The application of molecular markers to *Eucalyptus globulus* tree improvement

Jules S. Freeman

B. Sc. (Hons)

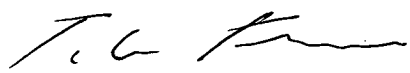
Submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy

School of Plant Science, University of Tasmania
September 2006

Declarations

This thesis does not contain any material which has been accepted for a degree or diploma by the University of Tasmania or any other institution. To the best of my knowledge and belief this thesis contains no material previously published or written by another person, except where due acknowledgement is made in the text.

This thesis may be made available for loan and limited copying in accordance with the Copyright Act 1968.

Signed:  Date: September 2006

Jules S. Freeman

Abstract

Molecular markers are being applied increasingly to various aspects of tree improvement. This thesis reports two such applications. The first involves tracing the natural origins and assessing the genetic diversity of the Portuguese Landrace of *E. globulus*, which is a core component of the Portuguese breeding program. In order to achieve this, variation in chloroplast DNA (cpDNA) sequence and microsatellite allele frequency in 47 individuals of the Portuguese Landrace were compared with native stand material. The native gene pool of *E. globulus* is genetically diverse, with substantial geographically structured genetic variation in molecular markers, which makes it possible to trace the origin of germplasm of unknown pedigree, such as the Portuguese Landrace. Similar affinities suggested by the two independent marker systems provided strong evidence that the Portuguese Landrace was predominantly derived from south-eastern Tasmania and to a lesser extent south-eastern Victoria. Ascertaining the origin of and diversity in such populations will help to avoid inbreeding and will capture the most favourable traits from native populations when they are used in breeding programs.

The second application of molecular markers addressed in this thesis involved the construction of a genetic linkage map and searching for molecular markers associated with quantitative trait loci (QTL) in *E. globulus*. This study focussed on identifying QTL for susceptibility to *Mycosphaerella* leaf blotch (MLB), a fungal disease affecting the growth of eucalypt plantations worldwide, and other potentially correlated traits. In order to search for QTL, parental and consensus linkage maps were constructed in an F_2 inter-provenance cross of *Eucalyptus globulus* with clonal replication. The consensus map used for QTL detection contained 169 markers (34 SSR and 135 AFLP loci) in 11 linkage groups. The inclusion of SSR markers allowed comparison of linkage and QTL information to other mapping and QTL studies conducted in different mapping pedigrees of *E. globulus* and other species in the subgenus *Symphyomyrtus*. The F_2 was genetically variable, with significant differences detected between genotypes for all recorded traits. Significant QTL were detected for susceptibility to MLB, herbivory by the autumn gum moth, growth and vegetative phase change. In the case of MLB susceptibility, two major unlinked QTL

explained over 70% of the phenotypic variance. This study reports the first QTL affecting height to vegetative phase change and disease resistance in *E. globulus*. These findings will help to elucidate the genetic control of these complex traits in *Eucalyptus*, and provide information that will be useful in attempts to incorporate them into breeding programs.

Acknowledgements

This research was funded by a Tasmania Graduate Research Scholarship and assisted by the Cooperative Research Centre for Sustainable Production Forestry (CRC-SPF). Thanks to the many people that contributed to this study, in particular, my supervisors René Vaillancourt, Brad Potts and Dorothy Steane. Merv Shepherd generously devoted his time to assist in constructing the linkage map and QTL location. Adam Smolenski is acknowledged for his patient and thorough technical advice. Dario Grattapaglia and Rosani Brondani assisted greatly providing the EMBRA microsatellite primers, many of these before their publication. The Portuguese Landrace study would not have been possible without the assistance of our Portuguese collaborators: Cristina Marques, Victor Carocha, Fatima Cunha and Nuno Borrahlo. I would also like to thank Hamish Jackson for his contribution to the landrace study and Dorothy Steane for providing the database of microsatellite variation in native stands. Kelsey Joyce and Gunns Ltd. provided access to the field trial. Paul Tilyard was integral in managing the collection and storage of phenotypic data and assisting with the quantitative analysis. Andrew Milgate and Tim Jones also provided phenotypic data. Thanks to my office mates during the course of this project and all those in the molecular lab for providing a helpful and amusing work environment. James Marthick deserves special mention for his skilled proof reading. Thanks to my family and friends for their encouragement and support, especially Di, my parents and Gram for sharing her botanical knowledge.

Publications arising from this project

Freeman JS, Milgate AW, Vaillancourt RE, Mohammed CL, Potts BM. Genetic control of resistance to *Mycosphaerella* leaf disease in *Eucalyptus globulus. Poster presented at* XIX International Congress of Genetics, Melbourne, July 6-11. 2003.

Potts BM, Milgate A, Joyce K, Mohammed C, Vaillancourt RE, Freeman JS, Dutkowski GW (2004) Quantitative genetic control of *Mycosphaerella* resistance in *Eucalyptus globulus* and impact on growth. Talk presented at *'Eucalyptus* in a changing world'. IUFRO Conference. (Aveiro, Portugal) 11-15 Oct. 2004.

Potts BM, McKinnon GE, Steane DA, Jones T, Foster S, Vaillancourt RE, Freeman JS (2005) Molecular insights into the gene pool of *Eucalyptus globulus*. Abstract of invited talk at 'IUFRO Tree Biotechnology Meeting, Pretoria, South Africa', S5.3. (Forestry and Agricultural Biotechnology Institute [FABI], University of Pretoria), 6-11th November 2005.

Freeman JS, Vaillancourt RE, Potts BM (2005) QTL detection for *Mycosphaerella* leaf disease in *Eucalyptus globulus*. Poster presented at 'IUFRO Tree Biotechnology Meeting, Pretoria, South Africa', S5.3. (Forestry and Agricultural Biotechnology Institute [FABI], University of Pretoria), 6-11th November 2005.

Freeman JS, Potts BM, Shepherd M, Vaillancourt RE. (2006) Parental and consensus linkage maps of *Eucalyptus globulus* using AFLP and microsatellite markers. *Silvae Genetica* **55**, 4-5.

Freeman JS, Marques CMP, Carocha Victor, Borralho N, Potts BM, Vaillancourt RE. Origins and diversity of the Portuguese Landrace of *Eucalyptus globulus*. (Submitted to: *Annals of Forest Science*, July 2006).

Whittock SP, Freeman JS, Potts BM, Vaillancourt RE (2006) Molecular dissection of growth, vegetative phase change, pest resistance, wood properties, and their interactions in *Eucalyptus globulus*. In Proceedings: 11th International Congress of Human Genetics. (Ed Griffiths L). Brisbane, Australia, (Human Genetics Society of Australia). August 6-10, 2006.

Table of contents

Chapter 1: General introduction

Domestication and breeding in forest trees usually involve recurrent selection and mating in order to bring about changes in gene frequencies (Eldridge *et al.* 1993). In comparison to crop plants, traditional breeding approaches, such as selfing and backcrossing, are hampered in forest trees because most are outcrossing, suffering from a high genetic load and inbreeding depression (Potts and Wiltshire 1997). These factors make the fixation of desirable alleles in a particular genetic background difficult (Campbell *et al.* 2003). The long generation time characteristic of forest trees, and often poor juvenile-mature trait correlations, also slow the rate of genetic improvement (Grattapaglia *et al.* 1996). However, a vast amount of genetic variation exists in many natural populations and is potentially available for domestication (Loveless and Hamrick 1984; Burley and Kanowski 2005). Furthermore, forest trees are at a very early stage of domestication in comparison to crop plants, most being only a few generations removed from wild populations (Moran *et al.* 2000; Potts *et al.* 2004), providing ample opportunity for the incorporation of germplasm from native populations.

Eucalyptus globulus domestication and improvement

Eucalyptus globulus (Myrtaceae; *sensu* Brooker 2000) is a forest tree native to the island of Tasmania, the Bass Strait Islands and the south-east of continental Australia (Williams and Potts 1996). Following the discovery of the species by Labillardière in 1799, it was rapidly distributed far from its natural origin and by the late 19th century plantings were well established in southern Europe and northern Africa (Doughty 2000). Today, *Eucalyptus globulus* is the principal temperate region hardwood plantation species, with extensive plantations world-wide in areas as diverse as India, Chile, Portugal and Australia (Potts *et al.* 2004). Its success as a plantation species is primarily based on its short fibre pulp, which is favoured by the pulp and paper industry. Favourable pulping characteristics, in combination with a broad adaptability, have led to the continued expansion of the global plantation estate (from an estimated 800 000 ha in 1973 [Poynton 1979], to 1.7 million ha in 1995 [Tibbits *et al.* 1997] to around 2.5

million ha to date [Potts *et al.* 2004]) making *E. globulus* a tree species of great economic importance.

Formal breeding of *E. globulus* did not commence until the late 1960s, with small breeding programmes in Portugal and Australia (Dillner *et al.* 1971; Orme 1977). The first major provenance collection of *E. globulus* was undertaken in 1975-76 ("Orme collection"; Orme 1977), followed by a more extensive collection in 1987-88 by the CSIRO Australian Tree Seed Centre (Gardiner and Crawford 1987, 1988). The late 1980s and early 1990s saw the establishment of the majority of the *E. globulus* breeding programmes as global demand for pulpwood increased (Tibbits *et al.* 1997). Most of these breeding programmes concentrated on provenance testing, followed by family tests and the assessment of genetic parameters. *Eucalyptus globulus* breeding programmes are currently active in at least seven countries including Australia, France, Spain, Portugal, Chile, Uruguay and Ethiopia (Potts *et al.* 2004), with most based on the Orme and CSIRO collections and in some cases selections from local landraces (e.g. India, Venkatesan *et al.* 1984; Portugal, Dillner *et al.* 1971; Chile, Griffin 2001).

The application of molecular markers to forest tree improvement

The last two decades has seen the rapid development of molecular marker technologies, to the point where they are now used routinely in many crop breeding programmes (e.g Gupta *et al.* 1999; Lorz and Wendel 2005). In comparison, the application of molecular markers to tree breeding has progressed more slowly. However, interest in their use is increasing, with research aimed at molecular breeding being undertaken in several regions of the world, mainly in pines, poplar and eucalypts (see reviews by Walter *et al.* 1998; Butcher *et al.* 1999; Jain and Minocha 2000; Campbell *et al.* 2003; Shepherd and Jones 2004). The use of molecular markers has numerous potential applications for the identification and management of genetic variability, uncovering the genetic control of complex traits and potentially helping the selection of superior genotypes (Butcher *et al.* 1999). In this chapter a brief review is provided of some of the applications in which molecular markers have, and potentially will, be useful in tree improvement in the

general areas of: characterisation of base populations and breeding programmes; seed production; mapping quantitative trait loci (QTL); and gene identification.

Base population characterisation

In terms of characterising base populations (i.e. natural populations of a domesticated species), molecular markers can provide novel insights into the extent and distribution of genetic variability that is available to breeding populations as well as evolutionary processes contributing to this variability (e.g. Moran *et al.* 2000). Specifically, genetic diversity can be quantified in terms of partitioning of its components, allelic richness, observed and expected heterozygosity and Wrights inbreeding coefficient (e.g. Jones *et al.* 2006b). Genetic diversity has been studied extensively in native populations of most economically important forest trees including *Swietenia macrophylla* (mahogony; Gillies *et al.* 1999), *Alnus rubra* (red alder; Xie *et al.* 2002), *Cedrela odorata* (Spanish cedar; Gillies *et al.* 1997), *Populus nigra* (black poplar; Winfield *et al.* 1998), *Acacia mangium* (Butcher *et al.* 1998), *Pinus radiata* (Moran and Bell 1987), and *Eucalyptus* species as such *E. nitens* (Byrne *et al.* 1998), *E. urophylla* (House and Bell 1994), *E. grandis* (Jones *et al.* 2006a), and *E. cladocalyx* (Macdonald *et al.* 2003). *Eucalyptus globulus* is perhaps the most thoroughly studied of any eucalypt species (reviewed by Potts *et al.* 2004). The base population has been studied extensively using random amplified polymorphic DNA (RAPDs; Nesbitt *et al.* 1995), nuclear microsatellites or simple sequence repeats (SSR; Jones *et al.* 2002a; Steane *et al.* 2006), functional genes (McKinnon *et al.* 2005; Poke *et al.* 2003) and chloroplast DNA (cpDNA; Freeman *et al.* 2001; McKinnon *et al.* 2004). With the exception of the functional genes studied, these markers are considered neutral to selection. Hence, the quantified variability within and between populations is more likely to reflect time since isolation, bottlenecks and gene flow as opposed to morphological traits that are influenced by selection (Jones *et al.* 2002a; Steane *et al.* 2006). In particular, studies of cpDNA variation have provided a completely new perspective into past evolutionary processes operating in *E. globulus* (Freeman *et al.* 2001; McKinnon *et al.* 2004). The cpDNA genome is inherited uniparentally and maternally in *Eucalyptus* (Byrne *et al.* 1993; McKinnnon *et al.* 2001a) and is therefore dispersed by seed only. Very strong spatial structuring is evident among

cpDNA genomes of *E. globulus*. In combination with concordant spatial structuring in the majority of the Tasmanian species in the same subgenus (*Symphyomyrtus*), this provided strong evidence for widespread interspecific hybridisation in the past, resulting in one of the largest scale occurrences of 'chloroplast capture' ever reported (McKinnon *et al.* 2001a). The adaptive significance of such past hybridisation on the nuclear genome is the subject of further study (McKinnon *et al.* 2005).

Characterising breeding populations

The geographic origins of breeding populations can potentially be determined from knowledge of the extent and distribution of genetic variability in molecular markers amongst natural populations. For example, the aforementioned strong spatial structuring that is evident in the cpDNA genome in natural populations of *E. globulus* provides a reference with which to identify the maternal origin of unpedigreed material, and for seed certification. Knowledge concerning the genetic variability in natural populations also provides a benchmark for comparison of the genetic diversity and inbreeding levels in breeding populations. Selectively neutral molecular markers provide ideal tools for these applications in contrast to morphological traits that may be influenced by selection. Such tools have been used in tree species including *Pinus radiata* (Moran and Bell 1987; Richardson *et al.* 1997), *Pinus caribaea* (Zheng and Ennos 1999), *Acacia mangium* (Butcher *et al.* 1998) and *Eucalyptus globulus* (Astorga *et al.* 2004; Gemas *et al.* 2004; Jones *et al.* 2006b). For example, Moran and Bell (1987) compared allozyme frequencies between Australian breeding programs and native stand material of *Pinus radiata*. The results suggested that the Australian material was derived mostly from populations in Ano Nuevo and Monterey, from which most of the genetic diversity had been captured. The three remaining native populations were identified as areas of high priority for screening and inclusion in breeding programmes (Moran and Bell 1987). In the case of *Acacia mangium*, genetic diversity was examined using SSR markers in an exotic seed orchard in Sumatra, Indonesia, which was used to stock Indonesian plantations (Butcher *et al.* 1998). The genetic diversity in the seed orchard was compared to its founder population in Australia (the Daintree, Queensland) and other native populations occurring in Papua New Guinea and the Moluccas. The results

demonstrated that the founding Daintree population featured low genetic diversity relative to other native populations, resulting in very high levels of inbreeding (70%). Hence, the Sumatran seed orchard and plantations throughout Indonesia sourced from this seed orchard also exhibited high levels of inbreeding (Butcher *et al.* 1998).

Molecular fingerprinting

The correct identification of elite genotypes is extremely important in breeding programs, particularly those which rely on controlled crosses or on the correct identification of clones for mass propagation (Walter *et al.* 1998; Butcher *et al.* 1999). The cost of mislabelling can be substantial and will compound as improvement programs progress through generations (Vaillancourt *et al.* 1998), especially if mislabelled trees are used as mothers for clonal propagation. Traditionally, clone and cultivar identification has been based on a combination of morphological and phenological characteristics. For example, the method adopted by the International Poplar Commission for identification, registration and certification of poplar clones is based on 64 morphological, phenological and floral characters (IUPOV 1981, cited by Rajora and Rahman 2003). However, this method of clonal identification is difficult, ambiguous and time consuming (Rajora and Rahman 2003). In contrast, molecular fingerprinting is a very powerful tool for rapidly and unambiguosly determining genotype identity. This technology has been used to register new cultivars in annual and perennial crops (Diwan and Cregan 1997; Dore *et al.* 2001) and will also be useful for this purpose in forest trees (Rajora and Rahman 2003). RAPD markers (Castiglione *et al.* 1993; Keil and Griffin 1994; Sigurdsson *et al.* 1995; Heinze *et al.* 1996; Vaillancourt *et al.* 1998; Tripathi *et al.* 2006) and isozymes (Adams 1983; Harju and Muona 1989) have been used successfully to identify clones. However, SSR markers, due to their multi-allelic nature, can fingerprint individuals with a much lower probability of error (Rajora and Rahman 2003; Kirst *et al.* 2005). Molecular markers have been used successfully to detect errors in genotype identity in improvement programs in various conifers (Adams 1983; Harju and Muona 1989; Wheeler and Jech 1992; Heinze *et al.* 1996; Bell *et al.* 2004), *Populus* (Sigurdsson *et al.* 1995; Rajora and Rahman 2003) and *Eucalyptus* (Keil and Griffin 1994; Vaillancourt *et al.* 1998; Rocha *et al.* 2002; Kirst *et*

al. 2005; Tripathi *et al.* 2006). Error rates in genotype identification are often disturbingly high. For example, on average about 10% of ramets were found to have the incorrect genotype in breeding programmes of species such as Douglas fir, Scots pine and loblolly pine (Adams 1983; Harju and Muona 1989; Wheeler and Jech 1992).

Seed production

Molecular markers also have many potential applications for quality control in seed production. These include verifying controlled pollination, quantifying outcrossing rates in open pollinated seed orchards, studying gene flow patterns and rates of pollen contamination in open pollinated progeny, all of which rely on paternity testing (Adams *et al.* 1988, 1997; Neale *et al.* 1992; Wheeler and Jech 1992). A quantification of the rates of outcrossing *versus* selfing is important to avoid inbreeding and maintain adequate levels of genetic diversity for continuous gains. A combination of RAPD and AFLP marker systems has been used to estimate outcrossing rates in an open pollinated breeding population of *Eucalyptus urophylla* (Gaiotto *et al.* 1997). Similarly, Patterson *et al.* (2004b) used isozymes to investigate factors affecting outcrossing rate in a clonal seed orchard of *E. globulus*, finding that the degree of self incompatibility of individual genotypes and the position of the flowers in the canopy had a greater correlation with outcrossing rates than flowering time or flower abundance.

Open pollination is commonly used in seed orchards. In this case, paternity testing can be useful for determining the pollen parent and estimating rates of pollen contamination from external sources. Such findings can have important implications for orchard design. For example, paternity assignment by SSR markers was used to study pollen flow in a seed orchard of *Eucalyptus grandis* and demonstrated that nearly half of the progeny analysed (46%) resulted from pollen from outside the seed orchard, most likely from nearby *E. grandis* plantations (Shepherd and Jones 2004). Similarly, in a small *E. grandis* seed orchard in Madagascar, an estimated 39% of pollination was occurring from trees outside the seed orchard (Chaix *et al.* 2003). In both these and a similar study in a *Eucalyptus regnans* seed orchard (Burczyk *et al.* 2002), paternity analysis suggested extensive pollen flow throughout the orchards, with roughly 50% of effective pollen

coming from 40-50 m from the maternal parent. Another novel application of molecular markers involved retrospective selection of elite parent trees by paternity testing the paternity of their progeny using SSR markers (Grattapaglia *et al.* 2004b). Superior individuals derived from a hybrid seed orchard with a single *Eucalyptus grandis* seed parent and six *E. urophylla* pollen parents were tested for their paternity. The results of this work suggested that three of the pollen parents essentially sired all of the selected offspring. In light of these findings the pollen parents with poor reproductive success were culled, which led to a significant realized gain in mean annual increment (Grattapaglia *et al.* 2004b).

Controlled crossing of selected parents is an important aspect of many tree improvement programmes (Adams 1988; Walter *et al.* 1998). However, genetic gains can be compromised by factors such as pollen contamination and mislabelling. Molecular markers have been used to quantify error rates and, in some instances, identify the source of error. For example, Adams *et al.* (1988) used isozymes to determine the accuracy of controlled crossing in operational programs of conifers. They found surprisingly high error levels (~30%) in controlled crossing of Douglas fir and loblolly pine, occurring primarily on the paternal side by pollen contamination. In contrast, in *Pinus radiata* SSR markers have been applied to test the parentage of individuals in two controlled cross pedigrees that were planted at multiple sites (Butcher *et al.* 1999). Two percent of progeny were labelled incorrectly in one pedigree while 20% of progeny were labelled incorrectly in another. Most of the incorrectly labelled individuals occurred in a single trial, suggesting the error occurred at the stage of trial establishment, rather than during crossing (Butcher *et al.* 1999). Similarly, Keil and Griffin (1994) were able to detect errors in parentage of *Eucalyptus* clones supplied by a commercial organization. Samples that had been supplied as a single clone were found to be derived from two separate crosses involving different pairs of parents.

The efficiency of new techniques for controlled pollination can also be tested using molecular techniques (Walter *et al.* 1998; Butcher *et al.* 1999). Seed orchard managers in New Zealand have experimented with various methods to produce control pollinated

seed in conifers, without the costly isolation of female cones. Paternity testing of experimental seed lots was carried out with nuclear SSR and paternally inherited chloroplast markers to determine the level of undesirable pollen parentage (Walter *et al.* 1998). Similarly, in *Eucalyptus nitens*, an SSR study demonstrated that the period for which flowers were bagged following pollination had a significant effect on pollen contamination. Contaminants increased from 0% where bags were left in place for two weeks following pollination, to 20% where bags were removed one week after pollination (Butcher *et al.* 1999). Patterson *et al.* (2004a) used a rare isozyme allele to study the efficiency of supplementary mass pollination techniques for large scale production of elite *E. globulus* seed. The results of this study suggested that it was possible to produce seed with low levels of contamination or selfing using a cut-style technique that does not require the costly steps of flower emasculation, isolation and labelling (Patterson *et al.* 2004a), and thus potentially having enormous benefits to seed producers.

Mapping quantitative trait loci

The development of molecular markers such as restriction fragment length polymorphism (RFLP; Helentjaris *et al.* 1985), RAPD (Williams *et al.* 1990), amplified fragment length polymorphism (AFLP; Vos *et al.* 1995) and SSR (Tautz 1989) has led to the availability of numerous polymorphisms, that are required for the construction of genetic linkage maps. Linkage maps based on molecular markers have been constructed for many important timber and pulp wood species, including species of the genera *Pinus*, *Populus*, *Picea*, *Castanea*, *Pseudotsuga* and *Eucalyptus* (see reviews by Cervera *et al.* 2000; Ahuja 2001). In many cases, these linkage maps have been used in attempts to locate the specific loci responsible for variation in quantitative traits, i.e. QTL (reviewed by Sewell and Neale 2000), by searching for marker-trait associations. Most QTL studies have focused on traits related to growth, stress resistance and wood properties in *Eucalyptus* (reviewed by Poke *et al.* 2005) and other forest trees (Sewell and Neale 2000). Potential applications of linkage mapping and QTL analysis include marker-assisted selection (MAS; Strauss *et al.* 1992; Wu 2002), providing a framework for map based cloning (Grattapaglia *et al.* 2004a) and candidate gene screening (Thamarus *et al.*

2002). Although QTL have not been directly applied to tree breeding, a benefit of QTL studies that is often overlooked is their contribution to a fundamental understanding of the genetic architecture of quantitative traits and relating specific genetic loci to the biological mechanisms associated with desirable phenotypes (Strauss *et al.* 1992; Bradshaw and Stettler 1995). This information will be useful for both traditional and marker-assisted breeding (Sewell and Neale 2000).

Marker-assisted selection involves indirect selection for specific traits in breeding programs using molecular markers, as opposed to selection based on phenotypic assessment. Marker-assisted selection has been successfully applied to breeding programs in some crop plants (e.g. Lorz and Wenzel 2005; Eglinton *et al.* 2006; Gardiner *et al.* 2006). Early selection based on molecular markers has the potential to yield even greater gains in forest trees than in annual crops because of their long generation time and the fact that most traits are assessed near rotation age. Hence, the possibility for QTL information to be used for MAS has been proposed for tree improvement in numerous taxa including *Pinus* (Devey *et al.* 2003; Wu 2002), *Populus* (Bradshaw and Stettler 1995) and *Eucalyptus* (Grattapaglia 2000; Missiagia *et al.* 2005). However, the highly heterozygous and outbred nature of forest trees makes MAS more challenging than in crop plants, due to factors such as large amounts of linkage equilibrium, unknown linkage phase and a lack of QTL stability in different genetic backgrounds, environments and throughout plant ontogeny (Strauss *et al.* 1992; Wu 2002). As a result, despite initial high expectations for the use of MAS in forest trees, to date little progress has been made.

In order for MAS to be widely accepted, the potential for it to yield substantial gains relative to conventional phenotypic selection must be demonstrated (Lande and Thompson 1990; Strauss *et al.* 1992; Wu 2002). Situations in which this may be possible include small elite breeding programs and hybrid breeding programs, particularly where clonal propagation is used. It would be possible to practise within-family selection in small elite breeding populations, if the phases of parent tree QTL-marker genotypes were known (Neale *et al.* 1992). Hybridisation and clonal propagation of selected

individuals are being used increasingly for breeding tropical eucalypts (Grattapaglia 2000) and some conifers (Walter *et al.* 1998). Hybridisation generates substantial linkage disequilibrium, while non-additive genetic variation can be captured by clonal propagation. Both linkage disequilibrium, and non-additive genetic vartiation in a population facilitate the application of MAS (Strauss *et al.* 1992). Nonetheless, the potential gains from MAS should be considered for each specific trait individually (Grattapaglia 2000). In general, MAS is likely to yield greater gain relative to traditional phenotypic selection for traits of low heritability (Wu 2002). However, precise QTL location for traits of low heritability is difficult, requiring very large progeny sizes (Beavis 1998), clonal replication, representative genetic backgrounds and multiple environments (Grattapaglia 2000). For traits with high heritability, where QTL detection is most reliable, phenotypic selection is more efficient (Wu 2002). Exceptions are traits which are highly heritable, but in which phenotypic assessment is difficult or costly, such as wood properties or disease resistance (Strauss *et al.*1992; Wu 2002). For example, the assessment of wood properties is costly, potentially destructive and requires trees to reach rotation age. Similarly, the accurate quantification of disease severity is challenging and requires an outbreak in the pedigree of interest, while factors such as disease escape within a field trial can complicate the assessment of true genetically governed resistance or tolerance.

In conclusion, despite the limited application of QTL for MAS in forest trees to date, the potential for their future application remains, given the rapid develpoment of high-throughput markers such as array based markers (e.g. diversity array technology; Lezar *et al.* 2004). Such markers will susbstantially reduce the cost of genotyping, thereby facilitating the use of larger mapping pedigrees and the investigation of QTL stability in different genetic backgrounds, environments and developmental stages.

Gene location

A limitation of the QTL approach is that the specific genes underlying quantitative traits are not identified. Given relatively small mapping populations, confidence intervals from QTL identification in forest trees are typically 10-20 cM and may contain hundreds

of functional genes. Selection based on the specific genes underlying quantitative traits would be of great value, since attempts to use QTL for selection will rely on minimising recombination between QTL and marker loci (Sewell and Neale 2000). This is particularly relevant for outcrossed forest trees, in which linkage equilibrium can result in recombination between even tightly linked markers and QTL.

The candidate gene approach is one potential method of gene identification. Candidate genes are known genes that putatively affect trait expression (Sewell and Neale 2000). Candidate genes may be identified from known genes that map to the same position as QTL, or through *a priori* knowledge of the regulatory or biochemical pathways involved in trait expression (Thamarus *et al.* 2002). The role that a candidate gene plays in the expression of a quantitative trait can then be estimated by QTL methods, where the candidate gene serves as an additional phenotypic marker (e.g. Thamarus *et al.* 2004). Candidate genes have been mapped in various forest trees including *Eucalyptus* (Gion *et al.* 2000; Thamarus *et al.* 2002; Missiagia *et al.* 2005), *Pinus* (Brown *et al.* 2003; Pot *et al.* 2006) and a *Populus* hybrid (Bradshaw *et al.* 1994). For example, genes of known function involved in monolignol biosynthesis and floral expression in addition to 31 cambium-specific expressed sequence tags were mapped on an RFLP based linkage map in *E. globulus* (Thamarus *et al.* 2002). Some of these genes co-located with QTL for wood properties (Thamarus *et al.* 2004) suggesting that these loci may contribute to variation in wood properties.

Another approach for identifying the specific genes underlying quantitative traits is through expression profiling (Boerjan 2005). Microarrays can be used to determine gene expression levels in segregating populations thereby identifying genomic regions that explain transcript variation in co-regulated genes (Kirst *et al.* 2004). When correlated with phenotypic data from a quantitative character this approach has successfully identified positional candidate genes by co-localising gene expression QTL with conventional QTL (Kirst *et al.* 2004; Plomion *et al.* 2004). The first application of this approach in forest trees was in an interspecific backcross in *Eucalyptus*, where QTL for growth co-localised with expression QTL for lignin-related genes (Kirst *et al.* 2004),

suggesting that growth and lignin characteristics were controlled by the same loci. Only one of the lignin-related genes mapped at the growth QTL, the gene encoding the enzyme S-adenosylmethionine synthase. A similar strategy to identify candidate genes for yield was undertaken at the proteome level in maritime pine (Plomion *et al.* 2004). While these results were promising, candidate genes identified by these methods could be associated with the trait in question simply because of linkage in the mapping pedigree (Boerjan 2005).

To verify that a candidate gene is influencing a quantitative trait, association studies can be used to link the variation in a trait with single nucleotide polymorphisms (SNPs) within a gene (Boerjan 2005). This technique takes advantage of the linkage disequilibrium generated in natural populations by many generations of recombination and random mating, which ensures that only tightly linked markers will show an association, allowing greater precision than standard QTL mapping procedures (González-Martínez *et al.* 2006). Furthermore, because associations are studied in natural populations rather than in isolated pedigrees, they will be far more likely to apply across different genetic backgrounds. Therefore, association mapping can potentially circumvent many of the problems involved with traditional QTL mapping. However, association mapping does have some potential limitations. For example, population structure (present in most widely distributed forest trees) is the most common cause of systematic bias in association studies (Marchini *et al.* 2004; Hirschon and Daly 2005), although methods have been developed that correct for population structure (Pritchard *et al.* 2000b) and have been applied successfully to plant species (Thornsberry *et al.* 2001). The large genome sizes characteristic of tree species (poplar and pine genomes are approximately four and 160 times larger than *Arabidopsis*, respectively) may prevent genome wide association studies because of the vast numbers of SNPs that would be required to adequately cover the genome (González-Martínez *et al.* 2006). As a result, a more realistic approach to association mapping, based on candidate genes and flanking promoter regions has been suggested for forest trees (Neale and Savolainen 2004). The genomics projects underway in *Populus*, *Eucalyptus*, *Pinus* and *Picea* have produced a vast number of candidate genes suitable for association mapping (Nehra *et al.* 2005),

while variants involving SNPs and indels are common in these candidate genes (Poke *et al.* 2003; Gill *et al.* 2003; Brown *et al.* 2004; Ingvarsson 2005). A further prerequisite for the successful use of association mapping in forest trees is a quantification of linkage disequilibrium (Ingvarsson 2005). Although linkage disequilibrium has been well studied in crop plants such as maize, it has been examined little in forest trees. However, recent studies investigating this issue have demonstrated the rapid decay of linkage disequilibrium within candidate genes in *Populus* (Ingvarsson 2005), various conifers (Brown *et al.* 2004; González-Martínez *et al.* 2006) and *Eucalyptus* (see Poke *et al.* 2005), indicating the potential for association genetics to identify genes responsible for trait variation. Consequently, association mapping projects are currently well underway in various conifer genera (Neale and Savolainen 2004), *Populus* and *Eucalyptus* (Thamarus *et al.* 2002; Grattapaglia 2004a). In the case of *Eucalyptus*, Thumma *et al.* (2005) identified 25 SNPs in cinnamoyl CoA reductase, a key lignin gene, in *E. nitens*. Two SNPs were significantly associated with microfibril angle. These were confirmed in two segregating full-sib families of *E. nitens* and *E. globulus*, indicating the potential for their use in gene assisted selection.

