial chromosomes using the snapshot labeling kit and sizing of restriction fragments by capillary electrophoresis. Genomics 82:378–389

MacDonald MC, Borralho NMG, Potts BM (1997) Genetic variation for growth and wood density in *Eucalyptus globulus* ssp. *globulus* in Tasmania (Australia). Silvae Genet 46:236–241

Mangin B, Goffinet B, Rebaï A (1994) Constructing confidence intervals for QTL location. Genetics 138:1301–1308

Marques CM, Araujo JA, Ferreira JG, Whetten R, O'Malley DM, Liu BH, Sederoff R (1998) AFLP genetic maps of *Eucalyptus globulus* and *E. tereticornis*. Theor Appl Genet 96:727–737

Marques CM, Vasquez-Kool J, Carocha VJ, Ferreira JG, O'Malley DM, Liu BH, Sederoff R (1999) Genetic dissection of vegetative propagation traits in *Eucalyptus tereticornis* and *E. globulus*. Theor Appl Genet 99:936–946

Marques CM, Brondani RPV, Grattapaglia D, Sederoff R (2002) Conservation and synteny of SSR loci and QTLs for vegetative propagation in four *Eucalyptus* species. Theor Appl Genet 105:474–478

Marra MA, Kucaba TA, Dietrich NL, Green ED, Brownstein B, Wilson RK, McDonald KM, Hillier LW, McPherson JD, Waterston RH (1997) High throughput fingerprint analysis of large-insert clones. Genome Res 7:1072–1084

Martin B, Quillet J (1974) The propagation by cuttings of forest trees in the Congo. Bois et Forets des Tropiques 155:15–33

Martins-Corder MP, Lopez CR (1997) Isozyme characterization of *Eucalyptus urophylla* (S.T. Blake) and *E. grandis* (Hill ex Maiden) populations in Brazil. Silvae Genet 46:192–197

McDonald MW, Rawlings M, Butcher PA, Bell JC (2003) Regional divergence and inbreeding in *Eucalyptus cladocalyx* (Myrtaceae). Aust J Bot 51:393–403

McGowen MH, Wiltshire RJE, Potts BM, Vaillancourt RE (2001) The origin of *Eucalyptus vernicosa*, a unique shrub eucalypt. Biol J Linn Soc 74:397–405

McGowen MH, Potts BM, Vaillancourt RE, Gore P, Williams DR, Pilbeam DJ (2004) The genetic control of sexual reproduction in *Eucalyptus globulus*. In: Tomé M (ed) IUFRO Conference: *Eucalyptus* in a Changing World, 11-15 Oct 2004, RAIZ, Instituto Investigação de Floresta e Papel, Aveiro, Portugal, pp 104-108

McKinnon GE, Vaillancourt RE, Jackson HD, Potts BM (2001a) Chloroplast sharing in the Tasmanian eucalypts. Evolution 55:703–711

McKinnon GE, Vaillancourt RE, Tilyard PA, Potts BM (2001b) Maternal inheritance of the chloroplast genome in *Eucalyptus globulus* and interspecific hybrids. Genome 44:831–835

McKinnon GE, Vaillancourt RE, Steane DA, Potts BM (2004) The rare silver gum, *Eucalyptus cordata*, is leaving its trace in the organellar gene pool of *Eucalyptus globulus*. Mol Ecol 13:3751–3762

McKinnon GE, Potts BM, Steane DA, Vaillancourt RE (2005) Population and phylogenetic analysis of the cinnamoyl CoA reductase gene in *Eucalyptus globulus* Labill. (Myrtaceae). Aust J Bot 53:827–838

McPherson JD, Marra M, Hillier L, Waterston RH, Chinwalla A, Wallis J, Sekhon M, Wylie K, Mardis ER, Wilson RK, Fulton R, Kucaba TA, Wagner-McPherson C, Barbazuk WB, Gregory SG, Humphray SJ, French L, Evans RS, Bethel G, Whittaker A, Holden JL, McCann OT, Dunham A, Soderlund C, Scott CE, Bentley DR, Schuler G, Chen HC, Jang W, Green ED, Idol JR, Maduro VV, Montgomery KT, Lee E, Miller A, Emerling S, Kucherlapati, Gibbs R, Scherer S, Gorrell JH, Sodergren E, Clerc-Blankenburg K, Tabor P, Naylor S, Garcia D, de Jong PJ, Catanese JJ, Nowak N, Osoegawa K, Qin S, Rowen L, Madan A, Dors M, Hood L, Trask B, Friedman C, Massa H, Cheung VG, Kirsch IR, Reid T, Yonescu R, Weissenbach J, Bruls T, Heilig R, Branscomb E, Olsen A, Doggett N, Cheng JF, Hawkins T, Myers RM, Shang J, Ramirez L, Schmutz J, Velasquez O, Dixon K, Stone NE, Cox DR, Haussler D, Kent WJ, Furey T, Rogic S, Kennedy S, Jones S, Rosenthal A, Wen G, Schilhabel M, Gloeckner G, Nyakatura G, Siebert R, Schlegelberger B, Korenberg J, Chen XN, Fujiyama A, Hattori M, Toyoda A, Yada T, Park HS, Sakaki Y, Shimizu N, Asakawa S, Kawasaki K, Sasaki T, Shintani A, Shimizu A, Shibuya K, Kudoh J, Minoshima S, Ramser J, Seranski P, Hoff C, Poustka A, Reinhardt R, Lehrach H; International Human Genome Mapping Consortium (2001) A physical map of the human genome. Nature 409:934–41

