GENOME MAPPING

of

Eucalyptus globulus

by

Peter C. Bundock

Submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

University of Tasmania

November 2002

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 acknowledgment is made in the text of this thesis.

Peter C. Bundock

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

Peter C. Bundock

ABSTRACT

This thesis covers the construction of genetic linkage maps of two *Eucalyptus globulus* parental trees (chapter 2), the detection of quantitative trait loci (QTL) carried out using these linkage maps (chapter 3) and a genetic analysis of rooting ability of micropropagated cuttings from two *E. globulus* families (chapter 4). Genetic linkage maps of two parent trees were constructed using 326 RAPD and 21 microsatellite markers. At a LOD score threshold of 4.9 for grouping and 3.0 for marker ordering, the male parent had 13 linkage groups consisting of 101 framework markers and the female parent had 11 linkage groups with 97 framework markers. Even though both parents originated from the same provenance it was found that polymorphic RAPD markers were readily detectable. Linkages between microsatellites previously reported for *E. grandis/E. urophylla* were conserved in the *E. globulus* cross. Segregation distortion of markers was found to be more prevalent than expected by chance.

The linkage maps were used to detect QTLs based on 155 progeny grown in field trials. Twelve traits were included in the QTL analysis. These included: wood density (Pilodyn penetration), extent of early flowering (bud abundance), and growth (height at years one and two, stem diameter at years two, three, four and six, and relative incremental growth between years one and two, two and four, and four and six). Using interval mapping a total of eight QTL with LOD score peaks over 2.0 were detected, corresponding to seven map intervals. QTL detected included: two for cumulative growth; two for wood density; one for early flowering and three for relative incremental growth. Since the 155 progeny trees were grown at seven trial sites, an analysis of marker by site interaction was

carried out with more marker by site interactions being found than expected by chance. Markers with significant QTL effects were examined for interaction with site with both of the cumulative growth QTL found to have significant site interaction. QTL stability with age was also analysed for these QTL, and both growth QTL were found to have a detectable association with the first measurements of height and diameter.

A study of the rooting ability of cuttings grown in tissue culture from two families of *E. globulus* was undertaken with the aim of detecting QTLs for this trait. The two families were found to have significantly different rooting abilities indicating that genetic variation for the trait was present. Estimates of variance components relating to genetic and environmental effects indicated that within both families the variance due to genetic effects was small compared to the environmental variance with clonal repeatabilities of 0.17 and 0.14 for the two families. These estimates were used in power calculations based on t-tests of single markers under some simple assumptions. However even under optimistic circumstances the power for QTL detection was found to be too low to warrant the time and expense of genotyping required for QTL analysis.

ACKNOWLEDGMENTS

I would like to acknowledge the help of Dr René Vaillancourt who supervised this project through thick and thin and without whose encouragement this thesis would not have been written. For the linkage mapping work I would like to thank all those in the Molecular Genetics laboratory for their friendship and tolerance and Matt Hayden for carrying out some of the RAPD work. For assistance in collecting *Eucalyptus globulus* leaves for DNA extraction and showing me how to shoot a Pilodyn I would like to thank Paul Tilyard. I would also like to acknowledge others who assisted in leaf collection at various sites across southern Australia including: Greg Dutkowski, John Owen, Ian Bail and Carl Grosser. CSIRO and Gunns are also acknowledged for permission to collect leaf material. Dr Dorothy Steane and Dr Gavin Moran are acknowledged for providing information on primers and amplification of microsatellites.

I would like to acknowledge the assistance of Assoc. Prof Brad Potts in the analysis of the quantitative trait data using SAS and ASREML, and single marker ANOVA and other SAS procedures. Also for his encouragement in the early years of this project I would like to thank Dr Nuno Borralho.