In summary, molecular markers have already provided a substantial contribution to improvement programmes in many forest trees. With genome sequencing efforts for many forest trees, including *Eucalyptus* (see Poke *et al.* 2005), well underway and the development of novel techniques for identifying specific genes underlying quantitative traits and ultimately allowing gene assisted selection, we are now at the brink of developing far more sophisticated applications for molecular markers. Hence, the application of molecular markers to forest tree improvement will no doubt continue to increase in the future. The greatest challenge will be to integrate the growing body of molecular research with traditional breeding practices.

Thesis outline

This thesis reports two diverse applications of molecular markers to tree improvement in *E. globulus*. The first involves tracing the geographic origin of the Portuguese Landrace

of *E. globulus* and assessing its genetic diversity. The second involves the construction of a linkage map and the detection of QTL in *E. globulus*, with a focus on susceptibility to *Mycosphaerella* leaf blotch.

The specific aims of this PhD study were:

-To use chloroplast DNA and nuclear SSR markers to find the Australian origins, and to compare the amount of genetic diversity in the Portuguese Landrace with that of native populations of *E. globulus* (Chapter 2).

-To produce parental and consensus linkage maps in an outbred F_2, inter-provenance cross of *E. globulus*, using AFLPs and SSR markers (Chapter 3).

-To allow comparison of linkage information and QTL between maps previously produced in the subgenus *Symphyomyrtus*, by mapping fully informative SSR loci derived from a variety of sources, thereby providing a link and contributing to the development of reference maps for *E. globulus* and the subgenus (Chapter 3).

-To use the consensus linkage map and quantitative genetic analysis to investigate the genetic control of resistance to *Mycosphaerella* leaf blotch and other potentially correlated traits (Chapter 4).

The three experimental chapters (2-4) are presented as self-contained units in the style of scientific journal articles. Each chapter contains an introduction to the relevant literature, outlining the potential contribution of the study undertaken to the field of research. A discussion of the findings of each study in relation to the relevant literature and conclusions from each chapter are also presented. Due to the level of discussion in each experimental chapter, Chapter 5 presents a brief general conclusion to summarise the major findings, their implications and directions for future research.

Chapter 2: Origins and diversity of the Portuguese Landrace of *Eucalyptus globulus*

Submitted to *Annals of Forest Science* as: Freeman JS, Marques CMP, Carocha V, Borralho N, Potts BM, Vaillancourt RE. Origins and diversity of the Portuguese Landrace of *Eucalyptus globulus*.

Introduction

Numerous studies have investigated the geographic origins of crop plants (e.g. Vaillancourt and Weeden 1992; Abbo *et al.* 2001; Ozkan *et al.* 2005) and their genetic diversity relative to wild progenitors (e.g. Beebe *et al.* 2000; Votava *et al.* 2005). On a global scale, the domestication of many important crops including wheat, barley, lentil and chick pea has been traced back to the fertile crescent of the near east around 10,000 years ago (Lev-Yadun *et al.* 2000; Salamini *et al.* 2002). In comparison, the domestication of forest trees grown for timber and fibre is in its infancy for most taxa (Campbell *et al.* 2003), with intensive domestication programs applied to only a few genera, including eucalypts, pines, poplars and acacias. However, despite their more recent domestication, in many cases historical records concerning the origin and genetic composition of founding populations of domesticated forest trees are uncertain and more detailed knowledge would be useful (Moran and Bell 1987; Moran *et al.* 2000). For example, spatially structured genetic variation has been demonstrated in many forest trees with widespread distribution (Moran 1992; Dumolin-Lapegue *et al.* 1997; Leonardi and Menozzi 1996; reviewed by Loveless and Hamrick 1984) and genetic material from some regions is usually preferred over others for breeding purposes (Moran *et al.* 2000). Hence, studies into the geographic origins of domesticated forest trees can identify the genetic resources captured during the domestication process and those that remain untapped (e.g. Moran and Bell 1987; Moran *et al.* 2000). Additionally, during the domestication process, tree breeders face the challenge of incorporating specific commercial traits while maintaining overall genetic diversity (Moran *et al.* 2000), making knowledge of the origin and genetic diversity of germplasm used in the

domestication of forest trees important for effective management of genetic resources (Richardson *et al.* 1997; Zheng and Ennos 1999).

Eucalyptus globulus is widely grown for pulpwood plantations in temperate regions of the world (Eldridge *et al.* 1993; Potts *et al.* 2004). The natural distribution of *E. globulus* (sensu Brooker 2000) is restricted to south-eastern Australia, including the island of Tasmania, southern Victoria and the Bass Strait Islands (Figure 2.1). However, *E. globulus* seed was rapidly spread throughout the world in the 19th century and landraces are now established in many countries (Doughty 2000). The first formal domestication of the species began in Portugal in 1966, based on phenotypic selections from local landrace populations (Dillner *et al.* 1971; Potts *et al.* 2004). Breeding programs for *E. globulus* have since been established in other countries including Australia, Chile and Spain (Eldridge *et al.* 1993; Potts *et al.* 2004). In many cases the Australian origin of these exotic populations have not been well-documented and are often complicated by multiple introductions (Potts *et al.* 2004). There is also concern that some of these landraces have originated from a narrow genetic base that could, for example, have contributed to the poor performance of *E. globulus* in South Africa (Eldridge *et al.* 1993). In addition, the area of origin within the native gene pool is important as the species is highly variable and germplasm from some areas have greater economic value for pulpwood plantations than others (Griffin 2001; Apiolaza *et al.* 2005).

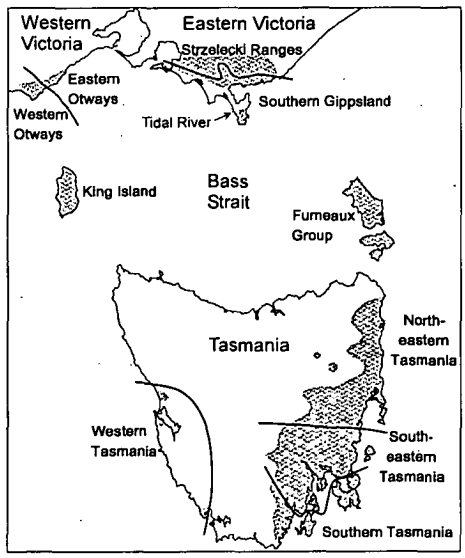


Figure 2.1 The natural distribution of *Eucalyptus globulus* and its major races as used in this study. Grey area indicates the natural distribution of the species. Geographical regions are shown in large font, while race names are shown in smaller font, with their boundaries defined by solid lines. Figure modified from Dutkowski and Potts (1999), based on new information from Lopez *et al.* (2001).

It is believed that *E. globulus* was first introduced into Portugal in 1829 (Doughty 2000), possibly via southern France, which is believed to have been an important secondary distribution point (Zacharin 1978). Later introductions of genetic material have no doubt taken place but, again, the Australian origin of these introductions is unrecorded. However, it is thought that the original Portuguese Landrace is probably derived from a narrow genetic base (Almeida 1993; Eldridge *et al.* 1993), which may have led to inbreeding (Eldridge *et al.* 1993). *Eucalyptus globulus* landrace material is now a major component of the breeding and deployment populations in Portugal (Araújo *et al.* 1997; Eldridge *et al.* 1993; Gemas *et al.* 2004). Such programs usually combine the landrace material with more recently introduced germplasm of known Australian origin (e.g. Araújo *et al.* 1997).

The phenotypic expression of quantitative traits has been used to estimate the population structure and genetic variation in landrace and native populations of *E. globulus* (e.g. Almeida *et al.* 1995; Dutkowski and Potts 1999; Lopez *et al.* 2002). Such analyses show that considerable spatially structured quantitative genetic variation exists in *E. globulus* and that the Portuguese Landrace appears to have affinities with the South-eastern Tasmanian race (Lopez *et al.* 2001; Orme 1977). However, many morphological traits are subject to selection, potentially giving an inaccurate picture of the genetic diversity and affinities of a given landrace (Jones *et al.* 2002a; Steane *et al.* 2006). The advent of selectively neutral molecular markers offers powerful tools to more accurately investigate these issues (e.g. Astorga *et al.* 2004; Gemas *et al.* 2004). Strong genetic differentiation exists within *E. globulus* at the molecular level in nuclear (Nesbitt *et al.* 1995; Jones *et al.* 2002a; Steane *et al.* 2006) and chloroplast (Freeman *et al.* 2001; Soto *et al.* 2004) DNA markers. Two such markers developed for *E. globulus* are the J_{LA+} (Freeman *et al.* 2001) region of chloroplast DNA (cpDNA) and nuclear SSR (Steane *et al.* 2001). CpDNA is inherited uniparentally and maternally in *Eucalyptus* (Byrne *et al.* 1993; McKinnon *et al.* 2001b), so will reflect the matrilineal component of an individual's pedigree, while nuclear SSR will reflect the overall genetic composition of an individual because they recombine in each generation. Several studies have

investigated the genetic diversity of selections from the *E. globulus* Portuguese Landrace compared to native material, based on morphology (Almeida *et al.* 1995; Lopez *et al.* 2001) and ISSR markers (Gemas *et al.* 2004); however, this study is the first to use cpDNA and nuclear SSR markers in an attempt to find the Australian origins and to compare the amount of genetic diversity in the Portuguese *E. globulus* landrace with that of native populations of *E. globulus*.

Materials and methods

Forty-seven trees were collected from 29 different localities from throughout the regions where *E. globulus* is grown in Portugal (Table 2.1; Fig 2.2). These trees were part of the initial RAIZ plus tree selection program, carried out during the late 1980s. The plantations were established using unimproved genetic material collected and produced in Portugal, hence representing the local Landrace. In addition, two trees of know Australian origin were also collected (see blind controls in Table 2.1), but the origins of these trees were masked until after the analysis was finished. DNA was extracted by the Doyle & Doyle (1990) method as modified by Grattapaglia & Sederoff (1994).

The J_{LA+} region of the chloroplast genome was amplified and sequenced in both the forward and reverse directions, following the methods of Freeman *et al.* (2001), except that sequencing was performed on a CEQ8000 (Beckman Coulter) automated sequencer. Sequences were aligned manually using Sequence Navigator software (ABI PRISM/Perkin-Elmer). Haplotypes were classified by comparing the cpDNA sequence in each of the 47 Portuguese individuals with our extensive database of J_{LA+} variation, comprising 122 variable characters scored in 292 trees from native populations of *E. globulus*. The database incorporates 225 trees genotyped by Freeman *et al.* (2001), 37 by McKinnon *et al.* (2004) and an additional 30 native trees which were genotyped for this study, including two individuals from the Furneaux group, five from the Otways and 23 from South-eastern Victoria. The sequence characters were based on those outlined by Freeman *et al.* (2001), with the addition of new characters discovered in *E. globulus* since that study. Portuguese Landrace individuals with cpDNA sequence identical to trees in the native range were identified as possessing the same (chloroplast genotype or) haplotype. CpDNA sequence affinities were used to assign haplotypes to groups that have been defined by phylogenetic analysis (Freeman *et al.* 2001; McKinnon *et al.* 2004). Some of these groups were monophyletic, and are referred to as clades (such as the central and southern clades), while the central clade was further subdivided into the Cg and Cc groups.

Thirty-four Portuguese Landrace individuals were fingerprinted using nuclear SSR markers. PCRs for SSR amplification used a total volume of 10 µl, containing 20 ng DNA, PCR buffer (67 mM Tris-HCl, pH 8.8, 16.6 mM $(NH_4)_2SO_4$, 0.45% Triton X-100, 0.2

Table 2.1 Identity of Portuguese Landrace samples, their chloroplast DNA (cpDNA) haplotype and their cpDNA and SSR assignment to various regions of the natural distribution of *E. globulus*. CpDNA assignment is based on the natural distribution of the clade (Figure 2.3) or haplotype (Figure 2.4). The first letter of the cpDNA haplotype code indicates the clade (C or S) while the second letter, where present, indicated the group within the central clade (c or g). SSR assignment indicates the native race (see Figure 2.1) with the highest and second highest probabilities of assignment, derived from analysis using *Structure* software (Pritchard *et al.* 2000a). * = Haplotypes endemic to the Portuguese Landrace. OP stands for open pollinated. CG741, and CG863 were 'blind controls' from known locations in Australia.

Identity	Locality (or pedigree)	CpDNA haplotype	CpDNA assignment	SSR assignment (prob%)
Portuguese Landrace				
VF18	Azambuja	Cc41	Tas (incl. King Is.)	SE Tas (64): W Tas (17)
TB43	Rio Maior	Cc41	Tas (incl. King Is.)	NE Tas (71): S Tas (17)
LP32	Castelo Paiva	Cc41	Tas (incl. King Is.)	SE Tas (54): NE Tas (14)
PL133	Povoa do Lanhoso	Cc41	Tas (incl. King Is.)	SE Tas (72): S Tas (15)
CT47	Coruche	Cc56	SE Tas	SE Tas (90): Tidal River (2)
LB250	Serra Monchique	Cc56	SE Tas	SE Tas (94): W Tas (1)
CN5	Abrantes	Cc56	SE Tas	SE Tas (89): S Tas (3)
AF12	Arouca	Cc56	SE Tas	
CA19	Mesao Frio	Cc56	SE Tas	SE Tas (80): S Tas (4)
AL12	Nisa	Cc56	SE Tas	
EST7	Ponte Lima	Cc56	SE Tas	
SN10	Sanguinhal	Cc56	SE Tas	
AL10	Nisa	Cc56	SE Tas	SE Tas (89): S Tas (3)
PL30	Constancia	Cc56	SE Tas	SE Tas (94): NE + W Tas (1)
RE25	Santa Tirso	Cc70 *	Widespread	SE Tas (83): S Tas (10)
CN32	Abrantes	Cg33	Gippsland	
FV19	Barcelos	Cg33	Gippsland	SE Tas (77): Strezlecki (6)
LB244	Monchique	Cg33	Gippsland	Tidal River: (29): KI (21)
SMC3	Barcelos	Cg33	Gippsland	Furneaux (30): SE Tas (25)
SPR7	Lousada	Cg33	Gippsland	

AV6	Castelo Branco	Cg33	Gippsland	King Is. (50): SE Tas (17)
ME70	Penamacor	Cg33	Gippsland	SE Tas (42): Gippsland (21)
JG2	Santo Tirso	Cg33	Gippsland	
VC9	Valongo	Cg33	Gippsland	SE Tas (71): Tidal River (9)
FC22	Geres Evora	S4	SE Tas	
CC4	Nisa	S43	SE Tas	SE Tas (76): S Tas (9)
PC10	Paredes de Coura	S43	SE Tas	
RE42	Santo Tirso	S43	SE Tas	SE Tas (95): S Tas (1)
MN43	Montemor Novo	S43	SE Tas	SE Tas (95): S Tas (1)
QG15	Constancia	S43	SE Tas	NE Tas (75): S Tas (7)
MP11	Penamacor	S43	SE Tas	NE Tas (42): S Tas (17)
MN35	Vendas Novas	S43	SE Tas	SE Tas (91): S Tas (1)
ST51	Santo Tirso	S68 *	SE Tas	SE Tas (81): NE Tas (4)
MB238	Bracal	S69 *	SE Tas	
CM7	Celourico	S69 *	SE Tas	
AM7	Arouca	S70 *	SE Tas	SE Tas (36): KI (22)
TC1	Celourico	S70 *	SE Tas	SE Tas (63): S Tas (23)
CR54	Chamusca	S70 *	SE Tas	
AJ1	Azambuja	S70 *	SE Tas	
CN44	Abrantes	S104 *	SE Tas	SE Tas (65): E Otways (15)
BN22	Ferreira Zezere	S105 *	SE Tas	SE Tas (70): S Tas (7)
CCD2	Nisa	S106 *	SE Tas	SE Tas (54): KI (19)
PL139	Constancia	S107 *	SE Tas	King Isld (51): SE Tas (33)
CH3	Amarante	S110 *	SE Tas	SE Tas (76): Furneaux (11)
VZ3	Nisa	S110 *	SE Tas	SE Tas (32): NE Tas (19)
Q7	Not given	S111 *	SE Tas	W. Otways (47): W Tas (14)
MN303	Vendas Novas	S112 *	SE Tas	SE Tas (54): Furneaux (14)
Blind controls				
CG863	From OP seed collected on Furneaux	JCg39 *	Gippsland or Furneaux	Furneaux (76): Tidal River (7)
CG741	From OP seed collected on Furneaux	JCg43 *	Gippsland or Furneaux	Furneaux (80): SE Tas (9)

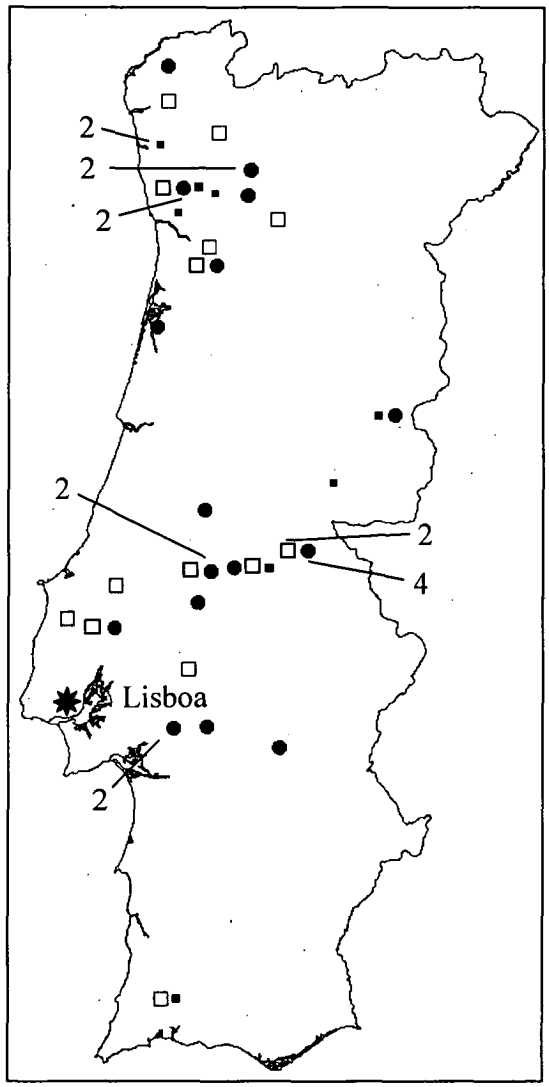


Figure 2.2 Map of Portugal, showing the localities where samples were taken and the distribution of the major chloroplast DNA clades and groups of haplotypes in Portugal. Numbers indicate the number of trees at localities with multiple samples. See Figure 2.3 for key to symbols.

mg/mL gelatine), 200 μM dNTPs, 2 mM $MgCl_2$, 5% dimethyl sulphoxide (DMSO), 100 nM of each forward and reverse primer (EMCRC 1a, 3, 7, 11) or 150 nM of each forward and reverse primer (EMCRC 2, 10, 12), 0.5 U *Taq* DNA polymerase. Sterile distilled water was added to achieve 10 μL final volume. PCR conditions (using a PTC-100, MJ Research, Inc. or Eppendorf Master Cycler Gradient, Eppendorf®) were: denaturation at 94°C for 30 sec; 15 cycles of denaturation at 94°C for 30 sec, annealing (at 56°C decreasing by 0.2°C each cycle) for 30 sec, and extension at 72°C for 45 sec. Followed by 20 cycles with conditions as above, except annealing at 53°C and a final extension step at 72°C for 7 min. A LI-COR 4200 sequencer was used to separate SSR alleles, using 6% acrylamide gels with L4000-448 as a size standard; electrophoretic output was recorded and alleles were sized using RFLPscan software (Scanalytics).

In order to assign the Portuguese Landrace individuals to native races of *E. globulus* (as defined by Steane *et al.* [2006]; Figure 2.1), the allele composition at 7 SSR loci in the Portuguese Landrace individuals was compared to that from 372 native trees representing 11 quantitative races (Eastern Otways, Western Otways, Southern Gippsland, Strzelecki Ranges, Furneaux Group, North-eastern Tasmania, South-eastern Tasmania, Southern Tasmania, Western Tasmania and King Island; Steane *et al.* 2006) and Tidal River in Wilsons Promontory National Park (Steane *et al.* unpubl. data), using the software *Structure* (Pritchard *et al.* 2000a). The loci used were chosen because they had already been genotyped in the aforementioned native trees (Steane *et al.* 2006 and unpubl. data). All loci were unlinked, with the possible exception of EMCRC 7 and 12 for which there was some evidence of linkage, although it was not conclusive (Steane *et al.* 2006). *Structure* employs a Bayesian clustering method to assign multi-locus genotypes of individuals to specific populations on the basis of allele frequencies estimated for each pre-defined population (i.e. native race). In order to assign individuals to races, a model was used that incorporated admixture and independent allele frequencies between populations. A burn-in of 100,000 iterations was followed by a run length of 100,000 iterations. Portuguese Landrace individuals were allocated to native races based on their probability of assignment from *Structure*. POPGENE (Version 1.31; Yeh *et al.* 1997) software was used to calculate the observed (N_a) and effective (N_e)

number of alleles, as well as the observed and expected heterozygosities (H_o and H_e) for the Portuguese Landrace sample, allowing comparison of these parameters with those obtained for the total native population, the mean of all the races, or the individual races.

Results

Chloroplast DNA diversity

All 47 Portuguese Landrace samples belonged to either of the two major clades found in native *E. globulus*, designated central (C; 24 samples) and southern (S; 23 samples) after their natural distribution (Fig 2.3, Table 2.2). The Portuguese Landrace collection was quite diverse at the haplotype level, with 16 haplotypes present in the 47 Portuguese Landrace samples for which complete J_{LA+} sequence was obtained. Despite the evident haplotype diversity, haplotype sharing was common, with 30 individuals represented by 4 common haplotypes (Cc41, Cc56, Cg33, S43). Eleven of the 16 haplotypes (one C and ten S) were unique to the Portuguese Landrace, while the remaining five (Cc41, Cc56, Cg33, S4, S43) have been found in natural stands. Within the major clades found in the Portuguese Landrace, a greater haplotype diversity (h = number of haplotypes/number of individuals) was evident in the S clade (0.52) than the C clade (0.17). Native trees from throughout the natural distribution of *E. globulus* have similar haplotype diversity within the S clade (0.52), but have more diversity within the C clade (0.33). Within the Central clade, the Cc group was more common (15 individuals) than the Cg group (nine individuals) in the Portuguese Landrace, which was also the case in the native trees. However, only three different Cc haplotypes were detected compared to one Cg haplotype, clearly indicating reduced genetic diversity in the Cg group in the Portuguese Landrace (h = 0.11) compared to the native gene pool (h = 0.36; Table 2.2).

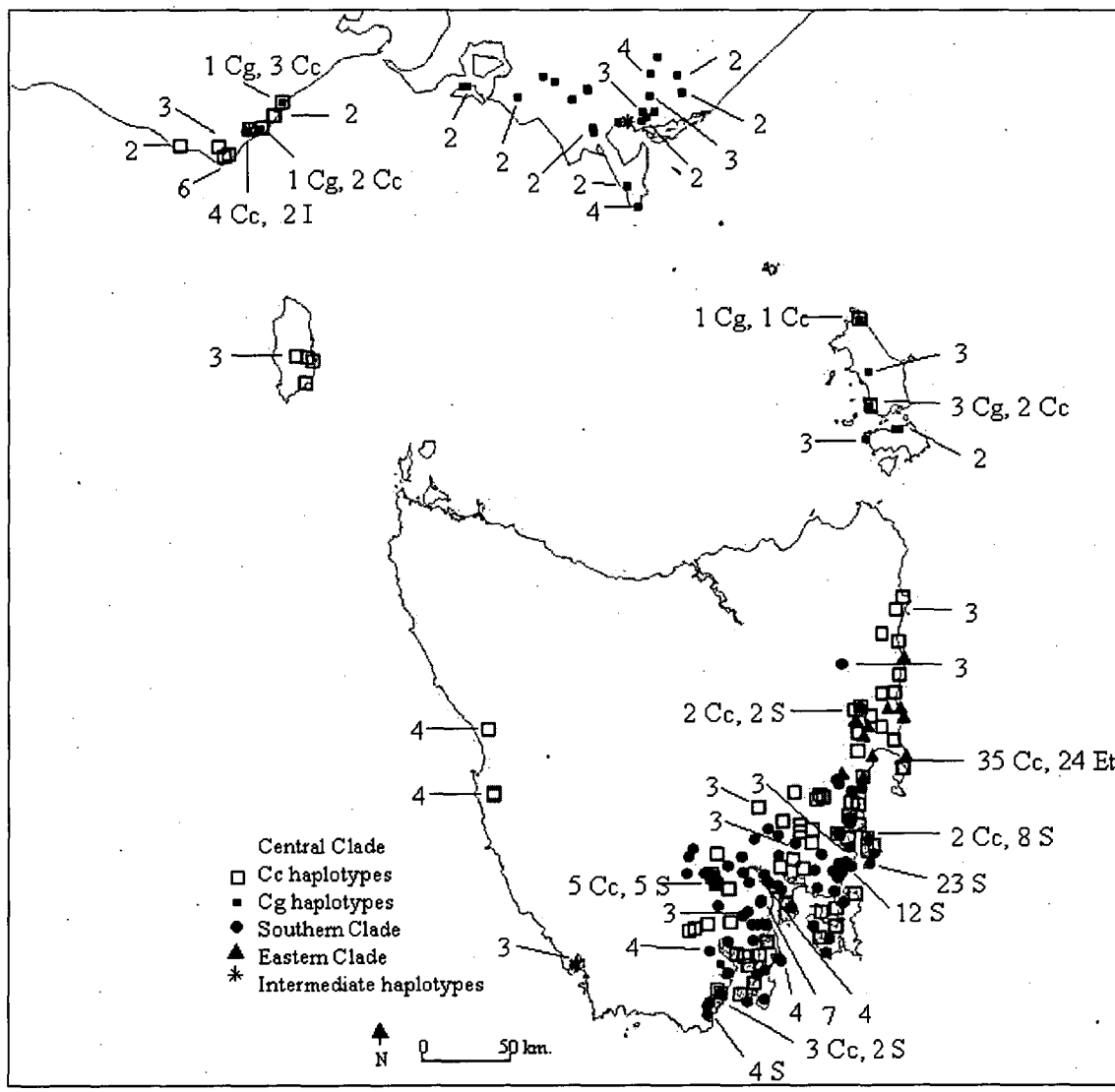


Figure 2.3 Broad-scale distribution of the major chloroplast DNA clades and groups of haplotypes across the native distribution of *Eucalyptus globulus* in south-eastern Australia. Numbers indicate the number of trees at localities with multiple samples.

Table 2.2 The number of cpDNA haplotypes and samples assigned to each major cpDNA clade or haplotype group of *Eucalyptus globulus*, in the native range and in the Portuguese Landrace.

		Native range				Portuguese Landrace			
Clade or haplotype group	Natural distribution	No. of haplotypes	No. of samples	h[a]	% of samples	No. of haplotypes	No. of samples	h[a]	% of samples
Central (Cc)	Widespread, but infrequent in Furneaux Group and not found in south-eastern Victoria	34	108	0.31	37.0	3	15	0.20	32
Central (Cg)	Not in Tasmania or King Island, most frequent in Furneaux Group and south-eastern Victoria	21	58	0.36	20.0	1	9	0.11	19
Southern (S)	Only southern and eastern Tasmania	57	109	0.52	37.0	12	23	0.52	49
Eastern (Et)	Only north-eastern Tasmania	6	12	0.50	4.0	0			
Intermediate (I)	Rare over whole range	2	2	1.00	0.7	0			
Western (W)	Only in south-western Tasmania	1	3	0.33	1.0	0			
Totals		122	292	0.42		16	47	0.34	

[a] Number of haplotypes per sample for each clade or group.

Chloroplast DNA affinities

Thirty-one of the 47 Portuguese Landrace individuals possessed haplotypes previously found in the native population of *E. globulus* (Table 2.1). The majority (26/31) of individuals with known cpDNA haplotypes in the Portuguese Landrace had haplotypes that in our database were restricted to two broad regions of the natural range of *E. globulus*, south-eastern Tasmania (Cc56, S4, and S43; Figure 2.4) and south-eastern Victoria (Cg33; Figure 2.4). The disproportionate representation of the S clade in the Portuguese Landrace (49% of samples) compared to that in the native range (37%) is evidence for a substantial south-eastern Tasmanian contribution to the Portuguese Landrace, since the S haplotypes have been found only in Tasmania and most come from the south-east (Figure 2.3). Clear affinities of some Portuguese Landrace samples to south-eastern Victoria (including Gippsland and the foothills of the Strezlecki Ranges) are suggested by the common occurrence of haplotype Cg33 (nine Portuguese Landrace individuals), which is widespread in this region and has not been found elsewhere to date (Figure 2.4). Haplotype Cc41, found in four Portuguese Landrace samples, predominantly occurs in south-eastern Tasmania, but has also been found on King Island (Figure 2.4). No clear spatial pattern was evident in the cpDNA haplotype or clade distribution in the 29 different localities sampled within Portugal (Figure 2.2). Similarly, most Portuguese localities with multiple trees sampled featured a mix of the major cpDNA clades or groups in *E. globulus*.

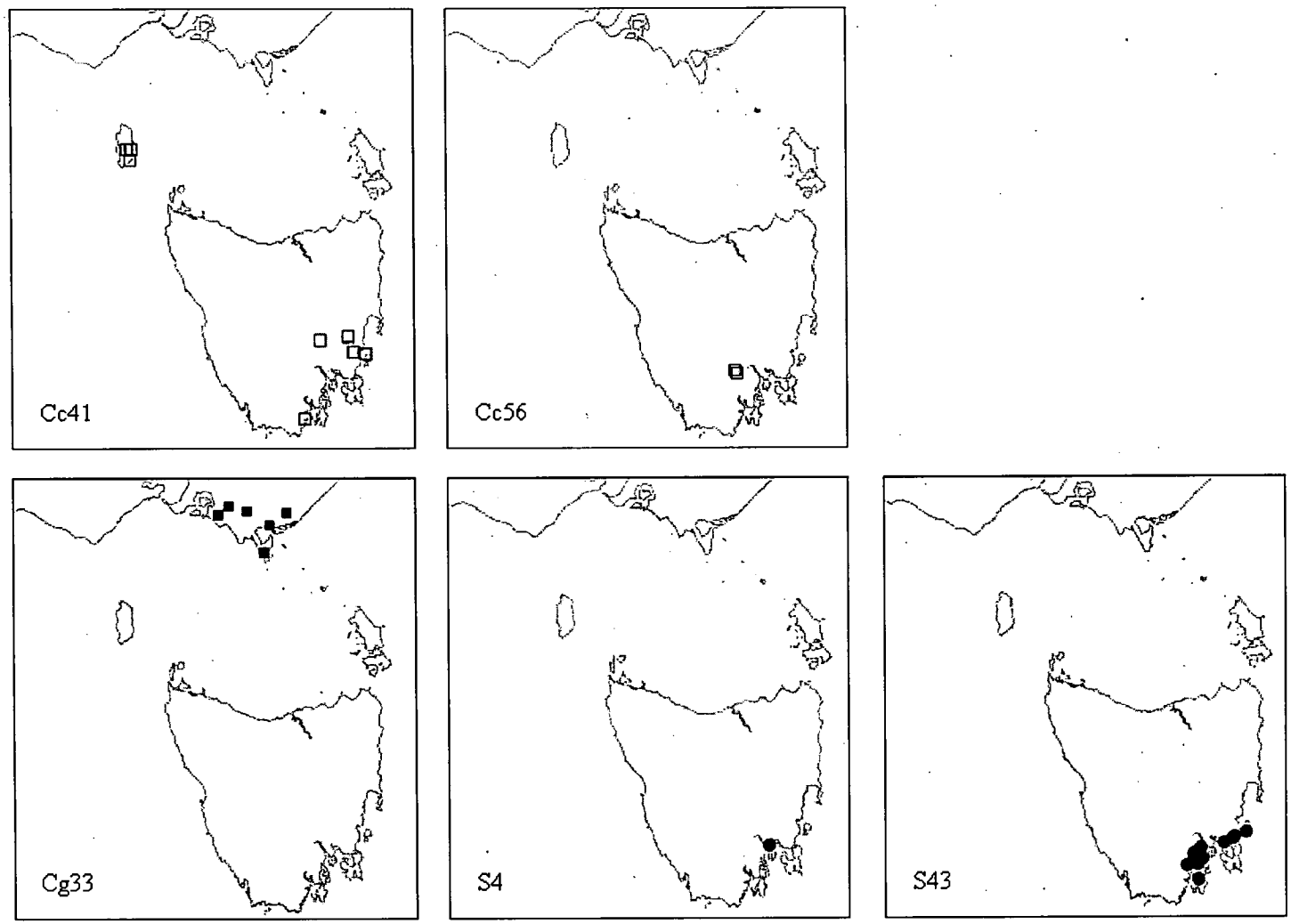


Figure 2.4 The native Australian distribution of individual haplotypes found in the Portuguese Landrace of *Eucalyptus globulus*.