McRae TA, Dutkowski GW, Pilbeam DJ, Powell MB, Tier B (2004a) Genetic evaluation using the TREEPLAN® System. In: McKeand S (ed) Forest Genetics and Tree Breeding in the Age of Genomics: Progress and Future, IUFRO Joint Conf Div 2, 1-5 Nov 2004, North Carolina State University, Charleston, SC

McRae TA, Pilbeam DJ, Powell MD, Dutkowski GW, Joyce K, Tier B (2004b) Genetic evaluation in Eucalypt breeding programs. In: Tomé M (ed) IURO Conf: *Eucalyptus* in a Chang-

ing World, 11-15 Oct 2004, RAIZ, Instituto Investigação de Floresta e Papel, Aveiro, Portugal, pp 189-190

Meddings RLA, McComb JA, Bell DT (2001) The salt-waterlogging tolerance of *Eucalyptus camaldulensis* × *E. globulus* hybrids. Aust J ExpAgric 41:787–792

Meyers BC, Scalabrin S, Morgante M (2004) Mapping and sequencing complex genomes: let's get physical! Nat Rev Genet 5:578–588

Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci USA 88:9828–9832

Milgate AW, Potts BM, Joyce K, Mohammed CL, Vaillancourt RE (2005) Genetic variation in *Eucalyptus globulus* for susceptibility to *Mycosphaerella nubilosa* and its association with tree growth. Aust Plant Pathol 34:11–18

Missiaggia AA, Grattapaglia D (2005) Genetic mapping of *Eef1*, a major effect QTL for early flowering in *Eucalyptus grandis*. Tree Genet Genom 1:79–84

Moran GF (1992) Patterns of genetic diversity in Australian tree species. New For 6:49–66

Moran GF, Bell JC (1983) *Eucalyptus*. In: Tanksley SD, Orton TJ (eds) Isozymes in Plant Genetics and Breeding. Elsevier, Amesterdam, pp 423–441

Moran GF, Thamarus KA, Raymond CA, Qiu D, Uren T, Southerton SG (2002) Genomics of *Eucalyptus* wood traits. Ann For Sci 59:645–650

Morgan TH (1910) Chromosomes and heredity. Am Nat 44:449–496

Morgante M, Salamini F (2003) From plant genomics to breeding practice. Curr Opin Biotechnol 14:214–219

Mullis KB, Faloona FA (1987) Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. Meth Enzymol 155:335–350

Myburg AA (2001) Genetic architecture of hybrid fitness and wood quality traits in a wide interspecific cross of *Eucalyptus* tree species. PhD thesis, North Carolina State University, Raleigh, NC: http://www.lib.ncsu.edu/theses/available/etd-20010723-175234/

Myburg AA (2004) The International *Eucalyptus* Genome Consortium (IEuGC): Opportunities and Resources for Collaborative Genome Research in *Eucalyptus*. In: Li B, McKeand S (eds) Forest Genetics and Tree Breeding in the Age of Genomics: Progress and Future, IUFRO Joint Conf of Division 2, Conf Proc, 1-5 Nov 2004, Double Tree Guest Suites, Charleston, SC

Myburg AA, Remington DL, O'Malley DM, Sederoff RR, Whetten RW (2001) High-throughput AFLP analysis using infrared dye-labeled primers and an automated DNA sequencer. Bio/Techniques 30:348–357

Myburg AA, Griffin AR, Sederoff RR, Whetten RW (2003) Comparative genetic linkage maps of *Eucalyptus grandis*, *Eucalyptus globulus* and their $F_1$ hybrid based on a double pseudo-backcross mapping approach. Theor Appl Genet 107:1028–1042