Gunns Limited (formally North Forest Products) is acknowledged for the use of their tissue culture facilities at West Ridgley. I would like to acknowledge the assistance of the technicians in the tissue culture laboratory at West Ridgley and Rebecca Brown from the University of Tasmania, for assisting with the maintenance of tissue cultures and carrying out much of the harvesting of shoots for testing of rooting ability. Gillian Rasmussen is acknowledged for technical

advice relating to tissue culture in the rooting ability study. For assisting in the statistical analyses on the rooting ability study I would like to thank Greg Dutkowski (University of Tasmania) and Dr Lyndon Brooks (Southern Cross University). The Australian Research Council and the CRC for Sustainable Production Forestry are acknowledged for financial assistance.

In addition I would like to thank my parents for their support during the process of studying for this thesis and finally and not least I must thank my wife Alison and our children Hugh and Patrick for putting up with someone who undertook a PhD, especially since much of the writing up was in 'spare time', a lot of which was really theirs.

TABLE OF CONTENTS

CHAPTER ONE

GENERAL INTRODUCTION

1.1 Molecular marker systems and their application to linkage mapping of forest trees

Developments in the marker systems used to construct genetic linkage maps in forest trees have generally followed their application in other plant and animal species. Marker systems tested and proven in animals, particularly model species such as *Drosophila* and mouse but also in humans, have been applied in many plant species including forest trees.

Before the introduction of molecular markers, linkage maps were composed of visible phenotypic markers that had Mendelian inheritance patterns. In plants this restricted the construction of genomic maps to genetically well-studied species such as tomato (*Lycopersicon esculentum*) and maize (*Zea mays*) where even there the maps constructed were sparse in their coverage of the genome and were usually composite maps resulting from many crosses between different marker lines (O'Brien 1987).

The introduction of techniques to detect allozymes, based on differential mobility during electrophoresis, enabled a new set of polymorphic loci to be utilised for

linkage mapping in plants (Tanksley and Orton 1983). The detection of allozymes (sometimes broadly referred to as isozymes) was exploited in forest trees and utilised for studies of genetic relatedness to examine mating systems (Fripp *et al.* 1987) and extent of genetic diversity in populations (Moran 1992). Linkage relationships between the relatively small number of allozyme loci were studied in a number of tree species (Adams and Joly 1980; El-Kassaby *et al.* 1982; Cheliak and Pitel 1985) including *Eucalyptus* (Moran and Bell 1983) although there were insufficient markers for construction of extensive linkage maps.

With the development of methods to detect restriction fragment length polymorphisms (RFLPs), the construction of dense linkage maps in humans (and thus other eukaryotes) using these DNA marker loci was hypothesised over two decades ago (Botstein *et al.* 1980). RFLP based maps have been constructed for many organisms of varying genome complexity (eg. bacteria, yeast, nematode, fish and mammals). This has included many species of crop plants (Helentjaris *et al.* 1986; Bonierbale *et al.* 1988; McCouch *et al.* 1988; Gebhardt *et al.* 1989; Graner *et al.* 1991; Vallejos *et al.* 1992) with maps soon being utilised for the detection of quantitative trait loci (QTL) (Patterson *et al.* 1988; Stuber *et al.* 1992) and qualitative loci (Sarfatti *et al.* 1989; Barone *et al.* 1990).

Early reports on the detection and inheritance of RFLP markers in tree species suggested map construction using these markers would be possible (Devey *et al.* 1991; Byrne *et al.* 1994). Several linkage maps were published for forest tree species based on RFLPs (with the inclusion of a limited number of isozyme loci) (Liu and Furnier 1993; Devey *et al.* 1994; Groover *et al.* 1994).

RFLPs are codominant markers, which facilitates the integration of mapping information from different individuals and they are often transferable between related species (Gebhardt *et al.* 1991). However the technique of RFLP detection is highly labour intensive, costly, requires large amounts of genomic DNA, and commonly radioactive labelling is used. Even the initial screening process requires the cloning of genomic or cDNA for use as probes, a proportion (sometimes large) of which will fail to detect polymorphisms. Thus the development of PCR based marker systems that are more amenable to high throughput methods became attractive for linkage map construction.