Microsatellites

Variation at seven SSR loci was examined in 34 individuals of the Portuguese Landrace. This collection of the Portuguese Landrace was highly polymorphic, with a mean of 9.7 alleles per locus. However, the mean effective number of alleles per locus ($N_e = 4.8$) was close to half the observed number of alleles per locus, indicating the presence of numerous rare alleles in the Portuguese Landrace. The high number of alleles was reflected in the high observed and expected heterozygosity ($H_o = 0.62$ and $H_e = 0.75$, respectively). Analysis of these SSR data using *Structure* suggested that the majority (26/34) of the Portuguese Landrace individuals have their closest affinities to the South-east Tasmanian race (Table 2.1). Other Portuguese Landrace individuals (3/34) had closest SSR affinities to the North-east Tasmanian race, King Island (2/34), the Furneaux Islands, the Otway region of Western Victoria and Tidal River in the south-east of Victoria (one individual each; Table 2.1).

In most cases, the SSR data confirmed the affinities suggested by cpDNA, in some instances, with a greater resolution. For example, those bearing haplotypes Cc41, which is found in both King Island and south-eastern Tasmania, all had closest SSR affinities to south-eastern Tasmania (Table 2.1). In agreement with the cpDNA data, the Portuguese Landrace had SSR affinities to eastern Tasmania in the bulk (15/17) of the trees (with SSR data available) bearing the S haplotype (13 South-eastern Tasmania, 2 North-eastern Tasmania) and all individuals bearing haplotypes Cc41 and Cc56 (9 South-eastern Tasmania, 1 North-eastern Tasmania). However, for seven individuals there was a clear discrepancy between the affinities suggested by SSR and cpDNA (Table 2.2). The majority of these discrepancies arose in individuals bearing C haplotypes (particularly Cg), with cpDNA affinities to mainland Australia, but SSR affinities to south-eastern Tasmania.

DISCUSSION

Origins of the Portuguese Landrace

The ability to locate where the original Australian collections for the Portuguese Landrace were made relies on the pronounced spatial structure in native populations of *E. globulus* that is evident in chloroplast (Freeman *et al.* 2001) and, to a lesser degree, in nuclear (Jones *et al.* 2002a; Steane *et al.* 2006) DNA markers. In terms of cpDNA, spatial structuring of variability in the native stand is evident in the distribution of clades and specific haplotypes (Freeman *et al.* 2001; McKinnon *et al.* 2004), both of which can be useful in assessing the matrilineal origin of the Portuguese Landrace. At the broadest spatial scale, cpDNA types are divided into clades or groups on the basis of phylogenetic analysis (Freeman *et al.* 2001; McKinnon *et al.* 2004). Three groups are especially informative because of their regional distribution: the S and Eastern (Et) clades and the Cg haplotypes (Table 2.1 and Figure 2.3). In addition, some haplotypes have been found only in very localised areas (Figure 2.4), hence in many cases the haplotype of an individual can provide a more precise estimate of the likely matrilineal origin than the clade alone.

Numerous studies have successfully used *Structure* software to assign individuals to groups using multi-locus genotype data (e.g. Hansen *et al.* 2001; Harter *et al.* 2004; Yang *et al.* 2005). Similarly, the frequency of nuclear SSR alleles in native *E. globulus* exhibits substantial geographic structuring, such that individuals can be readily assigned to the races identified in the native stand (Dutkowski and Potts 1999) by their allele frequencies at 7 SSR loci (Steane *et al.* 2006). However, because the nuclear genome is diploid and undergoes recombination in each generation, the origin suggested by SSR markers may be complicated by (possibly several generations of) inter-breeding between individuals originating from different native races of *E. globulus* since their introduction to Portugal. Despite differences between the two marker classes in modes of inheritance and genetic architecture in the native stand, the SSR affinities of the Portuguese Landrace individuals were largely congruent with the cpDNA evidence. The two blind

controls demonstrated the power of combining SSR and cpDNA analysis by independently identifying the same area of origin (Table 2.1).

The similar affinities suggested by the two independent marker systems provide strong evidence that the Portuguese Landrace individuals herein were predominantly derived from two broad regions, south-eastern Tasmania and to a lesser extent south-eastern Victoria. In particular, the frequent occurrence of cpDNA haplotypes from the S clade in the Portuguese Landrace individuals is evidence for a south-eastern Tasmanian matrilineal origin for many of the Portuguese Landrace samples; in our extensive sampling of *E. globulus* to date this clade has not been found outside eastern Tasmania and is most common in the south-east of the state (Figure 2.3). A south-eastern Tasmanian origin is supported by the high frequency of individuals with their closest SSR affinities to the South-eastern Tasmanian race. Furthermore, a strong representation of haplotype Cg33 in the Portuguese Landrace is almost certain to represent a south-eastern Victorian origin because it is common and restricted to that region in our studies of the native population to date.

The Otways region of western Victoria and King Island remain as possible, but not likely, areas of origin for some Portuguese Landrace individuals. However, in all three cases where these regions are suggested there is disagreement between the origin inferred from cpDNA and SSR markers. For example, haplotype Cc41 has been found in native trees from King Island (Figure 2.4), as well as eastern Tasmania, however the SSR affinities of all Portuguese Landrace individuals bearing Cc41 are to eastern Tasmania (3 south-eastern, 1 north-eastern), making King Island an unlikely origin for these individuals. Two individuals (AV6 and PL139) have their closest SSR affinities to King Island and one (Q7) to the Otways. However, in each case the probabilities of assignment are all close to 50%, well below the mean for this study (65.5%), with a substantial contribution from the race with the second highest probability of assignment, which in each case was from Tasmania (see Table 2.1). Such probabilistic assignment to multiple groups using *Structure* software has been used to infer admixture between differentiated groups in trout (Hansen *et al.* 2001), sunflower (Harter *et al.* 2004) and

between teosinte and maize (Fukunaga *et al.* 2005). The three Portuguese Landrace samples (AV6, PL139 and Q7) for which the SSR data suggests a substantial contribution of two different races, representing Tasmania and mainland Australia, are also likely to represent admixture between trees originating from different races. Similarly, in other trees the assignment probability suggests a substantial contribution of two different races (e.g. AM7 and SMC3; Table 2.1) and this is likely to be indicative of admixture.

Despite the general agreement about origins suggested by cpDNA and SSR, some obvious discrepancies exist between the data sets. Admixture between trees originating from different native races since their introduction to Portugal could account for the observed cytonuclear discordance, because progeny of such events would have the maternal cpDNA genotype, but a nuclear genotype reflecting the contribution of each parent (see Steane *et al.* 1998). In most cases the discrepancy arose where an individual with a cpDNA haplotype characteristic of mainland Australia is found in an individual with closest SSR affinities to South-east Tasmania. This result is consistent with the introgressive displacement of the nuclear genome of minor races by the most common native race represented in the Portuguese Landrace (South-eastern Tasmania). This would be likely to occur where seed is derived from open pollination of trees with a mainland Australian maternal ancestry, due to pollen swamping by the most common (South-east Tasmanian) race. The co-occurrence of trees with haplotypes from south-eastern Tasmania and mainland Australia in numerous locations in Portugal (Figure 2.2) is consistent with this hypothesis, since it shows that hybridisation between trees originating from various native races is possible.

Clearly, admixture between Portuguese Landrace trees originating from different native races of *E. globulus* will make identification of the native origin difficult on the basis of quantitative traits or nuclear markers alone. However, in agreement with the origins suggested in this study, a predominantly southern or south-eastern Tasmanian origin of *E. globulus* plantations in Portugal was suggested by Orme (1977) based on morphological observations, while Lopez *et al.* (2001) found that the closest quantitative

genetic affinities of the Portuguese Landrace were to southern Tasmania. A southern Tasmanian origin was also suggested by morphological (Orme 1977) and molecular (Astorga 2004) affinities for the Spanish *E. globulus* Landrace, as well as quantitative genetic affinities of landrace material from southern China (Zang *et al.* 1995) and Chile (Lopez *et al.* 2001). These findings are consistent with southern Tasmania being an early source of seed that was distributed around the world. France may have been a key distribution source of this germplasm and was active in spreading seeds to other countries (Doughty 2000). The first *E. globulus* introduced into France may have been derived from seed collected by French explorers when they visited southern Tasmania during the late 18th (D'Entrecasteaux expedition) or early 19th (Baudin expedition) century (Doughty 2000).

Genetic diversity of the Portuguese Landrace

While less than native *E. globulus* (122 haplotypes in 292 samples = 0.42), the cpDNA diversity in the Portuguese Landrace was substantial (16 haplotypes among 47 samples = 0.34). Assuming that the cpDNA haplotypes of the Portuguese samples have undergone few mutations since the original collection(s) was(were) made in Australia, the discovery of 16 haplotypes means that a minimum of 16 Australian trees is likely to have formed the basis of the Portuguese Landrace. However, it is quite possible that more trees were originally sampled since, in native stands, trees in close proximity may possess the same J_{LA+} haplotype (McKinnon *et al.* 2004). The fact that many of the Portuguese localities from which multiple individuals were sampled featured a mix of the major clades and haplotypes within *E. globulus* suggests that Portuguese plantations are likely to be genetically diverse, even though only two major regions of the native distribution are represented.

The measures of (SSR) genetic diversity in this sample of the Portuguese Landrace are comparable to those found within single races of *E. globulus* (Steane *et al.* 2006). Although there was a high number of alleles per locus (N_a = 9.7), a reduction in the effective number of alleles per locus (N_e = 4.8) was evident, indicating the presence of

numerous rare alleles, as is the case in the natural population (mean per race N_a= 9.5, N_e= 4.9; Steane *et al.* 2006). However, both N_a and N_e were lower in the Portuguese Landrace than across the entire species (N_a= 19.4, N_e =6.06). The observed and expected heterozygosities in this sample of the Portuguese Landrace (H_o = 0.62, H_e = 0.75) were very similar to those found in single races of *E. globulus* (mean diversity from 10 races encompassing the natural range of *E. globulus*, H_o = 0.65, H_e = 0.75; Steane *et al.* 2006), but the expected heterozygosity in the Portuguese Landrace was lower than the overall expected heterozygosity in the species (H_e = 0.82; Steane *et al.* 2006).

The finding of significant genetic diversity within the Portuguese Landrace is supported by quantitative genetic evidence that different provenances from the Portuguese Landrace appear to be as variable as a selection of 12 native provenances (Orme 1977) when assessing growth, wood density and frost tolerance (Almeida *et al.* 1995). Gemas *et al.* (2004), also found acceptable genetic diversity in selections from the Portuguese Landrace versus native stand material using ISSR markers. The aforementioned admixture between trees originating from different native races may have increased heterozygosity in the Portuguese Landrace and, in combination with natural and artificial selection, contributed to genetic differentiation since introduction. This suggestion is supported by observations of quantitative traits such as frost tolerance (Almeida *et al.* 1995) and growth form (Lopez *et al.* 2001).

Conclusion

Molecular profiles of the Portuguese *E. globulus* Landrace suggest that south-eastern Tasmania and, to a lesser extent, south-eastern Victoria, were very likely collection points. Although we argue against other potential areas of origin (e.g. King Island and the Otway Ranges) suggested by some of the molecular data, it remains a possibility that these regions had a minor contribution to the Portuguese Landrace. The relatively high level of genetic diversity (in cpDNA sequence and nuclear SSR) found in the Portuguese Landrace, and the fact that original collections appear to have been taken from at least two widely separated locations in the native range, are not consistent with previous suggestions that the Portuguese Landrace may have been derived from a very narrow original collection. However, the molecular evidence suggests the Portuguese Landrace is dominated by genetic material from south-eastern Tasmania, consistent with evidence from quantitative genetic data. Recently, selections for pulpwood breeding objectives derived from *E. globulus* base populations have favoured germplasm from races such as the Stezlecki Ranges, Otways and Furneaux (Griffin 2001; Apiolaza *et al.* 2005), which appear to be under represented in the Portuguese Landrace.

Chapter 3: Parental and consensus linkage maps of *Eucalyptus globulus* using AFLP and microsatellite markers.

Published as: Freeman JS, Potts BM, Shepherd M, Vaillancourt RE (2006) Parental and consensus linkage maps of *Eucalyptus globulus* using AFLP and microsatellite markers. *Silvae Genetica* **55**, 202-217.

Introduction

Eucalyptus globulus is one of the most commercially important plantation tree species in temperate regions of the world (Eldridge *et al.* 1993). As a result, breeding programs have been established in order to produce genetically superior *E. globulus* trees (Potts *et al.* 2004). Linkage maps and associated molecular markers can be useful in tree improvement programs as prerequisites for the study of quantitative traits using the quantitative trait loci (QTL) approach (Grattapaglia 1997, 2000; Ahuja 2001). Consequently, linkage maps have been produced for many forest tree genera of economic importance including *Eucalyptus* (e.g. Byrne *et al.* 1995), *Pinus* (e.g. Plomion *et al.* 1995), *Poplar* (e.g. Bradshaw and Stettler 1994) and *Acacia* (Butcher *et al.* 2002).

Molecular maps produced in tree species have employed a variety of DNA marker classes, each having their merits (reviewed by Cervera *et al.* 2000). For example, AFLP (Vos *et al.* 1995) are anonymous and dominant markers, based on arbitrarily selected PCR primers that require no *a priori* sequence knowledge, primer or probe development. The AFLP technique has the highest multiplex ratio of any PCR based technique (Myburg *et al.* 2001), allowing the production of high coverage parental maps in various plants including interspecific crosses of eucalypts such as *E. tereticornis* x *E. globulus* (Marques *et al.* 1998) and *E. grandis* x *E. globulus* (Myburg *et al.* 2003). However, dominant markers will predominantly be mapped to one parent, hence often co-dominant markers must be included to merge parental maps (Grattapaglia 2000). Similarly, the transferability of AFLP is often low, even within species (Cervera *et al.* 2000).

SSR markers have many attributes suitable for map construction (Cervera *et al.* 2000). They are multi-allelic, usually highly polymorphic and co-dominant in their inheritance (Avise 1994). However SSR markers require considerable investment in primer development, limiting their availability. For example, the limited availability of SSR primers for *Eucalyptus* has, until recently, made the construction of maps based solely on these markers unfeasible. In order to remedy this problem, AFLP markers can be used to provide high density maps on which to map co-dominant markers, such as SSRs (e.g. Cervera *et al.* 2001).

Importantly, SSRs are often highly transferable between pedigrees and species (Jones *et al.* 2001; Steane *et al.* 2001; Marques *et al.* 2002), allowing integration of parental maps, comparative mapping and ultimately the construction of species consensus maps (Grattapaglia 2000). The comparison of genetic linkage maps based on a common set of markers allows the identification of homologous loci and collinear chromosome segments between species. Comparative maps have demonstrated a high degree of co-linearity for numerous plant species including tomato and pepper, wheat and rice, and pine species *Pinus taeda* and *Pinus radiata* (Ahuja 2001). Similarly, in eucalypts, the mapping of interspecific hybrid progeny has allowed the demonstration of a high degree of synteny between maps of *E. grandis* and *E. urophylla* (Brondani *et al.* 1998, 2002), *E. grandis* and *E. globulus* (Myburg *et al.* 2003), and *E. grandis, E. urophylla, E. tereticornis* and *E. globulus* (Marques *et al.* 2002). The apparent genome conservation between many of the economically important eucalypt species within the subgenus *Symphyomyrtus* suggests that the construction of a consensus map for the subgenus will be attainable (Brondani *et al.* 2002; Marques *et al.* 2002).

Comparative mapping and ultimately the construction of a consensus map for the genus will be invaluable for molecular breeding and more fundamental research, such as studying genome evolution and genetic differentiation in *Eucalyptus*. In the area of molecular breeding, comparative mapping would enable the exchange and comparison of information between studies world-wide in areas such as QTL mapping, candidate gene mapping and marker-assisted selection (Brondani *et al.* 2002). In particular,

comparative mapping will allow the validation of putative QTL over different genetic backgrounds (e.g. Marques *et al.* 2002), providing a basis for effective marker-assisted selection (MAS) (Brondani *et al.* 2002).

Forest trees such as *Eucalyptus* are generally characterised by long generation times, high levels of genetic diversity and an absence of multi-generation pedigrees or inbred lines (Sewell and Neale 2000). Long generation times and an absence of multi-generation pedigrees in particular, combined with the large areas required for progeny trials, have limited the establishment of large advanced generation crosses which are most useful for the production of linkage maps. Most maps produced have utilised interspecific F_1 material and markers segregating in a double pseudo-testcross configuration, taking advantage of the high levels of heterozygosity found in eucalypt species (Grattapaglia and Sederoff 1994; Verhaegen and Plomion 1996; Marques *et al.* 1998; Agrama *et al.* 2002). Fewer studies have utilised F_2 crosses with grandparents from widely disjunct localities (Byrne *et al.* 1995) or interspecific hybrids (Myburg *et al.* 2003). In contrast to the majority of previous mapping experiments in eucalypts, this study utilised markers segregating in an F_2 progeny to map the F_1 parents from a wide intraspecific cross. The widely separated provenances of origin and three generation design of this cross maximises polymorphism by capturing the diversity between eucalypt populations. The increased polymorphism increases the likelihood of QTL segregating in this pedigree.

The aim of this study was to produce parental and consensus maps in an outbred, F_2 inter-provenance cross of *E. globulus,* using AFLPs and predominantly fully informative SSR loci. The mapping of fully informative SSR loci derived from a variety of sources (EMBRA, CSIRO, EMCRC), allowed comparison of synteny and collinearity of shared markers between this study and maps previously produced in a range of eucalypts from the subgenus *Symphyomyrtus* (Brondani *et al.* 1998, 2002; Bundock *et al.* 2000; Marques *et al.* 2002; Thamarus *et al.* 2002). As such, these maps provide a contribution toward the production of reference maps for *E. globulus* and the subgenus. The linkage maps produced in this study will also be used to investigate marker trait associations

(QTL) for the genetic control of important economic traits, including resistance to fungal pathogens from the genus *Mycosphaerella*.

Materials and Methods

Plant material

The mapping population consisted of an outbred, F_2 full sib family derived from an inter-provenance cross. The four grandparents were from two widely separated locations in Tasmania – King Island, situated to the northwest of Tasmania and Taranna in the far southeast (Figure 3.1). Trees from these locations are classified into different geographical races, differentiated by numerous quantitative traits (Dutkowski and Potts 1999). Crossing between unrelated grandparents from King Island and Taranna produced each F_1, which were then crossed to produce the F_2 (Figure 3.2). These grandparents were chosen in order to maximise the segregation of quantitative characters that differ between provenances, such as resistance to *Mycosphaerella sp.* (Carnegie *et al.* 1994). Seedlings were planted in a randomised incomplete block design, at Woolnorth in northwest Tasmania in May 1998. Originally 240 genotypes from the cross were planted with 160 genotypes represented by two clonal replicates. This cross suffered a high mortality (28%), and leaf tissue was collected in April 2001, only from the 121 different F_2 genotypes that had both clonal replicates surviving.

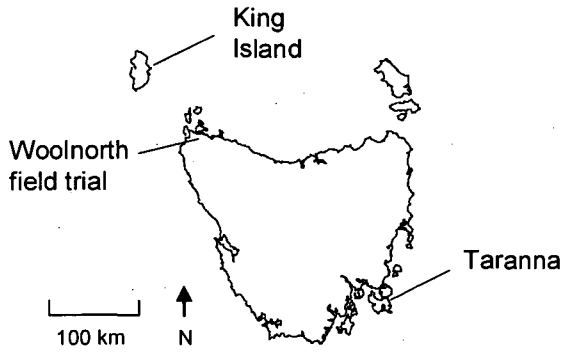


Figure 3.1 Origins of the grandparents of the mapping Pedigree in Tasmania, Australia.

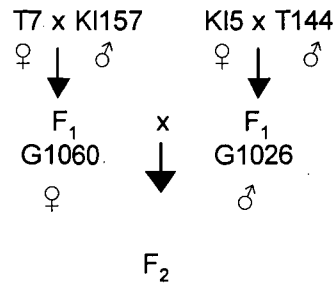


Figure 3.2 Pedigree of the mapping population. 'T' indicates grandparents of Taranna origin, 'KI' indicates King Island (see Figure 3.1). (The grandparents correspond to those used by Hodge *et al.* [1996]; T7 = h, KI157 = 19, KI5 = c, T144 = 6).

Total genomic DNA was extracted from 1 g of fresh leaf tissue, using a CTAB extraction protocol (Doyle and Doyle 1990). Following extraction, DNA samples were cleaned using Prep-a-Gene matrix purification (Bio-Rad). DNA concentration and purity were estimated using agarose gel electrophoresis with ethidium bromide staining, by comparison with a standard molecular weight marker (lambda *Hin*d III). Samples were diluted 1:4 in de-ionised water for PCR amplification.

SSR analysis

The parents of the mapping cross were initially screened with three sets of SSR primers: 233 from EMBRA (Brondani *et al.* 1998, 2002), 13 from EMCRC (Steane *et al.* 2001) and 34 from CSIRO (Byrne *et al.* 1996 and http://www.ffp.csiro.au/tigr/molecular/eucmsps.html). Loci segregating in both parents were preferentially selected to screen the progeny of the mapping cross. The details of primer sequences and reagents for the EMCRC SSRs are those given by Steane *et al.* (2001) and conditions for the CSIRO SSRs are given by Byrne *et al.* (1996). All SSR PCR reactions were carried out in 12.5 µl (final volume) reactions. The volume of each primer was elevated from 0.2 µM to 0.3 µM/reaction for the CSIRO SSRs. The PCR reagents used for the amplification of the EMBRA SSR are essentially those of Brondani *et al.* (1998) with the following modifications: 0.1% Triton X-100, 0.1 mg/mL BSA and no DMSO. Primer sequences and recommended annealing temperatures for EMBRA loci 1-20 are given by Brondani *et al.* (1998), for EMBRA 21-70 by Brondani *et al.* (2002) and for EMBRA 75 – 240 by Brondani (2001). EMCRC primers were synthesised with a fluorescent label (either HEX or TET) attached to the 5' end to allow detection of PCR products by a Gel Scan 2000 real-time acrylamide gel system (Corbett Research). The EMBRA and CSIRO markers incorporated a fluorescently labelled dUTP at 0.6 µmol/PCR reaction (Chromatide™ tetramethylrhodamine-6-dUTP, Molecular Probes) for Gel Scan 2000 detection. PCR reaction conditions of SSR loci often required optimisation for detection using the Gel Scan 2000, by slight modification of the recommended annealing temperature or $MgCl_2$ content (Appendix 1).

All SSR loci were amplified using the same amplification conditions; an initial denaturation step at 94°C for 2 minutes, followed by 30 cycles of 94°C for 1 minute, annealing temperature for 1 minute, and an extension step at 72°C for 1 minute. Following the 30 cycles there was a final extension step of 72°C for 5 minutes. A 'touchdown' PCR was used to enhance the quality of PCR products, with the annealing temperature dropping by 2°C every 10 cycles. Annealing temperature varied within and between primer sets (Appendix 1).

SSR fragments were separated on 15 cm 4% polyacrylamide gels, using a Gel Scan 2000. Gel analysis was performed using Gene Profiler™ software (Scanalytics, Inc.) by comparison to a Promega (CXR) 60-400 base fluorescent ladder, allowing a semi-automated scoring of individuals.

AFLP analysis

The method adopted for AFLP generation essentially followed the protocol of Vos *et al.* (1995), with minor adjustment for eucalypt DNA after Marques *et al.* (1998). The standard procedures were modified to allow restriction of genomic DNA and ligation of adaptors in a single step, with detection of fragments by a Gel Scan 2000 (see below). The DNeasy plant mini kit (Qiagen) was used to extract total genomic DNA (from 0.1 g of leaf tissue/sample) in order to provide the high purity required for thorough and even restriction of each sample. DNA concentration was estimated by comparison with lambda *Hind* III DNA standard on 1.5% agarose gels. Samples were diluted in de-ionised water to 20 ng/µl.

Initially, test cuts were performed with each restriction enzyme (*Eco*RI and *Mse*I) separately to ensure each achieved complete digestion across all samples, assessed by electrophoresis on 1.5% agarose gels with ethidium bromide staining. The restriction of genomic DNA and ligation of adaptors were performed in a single reaction. In 96 well polycarbonate plates 250 ng of genomic DNA was added to the restriction-ligation mixture and incubated at 37°C overnight. The reagents for the restriction-ligation reaction (30 µl final volume) were as follows: 3.9 U *Eco*RI (ie. 5 U/40 µl) (New

England Biolabs), 3.9 U *Mse*I (New England Biolabs), 0.6 U T4 Ligase (Promega), 1 x T4 Ligase buffer (Promega), 1.5 µg BSA, 0.05 M NaCl, 3 pM *Eco*RI adaptor, 30 pM *Mse*I adaptor.

The reagents for pre-selective and selective amplification were as described by Vos *et al.* (1995). The thermo-cycling conditions for the pre-selective amplification were adapted from Vos *et al.* (1995) to promote selectivity (Marques *et al.* 1998). The selective thermo-cycling conditions followed Vos *et al.* (1995). Primers with the optimum number of selective bases for pre-selective and selective amplification of eucalypt DNA (one and three respectively; Marques *et al.* 1998) were employed for AFLP reactions. Pre-selective primers were from Marques *et al.* (1998) (*Eco*RI +A, *Mse*I +C). Substantial variability in the number of bands produced by different primer combinations (*Eco*RI +3, *Mse*I +3) in the selective amplification warranted a preliminary screening of 40 primer combinations, including 22 from Marques *et al.* (1998), to allow the most informative to be employed for progeny screening (Table 3.1). The selective *Eco*RI (+3) primers were fluorescently labelled with TET for detection by the Gel Scan 2000.

AFLP samples from the selective amplification were combined with two volumes of formamide B-blue loading buffer (formamide with the addition of 0.1 mg/mL bromophenol blue sodium salt) and denatured for 2 minutes at 94°C, then rapidly cooled on ice. Subsequently, AFLP fragments were separated by electrophoresis on 15 cm 6% denaturing acrylamide gels, using a Gel-Scan 2000 real-time acrylamide gel system. Electrophoresis at 40°C and 1200 volts for 90 minutes allowed separation of bands ranging from approximately 50-450 bp in size. Bands were scored manually by comparison to a Promega (CXR) 60-400 base fluorescent ladder, utilising Gene Profiler™ software (Scanalytics). Only bands that were clear and unambiguous across the entire mapping population were scored. Subsequently, the 'Match bins' function of the Database manager within the Gene Profiler™ software was used to automatically generate presence/absence binary scores for each sample across all scored bands with

match tolerance set at 0.5%. AFLP fragments were given a three-point confidence rating denoting their quality and ease of scoring as described by Shepherd *et al.* (2003).

Table 3.1 The number of polymorphic AFLP amplification products generated over 21 different primer combination. *Eco* + 3 and *Mse* + 3 columns indicate selective nucleotides at the ends of the oligonucleotide primers in the selective PCR. All primer combinations except those indicated by * are from Marques *et al.* (1998).

Primer combination	*Eco* + 3	*Mse* + 3	Total number of polymorphic bands
1	ACA	CCA	6
2	ACA	CTC	5
3	ACA	CAG	18
4	ACA	CAC	11
5	ACA	CCC	5
6	ACA	CCT	13
7	AAA	CCG	7
8	AAA	CGA	28
9	AAC	CCC	12
10	AAC	CCT	14
11	AAG	CCG	13
12	AAG	CGG	12
13	AAG	CTG	20
14	ATG	CCA	11
15	ACC	CCA	12
16	AAG	CCT	12
17	AAC*	CGA*	7
18	ACA*	CGA*	10
19	ATC*	CGA*	10
21	AAC*	CAG*	6
22	ACT	CCA	11
		mean	11.6

Linkage analysis

Following scoring, all SSR and AFLP gels were checked visually for scoring errors. All loci were recoded by their segregation type, according to the cross pollinated coding scheme and separated into male and female datasets for analysis by Joinmap 3.0 (Van Ooijen and Voorrips 2001). The Joinmap 3.0 package allows the integration of data with various segregation types, and recombination estimates from a variety of sources into a single map (Van Ooijen and Voorrips 2001). Each data set included the parental genotypes at all loci (grandparental genotypes and linkage phase information were not

included), with Joinmap 3.0 automatically calculating the linkage phase of each marker when calculating pair-wise recombination frequencies.

All loci were tested for goodness-of-fit to expected Mendelian segregation ratios using Chi-square goodness-of-fit tests. AFLP loci were expected to follow segregation ratios of 3:1 (each parent heterozygous for the presence of the band) or 1:1 (band heterozygous in one parent only). The 3:1 segregating loci were given a classification code in the locus file to instruct the program to test for this segregation ratio, allowing for dominance. Most mapped SSR loci were fully informative, segregating in approximate ratios of 1:1:1:1 (either heterozygous in both parents with 3 alleles in total, or heterozygous in both parents with 4 alleles in total). The only mapped SSR loci that followed approximate segregation ratios of 1:1 (heterozygous in only one parent), were EMBRA 6, 78, 197 and EMCRC 2 from the male parent and EMBRA 208 from the female parent.

The Joinmap 3.0 programme initially groups loci that are significantly associated (linked) on the basis of a LOD score. Groupings are derived from the test for independence of loci in a contingency table, which are then converted to a LOD score (Van Ooijen and Voorrips 2001). There is a theoretical basis for selecting associations at different LOD values, based on genome size and the number of linkage tests (Keats *et al.* 1991), however, an intuitive approach is often used to select a LOD score at which the contents of most groups are relatively stable. Joinmap 3.0 is particularly suited to such an approach, by allowing the user to view the stability of groupings over a range of LOD thresholds simultaneously in tree format.

During the mapping procedure, Joinmap 3.0 uses a weighted least squares approach to sequentially build a map from pair-wise data. Within linkage groups, the maximum recombination frequency, and the minimum LOD score determines the stringency of marker ordering. Selecting a lower recombination frequency and/or higher LOD will increase the stringency of mapping by excluding more loci from the map calculations. Within these parameters, the chi-square jump threshold restricts the allowable increase

in 'goodness-of-fit' after the inclusion of a particular marker to an existing order. The lower the chi-square jump threshold the greater the stringency of marker ordering, with values of 3-5 recommended (Van Ooijen and Voorrips 2001). The genotype probability function indicates the number of genotypes with a low probability, conditional on the map and surrounding markers, thereby highlighting markers with possible genotyping errors.

Linkage groups in the comprehensive parental maps were assigned with a minimum LOD threshold of 4.0 for the male parent and 3.0 for the female parent. Within linkage groups, marker order was determined using Joinmap 3.0 default values of a minimum LOD threshold of 1.0, and a maximum recombination threshold of 0.4. In order to produce a map with robust marker order, a stringent maximum Chi-square goodness-of-fit jump threshold of 1.2 was used for all loci with the exception of EMBRA 6 which was placed at 1.8. The Kosambi mapping function was used to determine distance between markers.

Linkage analysis was conducted on male and female datasets independently, before merging parental maps into an integrated map. Parental maps were constructed in several stages. The analysis was initially conducted with the highest quality markers to establish a 'trusted order' including most SSR and AFLPs segregating 1:1. Subsequently, the data were reanalysed adding the markers of lower quality in three steps: the lowest ranking AFLP loci segregating 1:1, dominant AFLPs and all loci with segregation distortion with $\alpha \leq 0.05$. At each stage the order of markers was checked against the established order from the best markers and each parent against the other, the lower quality markers were accepted if they fitted the trusted order within the designated parameters (above). Where marker order changed following the inclusion of lower quality data, markers were dropped from the data set on the basis of their quality ranking (AFLPs), Chi-square goodness-of-fit jump contribution and genotype probability ranking. Once all markers fitting the designated parameters had been added and homology was established between the orders of shared markers in parental maps, an integrated map was produced based on shared markers. Markers were included in the

integrated map where locus order agreed with parental maps. The fixed order option was used to reduce chi-square jumps in the construction of the integrated map for groups 7 and 8 in order to achieve agreement with the trusted marker order from parental maps. Final marker orders were accepted when there was agreement between all maps within the designated parameters, giving a high degree of confidence in locus ordering. The numbering of linkage groups 1-11 (where present), in all maps followed Brondani *et al.* (1998, 2002 and unpublished data), in order to facilitate a comparison of linkage group homology.