Myburg AA, Vogl C, Griffin AR, Sederoff RR, Whetten RW (2004) Genetics of postzygotic isolation in *Eucalyptus*: whole-genome analysis of barriers to introgression in a wide interspecific cross of *Eucalyptus grandis* and *E. globulus*. Genetics 166:1405–1418

Nehls U, Beguiristain T, Ditengou F, Lapeyrie F, Martin F (1998) The expression of a symbiosis-regulated gene in eucalypt roots is regulated by auxins and hypaphorine, the tryptophan betaine of the ectomycorrhizal basidiomycete *Pisolithus tinctorius*. Planta 207:296–302

Newcombe G, Bradshaw HD Jr (1996) Quantitative trait loci conferring resistance in hybrid poplar to leaf spot caused by *Septoria populicola*. Can J For Res 26:1943–1950

Newcombe G, Bradshaw HD, Chastagner GA, Stettler RF (1996) A major gene for resistance to *Melampsora medusae* f.sp. *deltoidae* in a hybrid poplar pedigree. Phytopathology 86:87–94

Orita M, Iwahana H, Kanazawa H, Hayashi K, Sekiya T (1989) Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. Proc Natl Acad Sci USA 86:2766–2770

Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD (1988) Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature 335:721–726

Paterson AH, Damon S, Hewitt JD, Zamir D, Rabinowitch HD, Lincoln SE, Lander ES, Tanksley SD (1991) Mendelian factors underlying quantitative traits in tomato: comparison across species, generations, and environments. Genetics 127:181–197

Patterson B, Gore PL, Potts BM, Vaillancourt RE (2004a) Advances in pollination technique for large-scale production of *Eucalyptus globulus* seed. Aust J Bot 52:781–788

Patterson B, Vaillancourt RE, Pilbeam DJ, Potts BM (2004b) Factors affecting outcrossing rates in *Eucalyptus globulus*. Aust J Bot 52:773–780

Paux E, Tamasloukht M, Ladouce N, Sivadon P, Grima-Pettenati J (2004) Identification of genes preferentially expressed during wood formation in *Eucalyptus*. Plant Mol Biol 55:263–80

Paux E, Carocha V, Marques C, Mendes de Sousa A, Borralho N, Sivadon P, Grima-Pettenati J (2005) Transcript profiling of *Eucalyptus* xylem genes during tension wood formation. New Phytol 167:89–100

Pelleschi S, Guy S, Kim JY, Pointe C, Mahe A, Barthes L, Leonardi A, Prioul JL (1999) *Ivr2*, a candidate gene for a QTL of vacuolar invertase activity in maize leaves. Gene-specific expression under water stress. Plant Mol Biol 39:373–380

Pinto GL, Loureiro J, Lopes T, Santos C (2005) Analysis of the genetic stability of *Eucalyptus globulus* Labill. somatic embryos by flow cytometry. Theor Appl Genet 109:580–587

Plomion C, Bahrman N, Costa P, Dubos C, Frigerio J-M, Gerber S, Gion J-M, Lalanne C, Madur D, Pionneau C (2004) Proteomics for genetics and physiological studies in forest trees: application in maritime pine. In: Kumar S, Fladung M (eds) Molecular Genetics and Breeding of Forest Trees, Haworth, Binghamton, NY, pp 53–80

Poke FS, Vaillancourt RE, Elliott RC, Reid JB (2003) Sequence variation in two lignin biosynthesis genes, cinnamoyl CoA reductase (*CCR*) and cinnamyl alcohol dehydrogenase 2 (*CAD2*). Mol Breed 12:107–118

Poke FS, Wright JA, Raymond CA (2004) Predicting extractives and lignin contents in *Eucalyptus globulus* using Near Infrared Reflectance Analysis. J Wood Chem Technol 24:55–67

Poke FS, Vaillancourt RE, Potts BM, Reid JB (2005) Genomic research in *Eucalyptus*. Genetica 125(1):79-101

Pole MS, Hill RS, Green N, Macphail MK (1993) The Oligocene Berwick Quarry flora - rainforest in a drying environment. Aust Syst Bot 6:399–427

Potts BM (2004) Genetic improvement of eucalypts. In: Burley J, Evans J, Youngquist JA (ed) Encyclopedia of Forest Science. Elsevier, Oxford, pp 1480–1490

Potts BM, Dungey HS (2004) Hybridisation of *Eucalyptus*: key issues for breeders and geneticists. New For 27:115–138

Potts BM, Jordan GJ (1994) Genetic variation in the juvenile leaf morphology of *Eucalyptus globulus* ssp. *globulus*. For Genet 1:81–95

Potts BM, Pederick LA (2000) Morphology, phylogeny, origin, distribution and genetic diversity of the eucalypts. In: Brown BN (ed) Diseases and Pathogens of Eucalypts. CSIRO, Collingwood, Australia, pp 11-34