The development of RAPD (random amplified polymorphic DNA) markers (Welsh and McClelland 1990; Williams *et al.* 1990) provided a simple method of marker generation, which has found wide application particularly in plant systems. Linkage maps utilising RAPD markers have been published for a number of forest tree species including white spruce (*Picea glauca*) (Tulsieram *et al.* 1992), slash pine (*Pinus elliotti*) (Nelson *et al.* 1993), longleaf pine (*Pinus palustris*) (Nelson *et al.* 1994), Norway spruce (*Picea abies*) (Binelli and Bucci 1994), maritime pine (*Pinus pinaster*) (Plomion *et al.* 1995a; Plomion *et al.* 1995b), pinus hybrids (Kubisiak *et al.* 1995), sugi (*Cryptomeria japonica*) (Mukai *et al.* 1995), radiata pine (*Pinus radiata*) (Wilcox *et al.* 2001b), oak (*Quercus robur*) (Barreneche *et al.* 1998) and a poplar hybrid (Bradshaw *et al.* 1994). In the genus *Eucalyptus* RAPD maps of *E. grandis* and *E. urophylla* were constructed using F1 progeny from the interspecific cross (Grattapaglia and Sederoff 1994; Verhaegen and Plomion 1996) and a map using RFLP, RAPD and isozyme markers was constructed for *E. nitens* (Byrne *et al.* 1995).

Although the generation of RAPD markers is a relatively quick and simple process, there are a number of problems that make them less than ideal. Foremost of these is their dominant inheritance pattern with usually only two alleles being detectable at a locus. This limits the usefulness of RAPDs with regard to cross type, within species transferability, map integration and determination of mode of action of QTL mapped using these markers.

In forest tree species RAPD markers have been amenable for mapping studies by either using haploid megagametophyte tissue in the case of some conifer species (Tulsieram *et al.* 1992) or more generally by screening for markers segregating 1:1 in a two-way pseudo testcross configuration using diploid tissue (Grattapaglia and Sederoff 1994). In the pseudo testcross configuration each RAPD marker locus is generally mapped in only one parent of the cross resulting in two parental maps. For the determination of QTL effects using these individual parental maps the segregating QTL 'alleles' from one parent are contrasted without regard to the QTL 'alleles' at the same locus inherited from the other parent. In other words only two genotypes are recognised and contrasted where potentially four genotypes are present. By contrast, fully informative codominant markers at a QTL locus will enable all four possible genotypes to be recognised and compared providing information on the mode of action (dominant, additive) of the QTL.

A further problem with RAPD markers is their limited transferability between related species and also within species (Kesseli *et al.* 1992; Thormann and Osborn 1992; Kesseli *et al.* 1994). Within species of *Eucalyptus* the transferability has been shown to vary depending on the population/species sampled. On average 61% of RAPD markers were calculated to be transferable between ten *E.*

urophylla parents sampled from the same population compared with 15% for *E. grandis* parents sampled across widely distinct origins (Brondani *et al.* 1997). However, these estimates were calculated based on a single interspecific cross where one parent was known to be homozygous null. Based on this data for crosses between the *E. urophylla* parents (ie. within the species), the proportion of RAPD markers transferable to another cross (ie. segregating 1:1) would be expected, on average, to be only 33% and for *E. grandis* 20%.

Despite the drawbacks of RAPD markers they have proven to be the most commonly utilised marker for the construction of linkage maps in forest tree species. More recently however, a marker system has been developed that has characteristics similar to RAPDs but with much more information obtainable per amplification. AFLP™ (amplified fragment length polymorphisms) are generated from restriction endonuclease digests of genomic DNA by ligation of adapter sequences followed by selective PCR amplification (Vos *et al.* 1995). AFLPs suffer from the dominant characteristic of RAPDs and they are also more technically exacting in their generation, however, they are claimed to be more repeatable (Jones *et al.* 1997) and potentially may enable even quicker map construction than RAPDs (Powell *et al.* 1996; Costa *et al.* 2000).