Genotyping errors can lead to difficulty in locus ordering and inflated map lengths. In order to detect mis-typed markers, numerous linked loci that exceeded the threshold Chi-square goodness-of-fit contribution, or displayed excessive double crossovers (evidenced by low genotype probability), were checked (using the original gel image for SSR loci, or quality rating for AFLP markers) for the robustness of scoring and either excluded from mapping analysis, or added as accessories to indicate a lower confidence in their placement relative to framework markers. Accessory markers were located on the parental maps next to their nearest framework marker by their minimum recombination frequency and LOD score. Accessory markers were not included in the integrated linkage map, unless placed as a framework marker in one parent.

Results

Marker analysis

Twenty-one of the 40 AFLP primer combinations screened were employed for genotyping, including 17 from Marques *et al.* (1998) (Table 3.1). These combinations amplified 243 polymorphic bands in total, an average of 11.6 polymorphic bands per primer combination. One hundred and eighty four bands segregated in an approximate 1:1 ratio, comprising 87 from the female parent and 97 from the male parent. Fifty-nine bands segregated from both parents in an approximate 3:1 ratio.

Table 3.2 provides a summary of the results of screening the SSR loci. In total 165 of the 280 loci screened were able to be amplified in this *E. globulus* cross. The relative success of each set of loci reflected taxonomic distance from the species in which they were developed. Of the EMCRC loci (from *E. globulus*) 92% were amplifiable, compared with only 71% of CSIRO loci (from *E. globulus* and the closely related *E. nitens*), and only 55% of EMBRA loci, (developed from the more distantly related *E. grandis/E. urophylla)*. Of the 165 amplifiable SSR loci, 24 (15%) segregated exclusively from the male parent, 19 (12%) from the female parent only, 87 (53%) from both parents and were thus fully informative (Table 3.2), while 35 (21%) were not heterozygous in either parent.

Table 3.2. The amplification success and polymorphism of SSR loci by source.

Source	No. tested	No. amplified	Polymorphism			
			Neither	Male	Female	Both
EMBRA	233	129	31	19	14	65
EMCRC	13	12	1	3	3	5
CSIRO	34	24	3	2	2	17
Total	280	165	35	24	19	87

Linkage analysis

For the female parent (G1060; Figure 3.3a), twelve linkage groups (comprising one or more SSR or at least three AFLP loci) were defined by two-point linkage analysis at a minimum LOD score of 3.0. There were 106 framework markers, with 12 accessory markers added to make the comprehensive map, comprising 33 SSR and 85 AFLP loci (Figure 3a). Eleven markers were unlinked and 65 markers were excluded during map construction, because they did not fit the trusted order within the designated mapping parameters. Total map length was 698.3 cM. Linkage groups ranged in size from 2 cM (group 13), to 97.2 cM (group 1) with the average size being 58.2 cM. The 118 markers mapped to 117 positions separated by 105 intervals with an average size of 6.6 cM. The maximum interval between markers was 33.4 cM (group 5).

For the male parent (G1026; Figure 3.3b), fourteen linkage groups were defined at a minimum LOD score of 4.0 and comprised three or more loci. Two groups (group 2'; Figure 3.3b) correspond to linkage group 2 in the female parent, identified by fully informative SSR that were linked in the female parent and integrated map. These two groups united at LOD 3.0. However, the locus order of the resultant group departed from the established trusted order and was thus not accepted. The 14 linkage groups included 121 framework and 9 accessory markers, comprising 36 SSR and 94 AFLP loci (Figure 3.3b). Twenty markers were unlinked and 57 markers were excluded during map construction, because they did not fit the trusted order within the designated mapping parameters. Total map length was 775.5 cM, with linkage groups ranging in size from 113.7 cM (group 3) to 6.9 cM (group 9). The average size of linkage groups was 55.4 cM. The 130 markers mapped to 127 positions separated by 113 intervals with an average size of 6.9 cM. The maximum interval between markers was 28.6 cM (group 4).

SSR and AFLP loci segregating from both parents (Table 3.3) allowed the construction of an integrated map (Figure 3.3c). The integrated map featured 165 markers in 10 groups with a total map length of 701.9 cM. Thirty-seven of the markers in the integrated map had been placed in both parental maps, comprising 28 SSR and 9 AFLP,

allowing the identification of homology between parental maps for ten linkage groups. A small 11th group (9') was identified in the male parent. The 165 markers comprised 33 SSR and 132 AFLP loci. Markers that segregated bi-parentally were placed by the average of their recombination ratio. Despite the greater number of mapped markers segregating from the male parent, no significant difference was observed in recombination frequencies among consecutive pairs of fully informative loci at a threshold $\alpha = 0.05$.

Table 3.3 Number of loci mapped by segregation, marker type and source (for SSR). SSR markers that were fully informative could potentially be mapped in both parental maps as indicated by 'both' in the segregation column. Although, in some cases, a fully informative marker could only be mapped to one parent within the designated mapping parameters, causing some differences between the number of fully informative markers in parental and integrated maps.

		SSR			AFLP	Total
Map	Segregation	EMBRA	CSIRO	EMCRC		
Male	Male only	3	-	1	79	83
	Both	23	5	4	15	47
	Total	26	5	5	94	130
Female	Female only	1	-	-	68	69
	Both	22	6	4	17	49
	Total	23	6	4	85	118
Integrated	Male only	3	-	-	63	66
	Female only	1	-	-	55	56
	Both	21	4	4	14	43
	Total	25	4	4	132	165

Following pages:

Figures 3.3 (a-c) Linkage maps of *Eucalyptus globulus* using SSR and AFLP markers. Linkage groups are numbered following Brondani *et al.* (1998 and 2002). The distance between markers is indicated in Kosambi centimorgans. SSR markers are preceded by EMB (EMBRA), En or Eg (CSIRO, isolated from *E. nitens*, or *E. globulus* respectively) or CRC (EMCRC) to indicate their source. AFLP markers are labelled p#b# with numbers indicating the primer combination (p#) and band number (b#) respectively. Accessory markers are underlined, loci with distorted segregation ratios are indicated by asterisk; * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$,**** $P \leq 0.0001$

Figure 3.3(a) Linkage map of the female parent.
Figure 3.3(b) Linkage map of the male parent. Group 2' corresponds to group 2 in the female and integrated maps.
Figure 3.3(c) Integrated linkage map. The integrated linkage map includes only framework markers that fit the order established from parental maps. Linkage group 9' was identified in the male parent only.

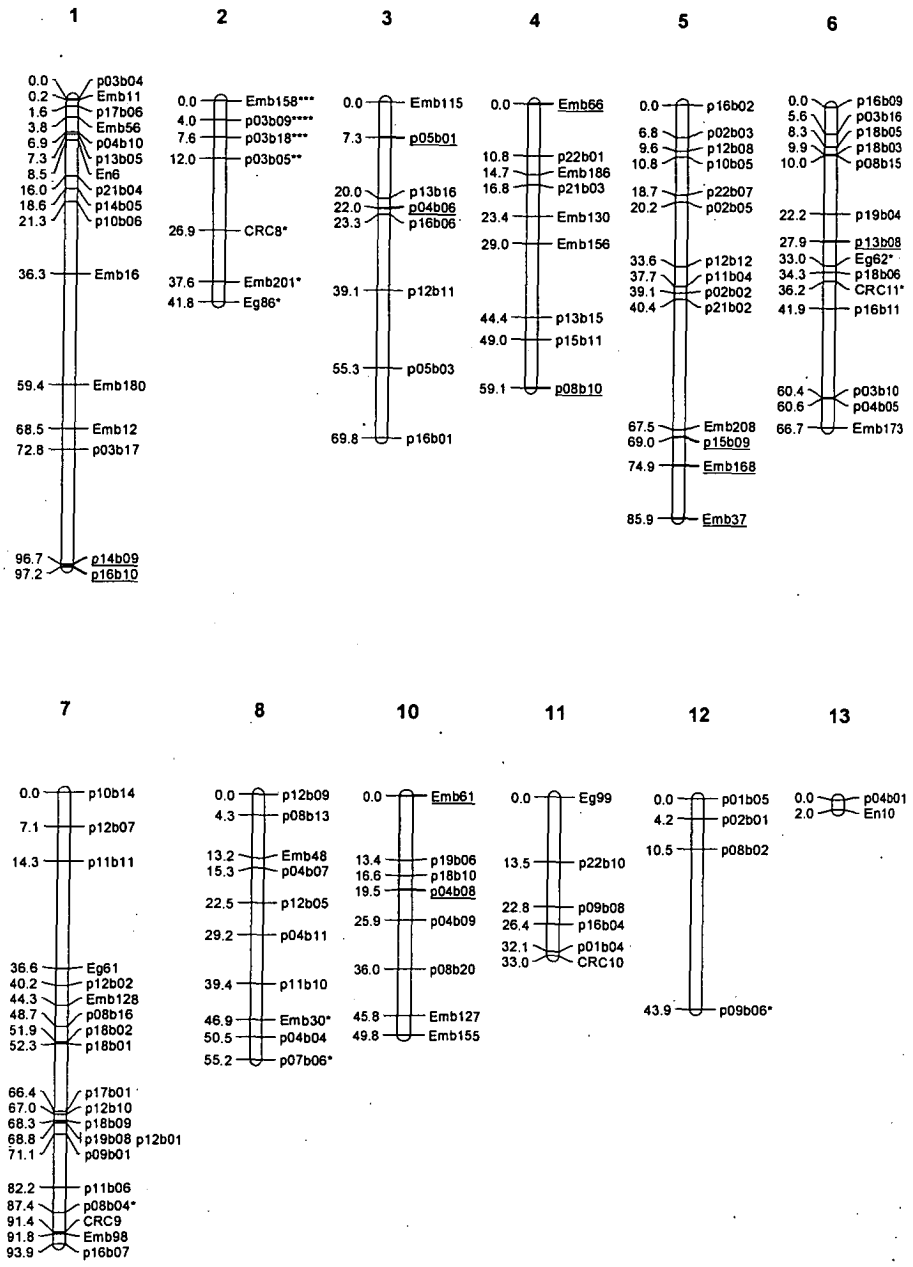

1
0.0 p03b04
0.2 Emb11
1.6 p17b06
3.8 Emb56
6.9 p04b10
7.3 p13b05
8.5 En6
16.0 p21b04
18.6 p14b05
21.3 p10b06
36.3 Emb16
59.4 Emb180
68.5 Emb12
72.8 p03b17
96.7 p14b09
97.2 p16b10
2
0.0 Emb158***
4.0 p03b09****
7.6 p03b18***
12.0 p03b05**
26.9 CRC8*
37.6 Emb201*
41.8 Eg86*
3
0.0 Emb115
7.3 p05b01
20.0 p13b16
22.0 p04b06
23.3 p16b06
39.1 p12b11
55.3 p05b03
69.8 p16b01
4
0.0 Emb66
10.8 p22b01
14.7 Emb186
16.8 p21b03
23.4 Emb130
29.0 Emb156
44.4 p13b15
49.0 p15b11
59.1 p08b10
5
0.0 p16b02
6.8 p02b03
9.6 p12b08
10.8 p10b05
18.7 p22b07
20.2 p02b05
33.6 p12b12
37.7 p11b04
39.1 p02b02
40.4 p21b02
67.5 Emb208
69.0 p15b09
74.9 Emb168
85.9 Emb37
6
0.0 p16b09
5.6 p03b16
8.3 p18b05
9.9 p18b03
10.0 p08b15
22.2 p19b04
27.9 p13b08
33.0 Eg62*
34.3 p18b06
36.2 CRC11*
41.9 p16b11
60.4 p03b10
60.6 p04b05
66.7 Emb173
7
0.0 p10b14
7.1 p12b07
14.3 p11b11
36.6 Eg61
40.2 p12b02
44.3 Emb128
48.7 p08b16
51.9 p18b02
52.3 p18b01
66.4 p17b01
67.0 p12b10
68.3 p18b09
68.8 p19b08 p12b01
71.1 p09b01
82.2 p11b06
87.4 p08b04*
91.4 CRC9
91.8 Emb98
93.9 p16b07
8
0.0 p12b09
4.3 p08b13
13.2 Emb48
15.3 p04b07
22.5 p12b05
29.2 p04b11
39.4 p11b10
46.9 Emb30*
50.5 p04b04
55.2 p07b06*
10
0.0 Emb61
13.4 p19b06
16.6 p18b10
19.5 p04b08
25.9 p04b09
36.0 p08b20
45.8 Emb127
49.8 Emb155
11
0.0 Eg99
13.5 p22b10
22.8 p09b08
26.4 p16b04
32.1 p01b04
33.0 CRC10
12
0.0 p01b05
4.2 p02b01
10.5 p08b02
43.9 p09b06*
13
0.0 p04b01
2.0 En10

Figure 3.3(a)

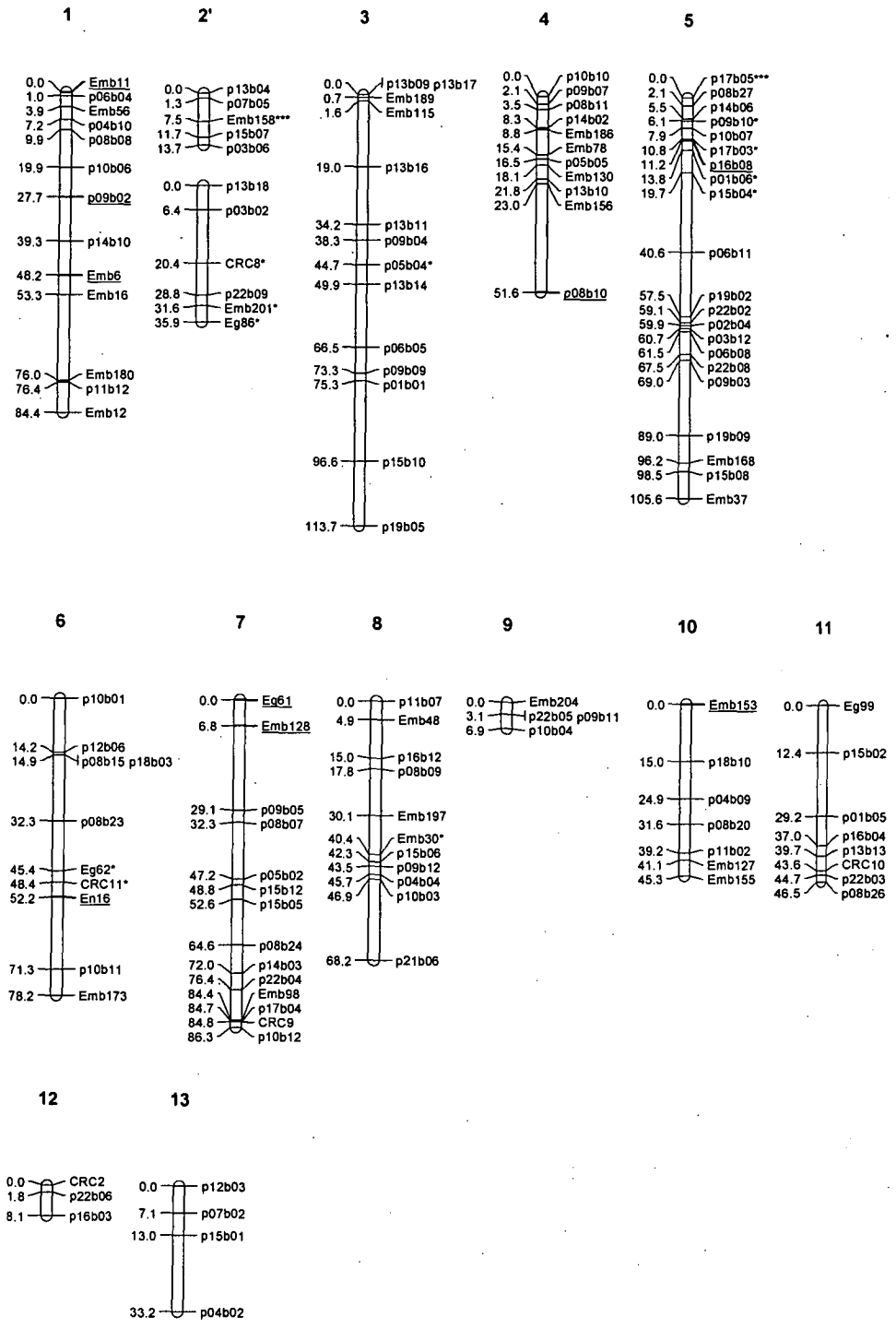

1
0.0 Emb11
1.0 p06b04
3.9 Emb56
7.2 p04b10
9.9 p08b08
19.9 p10b06
27.7 p09b02
39.3 p14b10
48.2 Emb6
53.3 Emb16
76.0 Emb180
76.4 p11b12
84.4 Emb12
2'
0.0 p13b04
1.3 p07b05
7.5 Emb158***
11.7 p15b07
13.7 p03b06
0.0 p13b18
6.4 p03b02
20.4 CRC8*
28.8 p22b09
31.6 Emb201*
35.9 Eg86*
3
0.0 p13b09 p13b17
0.7 Emb189
1.6 Emb115
19.0 p13b16
34.2 p13b11
38.3 p09b04
44.7 p05b04*
49.9 p13b14
66.5 p06b05
73.3 p09b09
75.3 p01b01
96.6 p15b10
113.7 p19b05
4
0.0 p10b10
2.1 p09b07
3.5 p08b11
8.3 p14b02
8.8 Emb186
15.4 Emb78
16.5 p05b05
18.1 Emb130
21.8 p13b10
23.0 Emb156
51.6 p08b10
5
0.0 p17b05***
2.1 p08b27
5.5 p14b06
6.1 p09b10*
7.9 p10b07
10.8 p17b03*
11.2 p16b08
13.8 p01b06*
19.7 p15b04*
40.6 p06b11
57.5 p19b02
59.1 p22b02
59.9 p02b04
60.7 p03b12
61.5 p06b08
67.5 p22b08
69.0 p09b03
89.0 p19b09
96.2 Emb168
98.5 p15b08
105.6 Emb37
6
0.0 p10b01
14.2 p12b06
14.9 p08b15 p18b03
32.3 p08b23
45.4 Eg62*
48.4 CRC11*
52.2 En16
71.3 p10b11
78.2 Emb173
7
0.0 Eg61
6.8 Emb128
29.1 p09b05
32.3 p08b07
47.2 p05b02
48.8 p15b12
52.6 p15b05
64.6 p08b24
72.0 p14b03
76.4 p22b04
84.4 Emb98
84.7 p17b04
84.8 CRC9
86.3 p10b12
8
0.0 p11b07
4.9 Emb48
15.0 p16b12
17.8 p08b09
30.1 Emb197
40.4 Emb30*
42.3 p15b06
43.5 p09b12
45.7 p04b04
46.9 p10b03
68.2 p21b06
9
0.0 Emb204
3.1 p22b05 p09b11
6.9 p10b04
10
0.0 Emb153
15.0 p18b10
24.9 p04b09
31.6 p08b20
39.2 p11b02
41.1 Emb127
45.3 Emb155
11
0.0 Eg99
12.4 p15b02
29.2 p01b05
37.0 p16b04
39.7 p13b13
43.6 CRC10
44.7 p22b03
46.5 p08b26
12
0.0 CRC2
1.8 p22b06
8.1 p16b03
13
0.0 p12b03
7.1 p07b02
13.0 p15b01
33.2 p04b02

Figure 3.3 (b)

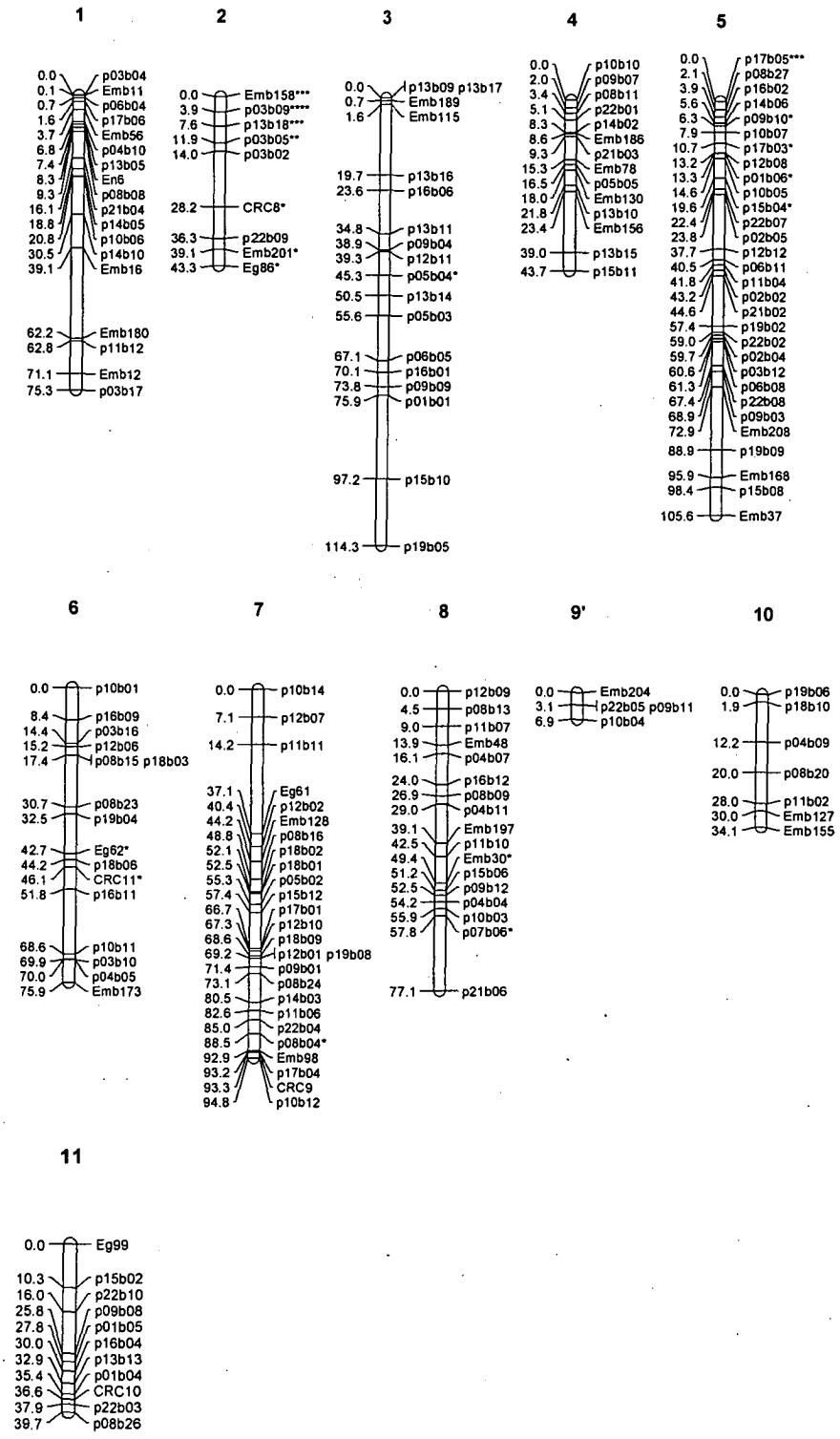

1
0.0 p03b04
0.1 Emb11
0.7 p06b04
1.6 p17b06
3.7 Emb56
6.8 p04b10
7.4 p13b05
8.3 En6
9.3 p08b08
16.1 p21b04
18.8 p14b05
20.8 p10b06
30.5 p14b10
39.1 Emb16
62.2 Emb180
62.8 p11b12
71.1 Emb12
75.3 p03b17
2
0.0 Emb158***
3.9 p03b09****
7.6 p13b18***
11.9 p03b05**
14.0 p03b02
28.2 CRC8*
36.3 p22b09
39.1 Emb201*
43.3 Eg86*
3
0.0 p13b09 p13b17
0.7 Emb189
1.6 Emb115
19.7 p13b16
23.6 p16b06
34.8 p13b11
38.9 p09b04
39.3 p12b11
45.3 p05b04*
50.5 p13b14
55.6 p05b03
67.1 p06b05
70.1 p16b01
73.8 p09b09
75.9 p01b01
97.2 p15b10
114.3 p19b05
4
0.0 p10b10
2.0 p09b07
3.4 p08b11
5.1 p22b01
8.3 p14b02
8.6 Emb186
9.3 p21b03
15.3 Emb78
16.5 p05b05
18.0 Emb130
21.8 p13b10
23.4 Emb156
39.0 p13b15
43.7 p15b11
5
0.0 p17b05***
2.1 p08b27
3.9 p16b02
5.6 p14b06
6.3 p09b10*
7.9 p10b07
10.7 p17b03*
13.2 p12b08
13.3 p01b06*
14.6 p10b05
19.6 p15b04*
22.4 p22b07
23.8 p02b05
37.7 p12b12
40.5 p06b11
41.8 p11b04
43.2 p02b02
44.6 p21b02
57.4 p19b02
59.0 p22b02
59.7 p02b04
60.6 p03b12
61.3 p06b08
67.4 p22b08
68.9 p09b03
72.9 Emb208
88.9 p19b09
95.9 Emb168
98.4 p15b08
105.6 Emb37
6
0.0 p10b01
8.4 p16b09
14.4 p03b16
15.2 p12b06
17.4 p08b15 p18b03
30.7 p08b23
32.5 p19b04
42.7 Eg62*
44.2 p18b06
46.1 CRC11*
51.8 p16b11
68.6 p10b11
69.9 p03b10
70.0 p04b05
75.9 Emb173
7
0.0 p10b14
7.1 p12b07
14.2 p11b11
37.1 Eg61
40.4 p12b02
44.2 Emb128
48.8 p08b16
52.1 p18b02
52.5 p18b01
55.3 p05b02
57.4 p15b12
66.7 p17b01
67.3 p12b10
68.6 p18b09
69.2 p12b01 p19b08
71.4 p09b01
73.1 p08b24
80.5 p14b03
82.6 p11b06
85.0 p22b04
88.5 p08b04*
92.9 Emb98
93.2 p17b04
93.3 CRC9
94.8 p10b12
8
0.0 p12b09
4.5 p08b13
9.0 p11b07
13.9 Emb48
16.1 p04b07
24.0 p16b12
26.9 p08b09
29.0 p04b11
39.1 Emb197
42.5 p11b10
49.4 Emb30*
51.2 p15b06
52.5 p09b12
54.2 p04b04
55.9 p10b03
57.8 p07b06*
77.1 p21b06
9'
0.0 Emb204
3.1 p22b05 p09b11
6.9 p10b04
10
0.0 p19b06
1.9 p18b10
12.2 p04b09
20.0 p08b20
28.0 p11b02
30.0 Emb127
34.1 Emb155
11
0.0 Eg99
10.3 p15b02
16.0 p22b10
25.8 p09b08
27.8 p01b05
30.0 p16b04
32.9 p13b13
35.4 p01b04
36.6 CRC10
37.9 p22b03
39.7 p08b26

Figure 3.3(c)

Segregation distortion

The proportion of markers exhibiting significant distortion from expected Mendelian segregation ratios were greater than would be expected by chance alone, for both parental maps. Loci that departed from expected segregation ratios were included in maps on the basis of fitting the 'trusted order' without substantial re-ordering, and conformation to the designated mapping parameters (see Materials and Methods). The proportion of mapped markers with distorted segregation ($\alpha \leq 0.05$) was greater for the female 13 (11%) than the male 12 (9%) parental maps, however this difference was not significant ($X^2_{1d} = 0.2178$, $P \leq 0.05$).

More significant distortion also occurred on the female map, where one marker (p03b09) was distorted at $\alpha \leq 0.0001$, one marker at $\alpha \leq 0.0005$ (p03b18), one marker at 0.005% (EMBRA 158) and one marker at $\alpha \leq 0.01$ (p03b05). The remaining nine markers were distorted at $\alpha \leq 0.05$. In contrast, only one marker was distorted at $\alpha \leq 0.005\%$ in the male map (EMBRA 158), with the remaining 11 markers distorted at $\alpha \leq 0.05$ (Figures 3a and 3b).

Distorted markers ($\alpha \leq 0.05$) mapped to 5 linkage groups in the female parent and 5 linkage groups in the male parent. The distribution of distorted markers was concentrated on two linkage groups (male 5 and female 2), other distorted markers occurred singly or scattered within linkage groups (Figures 3.3a and 3.3b). The most severe distortion occurred on linkage group 2, where all markers from the female parent exhibited segregation distortion (Figure 3.3a). A gradient was evident in the distortion ratios from a maximum $\alpha = 0.0001$ (0.01%), in the AFLP marker p03b09 near the chromosome end, to lesser distortion in surrounding markers in both directions. Markers inherited solely from the male parent in this group were not distorted. Shared SSR markers in the equivalent group from the male parent exhibit distortion, however this was due to the deficiency of a maternal allele for all but one marker (EMCRC 8) in which there was no deficiency of alleles from either parent. The distortion can therefore be considered specific to the female parent. In contrast, in linkage group 5, segregation

distortion was confined to markers inherited from the male parent (Figure 3.3b), with four markers near the groups extremity distorted at $\alpha \leq 0.05$.

Discussion

Marker polymorphism

Fifty three percent of amplifiable SSR loci screened in this study were fully informative (Table 3.2), compared with other reports in *E. globulus* of 60% in an F_1 inter-provenance cross (Thamarus *et al.* 2002) and 34% in an F_1 intra-provenance cross (Bundock *et al.* 2000). The intra-provenance nature of the mapping cross employed by Bundock *et al.* (2000) may explain the lower proportion of polymorphism, compared to the inter-provenance cross used herein.

The average of 12 polymorphic bands per AFLP primer combination in this study is low compared with other findings in eucalypts of 25 in an F_1 interspecific cross of *E. globulus* and *E. tereticornis* (Marques *et al.* 1998) and 35 in a pseudo-backcross of *E. globulus* and *E. grandis* (Myburg *et al.* 2003). Two factors may account for these differences. Firstly, the interspecific crosses used by Marques *et al.* (1998) and Myburg *et al.* (2003) are likely to produce a greater level of informative polymorphism than an intraspecific cross. Secondly, differences in the detection system employed for AFLP fragments may further account for differences in polymorphism between the studies. The sequencing gels employed for AFLP detection by Marques *et al.* (1998) and Myburg *et al.* (2003), may provide increased resolution, allowing the detection of a greater number of bands per gel than the Gelscan 2000 employed in this study.

Map construction

The total map length of 701.9 cM (integrated produced using 165 markers) in this study is low in comparison to other mapping efforts in *Eucalyptus* (see Thamarus *et al.* 2002). At the upper limit, Byrne *et al.* (1995) reported an integrated map length of 1462 cM in *E. nitens* using 335 predominantly RFLP and RAPD markers, while at the lower limit Marques *et al.* (1998) reported a total map length of 919 cM (female) and 967 cM (male) in maps comprising 268 and 200 AFLP markers, respectively. An important factor contributing to the low map length herein is that the pair-wise mapping algorithm employed by Joinmap software consistently produces lower map lengths compared to

multi-locus likelihood based packages (Butcher *et al.* 2002) such as Outmap (Williams and Whitaker 2001) and Mapmaker (Lander *et al.* 1987). Incomplete map coverage in some linkage groups, due to the relatively low number of mapped markers in this study, may also contribute to the observed differences in map length (see evidence below). Furthermore, the presence of genotyping errors can overestimate the proportion of double crossovers, severely inflating map lengths (Hackett and Broadfoot 2003). Therefore, the rigorous data checking and exclusion of numerous scored markers with an apparent excess of double crossovers may have had contributed to keeping map length to a minimum in this study.

There were more linkage groups than the haploid number of chromosomes in *Eucalyptus* (11), identified from cytological studies (reviewed in Potts and Wiltshire 1997) for both the male (14 groups) and female (12 groups) maps. The finding of a greater number of linkage groups than chromosomes is not unusual in linkage studies (Kesseli *et al.* 1994; Wu *et al.* 2001), including eucalypts (Byrne *et al.* 1995; Marques *et al.* 1998). Such a discrepancy between the number of chromosomes and linkage groups can occur even with large numbers of mapped loci because the number of informative meiotic events required to reliably place markers in linkage groups is usually many times larger than the number of loci (Edwards 1991). As a result, unlinked markers and small linkage groups that could not be united with larger groups, were no doubt the cause of the greater number of linkage groups than chromosomes in parental maps within this study.