Potts BM, Wiltshire RJE (1997) Eucalypt genetics and genecology. In: Wiliams J, Woinarski J (eds) Eucalypt Ecology: Individuals to Ecosystems. Cambridge University Press, Cambridge, UK

Potts BM, Barbour RC, Hingston AB, Vaillancourt RE (2003) Turner Review No. 6: Genetic pollution of native eucalypt gene pools - identifying the risks. Aust J Bot 51:1–25

Potts BM, Vaillancourt RE, Jordan GJ, Dutkowski GW, Costa e Silva J, McKinnon GE, Steane DA, Volker PW, Lopez GA, Apiolaza LA, Li Y, Marques C, Borralho NMG (2004) Exploration of the *Eucalyptus globulus* gene pool. In: Tomé M (ed) IUFRO Conf: *Eucalyptus* in a Changing World, 11-15 Oct 2004. RAIZ, Instituto Investigação de Floresta e Papel, Aveiro, Portugal, pp 46–61

Pound LM, Wallwork MAB, Potts BM, Sedgley M (2002a) Early ovule development following self- and cross-pollinations in *Eucalyptus globulus* Labill. ssp *globulus*. Ann Bot (London) 89:613–620

Pound LM, Wallwork MAB, Potts BM, Sedgley M (2002b) Self-incompatibility in *Eucalyptus globulus* ssp. *globulus* (Myrtaceae). Aust J Bot 50:365–372

Pound LM, Wallworth MAB, Potts BM, Sedgley M (2003) Pollen tube growth and early ovule development following self- and cross-pollination of *Eucalyptus nitens*. Sexual Plant Reprod 16:59–69

Prioul JL, Quarrie S, Causse M, De Vienne D (1997) Dissecting complex physiological functions through the use of molecular quantitative genetics. J Exp Bot 48:1151–1163

Pryor LD (1976) The Biology of Eucalypts. Edward Arnold, London

Pryor LD, Johnson LAS (1971) A Classification of the Eucalypts. Australian National University Press, Canberra, Australia

Pryor LD, Johnson LAS (1981) *Eucalyptus*, the universal Australian. In: Keast A (ed) Ecological Biogeography of Australia. Junk, The Hague, pp 499–536

Raga FR (2001) Perspectiva para el eucalipto Chileno. In: IUFRO Conf: Developing the Eucalypt of the Future, 10-15 Sept 2001. INFOR, Valdivia, Chile, pp 1–13

Raymond CA (2002) Genetics of *Eucalyptus* wood properties. Ann For Sci 59:525–531

Raymond CR, Apiolaza LA (2004) Incorporating wood quality and deployment traits in *Eucalyptus globulus* and *Eucalyptus nitens*. In: Carson M (ed) Plantation Forest Biotechnology for the 21st Century. Forest Research New Zealand, Rotorua, New Zealand, pp 87–99

Reddy KV, Rockwood DL (1989) Breeding strategies for coppice production in a *Eucalyptus grandis* base population with four generations of selection. Silvae Genet 38:148–151

Riccardi F, Gazeau P, Jacquemot MP, Vincent D, Zivy M (2004) Deciphering genetic variations of proteome responses to water deficit in maize leaves. Plant Physiol Biochem 42:1003–1011

Riesner DG, Steger U, Wiese M, Wulfert M, Heibey M, Henco K (1992) Temperature gradient electrophoresis for the detection of polymorphic DNA and for quantitative polymerase chain reaction. Electrophoresis 13:632–636

Rose JKC, Bashir S, Giovannoni JJ, Jahn MM, Saravanan RS (2004) Tackling the plant proteome: practical approaches, hurdles and experimental tools. Plant J 39:715–733

Rossetto M, Jezierski G, Hopper SD, Dixon KW (1999) Conservation genetics and clonality in two critically endangered eucalypts from the highly endemic south-western Australian flora. Biol Conserv 88:321–331

Rozefelds A (1996) *Eucalyptus* phylogeny and history: a brief summary. Tasforests 8:15–26

Saghai-Maroof MA, Solimen KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in Barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc Natl Acad Sci USA 81:8014–8018

Sale MM, Potts BM, West AK, Reid JB (1996) Relationships within *Eucalyptus* (Myrtaceae) using PCR-amplification and southern hybridisation of chloroplast DNA. Aust Syst Bot 9:273–282

Sanhueza R, Griffin AR (2001) FAMASA Fiber Yield Improvement Programme (F.Y.I.P.): 10 years experience breeding *Eucalyptus globulus*. In: Barros S (ed) Developing the Eucalypt of the Future. IUFRO Int Symp, 10-15 Sept 2001, INFOR, Chile