AFLPs have been used in map construction of a number of forest tree species including an F1 cross between *E. tereticornis* and *E. globulus* (Marques *et al.* 1998), in *Populus* species (Wu *et al.* 2000; Cervera *et al.* 2001), willow (*Salix*), (Tsarouhas *et al.* 2002) and with some modifications the technique was applied to map larch (*Larix*) (Arcade *et al.* 2000), *Pinus* (Travis *et al.* 1998; Remington *et al.*

1999; Costa *et al.* 2000; Lerceteau *et al.* 2000) and Norway spruce (*Picea abies*) (Paglia *et al.* 1998).

Of the major types of marker systems developed the one with the greatest potential for constructing consensus species maps and for enabling map integration and comparative mapping is microsatellite markers, also termed SSRs (simple sequence repeats) or STRs (short tandem repeats). Once developed for a species under study SSRs have a number of ideal attributes. They are commonly highly polymorphic in a population with many possible alleles and like RAPDs they are PCR based and require only small amounts of genomic DNA for generation but unlike RAPDs they are typically codominant and are potentially fully informative (Morgante and Olivieri 1993). Once available they can be used to generate linkage maps quickly and the integration of parental maps from a cross of heterozygous individuals from an outcrossing species is possible though not all loci may be fully informative. Allele differences are based on size differences in the PCR amplicon and differentiation of alleles requires separation on a gel system. If the differences in allele sizes are sufficiently large at a locus separations can be achieved using agarose gels (Brondani *et al.* 1998). However more commonly acrylamide gels are used and by co-loading PCR products amplified using different dye labelled primers a number of loci can be analysed per lane on slab gels or by capillary electrophoresis and detected using automated detection systems (Rafalski *et al.* 1996). This enables relatively high throughput of markers.

Saturated linkage maps composed of microsatellite markers have been constructed in animal systems (Dib 1996; Dietrich 1996) and although microsatellites are

reported to occur less frequently in plant genomes (Lagercrantz *et al.* 1993) they have been utilised extensively in forest trees for paternity analysis and estimation of outcrossing rates (Chase *et al.* 1996; Streiff *et al.* 1999; Lian *et al.* 2001), mating system analysis (Vogl *et al.* 2002) and gene flow and population genetic studies (Dow and Ashley 1996; Gonzalez-Martinez *et al.* 2002). Although microsatellites have been isolated from many forest tree species, few linkage maps published until recently have incorporated more than a smattering of microsatellite loci.

Recovery of large numbers of single locus polymorphic microsatellite loci from generally large and highly repetitive conifer genomes has proved difficult (Elsik and Williams 2001; Zhou *et al.* 2002) and the extent of transfer of microsatellite loci between conifer species has been low (Devey *et al.* 1999; Echt *et al.* 1999; Karhu *et al.* 2000; Mariette *et al.* 2001) although greater between closely related taxa (van de Ven and McNicol 1996; Fisher *et al.* 1998; Echt *et al.* 1999; Shepherd *et al.* 2002). However in Norway spruce (*Picea abies*), Paglia *et al.* (1998) mapped 61 microsatellites in a map of 413 loci and Devey *et al.* (1999) mapped nine microsatellite loci in comparative maps of two species of pine.

Angiosperm forest tree species generally have smaller genomes than conifers and repetitive DNA is less likely to create problems in isolating single locus microsatellite loci. Conservation of microsatellite loci has been shown between oak (*Quercus*) species (Isagi and Suhandono 1997) and Barreneche *et al.* (1998) mapped 18 microsatellites in *Quercus robur* from a total of 301 markers placed on two parental maps. In *Populus* microsatellites have been utilised to align homoeologous linkage groups and merge linkage groups from maps constructed

from crosses between a *Populus deltoides* parent with *P. nigra*, and *P. trichocarpa* (Cervera *et al.* 2001). Also in *Populus*, Trembling Aspen (*Populus tremuloides*) derived microsatellites have been shown to transfer to other members of the Salicaceae (Rajora and Rahman 2001). In *Eucalyptus*, 70 microsatellite loci were isolated and mapped in an *E. grandis* x *E. urophylla* cross (Brondani *et al.* 1998; Brondani *et al.* 2002).