Linkage group homology

Comparative mapping studies rely on a set of common transferable markers that segregate in the species or pedigree of interest (Ahuja 2001). Although various linkage maps have been constructed in eucalypts, the lack of transferability of markers commonly employed, such as RAPDs and AFLPs, means that linkage information and associated QTL data have been largely confined to the pedigrees used for map construction (Grattapaglia 2000). This deficiency of marker information has been a great hindrance to comparative mapping that would form the basis for the construction of a

species wide consensus map, and more powerful QTL analysis in *Eucalyptus* (Brondani *et al.* 2002).

The ongoing development of numerous EMBRA SSR markers from *Eucalypus grandis* and *E. urophylla* has begun to address the problem of marker transferability (Brondani *et al.* 1998, 2002). For example, within the section *Latoangulatae* (Brooker 2000), 46 fully informative (EMBRA) SSR markers are common to linkage maps of both *Eucalyptus grandis* and *E. urophylla* (Brondani *et al.* 2002). Between sections, comparisons have largely been limited to markers with less information content or fewer SSR loci. For example, linkage information from *Eucalyptus grandis* and *E. urophylla* (section *Latoangulatae*; Brondani *et al.* 1998; Gion *et al.* 2000) can be compared with *E. globulus* (section *Maidenaria*) based on 3 EMBRA loci and 5 lignin genes (Thamarus *et al.* 2002) or 7 fully informative EMBRA loci (Bundock *et al.* 2000). Similarly, linkage information from *Eucalyptus camaldulensis* (section *Exsertaria*; Agrama *et al.* 2002) can be compared with *E. grandis* and *E. urophylla* on the basis of 13 shared SSR markers (Brondani *et al.* 1998) and with *E. globulus* based on two (Thamarus *et al.* 2002) and five (Bundock *et al.* 2000) shared SSRs. The only study to date allowing comparison of synteny between sections in numerous SSR markers is that of Marques *et al.* (2002), which considered synteny of SSR loci and QTLs across *Eucalyptus grandis*, *E. urophylla*, *E. tereticornis* (section *Exsertaria*) and *E. globulus*. Marques *et al.* (2002) reported only the homologies supported by at least 3 SSR loci in two different species, comprising 22 SSR across 8 linkage groups in *E. globulus*.

In an attempt to allow comparison between this and previous studies, 37 of the 41 SSR mapped herein were in common with other linkage studies in *Eucalyptus*. Importantly, this study included SSR loci developed from a variety of sources, allowing a wider comparability to other studies than was possible based solely on EMBRA markers (Marques *et al.* 2002). This included 27 EMBRA markers in common with Brondani *et al.* (1998, 2002 and unpublished data), three EMCRC and four EMBRA markers in common with Bundock *et al.* (2000) and seven CSIRO markers in common with Thamarus *et al.* (2002).

The presence of 27 fully informative SSRs in common allowed all linkage groups identified by Brondani *et al.* (1998, 2002, and unpublished data) to be recognised in this pedigree (Table 3.4). Five linkage groups had three or more fully informative markers per group, allowing comparison of marker order. In agreement with past findings (Thamarus *et al.* 2002; Marques *et al.* 2002; Bundock *et al.* 2000), all linkage groupings of the EMBRA markers identified in a two way pseudo-testcross of *E. grandis/urophylla* (Brondani *et al.* 1998, 2002 and unpublished data), were preserved in the *E. globulus* maps in this study. However, a few minor inconsistencies in map distance and the order of markers were apparent. Permutations of marker order were evident in three of these five linkage groups. These involved mainly closely linked markers- EMBRA 208 and EMBRA 168 in linkage group 5 and EMBRA 127 and EMBRA 155, located at the end of linkage group 10. A more substantial difference occurs in linkage group 8, where EMBRA 197 and EMBRA 48 differ in order relative to EMBRA 30 between maps of *E. grandis/urophylla* and this study. Similarly, on linkage group 4, EMBRA 66 maps to the same position as EMBRA 130 in *E. grandis/urophylla* but is separated by 23 cM in this study and is in a different position relative to EMBRA 186.

SSR markers shared between this study and maps produced by Bundock *et al.* (2000), from an intra-provenance cross of *E. globulus*, and Thamarus *et al.* (2002), using an inter-provenance F_1 cross of *E. globulus*, both with parents originating from the same localities as those in this study (Table 3.4) allows tentative identification of homologous linkage groups. Homology between 5 linkage groups in this study and 7 linkage groups of Bundock *et al.* (2000) are proposed on the basis of shared (EMBRA, and EMCRC) SSR. Only two homologous groups (corresponding to 3 groups from Bundock *et al.* 2000) could be identified directly from shared markers, while the remaining homologies could be inferred by comparing EMBRA loci shared between each map and *E. grandis/urophylla*. This inference is justified as both this study and Bundock *et al.* (2000), found complete agreement between linkage groupings with *E. grandis/urophylla* (Brondani *et al.* 1998, 2002 and unpublished data). However, in the case of Bundock *et al.* (2000) there where more linkage groups than chromosomes. As a result some linkage

groups identified by Brondani *et al.* (1998) were equivalent to 2 linkage groups in the maps of Bundock *et al.* (2000). Additionally, there is a slight difference between the linkage groupings herein and those of Bundock *et al.* (2000). A single marker (EMCRC 9) maps to linkage group 7 in this study, but in the study of Bundock *et al.* (2000) it maps to a different linkage group (the equivalent to linkage group 9 in the present study).

Table 3.4 Linkage group homology. Linkage maps produced by Bundock *et al.* (2000) had more linkage groups than the haploid number of chromosomes in *Eucalyptus* (11), hence linkage groups 9 and 1 in this study were equivalent to two linkage groups from Bundock *et al.* (2000).

This study & Brondani *et al.* 1998, 2002	Bundock *et al.* (2000)	Thamarus *et al.* (2002)	Candidate genes reported by Thamarus *et al.* (2002)
1	4 and 5	7	COMT
2		4	EAP1, PAL
3		11	MsaS2
4			
5	3	9	AGE2
6		6	EXS1
7		2	CCoAOMT, ECA1
8	6		
9	1 and 2		
10			
11	13	3	4CL, AGE1, ELF1

The differences in marker order between this map and those produced by Brondani *et al.* (1998, 2002 and unpublished data) and Bundock *et al.* (2000) may represent actual differences caused by mechanisms such as chromosomal rearrangements between the mapping populations. Some evidence in support of chromosomal rearrangements between eucalypt species is provided by the fact that Agrama *et al.* (2002) also found a difference in marker order (in *E. camaldulensis*) compared with *E. grandis* and *E. urophylla* (Brondani *et al.* 1998) in the equivalent linkage group (8) to this study. Alternatively, amplification of non-target sites due to multiple copies of the same SSR loci, may have occurred as Erpelding *et al.* (1996) found when transferring SSR between wheat and barley. On the other hand, the discrepancies may be artefacts of mapping imprecision. Because very few apparent recombinants can change map order, accurate ordering of closely linked markers is often problematic, requiring very large mapping

populations (Remington *et al.* 1999). Indeed practical limitations on the number of individuals used in mapping populations of forest trees may account for the alternations in marker order which have occurred between parental maps produced in various tree species (Maliepaard *et al.* 1998; Devey *et al.* 1999), including *Eucalyptus* (Brondani *et al.* 2002; Marques *et al.* 1998, 2002). At present, the number of comparable linkage maps produced in *Eucalyptus* species is insufficient to conclusively confirm or refute the occurrence of chromosomal rearrangements between species. Hence, further mapping using larger populations and more markers is required in *E. globulus* and other *Symphyomyrtus* species to clarify the nature and extent of any differences in marker order at both the interspecific and intraspecific levels.

When compared with the map of Thamarus *et al.* (2002), all linkage groups of shared (CSIRO and EMBRA) SSR markers were conserved. Homology between 7 linkage groups of Thamarus *et al.* (2002) and this study are inferred on the basis of 1 or 2 SSR in common per linkage group. However, there are some inconsistencies in the relative position of markers within linkage groups between maps. Specifically, the CSIRO marker Eg86 maps to the end of linkage group 2 herein, but a third of the way along the equivalent group in Thamarus *et al.* (2002). Similarly, the CSIRO marker Eg99 maps to the end of the linkage group 11 herein but close to the middle of the equivalent linkage group in Thamarus *et al.* (2002). Similar discrepancies in the exact map location of specific markers have been noted previously in *Eucalyptus* (Marques *et al.* 1998; Myburg *et al.* 2003; Thamarus *et al.* 2002) and in the absence of evidence for chromosomal rearrangement are often attributable to sampling error (Myburg *et al.* 2003) or differences in map coverage between studies. Incomplete map coverage of some linkage groups in the present study is the most likely cause of the above discrepancies.

Segregation distortion

The proportion of markers in this study displaying distortion from expected Mendelian segregation ratios is greater than would be expected by chance alone. For example, 12 markers from the male parent and 13 markers from the female parent exhibited segregation distortion (at $\alpha \leq 0.05$), clearly exceeding expectations for the male (6.5/130) and female (5.9/118) parents. The concentration of many of the distorted markers in two groups (male 5 and female 2) suggests a biological cause, rather than other potential factors, such as genotyping errors, that might produce an apparent distortion of segregation ratios (Kearsey and Pooni 1996). The female group 2 is notable, in that all markers from this group are distorted and a gradient in the level of distortion from a maximum ($\alpha \leq 0.0001$) in p03b09 occurs in this linkage group. Such departures from expected Mendelian segregation ratios have been widely documented in plants and animals (e.g. Gillet and Gregorius 1992; Bradshaw and Stettler 1994), including *Eucalyptus* (Vaillancourt *et al.* 1995; Bundock *et al.* 2000; Myburg *et al.* 2003). Numerous mechanisms have been hypothesised to explain such distorted segregation ratios, operating at various stages of development (reviewed by Gillet and Gregorius 1992). These include prezygotic mechanisms such as gametophytic incompatibility systems (e.g. Gebhardt *et al.* 1991), meiotic drive (e.g. Buckler *et al.* 1999) and haploid expressed deleterious alleles and post-zygotic mechanisms such as the expression of genetic load (e.g. Gion *et al.* 2000; Bradshaw and Stettler 1994) and hybrid incompatibility systems (e.g. Rieseberg *et al.* 1995).

Gametophytic incompatibility systems reflect variable fertilisation success due to compatibility relations between gametic types and in most cases operate to exclude fertilisation by the male (Gillet and Gregorius 1992). Gametophytic incompatibility is thus unlikely to account for the observed segregation distortion in the female linkage group 2.

Ideally, a survey of allele frequencies before and after fertilisation would allow discrimination between segregation distortion due to pre-zygotic mechanisms such as meiotic drive and haploid expressed deleterious alleles and post-zygotic mechanisms

such as genetic load (Gillet and Gregorius 1992). Lacking such information, pre-zygotic mechanisms remain a possible cause of segregation distortion. However, the death of ~25% of the planted F_2 generation before they were sampled for this study suggests that some post-zygotic selection mechanism may well play a substantial role in distorting expected segregation ratios, as opposed to meiotic drive, haploid deleterious alleles or gametophytic incompatibility which would occur at the pre-zygotic stage.

The expression of genetic load involves the effects of deleterious or lethal alleles distorting segregation ratios of linked markers. Eucalypts are highly heterozygous and are known to carry many deleterious alleles in the heterozygous condition (Potts and Wiltshire 1997; Patterson *et al.* 2000). Such genes could possibly account for the observed segregation distortion from the female parent (group 2) in this study. Under this hypothesis, the distortion of the entire female linkage group 2 could be explained by one or several genes with a strong effect on viability segregating close to the markers where the distortion peaks, with diminishing effects toward the other extremity of the linkage group due to the effects of recombination. However, the expression of genetic load involves inheriting deleterious or lethal alleles from both parents. Consequently, this mechanism is unlikely to account for the observed distortion due to the lack of similar levels of distortion in equivalent linkage groups from each parent unless, for example, the deleterious effect of the female allele is only expressed in a specific male genetic background.

Another possible post-zygotic mechanism contributing to segregation distortion is hybrid incompatibility resulting from the broad inter-provenance origin of the grandparents in the mapping pedigree. Trees from these locations are differentiated by nuclear SSR (Potts *et al.* 2004) and as previously mentioned, are also classified into different geographical races on the basis of morphology (Dutkowski and Potts 1999). High levels of inviability have been noted in various interspecific hybrid combinations within *Eucalyptus* (reviewed by Potts and Dungey 2004). This process may also be applicable to broad intraspecific crosses such as the mapping pedigree in this study. For example, in an F_1 (intraspecific) inter-provenance cross between *E. globulus* from King

Island and Taranna, a reduced viability has been noted relative to an intra-provenance cross from King Island. However, the Taranna intra-provenance cross also performed poorly (Volker 2002). Although most studies have demonstrated hybrid incompatibility in the F_1 generation, barriers to hybridisation may also extend to further generations (outbreeding depression) as noted by Potts *et al.* (2000) where a reduced growth and survival of *E. globulus* by *E. nitens* affected F_1 and F_2 generations. Two general mechanisms are postulated to contribute to hybrid incompatibility, chromosomal rearrangements and genetic incompatibility between differentiated taxa (Rieseberg *et al.* 1995). Genetic incompatibility may occur between differentiated populations of the same species due to processes such as the disruption of coadaptive gene complexes and the introduction of adaptive, or maladaptive genes (Potts *et al.* 2003). As previously discussed, there is as yet no conclusive evidence for gross chromosomal rearrangements in *Eucalyptus* at the intra or interspecific level. Thus both general mechanisms remain possible causes of the observed segregation distortion in female linkage group 2.

Conclusion

The inclusion of numerous fully informative SSR, in conjunction with AFLP markers, has allowed the construction of parental and consensus maps in this pedigree of *E. globulus*. Mapping SSR markers derived from a variety of sources has demonstrated almost complete homology in linkage groupings across a range of *Symphyomyrtus* species, encompassing multiple sections, although some minor differences in the distance and order of markers within linkage groups were apparent. Further mapping within *E. globulus* using an expanded set of SSR markers will clarify the nature and extent of any differences between *Symphyomyrtus* species, paving the way for further comparative mapping and ultimately consensus maps of the subgenus. The low success in transfer of SSR loci from *E. grandis/urophylla*, from which most SSR in *Eucalyptus* have been developed, suggests that more SSRs will be required in order to produce high coverage maps in *E. globulus* based exclusively on these markers. Segregation distortion of some markers has been demonstrated, concentrated in two gender specific groups. In

the case of the severe distortion in the female group 2, post zygotic mechanisms are a likely cause.

Appendix 1 Reaction conditions and segregation, for mapped SSR.
Segregation: ab x cd = heterozygous in both parents with four alleles in total, ab x ac = heterozygous in both parents with three alleles in total, ab x cc = Male parent heterozygous, female parent homozygous with three alleles in total, ab x aa Male parent heterozygous, female parent homozygous with two alleles in total, aa x ab Female parent heterozygous, male parent homozygous with two alleles in total.

Locus	MgCl2 (mM)	Annealing temperature (°C)	Segregation (M x F)
EMBRA 6	1.5	58	ab x aa
EMBRA 11	1.5	56	ab x cd
EMBRA 12	1.5	56	ab x cd
EMBRA 16	1.8	56	ab x cd
EMBRA 30	1.5	58	ab x cd
EMBRA 37	1.5	56	ab x cd
EMBRA 48	1.5	54	ab x ac
EMBRA 56	1.8	58	ab x cd
EMBRA 61	1.5	58	ab x cd
EMBRA 66	1.5	56	ab x cd
EMBRA 78	1.5	58	ab x cc
EMBRA 98	1.5	56	ab x ac
EMBRA 115	1.5	56	ab x ac
EMBRA 127	1.5	56	ab x cd
EMBRA 128	1.5	57	ab x ac
EMBRA 130	1.5	57	ab x cd
EMBRA153	1.5	57	ab x ac
EMBRA 155	1.5	60	ab x ac
EMBRA 156	1.5	59	ab x cd
EMBRA 158	1.5	56	ab x ac
EMBRA 168	1.5	60	ab x ac
EMBRA 173	1.5	55	ab x cd
EMBRA 180	1.5	56	ab x ac
EMBRA 186	1.5	57	ab x ac
EMBRA 189	1.5	56	ab x cd
EMBRA 197	1.5	56	ab x aa
EMBRA 201	1.5	58	ab x ac
EMBRA 204	1.5	56	ab x ac
EMBRA 208	1.5	58	aa x ab
CSIRO En6	1.5	55	ab x cd
CSIRO En10	2.0	50	ab x ac
CSIRO En16	1.5	55	ab x cd
CSIRO Eg61	2.0	50	ab x ac
CSIRO Eg62	2.0	48	ab x cd
CSIRO Eg86	2.0	55	ab x cd
CSIRO Eg99	1.5	55	ab x ac
EMCRC 2	2.0	55	ab x aa
EMCRC 8	2.0	55	ab x cd
EMCRC 9	2.0	55	ab x cd
EMCRC 10	2.0	55	ab x cd
EMCRC 11	2.0	55	ab x cd

Chapter 4: The genetic dissection of pest susceptibility and growth traits in an F_2 inter-provenance cross of *Eucalyptus globulus*

Introduction

Variation in many traits of economic interest in forest trees is quantitative in nature. Under the infinitesimal (or polygene) model, continuous quantitative variation is assumed to reflect the collective action of numerous genes and environmental effects (Kearsey and Pooni 1996). Traditional approaches to quantitative genetic analysis use pedigree information to partition phenotypic variation into genetic and non-genetic variance components, which are used to estimate heritabilities and predict breeding values (Falconer and Mackay 1996; Lynch and Walsh 1998). However, this approach provides little information regarding the specific genetic factors underlying quantitative traits. The analysis of quantitative trait loci (QTL) is providing new insights into the number, chromosomal location and magnitude of effects of genetic entities controlling quantitative traits (Sewell and Neale 2000). In *Eucalyptus*, QTL have been located for traits including vegetative propagation (Grattapaglia *et al.* 1995; Marques *et al.* 1999), growth (Verhaegen *et al.* 1997; Grattapaglia *et al.* 1996), wood properties (Bundock 2003; Thamarus *et al.* 2004), early flowering (Missiagia *et al.* 2005), resistance to frost (Byrne *et al.* 1997) and disease (Junghans *et al.* 2003).

In recent years, breeding programmes for *E. globulus* have focussed primarily on selection for growth and wood property traits (e.g. Borralho *et al.* 1993; Miranda *et al.* 2001; Raymond and Apiolaza 2004). Of these traits, rapid growth directly affects plantation productivity and as such has historically been the main aim of selection. Variation in growth rate reflects a multitude of adaptive traits (Eldridge *et al.* 1993), including tolerance of abiotic and biotic stresses. Hence pests and pathogens that can limit *E. globulus* plantation growth and productivity in Australia have recently received increased attention. Such pests include possums (Dungey and Potts 2001; O'Reilly-Wapstra *et al.* 2002), sawfly (Jordan *et al.* 2002), southern eucalypt leaf beetle (Rapley

et al. 2004a), autumn gum moth (Jones *et al.* 2002b; Rapley *et al.* 2004b) and *Mycosphaerella* leaf blotch (Dungey *et al.* 1997; Carnegie and Ades 2003; Milgate *et al.* 2005a, 2005b). Similarly, in Brazil, studies are focussing on determining the genetic basis of resistance to diseases such as *Puccinia psiidi* (rust), *Cryphonectria cubensis* (canker), *Xanthomonas axonopodis* (spot) and *Ceratocystis fimbriata* (wilt) in *Eucalyptus grandis* and *E. urophylla* (Grattapaglia *et al.* 2004a).

Mycosphaerella leaf blotch (MLB) is a foliar fungal disease, affecting eucalypt plantations world-wide (Carnegie *et al.* 2001; Van Zyl *et al.* 2002; Mohammed *et al.* 2003). Over 30 species of *Mycosphaerella* have been detected on eucalypt leaves (Mohammed *et al.* 2003), of which *M. cryptica* and *M. nubilosa* are the most prevalent in southern Australia (Carnegie *et al.* 1998; Park *et al.* 2000). *Mycosphaerella cryptica* can occur on both juvenile and adult foliage and has been reported on 52 eucalypt species. In contrast, *M. nubilosa* appears to be confined to juvenile foliage and has been reported on only seven related species within the subgenus *Symphyomyrtus*, including *E. globulus* (Milgate *et al.* 2005b and references therein). MLB causes leaf necrosis and defoliation that can be highly detrimental to growth (Carnegie and Ades 2003; Milgate *et al.* 2005a). Crown damage from MLB can range from 10% necrosis of leaves to complete defoliation and tree death. However, even relatively minor damage can cause a significant loss in growth. For example, Carnegie and Ades (2003) reported that a loss of 10% of leaf area due to MLB resulted in reductions of up to 13% in tree height. The effect of MLB may be more severe in plantations than native forest because of plantings in sites outside the normal range of the species (Potts and Pederick 2000). In addition, plantations represent even-aged monocultures that may be more susceptible to disease than native forest, where epidemics can be restricted by the variable age structure and diversity of the plant community (Burgess and Wingfield 2002).

Genetic variation has been reported in the susceptibility of *E. globulus* to MLB both within and between provenances (Carnegie *et al.* 1994; Dungey *et al.* 1997; Milgate *et al.* 2005a, 2005b) and also within segregating families (Milgate *et al.* 2005b). These findings suggest selection for resistant genotypes of *E. globulus* has the potential to

reduce the impact of MLB on plantation productivity. However, the genetic control of resistance to these pathogens is not well understood. Effective management and breeding for disease resistance (using conventional or marker-assisted selection) will require a detailed understanding of the genetic variation within both the host and the pathogens and the genetic control of the host-pathogen interactions (Keane *et al.* 2000; Burdon 2001).

Complex inter-correlations have been detected between growth, height to vegetative phase change and the severity of damage caused by MLB. For example, a positive correlation has been observed between early growth and MLB severity (Dungey *et al.* 1997; Milgate *et al.* 2005a), suggesting that selection for rapid growth in the absence of the disease may inadvertently select for increased susceptibility (Dungey *et al.* 1997). However, Milgate *et al.* (2005a) reported severity of *M. nubilosa* was positively correlated with growth in the first year after planting, while from the second year onwards a significant negative correlation with growth traits was evident as the disease began to affect growth. Growth has also been positively correlated with height to phase change in some environments (Jordan *et al.* 2000). Correspondingly; a positive correlation has also been reported between height to phase change and the severity of MLB damage (Dungey *et al.* 1997). In light of these correlations, incorporating growth and height to phase change will be useful in dissecting the genetic control of resistance to MLB.

The main aims of this study were to investigate the quantitative genetic control of resistance to MLB, and other potentially correlated traits, in an F_2 inter-provenance cross of *E. globulus* with grandparents known to differ widely in their susceptibility to the disease (Dungey *et al.* 1997). QTL analysis was performed using a linkage map constructed from SSR and AFLP markers in this pedigree (see Chapter 3), using single marker tests, interval and multiple QTL model (MQM) analysis.

Materials and methods

Genetic material and trial design

In order to study genetic variation in *E. globulus* susceptibility to MLB, a large outcross F_2 family (family 1) was generated, with grandparents originating from King Island (KI) in Bass Strait and Taranna (T) in the south east of Tasmania (T7/KI157//T144/KI5). Family 1 was constructed from grandparents with divergent predicted breeding values for MLB susceptibility and resistance (Dungey *et al.* 1997) in order to promote segregation for disease susceptibility in the offspring. It contained 240 genotypes, of which 160 were replicated clonally (one ramet and one ortet), resulting in two trees per genotype. Clonal replication allowed an improved estimate of genetic variation, by accounting for environmental heterogeneity in the field trial. For example, clonal replication would reduce the chance of both individuals representing a single genotype having low disease severity solely due to disease escape, as opposed to genetically governed resistance. Family 1 was planted in a field trial at Woolnorth in north-west Tasmania in May 1998, as part of an experiment containing two other controlled cross families and four open pollinated families (see Milgate *et al.* 2005b for full description of the trial design). The trial featured an incomplete block design, consisting of two replicates, each with 20 incomplete blocks of 30 or 36 trees. The ortet and ramet of each cloned genotype were assigned to separate replicates at random.

Assessment of phenotypic traits

The Woolnorth field trial became infected with *Mycosphaerella cryptica* (identified by Milgate *et al.* 2005b) less than one year after planting. Disease severity was recorded after 12 months of growth when all trees were entirely in their juvenile phase (myco; Milgate *et al.* 2005b). The disease severity was quantified by estimating the percentage area of necrotic lesions present on each tree, on a ten-point scale (1 = 0-3%, 2 = 4-6%, 3 = 7-12%, 4 = 13-17%, 5 = 18-25%, 6= 26-38%, 7 = 39-50%, 8 =51-63%, 9 = 64-75%, 10 = 76-100%). The mid-point of each class was used for analysis. The field trial also suffered from herbivory by the autumn gum moth (*Mnesampela privata*) in the first year of growth. The level of herbivory was assessed on juvenile foliage, at 13 months of age

(Jones *et al.* 2002b), based on a visual assessment of damage across the whole tree (agm; as described by Jones *et a*l. 2002b). Various other phenotypic traits were recorded annually, following the establishment of the field trial, including growth assessed by total height (in m), at one to three years of age (ht1, ht2, ht3), diameter (in cm) at breast height (dbh) of the main stem at age two to six years (dbh2, dbh3, dbh4, dbh5, dbh6). Height at which vegetative phase change from the juvenile to the adult foliage took place was monitored annually for the first five years of growth. For each tree, the stem height (in m) at which the first petiolate leaf appeared was recorded, when phase change was first observed.

Analysis

In order to test for significant differences between genotypes within family 1, a mixed model incorporating the effects of replicate (fixed), genotype (random) and random error, was fitted for each trait using the PROC MIXED procedure of SAS. Only genotypes with clonal replication were included in this analysis. Phenotypic trait values used for QTL analysis were least-square means of two clonal individuals representing each genotype, calculated by fitting a mixed model including the effects of replicate (fixed), genotype (fixed), incomplete block within replicate (random) and random error. All families in the field trial were included in the mixed model and the least-square means for genotypes with clonal replication were selected for QTL analysis. Hence, the reported phenotypic variance explained by each putative QTL (see below) has had some of the environmental variance removed. The distribution of genotype least-square means was examined for each trait, in order to test the assumption of a normal distribution for interval and MQM mapping. Although the least-square means for several traits departed from normality (Kolomogorov-Smirnov $P < 0.05$), in most cases the data were evenly distributed about the mean and therefore no transformation was undertaken. A natural log transformation was applied where appropriate (agm and myco traits) to optimise normality, however in both cases even with transformation the distribution of these traits departed from normality (Kolomogorov-Smirnov $P < 0.01$). Pearson's correlation coefficients were calculated between traits using PROC CORR of SAS and genotype least square means.

In light of the correlations detected previously (between height and the severity of MLB [Dungey *et al.* 1997; Milgate *et al.* 2005a], weed competition, height and the severity of damage caused by the autumn gum moth [Jones *et al.* 2002b]) and in the present study (between height and height to phase change and between the severity of damage caused by *M. cryptica* and the autumn gum moth), covariates were fitted to remove their effects on the expression of each trait. The least-square means for height to vegetative phase change (htpc), severity of damage by the autumn gum moth (agm) and *M. cryptica* (myco) were used in the QTL analysis as raw values and residuals. The residuals were obtained by fitting the genotype least-square means for key covariates (ht1, myco and weed competition for agm, ht1 and weed competition for myco and ht1 for htpc) using PROC REG of SAS. Weed competition was assessed in a radius of one metre from the tree, using a five-point scoring system considering weed height and percentage cover (1 = no weeds; 5 = maximum cover overtopping tree; Jones *et al.* 2002b).

QTL analysis was conducted with MAPQTL 4.0 (Van Ooijen *et al.* 2002), using the consensus linkage map constructed in family 1 (described in detail in chapter three of this thesis) and least-square means for each trait. Putative QTL were declared at two different levels, significant (genome-wide type I error < 0.05) and suggestive (chromosome-wide type I error rate < 0.05). The LOD threshold for genome wide significance was determined by permutation testing (1000 replications; Churchill and Doerge 1994). An average chromosome-wide LOD threshold (LOD 3) for suggestive QTL, was determined by empirical simulations (Van Oijen 1999). Interval mapping was initially performed in order to scan the genome for map intervals significantly associated with each trait, using the default parameters of MAPQTL 4.0 (Van Ooijen *et al.* 2002). The suggestive threshold was used to select putative QTL as cofactors for the MQM procedure. Where map intervals exceeding this threshold existed, a single marker closest to the peak LOD was chosen as a cofactor for the multiple QTL model (MQM; Jansen 1993, 1994; Jansen and Stam 1994). During interval and MQM mapping procedures, the association at each map position is tested against the residual variance, the larger the genetic effect associated with a position is in relation to the residual variance, the more

significant the test (Van Oijen *et al.* 2002). MQM mapping removes the residual variance attributable to the selected cofactors, thereby increasing the power in subsequent scans for multiple QTL. As a result, MQM mapping can improve the precision in locating multiple QTL, for example by highlighting 'ghost QTL' (Van Oijen *et al.* 2002), as well as often identifying more QTL than interval mapping alone. An iterative approach, combining forward and backward selection of cofactors was used for QTL detection. Forward selection was used initially, by running the MQM procedure with the peaks exceeding the suggestive threshold from interval mapping selected as cofactors. This process was repeated until no further cofactors were discovered. Subsequently, a backward elimination procedure was performed. This involved selecting a subset of markers approximately every 20 cM (subject to computational limitations) along each chromosome sequentially as cofactors and running the MQM analysis including the cofactors discovered from forward selection (where present). The analysis for each trait was accepted when the selected cofactors were closest to each significant peak and no new putative QTL were discovered (Van Ooijen *et al.* 2002).

All methods of QTL discovery have their biases, so a combination of approaches can increase confidence in the detection of putative QTL (Asins 2002; Kearsey and Farquhar 1998). Because the distribution of the severity of *M. cryptica* and the autumn gum moth departed from normality, despite the log transformation, the single marker Kruskal-Wallis test was also used for QTL detection. This technique makes no assumptions regarding the distribution of the trait and is therefore robust to departures from normality (Van Ooijen *et al.* 2002).

Assessments of the effects of different allelic combinations for significant markers (from interval and MQM mapping) on the mean trait value were gained using PROC GLM of SAS. Contrast statements were used to test whether the effect segregated from the male or female parent for significant QTL. Where more than one unlinked marker significantly affected a quantitative trait (genome-wide type I error < 0.05 in interval mapping), they were tested for epistasis by testing for an interaction effect using PROC GLM of SAS.

Results

Family one exhibited significant variation between genotypes ($P < 0.05$) for all traits, with the exception of height after one and three years (ht1 and ht3; Table 4.1). The severity of *M. cryptica* damage (myco) and height to vegetative phase change (htpc) exhibited highly significant variation between genotypes ($P < 0.0001$). A high positive correlation was evident between growth measurements (diameter and height; dbh and ht) taken from one to six years of age (Table 4.2). Height to phase change (htpc) was correlated with early growth (diameter at two years and height at one year of age; dbh2, ht1; Table 4.2), indicating trees which changed phase at a greater height intially grew more rapidly. Damage due to the autumn gum moth (agm) was also positively correlated with height at one year of age (ht1), suggesting a tendency for taller plants to be damaged preferentially. Damage due to the autumn gum moth (agm) was negatively correlated with damage due to *M. cryptica* (myco; Table 4.2).

Table 4.1 The significance of differences between genotypes within family 1 for various traits scored in the *Eucalyptus globulus* Woolnorth field trial. The number of genotypes assessed, trait units, mean and its standard deviation, and Z value for the between genotype variance component and its significance from zero are indicated.