Sato S, Yamada N, Nakamoto S, Hibino T (2005) Expression profiling of the *Eucalyptus* transcription factor in differentiating xylem tissues. In: Plant & Animal Genome XIII conf, San-Diego, CA

Sax K (1923) The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. Genetics 8:552560

Schadt EE, Monks SA, Drake TA, Lusis AJ, Che N, Colinayo V, Ruff TG, Milligan SB, Lamb JR, Cavet G, Linsley PS, Mao M, Stoughton RB, Friend SH (2003) Genetics of gene expression surveyed in maize, mouse and man. Nature 422:297–302

Schein J, Kucaba TA, Sekhon M, Smailus D, Waterston RH, Marra MA (2004) High-throughput BAC fingerprinting. In Zhao S, Stodolsky M (eds) Methods in Molecular Biology. Vol 255: Bacterial Artificial Chromosomes: Library Construction, Physical Mapping and Sequencing. Vol 1. Humana, Totowa, NJ, pp 143–156

Schimleck LR, Michell AJ, Vinden P (1996) Eucalypt wood classification by NIR spectroscopy and principal components analysis. Appita J 49:319–324

Shepherd M, Chaparro J, Dale G, Jefferson L, Duong H, Vogel H, Walsh J, Gibbings M, Teasdale R (1995) Mapping insect resistance and essential oil traits in a tropical *Eucalyptus* hybrid. CRC-THF IUFRO Conf - Eucalypt Plantations: Improving Fibre, Yield and Quality, 19-24 Feb 1995, Hobart, Tamania, pp 421–423

Shepherd M, Chaparro JX, Teasdale R (1999) Genetic mapping of monoterpene composition in an interspecific eucalypt hybrid. Theor Appl Genet 99:1207–1215

Shepherd M, Jones M (2005) Molecular markers in tree improvement: characterisation and use in *Eucalyptus*. In: Lörz H, Wenzel G (eds) Biotechnology in Agriculture and Forestry, Vol 55: Molecular Marker Systems in Plant Breeding and Crop Improvement. Springer, Berlin Heidelberg New York, pp 399–409

Shield E (2004) Silviculture for sawlogs - a review of the key elements, with special reference to *Eucalyptus grandis*. In: Tomé M (ed) IUFRO Conf: *Eucalyptus* in a Changing World, 11-15 Oct 2004. RAIZ, Instituto Investigação de Floresta e Papel, Aveiro, Portugal, pp 62–67

Shizuya H, Birren B, Kim UJ, MancinoV, SlepakT, TachiiriY, Simon M (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. Proc Natl Acad Sci USA 89:8794–8797

Skabo S, Vaillancourt RE, Potts BM (1998) Fine-scale genetic structure of *Eucalyptus globulus* ssp. *globulus* forest revealed by RAPDs. Aust J Bot 46:583–594

Smith S, Hughes J, Wardell-Johnson G (2003) High population differentiation and extensive clonality in a rare mallee eucalypt: *Eucalyptus curtisii* - conservation genetics of a rare mallee eucalypt. Conserv Genet 4:289–300

Snedden CL, Verryn SD (2004) A comparative study of predicted gains for selection from a cloned breeding population and the implications for deployment. In: Tomé M (ed) IUFRO Conf: *Eucalyptus* in a Changing World, 11-15 Oct, RAIZ, Instituto Investigação de Floresta e Papel, Aveiro, Portugal, pp 137–144

Soderlund C, Humphray S, Dunham A, French L (2000) Contigs built with fingerprints, markers, and FPC V4.7. Genome Res 10:1772–1787

Soria F, Borralho NMG (1998) The genetics of resistance to *Phoracantha semipunctata* attack in *Eucalyptus globulus* in Spain. Silvae Genet 46:365–369

Soria F, Basurco F, Toval G, Silio L, Rodriguez MC, Toro M (1998) An application of Bayesian techniques to the genetic evaluation of growth traits in *Eucalyptus globulus*. Can J For Res 28:1286–1294

Southerton S, Birt P, Porter G, Ford H (2004) Review of gene movement by bats and birds and its potential significance for eucalypt plantation forestry. Aust For 67:44–53

Specht RL (1996) The influence of soils on the evolution of the eucalypts. In: Adams MA (ed) Nutrition of Eucalypts. CSIRO, Melbourne, Australia

Stam P (1993) Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. Plant J 3:739–744

Steane DA (2005) Complete nucleotide sequence of the chloroplast genome from the Tasmanian blue gum, *Eucalyptus globulus*. DNA Res 12(3):215-220