Transferability of microsatellite loci across *Eucalyptus* species and closely related genera has been demonstrated to be fairly high, but particularly so within subgenera (Byrne *et al.* 1996; Brondani *et al.* 1998; Jones *et al.* 2001; Steane *et al.* 2001; Brondani *et al.* 2002; Thamarus *et al.* 2002). In *Eucalyptus* at least, the possibility of utilising SSRs for species-wide and potentially genus-wide linkage mapping is feasible. This would allow QTL locations determined in one species to be analysed both within the species and also between species as demonstrated by Marques *et al.* (2002).

Another marker system that is being explored in a number of plant systems is based on single nucleotide polymorphisms (SNPs). SNPs have been discovered in the hundreds of thousands in human DNA sequences as part of the human genome sequencing efforts (Sachidanandam *et al.* 2001). Based on information from the sequencing of expressed sequence tags (ESTs), which are available for some species in public databases, SNP detection can be targeted to expressed sequences.

Some of the advantages of SNPs over other marker systems include their abundance in the genome, being many times more frequent than microsatellites, and also their analysis using high throughput systems. If found in expressed

genes, SNPs can become so called 'perfect' markers, since they are inherited in complete linkage with the gene of interest. SNP discovery is, however, expensive and time consuming, since at present it relies on direct sequencing information from two or more alleles of the SNP host locus. However ESTs can also be converted into mappable genetic markers where fragment length polymorphisms created by indels (insertion/deletions) occur between alleles (especially frequent outside the coding region), enabling size differences to be resolved on acrylamide gels.

A method of PCR amplifying regions 5' and 3' of ESTs has been used to detect indel and SNP polymorphism (without sequencing) in 45 ESTs from *Pinus* species in two pedigrees from *P. taeda* and one pedigree from *P. radiata*. Eighteen of these ESTs were mapped in a *Pinus radiata* cross (Cato *et al.* 2001).

Mapping EST sequences in preference to anonymous markers has the advantage that mapped loci can be chosen on the basis of known or putative function. These loci can be chosen to provide candidate genes for placement on maps utilised for QTL detection. Co-segregation of candidate loci and QTL may result from polymorphism at the locus influencing variance for the measured quantitative trait. These loci could be mapped potentially using SNPs or indels in or close to transcribed sequence (Cato *et al.* 2001) or using products amplified based on ESTs as RFLP probes (Thamarus *et al.* 2002).

In a recent report by Thamarus *et al.* (2002), 45 expressed genes were mapped in an interprovenance *E. globulus* cross using RFLP technology. Forty-one of the genes are known proteins and include enzymes involved in lignin and cell-wall

polysaccharide biosynthesis with the resulting map being used for QTL detection for wood and fibre traits (Thamarus *et al.* 2002).

A similar approach to mapping candidate genes but using single strand conformation polymorphisms (SSCP) has been reported by Gion *et al.* (2000) for eight lignin biosynthesis genes in an *E. grandis* x *E. urophylla* cross and the same research group mapped five lignin biosynthesis genes in Maritime pine (Plomion *et al.* 1999). The candidate gene approach in forest trees has yielded QTL for lignin chemistry co-segregating with mapped lignin biosynthesis loci (Gion *et al.* 2001).

1.2 QTL detection

Quantitative traits are generally considered to be traits that have a continuously variable numerical distribution (Falconer 1989). Quantitative trait loci (QTL) can be defined as chromosomal regions, harbouring one or more genes, that influence a quantitative trait (Gelderman 1975; Kearsey and Farquhar 1998; Patterson 1998). However evidence for QTLs comes from statistical associations between phenotypic trait values and genotype classes (or marker loci). Thus it can be envisaged that there are statistical QTLs, some of which may be artefacts, and biological QTLs that have a genuine genetic basis.

One of the main reasons for interest in QTLs is the potential to apply QTL information in the breeding of economically important animal and plant species. The most commonly envisaged application is to use genetic markers to select for superior QTL 'alleles' in the population – a process commonly termed marker-assisted selection (MAS). However in initial QTL detection experiments the

whole genome is scanned for associations between one or more traits and the possibility of a Type I error (chance association only) increases, requiring more stringent significance thresholds (Type I error threshold). Unfortunately more stringent significance thresholds increase the probability of Type II errors (real associations not being declared significant). Clearly a balance must be reached and this depends on the experimental aims (Beavis 1998).