	No.	Units	Mean	s.d.	Z	P
myco	115	%	5.8	0.64	5.94	<0.0001
agm	116	%	5.2	0.56	2.32	0.0102
htpc	117	m	4.3	0.66	5.67	<0.0001
ht1	119	m	1.7	0.30	0.05	0.4789
ht2	119	m	4.7	0.60	1.99	0.0231
ht3	116	m	6.4	0.67	1.61	0.0536
dbh2	119	cm	5.4	0.98	1.74	0.0412
dbh3	115	cm	7.0	1. 18	1.8	0.0362
dbh4	115	cm	9.4	1.47	1.75	0.0399
dbh5	115	cm	10.8	1.62	1.86	0.0317
dbh6	114	cm	12.5	1.93	1.71	0.044

Table 4.2 Matrix of Pearson correlation coefficients between the genotype least-square means for the various traits analysed in this study.

	myco	agm[a]	htpc	ht1	ht2	ht3	dbh2	dbh3	dbh4	dbh5
agm	-0.21*									
htpc	-0.03	0.00								
ht1	0.16	0.19*	0.23*							
ht2	-0.01	0.06	0.13	0.65****						
ht3	-0.12	0.14	0.11	0.57****	0.90****					
dbh2	-0.01	0.06	0.18*	0.67****	0.88****	0.8****				
dbh3	-0.09	0.15	0.14	0.60****	0.88****	0.90****	0.95****			
dbh4	-0.06	0.16	0.09	0.58****	0.85****	0.89****	0.91****	0.96****		
dbh5	-0.07	0.15	0.06	0.54****	0.83****	0.90****	0.88****	0.95****	0.97****	
dbh6	-0.07	0.14	0.04	0.53****	0.79***	0.87****	0.86****	0.92****	0.95****	0.98****

P > 0.05 = ns, * < 0.05, ** < 0.01, *** < 0.001 =, **** < 0.0001.
[a] Scored after myco.

Putative QTL were identified by interval and MQM mapping (Table 4.3), for severity of *M. cryptica* damage (myco) and autumn gum moth damage (agm). Putative QTL were also located for growth traits, height at two and three years of age (ht2 and ht3) and diameter at breast height at two, five and six years of age (dbh2, dbh5 and dbh6), in addition to height to vegetative phase change (htpc).

Interval mapping located putative QTL regions with significant (genome-wide type I error < 0.05) effect on the severity of damage by *M. cryptica* (myco) on linkage groups 1 (LOD 6.19) and 8 (LOD 11.39; Table 4.3). Another putative QTL region was located at the suggestive level (chromosome-wide type I error < 0.05) on linkage group 3 (LOD 3.08). The MQM mapping procedure increased the significance of the detected markers on linkage groups 1 and 8 (LOD 10.92 and 20.20, respectively) and allowed the detection of an additional significant QTL region on linkage group 7 (LOD 5.98) and suggestive QTL regions on linkage groups 2 and 3 (LOD 3.46 and 3.47, respectively; Table 4.3). In combination, the putative QTL regions explained 76.3% (interval mapping) and 71.8% (MQM mapping) of the phenotypic variance (i.e. the variation in genotype least-square means) in the severity of damage by *M. cryptica*. All putative QTL for *M. cryptica* damage segregated on the female side, except the one on linkage group 3 (Table 4.3). The difference between the mean for the severity of *M. cryptica* damage (5.8%; Table 4.4) and the genotype mean for individuals with the favourable allele for the putative QTL on linkage group 1 was 1.05%, while the difference between the genotype means for individuals with the positive (a) and negative (b) alleles from the female parent (i.e. the mean for ac plus ad versus the mean for bc plus bd) was 2.65% (Table 4.4). In the case of the putative QTL on linkage group 8, the difference between the family mean and the genotype mean for individuals with the favourable allele was 1.6%, while the difference between the genotype mean for individuals with the positive and negative alleles (from the female parent) was 4.3% (Table 4.4). The putative QTL discovered for severity of *M. cryptica* damage were stable when the slight positive correlation with height at one year of age (ht1; Table 4.2) was removed, in the analysis of covariance (interval mapping LOD; 5.60 and 11.65 for the putative QTL on linkage groups 1 and 8, respectively). The interaction effect between the two major QTL on

linkage groups 1 and 8 was non-significant, suggesting no epistasis between them and their effects are additive (data not shown).

A putative QTL region for susceptibility to the autumn gum moth was located, by interval mapping, on linkage group 8 at the suggestive level (LOD 3.15). Following the inclusion of putative QTL as cofactors in the MQM mapping procedure, the association became significant (LOD 5.50) and two more putative QTL regions were identified on linkage groups 5 and 6 at the suggestive level (LOD 3.22 and 4.15, respectively). In combination, the three putative QTL regions explained a total of 82.4% of the phenotypic variance in this trait (Table 4.3). The difference between the family mean (5.1%) and the genotype mean for individuals with the favourable allele was 0.5%, while the difference between genotype means for the individuals with the positive and negative alleles was 0.7% (Table 4.4). The putative QTL for susceptibility to the autumn gum moth were no longer present when the genotype least-square means were adjusted by fitting the correlated measures (Table 4.2) of height at one year of age (ht1), weed competition and severity of damage by *M. cryptica* (myco) as covariates (interval mapping LOD; 0.71, 2.33 and 2.42 for the putative QTL on linkage groups 5, 6 and 8, respectively).

A putative QTL region influencing height at 2 years of age (ht2) was identified on linkage group 1 at the suggestive level (LOD 3.26), explaining 23.0% of the phenotypic variance. The same region also affected height at three years of age (ht3), at the significant level (LOD 3.98), explaining 23.8% of the phenotypic variance (Table 4.3). The difference between the family mean for ht3 (6.5 m) and the genotype mean for individuals with the favourable allele was 0.1 m, while the difference between genotype means for the QTL alleles segregating from the female parent was 0.3 m (Table 4.4). Different QTL were found to influence diameter at breast height (dbh), but only at the suggestive level. A putative QTL region at two years of age (dbh2) was located on linkage group 5 at the suggestive level (LOD 3.06), explaining 14.2% of the phenotypic variance. A different suggestive QTL region was found to affect later age diameter (dbh5 and 6) on linkage group 10 at the suggestive level (LOD 3.07 and 3.31,

Table 4.3 Putative QTL identified from interval, MQM mapping and the (single marker) Kruskal-Wallis test.

Trait	Group	Segregation[a]	Adjacent Marker[b]	Map position cM[c]	Interval mapping LOD[d]	% exp[e]	MQM mapping[f] LOD[d]	% exp[e]	Single marker tests[g]
myco	1	F	p11b12	12.5	6.19*	24.6	10.92***	15.4	(Emb180, 13.1) ****
	2	F	CRC8	28.2	2.02	8.2	3.46[s]	4.9	*
	3	M	p09b04	75.4	3.08[s]	13.6	3.47[s]	6.2	
	7	F	p08b24	21.7	2.62	11.5	5.98*	9.0	
	8	F	p04b07	16.1	11.39***	38.1	20.20***	36.3	***
agm	5	M	p17b05	0.0	1.09	4.7	3.22[s]	14.2	*
	6	F	p12b06	15.2	2.08	24.3	4.15[s]	29.9	
	8	F	p10b03	55.9	3.15[s]	30.6	5.50*	38.3	
htpc	1	M	Emb12	4.2	4.06*	17.0			*
ht2	1	F	p21b04	59.2	3.26[s]	23.0			
ht3	1	F	p21b04	59.2	3.91*	23.8			*
dbh2	5	M	Emb37	105.6	3.06[s]	14.2			**
dbh5	10	M	p18b10	1.9	3.07[s]	17.2			
dbh6	10	M	p18b10	1.9	3.31[s]	19.3			

[a] The parent from which the QTL effect segregates: M = male segregation, F = female segregation.
[b] The marker closest to each QTL peak.
[c] Map position of the marker closest to each QTL peak.
[d] Peak LOD score for each QTL, genome wide significance is indicated by *.
[s] Suggestive QTL based on a chromosome-wide significance $P < 0.05$.
[e] The percentage of phenotypic variation explained at each QTL peak.
[f] The results of MQM mapping are presented only where more than one putative QTL is identified for a trait (myco and agm), otherwise the MQM results do not differ from those of interval mapping.
[g] Significance from the Kruskal-Wallis test. In cases where the closest marker to the QTL peak segregates from the opposite parent, the significance of the closest marker segregating from the same parent as the QTL is presented (marker, map position in cM). Significance level for the Kruskal-Wallis test and genome-wide LOD significance: * $P < 0.05$, P ** < 0.01, P *** < 0.001.

Table 4.4 Genotype and overall trait means for markers adjacent to the significant QTL, identified by MQM mapping[a].

Trait	LG[b]	Marker	Parental genotypes[c]		Genotype means				Family mean[d]
			Female	Male					
myco (%)	1	Emb12	ab***	cd	ac (5.4)	ad (4.1)	bc (6.9)	bd (7.9)	5.8
	7	p11b06	lm*	ll	lm (4.9)	ll (6.5)			5.7
	8	P04b07	lm***	ll	lm (4.0)	ll (8.3)			5.6
agm (%)	8	P04b04	hk	hk	hh 4.6	k_ 5.3			5.1
htpc (m)	1	Emb12	ab	cd**	ac (4.3)	ad (4.2)	bc (4.5)	bd (4.0)	4.2
ht3 (m)	1	p21b04	lm*	ll	Lm (6.6)	ll (6.3)			6.5

[a] Genotype means are given for the closest adjacent marker to each QTL peak, except where the closest marker segregates from the opposite parent to the QTL, in which case genotype means for the closest marker segregating from the same parent as the QTL are presented.
[b] LG = Linkage group
[c] Significance levels * < 0.05, ** < 0.01, *** < 0.001
[d] These family mean values differ slightly from those presented in Table 4.1 as they were calculated from the least-square means of each genotype in family 1, derived from the full trial analysis.

respectively) explaining 17.2 and 19.3% of the phenotypic variance for diameter at breast height, respectively (Table 4.3).

A putative QTL region for height to phase change (htpc) was also identified on linkage group 1 at the significant level (LOD 4.06), explaining 17% of the phenotypic variance in height to vegetative phase change. The difference between the family mean (4.2 m) and the means for genotypes bearing the c versus the d allele was 0.2 m and 0.1 m, respectively (hence there was a mean difference of 0.3 m in height to phase change between genotypes with each allele; Table 4.4). This QTL for phase change co-located with a more significant QTL for the severity of damage by *M. cryptica* (myco1; Table 4.3 and Figure 4.1), but segregation was from the opposite parent (htpc segregated from the male, myco1 from the female). The QTL for height to phase change (htpc) was still present, albeit with a reduced LOD, when the genotype least-square means were adjusted by removing the positive covariance with height at one year of age (interval

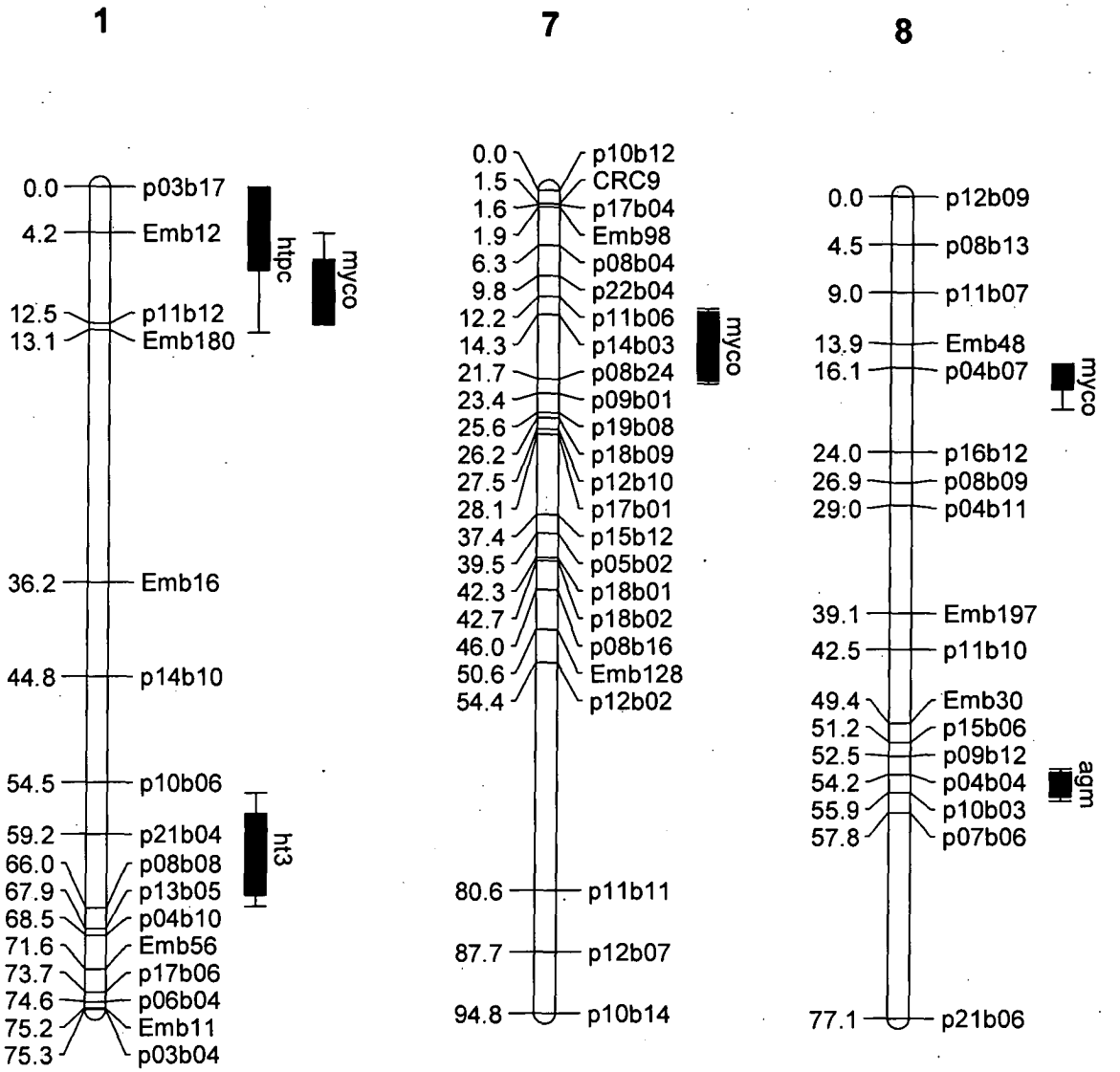


Figure 4.1 The location of each putative QTL exceeding the genome-wide significance level from interval and MQM mapping. Linkage groups and loci are as described in chapter 3. Solid bars and lines represent one- and two-LOD support intervals, respectively. The one-LOD support interval corresponds to an approximate 95% confidence interval (Lander and Botstein 1989). (htpc: height to vegetative phase change, myco: severity of *M. cryptica* damage, ht3: tree height at three years of age, agm: severity of damage by the autumn gum moth).

mapping LOD 3.22).

All putative QTL that exceeded the genome-wide significance threshold from interval mapping were confirmed by the single marker Kruskal Wallis test (Table 4.3). MQM mapping led to the detection of two further putative QTL exceeding the genome-wide significance level (for susceptibility to the autumn gum moth [agm] on linkage group 8 and *M. cryptica* severity [myco] on linkage group 7), however these QTL were not confirmed by the Kruskal Wallis test, reducing confidence in their validity. Further, several putative QTL at the suggestive level had a relatively low LOD in the interval mapping (those for myco on linkage groups 3 and 7, agm on linkage group 6, ht2 on linkage group 1, dbh5 and 6 on linkage group 10) and their flanking markers did not exceed the 0.05 threshold in single marker tests (Table 4.3), again reducing confidence in their validity. However, in the case of the putative QTL for height at two years of age (ht2; on linkage group 1), the same region was identified as influencing height in the third year and this was confirmed by the Kruskal Wallis test. In each case where more than one map region affected a trait (myco and agm), MQM mapping identified each QTL region more precisely, found more putative QTL and increased the significance of the putative QTL identified by interval mapping.

Discussion

The genetic variation within family 1 includes the segregation of variance due to differences between Taranna and King Island provenances of *E. globulus* and differences within these provenances. The highly significant genetic variation in height to vegetative phase change and the severity of damage caused by *M. cryptica* within this family are consistent with the large amount of intraspecific variation known to exist within *E. globulus* for a variety of morphological traits (Dutkowski and Potts 1999; Jordan *et al.* 1993). Specifically, genetic variation has previously been reported within and between provenances of *E. globulus* for damage caused by MLB (Carnegie *et al.* 1994; Dungey *et al.* 1997; Milgate *et al.* 2005a, 2005b) and the timing of vegetative phase change (Dutkowski and Potts 1999; Jordan *et al.* 1999). The significant differences detected between genotypes for all traits within this family demonstrate that numerous traits are segregating and have the potential for QTL discovery in this pedigree.

The lack of a significant correlation between the severity of *M. cryptica* (myco) and growth or height to phase change measurements, is in contrast to the findings of previous studies (Dungey *et al.* 1997; Milgate *et al.* 2005a). However, a slight, albeit not significant, positive correlation (r = 0.16) was evident between the severity of *M. cryptica* (myco) and height at one year of age (ht1), consistent with past findings (Dungey *et al.* 1997; Milgate *et al.* 2005a). The initial positive correlation may be due to the greater opportunity for disease increase within the larger juvenile canopy as suggested by Dungey *et al.* (1997). In the case of Milgate *et al.* (2005a), the severity of MLB was higher (mean leaf area damage = 34%) leading to a negative correlation with growth from the second year of measurement as the disease significantly affected growth. However, in the present study *M. cryptica* infection was far less severe (mean = 5.8%), which could explain the lack of a significant negative effect on growth, similar to the findings of Dungey *et al.* (1997). In the case of height to phase change, the positive correlations found by Dungey *et al.* (1997) are likely to reflect the time at which MLB damage was scored (see below).

The negative correlation ($r = -0.21$; $P < 0.05$) between susceptibility to the autumn gum moth (agm) and the severity of *M. cryptica* (myco) is likely to reflect the influence of *M. cryptica* damage on choice of oviposition by the autumn gum moth, since the outbreak of *M. cryptica* occurred prior to the main outbreak of defoliation by the autumn gum moth (Jones *et al.* 2002b). In contrast, susceptibility to the autumn gum moth (agm) was positively correlated ($r = 0.19$; $P < 0.05$) with height after one year of growth (ht1). This positive correlation may reflect the fact that taller trees had less weed competition and provided a more prominent target for herbivory by the autumn gum moth, at this early stage of growth.

Threshold levels for QTL identification

Numerous researchers have proposed adopting relaxed type I error rates in exploratory QTL studies, in order to facilitate comparative QTL mapping (Lander and Kruglyak 1995; Beavis 1998; Van Oijen 1999). Stringent type I error rates will result in an increased rate of type II errors and genuine QTL being overlooked. Hence, putative QTL at a 'suggestive' (chromosome-wide type I error < 0.05; Van Oijen 1999) level have also been presented. Although it is expected some of the suggested QTL will be false positives, others may represent QTL with small effects that may prove more significant in future QTL studies. Comparative QTL mapping will be particularly useful in a genus such as *Eucalyptus*, in which QTL mapping is in its infancy and in outcrossing taxa in general, where their heterogeneous genetic backgrounds will make QTL validation more difficult than in crop plants. Furthermore, the small sample size generally employed in QTL studies of forest trees reduces the power to identify statistically significant QTL at the genome-wide level (Beavis 1998). The inclusion of numerous transferable SSR markers in the linkage maps used for QTL location in this study presents a good opportunity for such comparative mapping.

QTL number and effects

The following discussion focuses on putative QTL that were significant at the genome-wide level (type I error < 0.05). Although one significant QTL has been reported for

susceptibility to the autumn gum moth (agm), it was not supported by the single marker Kruskal Wallis test (Table 4.3) reducing confidence in its validity. Further, the fact that this QTL was not present after fitting the correlated traits (height at one year of age [ht1], weed competition and the severity of damage by *M. cryptica* [myco]; Table 4.2), as covariates suggests the expression of susceptibility to the autumn gum moth (agm) is strongly influenced by variation in the covariates, rather than solely reflecting morphological or physiological based resistance *per se*. Hence, the QTL for susceptibility to the autumn gum moth (agm) will not be discussed at length.

Three significant QTL regions have been detected for *M. cryptica* severity (myco), of which two explain a large proportion of the phenotypic variance in this study. Less significant QTL have been identified for susceptibility to the autumn gum moth, height and height to phase change. Contrary to predictions from an infinitesimal (or polygenic) model, QTL studies in forest trees and indeed all plants have often suggested oligogenic or major gene control, with a few major genes responsible for much of the total phenotypic variability. For example, a survey of QTL experiments across various traits in 20 forest tree species, found the number of QTL identified per trait ranged from 0-7, with a mean of 2.7, each QTL explaining from 3.4-62.6% of phenotypic variance (Sewell and Neale 2000). Similarly, in *Eucalyptus* a few major loci or QTL with reasonably large effects have often been identified for quantitative traits, ranging from disease resistance (Junghans *et al.* 2003), to volume growth (Grattapaglia *et al.* 1996) and frost resistance (Byrne *et al.* 1997). However, the common finding of oligogenic control must be interpreted with caution in light of the fact that the power of QTL detection is influenced by numerous factors including the sample size, the magnitude of the QTL effect and the heritability of the trait (Beavis 1998; Kearsey and Farquhar 1998; Erickson *et al.* 2004). In combination, these factors commonly result in under estimates of the number of QTL affecting a trait and over estimates of their phenotypic effects.

Nonetheless, discovery of just two major QTL, explaining a large proportion of the phenotypic variance in *M. cryptica* severity in this cross suggests that susceptibility may be under oligogenic control. The fact that the QTL are stable after fitting height at one

year (ht1) and weed competition as covariates increases confidence that the expression of this trait is reflecting genetic based resistance, distinct from the influence of the covariates. Oligogenic control is consistent with the moderate to high heritability previously reported for damage from MLB within *E. globulus* (narrow sense heritability = 0.12 - 0.6; Dungey *et al.* 1997; Milgate *et al.* 2005a). However, the possibility that the regions with QTL of large effect represent clusters of closely linked genes cannot be dismissed, since resistance genes are commonly clustered in plants (Wang *et al.* 2001; Chu *et al.* 2004; Mondragon-Palomino and Gaut 2005). In contrast to crop plants, there have been few studies attempting to locate QTL affecting disease resistance in forest trees. This study reports the first QTL affecting disease resistance in *E. globulus* and is the second reported for a eucalypt species. Hence direct comparison is limited. However, Junghans *et al.* (2003) reported only two major QTL controlling a large proportion of the phenotypic variation for rust resistance in *Eucalyptus grandis*, which appeared to display qualitative (i.e. Mendelian) inheritance. In other forest trees, QTL studies have commonly found that few genomic regions explain a large proportion of the phenotypic variation in disease resistance. For example, one or two major QTL explain most of the variance in resistance to *Melampsora* and *Septoria* in poplar (Lefèvre *et al.* 1998; Newcombe *et al.* 1996; Jorge *et al.* 2005) and *Cronartium ribicola* in *Pinus* (Devey *et al.* 1995). Similarly, Young (1996) reviewed QTL studies of disease resistance in crop plants, finding that while there were some examples of many (i.e. > 10) loci influencing quantitative resistance, it was more common to find fewer (i.e. 3-5) QTL and often one or two predominated.

Attempts to investigate the genetic control of resistance to MLB are complicated by the many potential mechanisms contributing to different levels of host resistance, ranging from morphological traits such as stomatal wax coverage and leaf density (Smith *et al.* 2005) to physiological effects such as the hypersensitive response (Hammond-Kosack and Jones 1996). However, in general, the suggestion of oligogenic control is consistent with specialised associations that may exist between hemibiotrophic pathogens such as *Mycosphaerella* and their host. Hemibiotrophic pathogens invade living cells and subvert metabolism to favour their growth and reproduction, hence minor differences in

either organism can upset the balance (Hammond-Kosack and Jones 1996). The genetics of host resistance and pathogen virulence in biotrophic pathogens often fit the classic gene-for-gene model (Flor 1971), whereby complimentary pairs of dominant genes in both the pathogen and the host govern resistance. This model may apply to the control of resistance to *Mycosphaerella cryptica* in *E. globulus*. Milgate *et al.* (2005b) provided evidence for a close association between the host and pathogen, finding specialisation in the occurrence of two different biotypes of *M. cryptica* both between *E. globulus* and *E. nitens* and between resistant and susceptible individuals within *E. globulus,* in the same trial as was used in this study.

Growth has been one of the most commonly studied traits for QTL identification in *Eucalyptus* (Verhaegen *et al.* 1997; Grattapaglia *et al.* 1996; Bundock 2003) and forest trees in general (Bradshaw and Stettler 1995; Kaya *et al.* 1999), reflecting its economic importance and ease of measurement. Again, there is limited opportunity for direct comparison between studies, because most studies have employed different experimental designs, markers and analysis techniques. However, in comparison to disease resistance traits, studies have generally found growth to be explained by QTL with lesser phenotypic effect. For example, in *E. grandis* three QTL were identified which explained a total of 13.7% of phenotypic variation (in circumference at breast height) at 6.5 years of age (Grattapaglia *et al.* 1996). Similarly, Bundock (2003) found two QTL explaining a total of 18.5% of the phenotypic variance in stem diameter in *E. globulus*. These findings probably reflect the fact that growth is influenced by numerous adaptive traits and strong environmental effects (Eldridge *et al.* 1993), thus generally features moderate to low heritability. Hence, the location of just one significant putative QTL, influencing 23.8% of the phenotypic variance for height, is probably not indicative of oligogenic control in this instance, but may reflect bias in the estimation of the QTL magnitude and insufficient power to detect QTL due to a small sample size (Beavis 1998). Furthermore, in studies where shared markers allow direct comparison of QTL homology with the present study (Bundock 2003; Grattapaglia *et al.* 1996) a total of nine different QTL for growth traits (height, diameter at breast height and circumference at

breast height) have been located in seven different linkage groups (in *E. globulus* and *E. grandis* combined; Table 4.5), consistent with polygenic control.

Eucalyptus globulus, like many eucalypt species, is heteroblastic, i.e. it shows an abrupt and marked change in vegetative morphology. Apart from obvious changes in leaf morphology, the different leaf phases may also differ in physiology (Pederick 1979; James and Bell 2001), chemistry (Li *et al.* 1995, 1996) and resistance to pests and diseases (Farrow *et al.* 1994; Dungey *et al.* 1997; Steinbauer 2002). Phase change appears to be an adaptive trait, although the adaptive significance of the transition to adult foliage is complex and appears to differ across environments (Jordan *et al.* 2000). The timing of phase change is under moderate to strong genetic control (narrow sense heritability 0.2 - 0.9; Jordan *et al.* 1999). However, despite marked genetic differences in the timing of vegetative phase change within *E. globulus* (Dutkowski and Potts 1999; Jordan *et al.* 2000), QTL influencing this trait have not been identified to date in *E. globulus* or other species in the genus. Hence, direct comparison of QTL number and effects is not possible. However, QTL have recently been reported in *E. grandis*, for another developmental trait, flowering precocity (Missiaggia *et al.* 2005), for which a locus with major effect was identified segregating in close to a 1:1 ratio. Similarly, Bundock (2003), found one QTL affecting bud abundance in *E. globulus*. In light of the above findings and the moderate to high heritability for the timing of vegetative phase change (Jordan *et al.* 2000), it could be hypothesised that this trait may also be under oligogenic control, despite the discovery of only one QTL with moderate effects in this cross.

Direct comparison of the number and effects of QTL for various traits is limited by the different species, experimental designs, analysis techniques and significance thresholds employed by comparable studies. However, the number and effects of the QTL reported herein appear generally consistent with previous studies, although there are perhaps fewer significant QTL than would be expected. Small sample size, a departure from normality in the distribution of susceptibility to *M. cryptica* and autumn gum moth damage and insufficient map coverage may have contributed to the finding of only a few

Table 4.5 QTL homology in subgenus *Sympyomyrtus*. The numbering of linkage groups in this study is the same as that established in *E. grandis* by Grattapaglia *et al.* (1995, 1996).

group[a]	This study[b]	Bundock (2003)[c]	Thamarus *et al.* (2001)[d]	Marques *et al.* (1999)[e]	Grattapaglia *et al.* (1995)[e]	Grattapaglia *et al.* (1996)[f]	Missiaggia *et al.* (2005)
	E. globulus	*E. globulus*	*E. globulus*	*E. tereticornis* & *E. globulus*	*E. grandis* & *E. urophylla*	*E. grandis*	*E. grandis*
1	myco*, ht2, ht3*, htpc*	bud abundance3 (4 & 5)	MFA (7)		veg prop (g, u4)		
2	myco		pulp, cellulose (4)	veg prop (2 & 14)	veg prop (g)	wsg, cbh	Early flowering
3	myco		density, cellulose (11)		veg prop (g, u3)		
4					veg prop (u8)		
5	agm, dbh2	dbh6 (3)	density (9)	veg prop (10)	veg prop (g, u6)	wsg, cbh, %bark	
6	agm					wsg, %bark	
7	myco				veg prop (g, u2)	cbh, wsg	
8	myco*, agm	pilodyn6/1 (6)			veg prop (g)		
9		dbh3 (1 & 2)			veg prop (u1)		
10	dbh5, dbh 6			veg prop (4)	veg prop (g)		
11					veg prop (g)		

[a] Linkage group in this study, in maps of *E. grandis* by Grattapaglia *et al.* (1995, 1996) and Brondani *et al.* (1998, 2002).
[b]* Significant QTL in the present study, from interval and MQM mapping. Significance threshold determined by permutation testing (1000 replications; Churchill and Doerge 1994).
[c]dbh = Stem diameter at breast height, pilodyn = pilodyn penetration (an indirect measure of wood density).
For each trait # = age in years, (#) = linkage group, coded as reported in the study referenced.
[d]pulp = Pulp yield, cellulose = cellulose content, density = wood density, MFA = micro fibril angle.
[e]veg prop = Traits relating to vegetative propagation. u = *E. urophylla*, g = *E. grandis*, (#) = linkage group, for *E. urophylla*.
[f]wsg = Wood specific gravity, cbh = stem circumference at breast height, %bark = percentage weight of bark wood.

significant QTL in this study and an over estimate of their effects. Specifically, interval mapping assumes an approximately normal trait distribution. Therefore, non-normality may have led to an over-estimate of the QTL effects (Shepherd *et al.* 1999) for *M. cryptica* and particularly autumn gum moth susceptibility, where the estimated phenotypic effects appear inflated because two of the QTL were not confirmed by the Kruskal-Wallis test. However, in the case of the QTL for severity of damage by *M. cryptica* (myco), the fact that the two major QTL were confirmed by the non-parametric Kruskal-Wallis test provides some protection against a departure from normality. Despite the deficiencies in the present study, the power for QTL detection was enhanced by clonal replication of the mapping progeny and the use of F_2 mapping pedigree (Beavis 1998), as opposed to the pseudo-backcross design commonly used for QTL detection in eucalypts (Grattapaglia 2000). Further, susceptibility to MLB has been shown to be highly heritable in *E. globulus* (Milgate *et al.* 2005a), again increasing the power for QTL detection (Beavis 1998).

QTL location

Four of the putative QTL located in this study (one for *M. cryptica* severity, two for height and one for height to vegetative phase change) were found on linkage group 1 (Table 4.3 and Figure 4.1). The same map region affected height at two and three years of age (ht2 and ht3) and is likely to indicate control by the same loci. This putative QTL (ht2 and ht3) was located at the opposite end of the linkage group to the other two traits (55 and 46.7 cM from the QTL for height to phase change and *M. cryptica* severity, respectively; Figure 4.1), suggesting the locus (or loci) influencing height is not the same as those influencing the other two traits on this linkage group. However, the QTL peak for height to vegetative phase change (htpc) was located close (8.3 cM) to the QTL peak for *M. cryptica* severity (myco) on linkage group 1, with overlapping 95% confidence intervals (Table 4.3 and Figure 4.1). Co-location of QTL affecting different traits could be attributable to factors such as the direct effect of one trait on another, pleiotropy, or linkage of loci affecting both traits (Prioul *et al.* 2004). Support for the first two hypotheses was provided by co-localisation of QTL for (Prioul *et al.* 2004), and

an association between (Le May 2002), canopy architecture and the severity of *Mycosphaerella pinodes* in pea. Plants with denser canopies showed faster disease development by providing a more favourable microclimate for fungal development and disease spread (Le May 2002). Similarly, evidence for pleiotropy, or the direct effects of phase change on *M. cryptica* severity (or of disease on phase change), was also provided by the positive correlation detected between the severity of MLB and height to phase change in *E. globulus* (Dungey *et al.* 1997), when the disease was assessed after some trees had undergone phase change and others had not. In this case a mechanism to explain the correlation was proposed, whereby trees which undergo phase change later are more susceptible to MLB as a result of the greater opportunity for disease increase within the larger juvenile crown (Dungey *et al.* 1997). However, in the present study, no correlation between phase change and *M. cryptica* severity was observed, which is to be expected, since the disease was scored (at one year of age) before phase change had occurred. Thus in the present case, the co-location of QTL is probably due to linkage, since (i) there is no correlation between the traits, and (ii) the alleles affecting each trait segregate from different parents.