Steane DA, West AK, Potts BM, Overden JR, Reid JB (1991) Restriction fragment length polymorphisms in chloroplast DNA from six species of *Eucalyptus*. Aust J Bot 39:399–414

Steane DA, Vaillancourt RE, Russell J, Powell W, Marshall D, Potts BM (2001) Development and characterization microsatellite loci in *Eucalyptus globulus* (Myrtacea). Silvae Genet 50(2):89

Steane DA, Nicolle D, Vaillancourt RE, Potts BM (2002) Higher-level relationships among the eucalypts are resolved by ITS-sequence data. Aust Syst Bot 15:49–62

Steane DA, Jones RC, Vaillancourt RE (2005) A set of chloroplast microsatellite primers for *Eucalyptus* (Myrtaceae). Mol Ecol Notes 5:538–541

Stirling B, Newcombe G, Vrebalov J, Bosdet I, Bradshaw HD Jr (2001) Suppressed recombination around the MXC3 locus, a major gene for resistance to poplar leaf rust. Theor Appl Genet 103:1129–1137

Stokoe R, Shepherd M, Lee D, Nikles D, Henry R (2001) Natural inter-subgeneric hybridisation between *Eucalyptus acmenoides* Schauer and *Eucalyptus cloeziana* F. Muell (Myrtaceae) in south-eastern Queensland. Ann Bot (London) 84:563–570

Strauss SH, Lande R, Namkoong G (1992) Obstacles to molecular-marker-aided selection in forest trees. Can J For Res 22:1050–1061

Stuber CW, Lincoln SE, Wolff DW, Helentjaris T, Lander ES (1992) Identification of genetic factors contributing to het-

erosis in a hybrid from two elite maize inbred lines using molecular markers. Genetics 132:823–839

Sulston J, Mallett F, Staden R, Durbin R, Horsnell T, Coulson A (1988) Software for genome mapping by fingerprinting techniques. Comp Appl Biosci 4:125–132

Tanksley SD, Ganal MW, Martin GB (1995) Chromosome landing: a paradigm for map-based gene cloning in plants with large genomes. Trends Genet 11:63–68

Tao Q, Chang YL, Wang J, Chen H, Islam-Faridi MN, Scheuring C, Wang B, Stelly DM, Zhang HB (2001) Bacterial artificial chromosome-based physical map of the rice genome constructed by restriction fingerprint analysis. Genetics 158:1711–1724

Teulat B, Merah O, Souyris I, This D (2001) QTLs for agronomic traits from a Mediterranean barley progeny grown in several environments. Theor Appl Genet 103:774–787

Thamarus K, Groom K, Murrell J, Byrne M, Moran G (2002) A genetic linkage map for *Eucalyptus globulus* with candidate loci for wood, fibre and floral traits. Theor Appl Genet 104:379–387

Thamarus KA, Groom K, Bradley A, Raymond CA, Schimleck LR, Williams ER, Moran GF (2004) Identification of quantitative trait loci for wood and fibre properties in two full-sib pedigrees of *Eucalyptus globulus*. Theor Appl Genet 109:856–864

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796–815

Thumma R, F. NM, 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

Tibbits WN, Potts BM, Savva MH (1991) Inheritance of freezing resistance in interspecific $F_1$ hybrids of *Eucalyptus*. Theor Appl Genet 83:126–135

Toro MA, Silió L, Rodriguez MC, Soria F, Toval G (1998) Genetic analysis of survival to drought in *Eucalyptus globulus* in Spain. In: 6th World Congress on Genetics Applied to Livestock Production, Armidale, Australia, 27:499–502

Tournier V, Grat S, Marque C, El Kayal W, Penchel R, de Andrade G, Boudet AM, Teulieres C (2003) An efficient procedure to stably introduce genes into an economically important pulp tree (*Eucalyptus grandis* × *Eucalyptus urophylla*). Transgenic Res 12:403–411

Tragoonrung S, Kanizin V, Hayes PM, Blake TK (1992) Sequence tagged-site-facilitated PCR for barley genome mapping. Theor Appl Genet 84:1002–1008

Turnbull JW (1999) Eucalypt plantations. New For 17:37–52

Tuskan G, West D, Bradshaw HD, Neale D, Sewell M, Wheeler N, Megraw B, Jech K, Wiselogel A, Evans R, Elam C, Davis M, Dinus R (1999) Two high-throughput techniques for determining wood properties as part of a molecular genetics analysis of loblolly pine and hybrid poplar. Appl Biochem Biotechnol 77-9:55–65

Udovicic F, Ladiges PY (2000) Informativeness of nuclear and chloroplast DNA regions and the phylogeny of the eucalypts and related genera (Myrtaceae). Kew Bull 55:633–645