The classical method of QTL detection is to use t-tests (or equivalently ANOVA) to detect statistical associations between marker loci and the quantitative trait of interest (Soller *et al.* 1976; Edwards *et al.* 1987; Groover *et al.* 1994). There are several advantages to this approach. One is that the method is relatively simple, and readily available statistical packages can be used to carry out the analysis. Another advantage is that the distribution of the test statistic is well known, enabling straightforward evaluation of power and precision (Soller 1991; Beavis 1998). There are however valid concerns regarding the application of this method for QTL detection. Firstly, where QTL are located at some distance from the nearest marker loci, the QTL effect, as measured at the nearest marker locus, decreases according to the square of the recombination fraction (θ) between the marker and the QTL (Soller 1991) and thus the power of QTL detection can be low and the size of the QTL effect underestimated. The second concern arises from the fact that Type I error rates are calculated based on the number of independent tests undertaken. However, with genetic linkage between markers the assumption of independence is clearly incorrect.

There are however ways to address both of these problems. If the genetic map used for QTL detection is saturated with markers, the likelihood of QTL being

any great distance from a marker will be reduced. The general result will be that there will be little loss of QTL detection power due to distance between mapped marker and QTL locus, and estimates of QTL effect will not be greatly underestimated. The Type I error rate can be determined either based on the number of independently segregating locations in the genome or the data can be permuted to estimate empirically the correct experiment-wise Type I error threshold level (Churchill and Doerge 1994). All the same it is not always easy to saturate genetic maps. Also the t-test classical approach to QTL detection does not use all available information when it comes to determining the position of a QTL, the QTL is inferred as being closest to the marker with the greatest significance, but this does not take into account missing data.

A more complex method of QTL analysis, termed interval mapping, uses flanking markers and maximum likelihood algorithms in detection and also estimation of QTL location (Lander and Botstein 1989). This method has probably been the most commonly utilised method for QTL detection cited in studies of forest tree species (eg. Bradshaw and Stettler 1995; Grattapaglia *et al.* 1995; Verhaegen *et al.* 1997). The software to carry out interval mapping analyses, MAPMAKER/QTL, has been freely available since about 1989 (Lander and Botstein 1989). MAPMAKER/QTL examines intervals for QTL and calculates log likelihood ratio scores (LOD scores) for the presence of a QTL as the interval is scanned in stepwise units. There is a claimed increase in QTL detection power, where QTLs are located some distance from a marker locus (Lander and Botstein 1989). It has however been shown that this method has the same power as ANOVA where a QTL is perfectly linked to a marker locus and the power is similar up to about 20cM distant from the nearest marker (Darvasi *et al.* 1993).

For interval mapping, determining appropriate significance thresholds has perhaps proved more difficult than envisaged, however as for the classical t-test approach, experiment-wise significance thresholds can be determined empirically (Churchill and Doerge 1994).

A number of other QTL analysis methods have been developed and applied in QTL detection studies. Regression analysis (least squares) using multiple linked markers (Haley *et al.* 1994) enables greater flexibility than interval mapping in modelling multiple QTL effects, interactions and other effects and has been applied in forestry studies (Knott *et al.* 1997). Developments based on interval mapping, but using markers linked to QTL as cofactors, known as multiple QTL models (as opposed to single QTL models) or composite interval mapping, have been reported (Jansen 1993; Zeng 1993). Some increase in power over the classical approach and interval mapping has been claimed using all these methods, however decision rules on inclusion of markers as cofactors have not been adequately determined (Jansen 1994; Beavis 1998). Thus interval mapping and ANOVA based methods have been the mainstays in applied QTL detection.