QTL homology

Despite the publication of several QTL studies within *Eucalyptus*, the lack of transferability of the markers commonly employed, such as RAPD and AFLP, has limited the comparison of QTL information and validation in independent pedigrees (Grattapaglia 2000). The inclusion of numerous SSR in this study allowed comparison of homology between linkage groups (see chapter 2 of this thesis) and the general positions of QTL with other studies within the subgenus *Symphyomyrtus* (Table 4.5). For example, a QTL for bud abundance was located close to Emb12 in an intra-provenance cross of *E. globulus* (Bundock 2003), very close to the estimated location of the QTL for another developmental trait, height to phase change, on the equivalent linkage group in this study. Again, this broad co-location could be indicative of a locus with pleiotropic control on both traits, linkage of loci affecting both traits, or the direct effect of one trait on the other. However, past studies have demonstrated that the timing

of flowering and vegetative phase change are genetically independent in *Eucalyptus* (Ipinza *et al.* 1994; Wiltshire *et al.* 1998; Jordan *et al.* 1999) and other genera (e.g. maize; Abedon *et al.* 1996). Hence the hypothesis that these traits directly affect each other can be rejected. The only QTL detected for flowering time in *Eucalyptus* was a major effect QTL in *Eucalyptus grandis* (Missiaggia *et al.* 2005). However, over 80 genes have been implicated in the regulation of flowering time in *Arabidopsis* (Simpson *et al.* 1999) and sequence homologues for 21 of these have been located in *Eucalyptus* (Dornelas and Rodriguez 2005). Hence, many additional genes are likely to be influential in the regulation of flowering time in eucalypts. The suggestion of control by numerous genes does not necessarily contradict the hypothesis of oligogenic control for height to phase change mentioned previously, as even under oligogenic control there are often many genes with lesser effect contributing to variation in a phenotypic trait. Hence, the apparent co-location of QTL for the timing of flowering and vegetative phase change is likely to represent a genomic region with a small effect on each trait. It is likely that many other loci are influential, most of which are independent of each other, which would explain the genetic independence of flowering and vegetative phase change suggested by the aforementioned quantitative studies. Mapping additional SSR near the QTL region in each study would be one way to differentiate between the alternative hypothesis of pleiotropy or linkage.

Growth traits can also be compared between studies, as measured by the diameter in *E. globulus* (dbh2 and dbh6; this study and Bundock 2003, respectively), or circumference of the stem at breast height in *E. grandis* (cbh) (Grattapaglia *et al.* 1996). Linkage group 5 (or the equivalent), contains QTL for growth related traits in the present study (at the suggestive level), Bundock (2003) and Grattapaglia *et al.* (1996). However, in the case of Grattapaglia *et al.* (1996), the QTL (for cbh) is located at the opposite end of the chromosome to the comparable QTL in the present study and is therefore likely to indicate a different QTL. In the case of Bundock (2003), there are too few SSR in common on the linkage group to precisely determine homology. However, the peak LOD for the QTL in this study was located at a SSR (Emb37), which (if polymorphic)

could be screened in the mapping progeny used by Bundock (2003) to verify the QTL position.

QTL stability

Attempts to use QTL for selection or deployment will rely on QTL stability, in different pedigrees, across sites and in different developmental stages (Sewell and Neale 2000). Although QTL for disease resistance often differ with plant ontogeny and environment, in some instances stable QTL have been located. For example, Prioul *et al.* (2004) found that while some QTL for *Mycosphaerella pinodes* in pea varied with plant ontogeny and environment, a major QTL was located in both the seedling stage in a growth chamber and the adult stage in field conditions. Similarly, in conifers, a single gene has been located for resistance to western gall rust (Van Der Kamp 1991) and resistance was stable across environments (Wu and Ying 1998). In the case of *M. cryptica* resistance, the prospects for finding stable QTL are positive, as temporal (Carnegie et *al.* 1994) and environmental (Reinoso 1992) stability of disease expression has previously been demonstrated.

A lack of temporal stability of QTL expression for growth traits has been reported in forest trees including hybrid *Populus* (Bradshaw and Stettler 1995) and *Pinus* (Kaya *et al.* 1999), consistent with different abiotic and biotic factors, therefore different loci, affecting phenotypic variation in growth during different stages of development. Similarly, Verhaegen *et al.* (1997) investigated stability of QTL for growth and wood density over a three-year period in a hybrid *E. grandis* x *E. urophylla* cross. None of the QTL were significant for all three times of measurement, but 68% were significant at two ages. Similar results were found in this study, in that the QTL for height found at age two and three years, was different to that for the highly correlated measures of diameter at breast height at age five and six (Pearsons correlation coefficient = 0.90 and 0.87, respectively). However, this comparison is complicated by the different measurements of growth, as the QTL detected for height at age two is different to that for diameter at breast height at the same age, suggesting slightly different genetic control of these two measures of growth.

Conclusion

This study reports QTL for the severity of damage by *M. cryptica*, and potentially correlated traits, including the severity of damage by the autumn gum moth, height to vegetative phase change and growth traits. This is the first report of QTL for the severity of damage by *M. cryptica* and for vegetative phase change in *Eucalyptus* and only the second report of QTL for disease resistance in the genus. The discovery of two major QTL explaining a large proportion of the phenotypic variance in *M. cryptica* severity is consistent with past findings in *Eucalyptus* and other forest trees that suggest few major loci often explain a large proportion of the phenotypic variance in numerous quantitative traits. This finding may be indicative of oligogenic control for the severity of damage by *M. cryptica*, in agreement with the specialised associations that often exist between hemibiotrophic pathogens and their host, and the host specificity of different biotypes of *M. cryptica* previously demonstrated within *E. globulus*. One of the QTL for the severity of damage by *M. cryptica* co-located with a QTL for height to vegetative phase change, for which linkage is the most likely explanation. The inclusion of numerous SSR, that have been used in previous QTL analysis, in the present study also allowed a comparison of QTL homology between other studies conducted in the subgenus. The QTL for height to phase change in the present study co-localised with a QTL for another developmental trait, bud abundance, in an independent cross of *E. globulus*. However, two possible hypotheses to explain this co-location; pleiotropy or linkage cannot be separated with the available evidence. This study will contribute to a fundamental understanding of the genetic control of the traits investigated, while the inclusion of fully informative SSR will allow validation in independent pedigrees which may ultimately enable their use in marker-assisted selection.

Chapter 5: General conclusion

The Portuguese Landrace of *E. globulus* forms an integral part of breeding programs in Portugal. Until now, the precise Australian origins of this population were unclear. Using chloroplast and nuclear DNA markers, I have shown that the origins of the Portuguese Landrace of *E. globulus* lie in south-eastern Tasmania with a lesser contribution from south-eastern Victoria. Previous studies based on morphology or single marker systems have also indicated south-eastern Tasmania was a likely area of origin for the Portuguese Landrace, but the contribution of south-eastern Victoria had not been suggested before this study. The relatively high level of genetic diversity in the Portuguese Landrace compared to that in natural populations and clear evidence for the genetic contribution by races from two widely separated regions was not consistent with the previous suggestion that the Portuguese Landrace was derived from a very narrow original collection. The information provided by this study will be important for avoiding inbreeding within Portuguese breeding programs, in addition to highlighting the genetic resources captured by the Portuguese Landrace and those that remain untapped. For example, the infusion of germplasm from areas currently favoured by *E. globulus* breeders, such as the Strezlecki and Otway Ranges, would be useful for breeding programs incorporating the Portuguese Landrace. This study has demonstrated the strength of using independent marker systems, particularly uniparentally inherited cpDNA and biparentally inherited nuclear SSR markers, for studying the geographic origins and genetic diversity of germplasm used in breeding populations in comparison to native stands.

Genetic linkage maps provide important information regarding the structure, organisation and evolution of plant genomes, in addition to providing the basis for QTL location. Parental and consensus maps in an F_2, inter-provenance cross of *E. globulus* were produced using AFLP and SSR markers. The AFLP technique was effective in allowing the rapid generation of numerous polymorphic markers, providing a framework on which to place the more informative SSR markers. Linkage maps and QTL studies performed within *Eucalyptus* to date have included few transferable markers in

common, thereby limiting their application beyond the pedigrees employed for map construction. In the present study, the inclusion of numerous SSR markers derived from a variety of sources allowed the comparison of linkage and QTL information to numerous other studies within the subgenus. A high degree of synteny and collinearity with maps produced previously in *E. globulus*, *E. grandis* and *E. urophylla* was demonstrated. As a result, the maps produced in this study will contribute to the construction of linkage maps for the subgenus *Symphyomyrtus* that will provide a basis for QTL validation in different pedigrees and species. Validating the effects of different loci in different genetic backgrounds and environments, will be critical to the effective use of QTL for MAS.

Fungal pathogens from the genus *Mycosphaerella* can seriously affect the productivity *E. globulus* plantations world-wide. My discovery of two major QTL influencing the severity of damage by *Mycosphaerella* is a significant step forward in understanding the control of this disease. The data suggest that *Mycosphaerella* disease severity is under oligogenic control, a factor that may facilitate the incorporation of disease resistance traits into breeding programs. The results presented in this thesis are particularly significant because no other QTL for disease severity has ever been reported for *E. globulus*, and only one other has been reported for other *Eucalyptus* species (Junghans *et al.* 2003).

Future studies

The findings of this study highlight many potential areas for future research. In the case of the Portuguese Landrace study, the immense value of our databases of molecular variation in native stands of *E. globulus* for fingerprinting purposes was demonstrated. These data will be useful for fingerprinting unpedigreed material in breeding populations and for seed certification, both in Australia and internationally. For example, landraces of *E. globulus* have been established in many other countries, including Chile, India and Spain, often from seed of unknown geographic origin (Eldridge *et al.* 1993). Breeding programs in these countries are currently infusing new base population material with landrace selections (e.g. Chile; Griffin 2001). A knowledge of the native origin of

landrace material will be useful to avoid inbreeding and guide the infusion process by identifying germplasm from native populations that is lacking in the landraces.

The linkage and QTL studies also identified many potential areas for future research. The linkage map produced in this study will provide a solid framework for mapping additional markers. It would be particularly beneficial to place additional SSR markers in common with other studies, with the aim of identifying homology between all linkage groups in the maps produced in *E. globulus* by Thamarus *et al.* (2002) and Bundock *et al.* (2001), and further investigating synteny and collinearity with the map produced in *E. grandis* and *E. urophylla* by Brondani *et al.* (2002). Placement of candidate genes (e.g. for wood properties; Thamarus *et al.* 2002; Gion *et al.* 2000) - especially those that have been found to co-locate with QTL - onto the maps created in this study will facilitate planned attempts to locate QTL for wood properties and other commercially important traits.

Validating the QTL for the severity of damage by *M. cryptica* in other pedigrees and environments will be a priority for future research. Following validation, the linkage map would provide a robust framework for more precise QTL location, by methods such as association mapping or candidate gene screening, ultimately allowing their use for MAS. The linkage map produced in this study will also provide the basis for future QTL research. For example, it could be used in an attempt to locate QTL for some of the morphological mechanisms that influence *Mycosphaerella* severity, such as the degree of stomatal wax coverage and leaf density (Smith *et al.* 2005). Co-location of the QTL in this study, with QTL for these morphological traits would provide further evidence that these mechanisms influence the severity of damage caused by *M. cryptica*. Other traits that are being used for QTL analysis in this pedigree include wood properties and leaf chemistry (assayed by near infrared spectroscopy).

This thesis has demonstrated the diverse roles that molecular markers can play in tree breeding programs. Native eucalypt forests are unlikely to adequately meet increasing world-wide demands for timber and pulp products. Instead, plantation forestry using

improved genotypes will be required. With the decreasing cost of high throughput molecular technologies and the development of increasingly sophisticated techniques for genetic analysis, it is likely that molecular marker based technologies will in future form an essential part of all elite breeding programs.

References

Abbo S, Lev-Yadun S, Ladizinsky G (2001) Tracing the wild genetic stocks of crop plants. *Genome* **44**, 309-310.

Abedon BG, Revilla P, Tracy WF (1996) Vegetative phase change in sweet corn populations: genetics and relationships with agronomic traits. *Maydica* **41**, 77-82.

Adams WT (1983) Applications of isozymes in tree breeding. pp. 381-400 in 'Isozymes in plant genetics and breeding', Part A. (Eds Tanksley SD, Orton TJ) (Elsevier Science Publishers B.V. Amsterdam).

Adams WT, Hipkins VD, Burczyk J, Randall WK (1997) Pollen contamination trends in a maturing Douglas-fir seed orchard. *Canadian Journal of Forest Research* **27**, 131-134.

Adams WT, Neale DB, Loopstra CA (1988) Verifying controlled crosses in conifer tree-improvement programs. *Silvae Genetica* **37**, 147-152.

Agrama HA, George TL, Salah SF (2002) Construction of genome map for *Eucalyptus camaldulensis* DEHN. *Silvae Genetica* **51**, 201-206.

Ahuja MR (2001) Recent advances in molecular genetics of forest trees. *Euphytica* **121**, 173-195.

Almeida MH (1993) Estudo da variabilidade geográfica em *Eucalyptus globulus* Labill. Tese de Doutoramento ISA University Técnica de Lisboa.

Almeida MH, Pereira H, Miranda I, Tomé M (1995) Provenance trials of *Eucalyptus globulus* Labill. in Portugal. pp 195-198 In 'Eucalypt plantations: Improving fibre yield and quality'. (Eds. BM Potts, NMG Borralho, JB Reid, RN Cromer, WN Tibbits and CA Raymond). Proceedings CRC-IUFRO Conference, Hobart, 19-24 February 1995.

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

Araújo JA, Lemos L, Ramos A, Ferreira JG, Borralho NMG (1997) The RAIZ *Eucalyptus globulus* breeding program: A BLUP rolling-front strategy with a mixed clonal and seedling deployment scheme. pp. 371-376 In Proceedings of the IUFRO Conference on Silviculture and Improvements of Eucalypts. Salvador, Brazil.

Asins MJ (2002) Present and future of quantitative trait locus analysis in plant breeding. *Plant Breeding* **121**, 281-291.

Astorga R, Soria F, Bascuro F, Toval G (2004) Diversity analysis and genetic structure of *Eucalyptus globulus* Labill. In '*Eucalyptus* in a changing world' (Eds. NMG Barralho, JS Pereira, C. Marques, J. Coutinho, M. Madeira, M. Tomé) Proceedings IUFRO Conference, Aveiro, Portugal, 11-15 October 2004.

Avise JC (1994) Molecular markers, natural history and evolution. (Chapman and Hall International, New York).

Beavis WD (1998) QTL analysis: power, precision, and accuracy. In 'Molecular dissection of complex traits.' (Ed Patterson HA) pp. 145-162. (CRC Press: Boca Raton, Florida).

Beebe S, Skroch PW, Tohme J, Duque MC, Pedraza F, Nienhuis J (2000) Structure of genetic diversity among common bean landraces of middle American origin based on correspondance analysis of RAPD. *Crop Science* **40**, 264-273.

Bell JC, Powell M, Devey ME, Moran GF (2004) DNA profiling, pedigree lineage analysis and monitoring in the Australian breeding program of radiata pine. *Silvae Genetica* **53**, 130-134.

Boerjan W (2005) Biotechnology and the domestication of forest trees. *Current Opinion in Biotechnology* **16**, 159-166.

Borralho NMG, Cotterill PP, Kanowski PJ (1993) Breeding objectives for pulp production of *Eucalyptus globulus* under different industrial cost structures. *Canadian Journal of Forest Research* **23**, 648-656.

Bradshaw HD, Stettler RF (1994) Molecular genetics of growth and development in *Populus* II. Segregation distortion due to genetic load. *Theoretical and Applied Genetics* **89**, 551-558.

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

Bradshaw HD, Villar M, Watson BD, Otto KG, Stewart S, Stettler RF (1994) Molecular-genetics of growth and development in *Populus*. 3. A genetic-linkage map of a hybrid poplar composed of RFLP, STS, and RAPD markers. *Theoretical and Applied Genetics* **89**, 167-178.

Brondani RPV (2001) Development, characterization and mapping of microsatellite markers in *Eucalyptus*. PhD thesis, University of Brasília, Brazil.

Brondani RPV, Brondani C, Grattapaglia D (2002) Towards a genus-wide reference linkage map for *Eucalyptus* based exclusively on highly informative microsatellite markers. *Molecular Genetics and Genomics* **267**, 338-347.

Brondani RPV, Brondani C, Tarchini R, Grattapaglia D (1998) Development, characterization and mapping of microsatellite markers in *Eucalyptus grandis* and *Eucalyptus urophylla*. *Theoretical and Applied Genetics* **97**, 816-827.

Brooker MIH (2000) A new classification of the genus *Eucalyptus* L'Hér. (Myrtaceae). *Australian Systematic Botany* **13**, 79-148.

Brown GR, Bassoni DL, Gill, GP, Fontana JR, Wheeler NC, Megraw RA, Davis MF, Sewell MM, Tuskan GA, Neale DB (2003) Identification of quantitative trait loci influencing wood property traits in loblolly pine (*Pinus taeda* L.). III. QTL verification and candidate gene mapping. *Genetics* **164**, 1537-1546.

Brown GR, Gill GP, Kuntz RJ, Langley CH, Neale DB (2004) Nucleotide diversity and linkage disequilibrium in loblolly pine. *Proceedings of the National Academy of Sciences' USA* **101**, 15255-15260.

Buckler ES, Phelps-Durr TL, Keith Buckler CS, Kelly Dawe R, Doebley JF, Holtsford TP (1999) Meiotic drive of chromosomal knobs reshaped the maize genome. *Genetics* **153**, 415-426.

Bundock P (2003) Genome mapping of *Eucalyptus globulus*. PhD Thesis, University of Tasmania, Australia.

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

Burczyk J, Adams WT, Moran GF, Griffin AR (2002) Complex patterns of mating revealed in a *Eucalyptus regnans* seed orchard using allozyme markers and the neighbourhood model. *Molecular Ecology* **11**, 2379-2391.

Burdon RD (2001) Genetic diversity and disease resistance: some considerations for research, breeding, and deployment. *Canadian Journal of Forest Research* **31**, 596-606.

Burgess T, Wingfield MJ (2002) Impact of fungal pathogens in natural forest ecosystems: a focus on *Eucalyptus*. pp. 285-306 In 'Microorganisms in plant conservation and biodiversity'. (Eds. Sivasithamparam K, Dixon KW, Barrett RL) (Kluwer Academic Publishers).

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

Butcher PA, Glaubitz JC, Moran GF (1999) Applications for microsatellite markers in the domestication and conservation of forest trees. *Forest Genetic Resources* **27**, 3442-3454.

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

Byrne M, Marques-Garcia MI, Uren T, Smith, DS, Moran GF (1996) Conservation and genetic diversity of microsatellite loci in the genus *Eucalyptus*. *Australian Journal of Botany* **44**, 331-341.

Byrne M, Moran GF, Tibbits WN (1993) Restriction map and maternal inheritance of chloroplast DNA in *Eucalyptus nitens*. *Journal of Hereditary* **84**, 218-220.

Byrne M, Murrell JC, Allen B, Moran GF (1995) An integrated genetic linkage map for eucalypts using RFLP, RAPD and isozyme markers. *Theoretical and Applied Genetics* **91**, 869-875.

Byrne M, Murrell JC, Owens JV, Williams ER, Moran GF (1997) Mapping of quantitative trait loci influencing frost tolerance in *Eucalyptus nitens*. *Theoretical and Applied Genetics* **95**, 975-979.

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

Butcher PA, Williams ER, Whitaker D, Ling S, Speed TP, Moran GF (2002) Improving linkage analysis in outcrossed forest trees – an example from *Acacia mangium*. *Theoretical and Applied Genetics* **104**, 1185-1191.

Campbell MM, Brunner AM, Jones HM, Strauss SH (2003) Forestry's fertile crescent: the application of biotechnology to forest trees. *Plant Biotechnology Journal* **1**, 141-154.

Carnegie AJ, Ades PK (2003) *Mycosphaerella* leaf disease reduces growth of plantation-grown *Eucalyptus globulus*. *Australian Forestry* **66**, 113-119.

Carnegie AJ, Ades PK, Ford R (2001) The use of RAPD-PCR analysis for the differentiation of *Mycosphaerella* species from *Eucalyptus* in Australia. *Mycological Research* **105**, 1313-1320.

Carnegie AJ, Ades PK, Keane PJ, Smith IW (1998) *Mycosphaerella* diseases of juvenile foliage in a eucalypt species and provenance trial in Victoria, Australia. *Australian Forestry* **61**, 190-194.

Carnegie AJ, Keane PJ, Ades PK, Smith IW (1994) Variation in susceptibility of *Eucalyptus globulus* provenances to *Mycosphaerella* leaf disease. *Canadian Journal of Forest Research* **24**, 1751-1757.

Castiglione S, Wang G, Damiani G, Bandi C, Bisoffi S, Sala F (1993) RAPD fingerprints for identification and for taxonomic studies of elite poplar (*Populus* spp.) clones. *Theoretical and Applied Genetics* **87**, 54-59.

Cervera MT, Plomion C, Malpica C (2000) Molecular markers and genome mapping in woody plants. pp. 375-394 In 'Molecular biology of woody plants' Volume 1. (Eds. Jain SM and Minocha SC) (Kluwer Academic Publishers: Netherlands).

Cervera MT, Storme V, Ivens B, Gusmão J, Liu BH, Hostyn V, Slycken JV, Montagu MV, Boerjan W (2001) Dense genetic linkage maps of three *Populus* species (*Populus deltoides*, *P. nigra* and *P. trichocarpa*) based on AFLP and microsatellite markers. *Genetics* **158**, 787-809.

Chaix G, Gerber S, Razafimaharo V, Vigneron P, Verhaegen D, Hamon S (2003) Gene flow estimation with microsatellites in a Malagasy seed orchard of *Eucalyptus grandis*. *Theoretical and Applied Genetics* **107**, 705-712.

Chu Z, Ouyang Y, Zhang J, Yang H, Wang S (2004) Genome-wide analysis of defense-responsive genes in bacterial blight resistance of rice mediated by the recessive R gene xa13. *Molecular Genetics and Genomics* **271**, 111-120.

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. *Genetics* **138**, 963-971.

Devey ME, Carson SD, Nolan MF, Matheson AC, Te Riini C, Hohepa J (2003) QTL associations for density and diameter in *Pinus radiata* and the potential for marker aided selection. *Theoretical and Applied Genetics* **108**, 516-524.

Devey ME, Delfino-Mix A, Kinloch BB, Neale DB (1995) Random Amplified polymorphic DNA markers tightly linked to a gene for resistance to white pine blister rust in sugar pine. *Proceedings of the National Academy of Sciences USA* **92**, 2066-2070.

Devey ME, Sewell MM, Uren TL, Neale DB (1999) Comparative mapping in loblolly and radiata pine using RFLP and microsatellite markers. *Theoretical and Applied Genetics* **99**, 656-662.

Dillner B, Ljunger A, Herud OA, Thune-Larson E (1971) The breeding of *Eucalyptus globulus* on the basis of wood density, chemical composition and growth rate. *Timber Bulletin Europe* 23, 120-151.

Diwan N, Cregan PB (1997) Automated sizing of fluorescent-labeled simple sequence repeat (SSR) markers to assay genetic variation in soybean. *Theoretical and Applied Genetics* **95**, 723-733.

Dore C, Dosba F, Baril C (2001) ISHS International Symposium on 'Molecular Markers for Characterizing Genotypes and Identifying Cultivars in Horticulture'. March 6-9, Montpellier France.

Dornelas Mc, Rodriguez APM (2005) Identifying *Eucalyptus* expressed sequence tags related to *Arabidopsis* flowering-time pathway genes. *Brazilian Journal of Plant Physiology* **17**, 255-266.

Doughty RW (2000) 'The *Eucalyptus*: A natural and commercial history of the gum tree.' (Johns Hopkins University Press: Baltimore).

Doyle JJ and Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus* **12**, 13-15.

Dumolin-Lapègue S, Demesure B, Fineschi S, Le Corre V, Petit RJ (1997) Phylogeographic structure of white oaks throughout the European continent. *Genetics* **146**, 1475-1487.

Dungey HS, Potts BM (2001) Susceptibility of some *Eucalyptus* species and their hybrids to possum damage. *Australian Forestry* **65**, 23-30.

Dungey HS, Potts BM, Carnegie AJ, Ades PK (1997) *Mycosphaerella* leaf disease: genetic variation in damage to *Eucalyptus nitens*, *Eucalyptus globulus*, and their F1 hybrid. *Canadian Journal of Forest Research* **27**, 750-759.

Dutkowski GW, Potts BM (1999) Geographic patterns of genetic variation in *Eucalyptus globulus* ssp. *globulus* and a revised racial classification. *Australian Journal of Botany* **47**, 237-263.

Edwards JH (1991) The reliability of locus orderings. *Annals of Human Genetics* **55**, 315-320.

Eglinton J, Coventry S, Chalmers K (2006) Breeding outcomes from molecular genetics. Breeding for success: diversity in action. 13th Australasian plant breeding conference, Christchurch, New Zealand, 18-21 April 2006.

Eldridge K, Davidson J, Harwood C, Van Wyk G (1993) 'Eucalypt domestication and breeding' (Oxford University Press: New York).

Erickson DL, Fenster CB, Stenøien HK, Price D (2004) Quantitative trait locus analyses and the study of evolutionary processes. *Molecular Ecology* **13**, 2505-2522.

Erpelding JE, Blake NK, Blake TK, Talbert LE (1996) Transfer of sequence tagged site PCR markers between wheat and barley. *Genome* **39**, 802-810.

Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics, 4th Edition. (Longman, Harlow, Essex).

Farrow RA, Floyd RB, Newmann FG (1994) Inter-provenance variation in resistance of *Eucalyptus globulus* juvenile foliage to insect feeding. *Australian Forestry* **57**, 65-68.

Flor HH (1971) Current status of the gene-for-gene concept. *Annual Review of Phytopathology* **9**, 275-296.

Freeman J, Jackson HD, Steane DA, McKinnon GE, Dutkowski GW, Potts BM and Vaillancourt RE (2001) Chloroplast DNA phylogeography of *Eucalyptus globulus* *Australian Journal of Botany* **49**, 831-835.

Fukunaga K, Hill J, Vigouroux Y, Matsuoka Y, Jesus Sanchez Y, Liu K, Buckler ES, Doebley J (2005) Genetic diversity and population structure of teosinte. *Genetics* **169**, 2241-2254.

Gaiotto FA, Barmucci M, Grattapaglia D (1997) Estimation of outcrossing rate in a breeding population of *Eucalyptus urophylla* s.t. Blake with dominant RAPD and AFLP markers. *Theoretical and Applied Genetics* **95**, 842-849.

Gardiner S, Bus V, Volz R, Bassett H (2006) Marker assisted selection in apple breeding internationally. 'Breeding for success: diversity in action'. 13th Australasian plant breeding conference, Christchurch, New Zealand, 18-21 April 2006.

Gardiner CA, Crawford DA (1987) '1987 Seed Collections of *Eucalyptus globulus* subsp. *globulus* for tree improvement purposes' (CSIRO Division of Forest Research: Canberra).

Gardiner CA, Crawford DA (1988) '1988 Seed Collections of *Eucalyptus globulus* subsp. *globulus* for tree improvement purposes' (CSIRO Division of Forest Research: Canberra).

Gebhardt C, Ritter E, Barone E, Debener T, Walkemeier B, Schachtschabel U, Kaufmann H, Thompson RD, Boneirbale MW, Ganal MW, Tanksley SD, Salamini F (1991) RFLP maps of potato and their alignment with the homologous tomato genome. *Theoretical and Applied Genetics* **83**, 49-57.

Gemas VJV, Neves LO, Araújo C, Fevereiro P (2004) How can ISSR markers be useful in the management of genetic resources in a *Eucalyptus globulus* ssp. *globulus* baseline collection. In '*Eucalyptus* in a changing world'. (Eds. NMG Barralho, JS Pereira, C. Marques, J. Coutinho, M. Madeira, M. Tomé) Proceedings IUFRO Conference, Aveiro, Portugal, 11-15 October 2004.

Gill GP, Brown GR, Neale DB (2003) A sequence mutation in the cinnamyl alcohol dehydrogenase gene associated with altered lignification in loblolly pine. *Plant Biotechnology Journal* **1**, 253-258.

Gillet E and Gregorius HR (1992) What can be inferred from open pollinated progenies about the source of observed segregation distortion? – A case study in *Castanea sativa* Mill. *Silvae Genetica* **41**, 82-87.

Gillies ACM, Cornelius JP, Newton AC, Navarro C, Hernández M, Wilson J (1997) Genetic variation in Costa Rican populations of the tropical timber species *Cedrela odorata* L., assessed using RAPDs. *Molecular Ecology* 6, 1133-1145.

Gillies ACM, Navarro C, Lowe AJ, Newton AC, Hernández M, Wilson J, Cornelius JP (1999) Genetic diversity in Mesoamerican populations of mahogany (*Swietenia macrophylla*), assessed using RAPDs. *Heredity* 83, 722-732.

Gion JM, Rech P, Grima Pettenati J, Verhaegen D, Plomion C (2000) Mapping candidate genes in *Eucalyptus* with emphasis on lignification genes. *Molecular Breeding* **6**, 441-449.

González-Martínez SC, Konstantin VK, Neale DB (2006) Forest-tree population genomics and adaptive evolution. *New Phytologist* **170**, 227-238.

Grattapaglia D (1997) Opportunities and challenges for the incorporation of genomic analysis in *Eucalyptus* breeding. pp. 129-136 in Silviculture and Improvement of Eucalypts, Proceedings IUFRO Conference, Salvador, Brazil.

Grattapaglia D (2000) Molecular breeding of Eucalyptus. In 'Molecular biology of woody plants' (Eds. Jain SM and Minocha SC) Volume 1, pp. 451-474 (Kluwer Academic Publishers: Netherlands).

Grattapaglia D (2004) Integrating genomics into *Eucalyptus* breeding. *Genetics and Molecular Research* **3**, 369-379.

Grattapaglia D, Alfenas A, Coelho A, Bearzoti E, Pappas G, Pasquali G, Pereira G, Colodette J, Gomide JL, Bueno J, Cascardo J, Brondani R, Brommonschenkel S (2004a) Building resources for molecular breeding of *Eucalyptus*: The Genolyptus project in Brazil. pp 30-36 In '*Eucalyptus* in a changing world'. (Eds. Borralho N, Periera JS, Marques C, Coutinho J, Madeira M, Tome M). Proceedings IUFRO Conference.RAIZ, Instituto Investigação de Floresta e Papel, Aveiro, Portugal.

Grattapaglia D, Bertolucci FL, Penchel R, Sederoff RR (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, Bertolucci FL, Sederoff RR (1995) Genetic mapping of QTLs controlling vegetative propagation in *Eucalyptus grandis* and *E. urophylla* using a pseudo-testcross strategy and RAPD markers. *Theoretical and Applied Genetics* **90**, 933-947.

Grattapaglia D, Ribeiro GD, Rezende SP (2004b) Retrospective selection of elite parent trees using paternity testing with microsatellite markers: an alternative short term breeding tactic for *Eucalyptus*. *Theoretical and Applied Genetics* 109, 192-199.

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.

Griffin AR (2001) Deployment decisions-capturing the benefits of tree improvement with clones and seedlings. 'In developing the eucalypt of the future' (Ed. Barros S) Proceedings IUFRO international symposium, Valdivia, Chile 2001.

Gupta PK, Varshney RK, Sharma PC, Ramesh B (1999) Molecular markers and their application in wheat breeding. *Plant Breeding* **118**, 369-390.

Hackett CA, Broadfoot LB (2003) Effects of genotyping errors, missing values and segregation distortion in molecular marker data on the construction of linkage maps. *Heredity* **90**, 33-38.

Hammond-Kosack KE, Jones JDG (1997) Plant disease resistance genes. *Annual Reviews of Plant Physiology and Plant Molecular Biology* **48**, 575-607.

Hansen MH, Nielson EE, Bekkevold D, Mensburg KD (2001) Admixture analysis and stocking impact assessment in brown trout (*Salmo trutta*), estimated with incomplete baseline data. *Canadian Journal of Fish Aquatic Science* **58**, 1853-1860.