Vaillancourt RE, Jackson HD (2000) A chloroplast DNA hypervariable region in eucalypts. Theor Appl Genet 101:473–477

Vaillancourt RE, Potts BM, Manson A, Reid JB (1994) Detection of QTLs in a *Eucalyptus gunni* X *E. globulus* $F_2$ using a RAPD linkage map. In: Proc Int Wood Biotechnol Symp, 31 Aug-1 Sept, Hokutopia (Convention Hall) Tokyo, pp 63–70

Vaillancourt RE, Petty A, McKinnon GE (2004) Maternal inheritance of mitochondria in *Eucalyptus globulus*. J Hered 95:353–355

Varghese M, Nagarajan B, Nicodemus A, Bennet SSR, Subramanian K (2000) Hybrid breakdown in Mysore gum and need for genetic improvement of *Eucalyptus camaldulensis* and *E. tereticornis* in India. In: Nikles DG (ed) Hybrid Breeding and Genetics of Forest Trees. Proc QFRI/CRC-SPF Symp, 9-14 April 2000, Noosa, Queensland, Australia. Dept of Primary Industries, Brisbane, Australia, pp 519–525

Verhaegen D, Plomion C (1996) Genetic mapping in *Eucalyptus urophylla* and *E grandis* using RAPD markers. Genome 39:1051–1061

Verhaegen D, Plomion C, Gion JM, Poitel M, Costa P, Kremer A (1997) Quantitative trait dissection analysis in *Eucalyptus* using RAPD markers. 1. Detection of QTL in interspecific hybrid progeny, stability of QTL expression across different ages. Theor Appl Genet 95:597–608

Vigneron P, Bouvet J (2000) Eucalypt hybrid breeding in Congo. In: Nikles DG (ed) Hybrid Breeding and Genetics of Forest Trees. Proc QFRI/CRC-SPF Symp, 9-14 April 2000, Noosa, Queensland, Australia, Dept of Primary Industries, Brisbane, Australia, pp 14–26

Vincent D, Lapierre C, Pollet B, Cornic G, Negroni L, Zivy M (2005) Water deficits affect caffeate O-methyltransferase, lignification, and related enzymes in maize leaves. A proteomic investigation. Plant Physiol 137:949–960

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Fritjters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Wagner DB, Furnier GR, Saghai-Maroof MA, Williams SM, Dancik BP, Allard RW (1987) Chloroplast DNA polymorphisms in lodgepole and jack pines and their hybrids. Proc Natl Acad Sci USA 84:2097–2100

Wallis JW, Aerts J, Groenen MA, Crooijmans RP, Layman D, Graves TA, Scheer DE, Kremitzki C, Fedele MJ, Mudd NK, Cardenas M, Higginbotham J, Carter J, McGrane R, Gaige T, Mead K, Walker J, Albracht D, Davito J, Yang SP, Leong S, Chinwalla A, Sekhon M, Wylie K, Dodgson J, Romanov MN, Cheng H, de Jong PJ, Osoegawa K, Nefedov M, Zhang H, McPherson JD, Krzywinski M, Schein J, Hillier L, Mardis ER, Wilson RK, Warren WC (2004) A physical map of the chicken genome. Nature 432:679–680

Wardell-Johnson GW, Williams JE, Hill KD, Cumming R (1997) Evolutionary biogeography and contemporary distribution

of eucalypts. In: Woinarski JCZ (ed) Eucalypt Ecology: Individuals to Ecosystems. Cambridge University Press, Cambridge, pp 92–128

Waugh G (2004) Growing *Eucalyptus globulus* for high-quality sawn products. In: Tomé M (ed) IUFRO Conf: *Eucalyptus* in a Changing World, 11-15 Oct 2004. RAIZ, Instituto Investigação de Floresta e Papel, Aveiro, Portugal, pp 79–84

Wayne ML, McIntyre LM (2002) Combining mapping and arraying: an approach to candidate gene identification. Proc Natl Acad Sci USA 99:14903–14906

Wei X, Borralho NMG (1999) Objectives and selection criteria for pulp production of *Eucalyptus urophylla* plantations in South East China. For Genet 6:181–190

Welsh J, McClelland M (1990) Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res 18:7213–7218

Weng C, Kubisiak TL, Nelson CD, Stine M (2002) Mapping quantitative trait loci controlling early growth in a (longleaf pine × slash pine) × slash pine BC1 family. Theor Appl Genet 104:852–859

Whittock S, Steane DA, Vaillancourt RE, Potts BM (2003) Is *Eucalyptus* subgenus *Minutifructa* over-ranked? Trans R Soc South Aust 127:27–32

Wicking C, Williamson B (1991) From linked marker to gene. Trends Genet 7:288–293

Wilcox PL, Amerson HV, Kuhlman EG, Liu BH, O'Malley DM, Sederoff RR (1996) Detection of a major gene for resistance to fusiform rust disease in loblolly pine by genomic mapping. Proc Natl Acad Sci USA 93:3859–3864

Williams JE, Woinarski JCZ (1997) Eucalypt ecology: individuals to ecosystems. Cambridge University Press, Cambridge, UK

Williams JG, Kubelik AR, Livak KJ, Rafalski LA, Tingey SV (1990) DNA polymorphism amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531–6535

Williams DR, Potts BM, Black PG (1999) Testing single visit pollination procedures for *Eucalyptus globulus* and *E. nitens*. Aust For 62:346–352

Williams E, Matheson A, Harwood C (2002) Experimental design and analysis for tree improvement. CSIRO, Melbourne, Australia

Wiltshire RJE (2004) Eucalypts. In: Burley J, Evans J, Youngquist JA (ed) Encyclopedia of Forest Science. Elsevier, Oxford, pp 1687–1699

Wu C, Sun S, Nimmakayala P, Santos FA, Meksem K, Springman R, Ding K, Lightfoot DA, Zhang HB (2004) A BAC- and BIBAC-based physical map of the soybean genome. Genome Res 14:319–26

Xu Z, van den Berg MA, Scheuring C, Covaleda L, Lu H, Santos FA, Uhm T, Lee MK, Wu C, Liu S, Zhang HB (2005) Genome physical mapping from large-insert clones by fingerprint analysis with capillary electrophoresis: a robust physical map of *Penicillium chrysogenum*. Nucleic Acids Res 33:e50

Yang JM, Park S, Kamdem DP, Keathley DE, Retzel E, Paule C, Kapur V, Han KH (2003) Novel gene expression profiles define the metabolic and physiological processes characteristic of wood and its extractive formation in a hardwood tree species, *Robinia pseudoacacia*. Plant Mol Biol 52:935–956

Yang JM, Kamdem DP, Keathley DE, Han KH (2004a) Seasonal changes in gene expression at the sapwood-heartwood transition zone of black locust (*Robinia pseudoacacia*) revealed by cDNA microarray analysis. Tree Physiol 24:461–474

Yang SH, van Zyl L, No EG, Loopstra CA (2004b) Microarray analysis of genes preferentially expressed in differentiating xylem of loblolly pine (*Pinus taeda*). Plant Sci 166:1185–1195

Yu J, Hu SN, Wang J, Li SG, Wong KSG, Liu B, Deng Y, Dai L, Zhou Y, Zhang XQ, Cao ML, Liu J, Sun JD, Tang JB, Chen YJ, Huang XB, Lin W, Ye C, Tong W, Cong LJ, Geng JN, Han YJ, Li L, Li W, Hu GQ, Huang XG, Li WJ, Li J, Liu ZW, Li L, Liu JP, Qi QH, Liu JS, Li L, Wang XG, Lu H, Wu TT, Zhu M, Ni PX, Han H, Dong W, Ren XY, Feng XL, Cui P, Li XR, Wang H, Xu X, Zhai WX, Xu Z, Zhang JS, He SJ, Zhang JG, Xu JC, Zhang KL, Zheng XW, Dong JH, Zeng WY, Tao L, Chen XW, He J, Liu DF, Tian W, Tian CG, Xia HG, Li G, Gao H, Li P, Chen W, Wang XD, Zhang Y, Hu JF, Wang J, Liu S, Yang J, Zhang GY, Xiong YQ, Li ZJ, Mao L, Zhou CS, Zhu Z, Chen RS, Hao BL, Zheng WM, Chen SY, Guo W, Li GJ, Liu SQ, Huang GY, Tao M, Wang J, Zhu LH, Yuan LP, Yang HM (2001) A draft sequence of the rice (*Oryza sativa* ssp *indica*) genome. Chin Sci Bull 46:1937–1942

Yvert G, Brem RB, Whittle J, Akey JM, Foss E, Smith EN, Mackelprang R, Kruglyak L (2003) Trans-acting regulatory variation in *Saccharomyces cerevisiae* and the role of transcription factors. Nat Genet 35:57–64

Zacharin RF (1978) Emigrant eucalypts: gum trees as exotics. Melbourne University Press, Melbourne, Australia

Zeng ZB (1994) Precision mapping of quantitative trait loci. Genetics 136:1457–1468

Zobel BJ, Sprague JR (1998) Juvenile Wood in Forest Trees, Springer, Berlin Heidelberg New York, p 300