Harju A, Muona O (1989) Background pollination in *Pinus sylvestris* L. seed orchards. *Scandinavian Journal of Forest Research* **4**, 513-520.

Harter AV, Gardner KA, Falush D, Lentz DL, Bye RA, Rieseburg LH (2004) Origin of extant domesticated sunflowers in eastern North America. *Nature* **430**, 210-205.

Heinze B, Westcott R, Schmidt J, Glössl J (1996) Application of random amplified polymorphic DNA (RAPD) to detect genetic variation in Norway spruce. *New Forests* **11**, 173-184.

Helentjaris T, King G, Slocum M, Sidenstrang C, Wegman S (1985) Restriction fragment polymorphism as probes for plant diversity and their development as tools for applied plant breeding. *Plant Molecular Biology* **5**, 109-118.

Hirschhorn JN, Daly MJ (2005) Genome-wide association studies for common diseases and complex traits. *Nature Reviews Genetics* **6**, 95-108.

House APN, Bell JC (1994) Isozyme variation and mating system in *Eucalyptus urophylla* S. T. Blake. *Silvae Genetica* **43**, 167-176.

Ingvarsson PK (2005) Nucleotide polymorphism and linkage disequilibrium within and among natural populations of European aspen (*Populus tremula* L., Salicaceae). *Genetics* **169**, 945-953.

Ipinza RH, Garcia X, Apiolaza L, Paz Molina M, Chung P, Parra P (1994) Variacion juvenil de un ensayo de procedencias y familias de *Eucalyptus globulus* subsp. *globulus* labill, en la septima region, Chile. *Ecologia* **8**, 259-270.

IUPOV (1981) Guidelines for the conduct of tests of distinctness, homogeneity and stability – *Populus* L. International Union for the Protection of New Varieties of Plants, Geneva, Switzerland.

Jain SM, Minocha SC (eds) (2000) 'Molecular biology of woody plants' volume 1. (Kluwer Academic Publishers: Netherlands).

James SA, Bell DT (2001) Leaf morphological and anatomical characteristics of heteroblastic *Eucalyptus globulus* ssp. *globulus* (Myrtaceae). *Australian Journal of Botany* **49**, 259-269.

Jansen RC (1993) Interval mapping of multiple quantitative trait loci. *Genetics* **135**, 205-211.

Jansen RC (1994) Controlling the type I and type II errors in mapping quantitative trait loci. *Genetics* 138, 871-881.

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

Jones ME, Shepherd M, Henry RJ, Delves A (2006a) Chloroplast DNA variation and population structure in the widespread forest tree, *Eucalyptus grandis*. *Conservation Genetics* (DOI 10.1007/s10592-005-9104-7).

Jones ME, Stokoe RL, Cross MJ, Scott LJ, Maguire TL, Shepherd M (2001) Isolation of microsatellite loci from spotted gum (*Corymbia variegata*), and cross-species amplification in *Corymbia* and *Eucalyptus*. *Molecular Ecology Notes* **1**, 276-278.

Jones RC, Steane DA, Potts BM, Vaillancourt RE (2002a) Microsatellite and morphological analysis of *Eucalyptus globulus* populations. *Canadian Journal of Forestry Research* 32, 59-66.

Jones TH, Potts BM, Vaillancourt RE, Davies NW (2002b) Genetic resistance of *Eucalyptus globulus* to autumn gum moth defoliation and the role of cuticular waxes. *Canadian Journal of Forest Research* 32, 1961-1969.

Jones TH, Steane DA, Jones RC, Pilbeam D, Vaillancourt RE, Potts BM. (2006b). Effects of domestication on genetic diversity in *Eucalyptus globulus*. *Forest Ecology and Management* (in press).

Jordan GJ, Potts BM, Chalmers P, Wiltshire RJE (2000) Quantitative genetic evidence that the timing of vegetative phase change in *Eucalyptus globulus* ssp. *globulus* is an adaptive trait. *Australian Journal of Botany* **48**, 561 – 567.

Jordan GJ, Potts BM, Clarke AR (2002) Susceptibility of *Eucalyptus globulus* ssp. *globulus* to sawfly (*Perga affinis* ssp. *insularis*) attack and its potential impact on plantation productivity. *Forest Ecology and Management* **160**, 189-199.

Jordan GJ, Potts BM, Kirkpatrick JB, Gardiner C (1993) Variation in the *Eucalyptus globulus* complex revisited. *Australian Journal of Botany* **41**, 763-785.

Jordan GJ, Potts BM, Wiltshire RJE (1999) Strong, independent, quantitative genetic control of the timing of vegetative phase change and first flowering in *Eucalyptus globulus* ssp. *globulus* (Tasmanian Blue Gum). *Heredity* **83**, 179-187.

Jorge V, Dowkiw A, Faivre-Rampant P, Bastien C (2005) Genetic architecture of qualitative and quantitative *Melampsora larici-populina* leaf rust resistance in hybrid poplar: genetic mapping and QTL detection. *New Phytologist* **167**, 113-127.

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. *Theoretical and Applied Genetics* **108**, 175-180.

Kaya Z, Sewell MM, Neale DB (1999) Identification of quantitative trait loci influencing annual height- and diameter-increment growth in loblolly pine (*Pinus taeda* L.) *Theoretical and Applied Genetics* **98**, 586-592.

Keane PJ, Kile GA, Podger FD, Brown BN (Eds) (2000) Diseases and pathogens of eucalypts. (CSIRO: Melbourne).

Kearsey MJ, Farquhar AGL (1998) QTL analysis in plants; where are we now? *Heredity* **80**, 137-142.

Kearsey MJ, Pooni HS (1996) 'The genetical analysis of quantitative traits.' (Chapman and Hall: London).

Keats BJB, Sherman SL, Morton NE, Robson EB, Buetow KH, Cartwright PE, Chakravarti A, Francke U, Green PP, Ott J (1991) Guidelines for human linkage maps – an International System for Human Linkage Maps (ISLM 1990). *Annals Human Genetics* **55**, 1-6.

Keil M, Griffin AR (1994) Use of random amplified polymorphic DNA (RAPD) markers in the discrimination and verification of genotypes in *Eucalyptus*. *Theoretical and Applied Genetics* **89**, 442-450.

Kesseli R, Paran I, Michelmore RW (1994) Analysis of a detailed genetic linkage map of *Lactuca sativa* (lettuce) constructed from RFLP and RAPD markers. *Genetics* **136**, 1435-1446.

Kirst M, Cordeiro CM, Rezende GDSP, Grattapaglia D (2005) Power of microsatellite markers for fingerprinting and parentage analysis in *Eucalyptus grandis* breeding populations. *Journal of Heredity* **96**, 161-166.

Kirst M, Myburg AA, De Léon JPG, 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 Physiology* **135**, 2368-2378.

Labillardière JJHd (1799) 'Relation du voyage à la recherche de la Pérouse.' (Jensen: Paris).

Lande R, Thompson R (1990) Efficiancy of marker-assisted selection in the improvement of quantitative traits. *Genetics* **124**, 743-756.

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

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) Mapmaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* **1**, 174-181

Lander E, Kruglyak L (1995) Genetic dissection of quantitative traits: guidelines for interpreting and reporting linkage results. *Nature Genetics* **11**, 241-247.

Le May (2002) Effet de la structure du couvert végétal du pois protéagineux sur le développement spatio-temporel de l'anthracnose à *Mycospharela pinodes*. Consequences sur l'élaboration du rendement. PhD thesis, ENSA de Rennes, Rennes, France.

Lefèvre F, Goué-Mourier MC, Faivre-Rampant P, Villar M (1998) A single resistance gene cluster controls incompatibility and partial resistance to various *Melampsora larici-populina* races in hybrid poplar. *Phytopathology* **88**, 156-163.

Leonardi S, Menozzi P (1996) Spatial structure of genetic variability in natural stands of *Fagus sylvatica* L. (beech) in Italy. *Heredity* **77**, 359-368.

Lev-Yadun S, Gopher A, Abbo S (2000) The cradle of agriculture. *Science* **288**, 1602-1603.

Lezar S, Myburg AA, Berger DK, Wingfield MJ, Wingfield BD (2004) Assessment of microarray-based DNA fingerprinting in *Eucalyptus grandis*. *Theoretical and Applied Genetics* **109**, 1329 -1336.

Li H, Madden JL, Potts BM (1995) Variation in volatile leaf oils of the Tasmanian *Eucalyptus* species. I. Subgenus Monocalyptus. *Biochemical Systematics and Ecology* **23**, 299-318.

Li H, Madden JL, Potts BM (1996) Variation in volatile leaf oils of the Tasmanian *Eucalyptus* species. II. Subgenus *Symphyomyrtus*. *Biochemical Systematics and Ecology* **24**, 547-569.

Lopez GA, Potts BM, Dutkowski GW, Apiolaza LA, Gelid P (2002) Genetic variation and inter-trait correlations in *Eucalyptus globulus* base population trials in Argentina. *Forest Genetics* **9**, 223-237.

Lopez GA, Potts BM, Dutkowski GW, Rodriguez Traverso JM (2001) Quantitative genetics of *Eucalyptus globulus*: affinities of landrace and native stand localities. *Silvae Genetica* **50**, 5-6.

Lorz H, Wendel G (eds) (2005) 'Molecular marker systems in plant breeding and crop improvement' (Springer-Verlag: Heidelberg).

Loveless MD, Hamrick JL (1984) Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics* **15**, 65-95.

Lynch M, Walsh B (1998) 'Genetics and analysis of quantitative traits.' (Sinauer Associates Inc: Sunderland, MA, U.S.A.).

Macdonald MW, Rawlings M, Butcher PA, Bell JC (2003) Regional divergence and inbreeding in *Eucalyptus cladocalyx* (Myrtaceae). *Australian Journal of Botany* **51**, 393-403.

Maliepaard C, Alston FH, van Arkel G, Brown LM, Chevreau E, Dunemann F, Evans KM, Gardiner S, Guilford P, van Heusden AW, Janse J, Laurens F, Lynn JR, Manganaris AG, den Nijs APM, Periam N, Rikkerink E, Roche P, Ryder C, Sansavini S, Schmidt H, Tartarini S, Verhaegh JJ, Vrielink-van Ginkel M, King GJ (1998) Aligning male and female linkage maps of apple (*Malus pumila* Mill.) using multi-allelic markers. *Theoretical and Applied Genetics* **97**, 60-73.

Marchini J, Cardon LR, Phillips MS, Donnelly P (2004) The effects of human population structure on large genetic association studies. *Nature Genetics* **36**, 512-517.

Marques CM, Araujo JA, Ferriera JG, Whetten R, Omalley DM, Liu BH, Sederoff R (1998) AFLP genetic maps of *Eucalyptus globulus* and *E. tereticornis*. *Theoretical and Applied Genetics* **96**, 727-737.

Marques CM, Brondani RPV, Grattapaglia D, Sederoff R (2002) Conservation and synteny of SSR loci and QTLs for vegetative propagation in four *Eucalyptus* species. *Theoretical and Applied Genetics* **105**, 474-478.

Marques CM, Vasquez-Kool J, Carocha VJ, Ferreira JG, O'Malley DM, Liu BH, Sederoff RR (1999) Genetic dissection of vegetative propagation traits in *Eucalyptus tereticornis* and *E. globulus*. *Theoretical and Applied Genetics* **99**, 936-946.

McKinnon GE, Potts BM, Steane DA, Vaillancourt RE (2005) Population and phylogenetic analysis of the cinnamoyl coA reductase gene in *Eucalyptus globulus* (Myrtaceae). *Australian Journal of Botany* **53**, 827-838.

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

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*. *Molecular Ecology* **13**, 3751-3762.

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

Milgate AW, Potts BM, Joyce K, Mohammed C, Vaillancourt RE (2005a) Genetic variation in *Eucalyptus globulus* for susceptibility to *Mycosphaerella nubilosa* and its association with tree growth. *Australasian Plant Pathology* **34**, 11-18.

Milgate AW, Vaillancourt RE, Mohammed C, Powell M, Potts BM (2005b) Genetic structure of a *Mycosphaerella cryptica* population. *Australasian Plant Pathology* **34**, 345-354.

Miranda I, Almeida MH, Pereira H (2001) Variation of fibre biometry in different provenances of *Eucalyptus globulus* Labill. *Appita Journal* **54**, 272-275.

Missiaggia AW, Piacezzi AL, Grattapaglia D (2005) Genetic mapping of *Eef*1, a major effect QTL for early flowering in *Eucalyptus grandis*. *Tree Genetics and Genomes* **1**, 79-84.

Mohammed C, Wardlaw T, Smith A, Pinkard E, Battaglia M, Glen M, Tommerup I, Potts BM, Vaillancourt RE (2003) *Mycosphaerella* leaf diseases of temperate eucalypts around the Southern Pacific Rim. *New Zealand Journal of Forestry Science* **33**, 362-372.

Mondragon-Palomino M, Gaut BS (2005) Gene conversion and the evolution of three leucine-rich repeat gene families in *Aribidopsis thaliana*. *Molecular Biology and Evolution* **22**, 2444-2456.

Moran GF (1992) Patterns of genetic diversity in Australian tree species. *New Forest* **6**, 49-66.

Moran GF, Bell JC (1987) The origin and genetic diversity of *Pinus radiata* in Australia. *Theoretical and Applied Genetics* **73**, 616-622.

Moran GF, Butcher PA, Glaubitz JC (2000) Application of genetic markers in domestication, conservation, and utilisation of genetic resources of Australasian tree species. *Australian Journal of Botany* **48**, 313-320.

Myburg AA, Griffin AR, Sederoff RR, Whetten RW (2003) Comparative genetic linkage maps of *Eucalyptus grandis, Eucalyptus globulus* and their F1 hybrid based on a double pseudo-backcross mapping approach. *Theoretical and Applied Genetics* **107**, 1028-1042.

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. *BioTechniques* **30**, 348-357.

Neale DB, Devey ME, Jermstad KD, Ahuja MR, Alosi MC, Marshall KA (1992) Use of DNA markers in forest tree improvement research. *New Forests* **6**, 391-407.

Neale DB, Savolainen O (2004) Association genetics of complex traits in conifers. *Trends in Plant Science* **9**, 325-330.

Nehra NS, Becwar MR, Rottmann WH, Pearson L, Chowdhury K, Change S, Wilde HD, Kodrzycki RJ, Zhang C, Gause KC, Parks DW, Hinchee MA (2005) Forest biotechnology: innovative methods, emerging opportunities. *In Vitro Cellular and Developmental Biology-Plant* **41**, 701-717.

Nesbitt KA, Potts BM, Vaillancourt RE, West AK, Reid JB (1995) Partitioning and distribution of RAPD variation in a forest tree species, *Eucalyptus globulus* (Myrtaceae). *Heredity* **74**, 628-637.

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.

O'Reilly-Wapstra J, McArthur C, Potts BM (2002) Genetic variation in resistance of *Eucalyptus globulus* to marsupial browsers. *Oecologia* **130**, 289-296.

Orme RK (1977) *Eucalyptus globulus* provenances. In Proceedings of the Third World Consultation on Forest Tree Breeding. pp. 207-222 CSIRO, Canberra, Australia 1977.

Ozkan H, Brandolini A, Pozzi C, Effgen S, Wunder J, Salamini F (2005) A reconsideration of the domestication geography of tetraploid wheats. *Theoretical and Applied Genetics* **110**, 1052-1060.

Park RF, Keane PJ, Wingfield MJ, Crous PW (2000) Fungal diseases of *Eucalyptus* foliage. pp. 153-240 In 'Diseases and pathogens of eucalypts' (Eds. Keane PJ, Kile GK, Podger FD, Brown BN) (CSIRO: Melbourne).

Patterson B, Gore P, Potts BM, Vaillancourt RE, (2004a) Advances in pollination techniques for large-scale seed production in *Eucalyptus globulus*. *Australian Journal of Botany* **52**, 781-788.

Patterson B, Vaillancourt RE, Pilbeam DJ, Potts BM (2004b) Factors affecting variation in outcrossing rate in *Eucalyptus globulus*. *Australian Journal of Botany* **52**, 773-780.

Patterson B, Wolbang CM, Vaillancourt RE, Potts BM (2000) Inheritance of two chlorophyll mutants in *Eucalyptus globulus*. *Silvae Genetica* **49**, 290-291.

Pederick LA (1979) Natural variation in shining gum (*Eucalyptus nitens*). *Australian Forest Research* **9**, 41-63.

Plomion C, Bahrman N, Costa P, Dubos C, Frigerio J-M, Gion J-M, Lalanne C, Madur D, Pionneau C, Gerber S (2004) Proteomics for genetic and physiological studies in forest trees: application in maritime pine. In 'Molecular Genetics and Breeding of Forest Trees' (Eds. Kumar S, Fladung M) pp. 53-79 (Food Products Press, New York)

Plomion C, Bahrman N, Durel C, O'Malley DM (1995) Genomic mapping in *Pinus pinaster* (maritime pine) using RAPD and protein markers. *Heredity* **74**, 661-668.

Poke FS, Vaillancourt RE, Elliot RE, Reid JB (2003) Sequence variation in two lignin biosynthesis genes, cinnamoyl CoA reductase (CCR) and cinnamyl alcohol dehydrogenase 2 (CAD2). *Molecular Breeding* **12**, 107-118.

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

Pot D, Carlos-Rodrigues J, Rozenberg P, Chantre G, Tibbits J, Cahalan C, Pichavant F, Plomion C (2006) QTLs and candidate genes for wood properties in maritime pine (*Pinus pinaster* Ait.). *Tree Genetics and Genomes* **2**, 10-24.

Potts B M, Barbour RC, Hingston AB, Vaillancourt RE (2003) Turner Review no. 6, Genetic pollution of native eucalypt gene pools-identifying the risks. *Australian Journal of Botany* **51**, 1-25.

Potts BM, Dungey HS (2004) Interspecific hybridization of *Eucalyptus*: key issues for breeders and geneticists. *New Forests* **27**, 115-138.

Potts BM, Pederick (2000) Morphology, phylogeny, origin, distribution and genetic diversity of eucalypts. 'Diseases and pathogens of eucalypts' (Eds. Keane PJ, Kile GK, Podger FD, Brown BN) pp. 11-34. (CSIRO, Melbourne).

Potts B M, Vaillancourt RE, Jordan G, Dutkowski G, Costa e Silva J, McKinnon G, Steane D, Volker P, Lopez G, Apiolaza L, Li Y, Marques C, Borralho N (2004) Exploration of the *Eucalyptus globulus* gene pool. Plenary paper, pp 46-61 in '*Eucalyptus* in a changing world'. (Eds. Borralho N, Periera JS, Marques C, Coutinho J, Madeira M, Tome M) Proceedings IUFRO Conference.RAIZ, Instituto Investigação de Floresta e Papel, Aveiro, Portugal.

Potts B M, Volker PW, Tilyard PA, Joyce K (2000) The genetics of hybridisation in the temperate *Eucalyptus*. pp. 200-210 In 'Hybrid Breeding and Genetics of Forest Trees'. (Eds. Dungey HS, Dieters MJ, Nikles DG) Proceedings of QFRI/CRC-SPF Symposium, Noosa, Queensland, Australia.

Potts BM, Wiltshire RJE (1997) Eucalypt genetics and genecology. In 'Eucalypt ecology: Individuals to ecosystems' (Eds. Williams J, Woinarski J) pp. 56-91. (Cambridge University Press: Cambridge).

Poynton RJ (1979) Tree planting in South Africa Vol. 2 The eucalypts. Technical report, South Africa Forestry Research Institute, South Africa.

Prioul S, Frankewitz A, Deniot G, Morin G, Baranger A (2004) Mapping of quantitative trait loci for partial resistance to *Mycosphaerella pinodes* in pea (*Pisum sativum* L.) at the seedling and adult plant stages. *Theoretical and Applied Genetics* **108**, 1322-1334.

Pritchard JK, Stephens M, Donnelly P (2000a) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.

Pritchard JK, Stephens M, Rosenberg NA, Donnelly P (2000b) Association mapping in structured populations. *American Journal Human Genetics* **67**, 170-181.

Rajora OP, Rahman MH (2003) Microsatellite DNA and RAPD fingerprinting, identification and genetic relationships of hybrid poplar (*Populus* x *Canadensis*) cultivars. *Theoretical and Applied Genetics* **106**, 470-477.

Rapley LP, Allen GR, Potts BM (2004a) Genetic variation in *Eucalyptus globulus* in relation to susceptibility from attack by the southern eucalypt leaf beetle, *Chrysophtharta agricola*. *Australian Journal of Botany* **52**, 747-756.

Rapley LP, Allen GR, Potts BM (2004b) Susceptibility of *Eucalyptus globulus* to *Mnesampela privata* defoliation in relation to a specific foliar wax compound. *Chemoecology* **14**, 157-163.

Raymond CR, Apiolaza LA (2004) Incorporating wood quality and deployment traits in *Eucalyptus globulus* and *Eucalyptus nitens*. pp. 97-99 In 'Plantation Forest Biotechnology for the 21st Century' (Ed. Carson M). (Forest Research New Zealand: Rotorua, New Zealand).

Reinoso C (1992) Variation in *Eucalyptus globulus* in susceptibility to *Mycosphaerella* leaf diseases. Honours thesis, University of Melbourne Australia.

Remington DL, Whetten RW, Liu BH, O'Malley DM (1999) Construction of an AFLP genetic map with nearly complete genome coverage in *Pinus taeda*. *Theoretical and Applied Genetics* **98**, 1279-1292.

Richardson T, Mackenzie RCM, Lee JR, Kent J, Carson SD (1997) Effects of domestication on genetic diversity in *Pinus radiata*. Proceedings Plant and animal genome V conference, San Diego California.

Rieseberg LH, Linder CR, Seiler GJ (1995) Chromosomal and genic barriers to introgression in *Helianthus*. *Genetics* **141**, 1163-1171.

Rocha RB, Abad JIM, Pires IE, Araujo EF (2002) Fingerprint and genetic diversity analysis of *Eucalyptus* ssp. genotypes using RAPD and SSR markers. *Scientia Forestalis* **62**, 24-31.

Salamini F, Ozkan H, Brandolini A, Schafer-Pregi R, Martin W (2002) Genetics and geography of wild cereal domestication in the near east. *Nature Reviews Genetics* 3, 429-441.

Sewell MM, Neale DB (2000) Mapping quantitative traits in forest trees. In 'Molecular biology of woody plants' (Ed Jain, SM and Minocha SC) Volume 1, pp 407-423 (Kluwer Academic Publishers: The Netherlands).

Shepherd M, Chaparo JX, Teasdale R (1999) Genetic mapping of monoterpene composition in an interspecific eucalypt hybrid. *Theoretical and Applied Genetics* **99**, 1207-1215.

Shepherd M, Cross M, Dieters MJ, Henry R (2003) Genetic maps for *Pinus elliottii* var. *elliottii* and *P. caribaea* var. *hondurensis* using AFLP and microsatellite markers. *Theoretical and Applied Genetics* **106**, 1409-1419.

Shepherd M, Jones M (2004) Molecular markers in tree improvement: characterisation and use in *Eucalyptus*. pp. 399-409 In 'Biotechnology in Agriculture and Forestry' (Eds. Lorz H, Wenzel G) (Springer-Verlag: Heidelberg).

Sigurdsson V, Anamthawat-Jonsson K, Sigurgeirsson A (1995) DNA fingerprinting of *Populus trichocarpa* clones using RAPD markers. *New Forests* **10**, 197-206.

Simpson GG, Gendall AR, Dean C (1999) When to switch to flowering. *Annual Review of Cell and Developmental Biology* **15**, 519-550.

Smith AH, Gill WM, Pinkard EA, Hunter GC, Wingfield BD, Mohammed C (2005) Defence responses in eucalypts to infection by *Mycosphaerella* species. p. 10 In 'Proceedings of the 15th Australasian Plant Pathology Society Conference, Mycosphaerella workshop.' Geelong, Australia.

Soto A, Sánchez N, Astorga R, Alia R (2004) Geographic structure of *Eucalyptus globulus* ssp. *globulus populations* according to chloroplast DNA. 'In *Eucalyptus* in a changing world'. (Eds. NMG Barralho, JS Pereira, C. Marques, J. Coutinho, M. Madeira, M. Tomé) Proceedings IUFRO Conference, Aveiro, Portugal, 11-15 October 2004.

Steane DA, Byrne M, Vaillancourt RE, Potts BM (1998) Chloroplast DNA polymorphism signals complex interspecific interactions in *Eucalyptus* (Myrtaceae). *Aust Syst Bot* 11, 25-40.

Steane DA, Conod N, Jones RC, Vaillancourt RE and Potts BM (2006) A comparative analysis of population structure of *Eucalyptus globulus* (Myrtaceae) using microsatellite markers and quantitative traits. *Tree Genetics and Genomes* **2**, 30-38.

Steane DA, Vaillancourt RE, Russell J, Powell W, Marshall D, Potts BM (2001) Development and characterisation of microsatellite loci in *Eucalyptus globulus* (Myrtaceae). *Silvae Genetica* **50**, 89-91.

Steinbauer MJ (2002) Oviposition preference and neonate performance of *Mnesampela privata* in relation to heterophylly in *Eucalyptus dunnii* and *Eucalyptus globulus*. *Agricultural and Forest Entomology* **4**, 245-253.

Strauss SH, Lande R, Namkoong G (1992) Limitations of molecular-marker-aided selection in forest tree breeding. *Canadian Journal of Forest Research* **22**, 1050-1061.

Tautz D (1989) Hypervariable simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research* **17**, 6463-6471.

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*. *Theoretical and Applied Genetics* **109**, 856-864.

Thamarus KA, Groom K, Murrell J, Byrne M, Moran GF (2002) A genetic linkage map for *Eucalyptus globulus* with candidate loci for wood, fibre, and floral traits. *Theoretical and Applied Genetics* **104**, 379-387.

Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielson D (2001) *Dwarf8* polymorphisms associate with variation in flowering time. *Nature Genetics* **28**, 286-289.

Thumma BR, Nolan MF, 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, Boomsma DB, Jarvis S (1997) In 'Distribution, biology, genetics and improvement programs for *Eucalyptus globulus* and *E. nitens* around the world' pp. 1-15 Southern Tree Improvement Committee, Orlando, Florida.

Tripathi SB, Mathish NV, Gurumurthi K (2006) Use of genetic markers in the management of micropropagated *Eucalyptus* germplasm. *New Forests* **31**, 361-372.

Vaillancourt RE, Potts BM, Manson A, Eldridge T, Reid JB (1995): Using RAPDs to detect QTLs in an interspecific F2 hybrid of *Eucalyptus*. pp 430-433 In 'Eucalypt plantations: Improving fibre yield and quality'. (Eds. Potts BM, Borralho NMG, Reid JB, Cromer RN, Tibbits WN, Raymond CA) Proceedings paper CRC IUFRO conference, Hobart, Australia.

Vaillancourt RE, Skabo S, Gore P (1998) Fingerprinting for quality control in breeding and deployment. *Australian Forestry* 61, 207-210.

Vaillancourt RE, Weeden NF (1992) Chloroplast DNA polymorphism suggests Nigerian centre of domestication for the cowpea, *Vigna unguiculata* (Leguminosae). *American Journal of Botany* **79**, 1194-1199.

Van Der Kamp BJ (1991) Major gene resistance of Scots pine to western gall rust. *Canadian Journal of Forest Research* **21**, 375-378.

Van Ooijen JW (1999) LOD significance thresholds for QTL analysis in experimental populations of diploid species. *Heredity* **83**, 613-624.

Van Ooijen JW, Boer MP, Jansen C, Maliepaard C (2002) MapQTL® 4.0, software for the calculation of QTL position on genetic maps. Plant Research International, Wageningen, the Netherlands.

Van Ooijen JW, Voorrips RE (2001): Joinmap® 3.0, software for the calculation of genetic linkage maps. Plant Research International, Wageningen, the Netherlands.

Van Zyl LM, Coutinho TA, Wingfield MJ, Pongpanich K, Wingfield BD (2002) Morphological and molecular relatedness of geographically diverse isolates of *Coniothryium zuluense* from South Africa and Thailand. *Mycological Research* **106**, 51-59.

Venkatesan KR, Kumaravelu G, Somasundaram K (1984) 'Eucalypts in India. Past, present and future'. pp. 290-296 In Proceedings of the national seminar held at Kerala Forest Research Institute, January 30-31, 1984.

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

Verhaegen D, Plomion C, Gion JM, Poitel M, Costa P, Krmer A (1997) Quantitative trait dissection analysis in *Eucalyptus* using RAPD markers 1. Detection of QTL in hybrid progeny, stability of QTL expression across different ages. *Theoretical and Applied Genetics* **95**, 597-608.

Volker, P.W. (2002): Genetics of *Eucalyptus globulus, E. nitens* and F1 hybrid. PhD Thesis, University of Tasmania, Australia.

Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* **23**, 4407-4414.

Votava EJ, Baral JB, Bosland PW (2005) Genetic diversity of chile (*Capsicum annuum* var. *annuum L.*) landraces from northern New Mexico, Colorado, and Mexico. *Economic Botany* **59**, 8-17.

Walter C, Carson SD, Menzies MI, Richardson T, Carson M (1998) Review: Application of biotechnology to forestry – molecular biology of conifers. *World Journal of Microbiology and Biotechnology* **14**, 321-330.

Wang Z, Taramino G, Yang D, Liu G, Tingey SV, Miao G-H, Wang G-L (2001) Rice ESTs with disease-resistance gene- or defence-response gene-like sequences mapped to regions containing major resistance genes or QTLs. *Molecular Genetics and Genomics* **265**, 302-310.

Wheeler NC, Jech KS (1992) The use of electrophoretic markers in seed orchard research. *New Forests* **6**, 311-328.

Williams ER, Whitaker D (2001): Outmap© 1.0, a package for genetic mapping, CSIRO forestry and forest products, Canberra, Australia and University of Waikato, Hamilton, New Zealand.

Williams JGK, Kubelik AR, Livak KJ, Raflaski JA, Tinget SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* **18**, 6531-6535.

Williams KJ, Potts BM (1996) The natural distribution of *Eucalyptus* species in Tasmania. *Tasforests* **8**, 39-164.

Wiltshire RJE, Potts BM, Reid JB (1998) The genetic control of reproductive and vegetative phase change in the *Eucalyptus ridonii/E. tenuiramis* complex. *Australian Journal of Botany* **46**, 45-63.

Winfield MO, Arnold GM, Cooper F, Le Ray M, White J, Karp A, Edwards KJ (1998) A study of genetic diversity in *Populus nigra* subsp. *betulifolia* in the upper Severn area of the UK using AFLP markers. *Molecular Ecology* **7**, 3-10.

Wu HX (2002) Study of early selection in tree breeding 4. Efficiency of marker aided early selection (MAES). *Silvae Genetica* **51**, 5-6.

Wu HX, Ying CC (1998) Stability to western gall rust and needle cast resistance in lodgepole pine. *Canadian Journal of Forest Research* **28**, 439-449.

Wu RL, Han YF, Hu JJ, Fang JJ, Li L, Li ML, Zeng ZB (2001) An integrated genetic map of *Populus deltoides* based on amplified fragment length polymorphisms. *Theoretical and Applied Genetics* **100**, 1249-1256.

Xie CY, El-Kassaby YA, Ying CC (2002) Genetics of red alder (*Alnus rubra* Bong.) populations in British Columbia and its implications for gene resources management. *New Forests* **24**, 97-112.

Yang BZ, Zhao H, Kranzler HR, Gelernter J (2005) Practical population group assignment with selected informative markers: characteristics and properties of Bayesian clustering via Structure. *Genetic Epidemiology* **28**, 302-312.

Yeh FC, Yang RC, Boyle TJ, Ye ZH, JX Mao (1997) POPGENE, the user friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Edmonton, Canada.

Young ND (1996) QTL mapping and quantitative disease resistance in plants. *Annual Review of Phytopathology* **34**, 479-501.

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

Zang D, Wang H, You Y (1995) Performance and selection of a 4 year *Eucalyptus globulus* seedling seed orchard in Yunnan, China. pp 195-198 In Eucalypt plantations: Improving fibre yield and quality (Eds. BM Potts, NMG Borralho, JB Reid, RN Cromer, WN Tibbits and CA Raymond). Proceedings CRC-IUFRO Conference, Hobart, 19-24 February 1995.

Zheng YQ, Ennos RA (1999) Genetic variability and structure of natural and domesticated populations of Caribbean pine (*Pinus caribaea* Morelet) *Theoretical and Applied Genetics* **98**, 765-771.