1.3 Introduction to experimental work covered in this thesis

Eucalyptus globulus is the most significant temperate hardwood plantation tree species utilised for pulp and paper production. It is grown as a plantation species in many countries outside of Australia including Argentina, Chile, South Africa, China, Spain and Portugal with plantations covering vast areas (Eldridge *et al.* 1993). In many of these countries genetic improvement programs are underway to

select superior genotypes for breeding qualities required for productive plantation growth.

The most commonly reported traits that are targeted for genetic study and selection are growth, wood density and pulping quality (Borralho 2001; Lopez *et al.* 2002; Miranda and Pereira 2002; Wimmer *et al.* 2002). Since *E. globulus* has only recently been subjected to selective breeding, there is expected to be considerable potential for improvement in performance. However efforts at improvement are hampered by the long generation interval and any method of reducing lag time in selection is worthy of investigation.

At the time of commencement of the PhD project reported in this thesis, it had been established that molecular marker linkage maps could be constructed in forest tree species (Bradshaw *et al.* 1994; Devey *et al.* 1994) including *Eucalyptus* (Grattapaglia and Sederoff 1994; Byrne *et al.* 1995). Linkage maps for *Eucalyptus* had been based on either a hybrid cross between *E. grandis* and *E. urophylla* (Grattapaglia and Sederoff 1994) or a three-generation pedigree of *E. nitens* where the four grandparents originated from disjunct populations (Byrne *et al.* 1995). In both of these instances the genetic distance between the parents/grandparents would be expected to ensure a high frequency of polymorphic markers segregating in the progeny. One of the aims of the present study was to determine the feasibility of producing linkage maps for *Eucalyptus globulus* based on an intraprovenance cross, where the genetic distance between the parents/ grandparents would be expected to be much closer than in maps published to date. It was intended that the *E. globulus* linkage maps could eventually form the basis of a reference map for the species.

However the main purpose of map construction was to utilise the resultant linkage maps directly to detect quantitative trait loci (QTL) segregating in the family that constituted the mapping population. There were no published reports of QTL detection in *E. globulus* and to date only one study examining QTL for vegetative propagation has been published for a hybrid cross with *E. tereticornis* (Marques *et al.* 1999; Marques *et al.* 2002). Other applications for the linkage maps had been envisaged for future related studies on *E. globulus*, such as comparative genomic studies and population and paternity studies using mapped markers (Steane *et al.* 2001; Jones *et al.* 2002a). For these studies it was seen as essential to place microsatellite markers on the linkage map.

Early studies in crop plant species had indicated the potential for discovering quantitative trait loci (QTL) for complex traits (Edwards *et al.* 1987; Patterson *et al.* 1988). Early work on forest tree species indicated that QTL could be detected in these species also (Bradshaw and Grattapaglia 1994; Bradshaw and Stettler 1995; Grattapaglia *et al.* 1995), even though the pedigrees and crosses that were theoretically highly suitable for such studies (crosses between inbred lines divergent for the trait of interest with QTL segregation in the F2) were not available in forest tree species. The main attraction of detecting QTL in forest tree species such as *E. globulus* was basically the same as for crop species – the possibility of selecting for advantageous QTL with the assistance of markers. However in forest tree species the time required before traits can be evaluated is excessively long, as is the generation interval. The extra appeal of MAS for forest tree breeding is its potential to reduce the time lag for trait evaluation by carrying out marker-based evaluations at the seedling stage, long before rotation age (Williams and Neale 1992; Bradshaw and Grattapaglia 1994; O'Malley and

McKeand 1994; Wilcox *et al.* 2001a). Since in the present study linkage maps were based on a full sib family that was already planted out in a field trial, the aim was to utilise the constructed maps to detect and map QTL for traits of commercial importance. Measurements of commercially important traits that could be included in QTL analyses were growth (height and stem diameter), scores of bud abundance (early flowering), wood density and bark thickness.

Another trait in *E. globulus* that was of potential interest to tree breeders was rooting ability – the production of adventitious roots from the stem base of cuttings. The interest in this trait is due to its importance in clonal forestry. In tropical eucalypt species such as *E. grandis* and *E. urophylla*, superior genotypes can be selected and propagated in most cases by vegetative cuttings for deployment in plantations. The emulation of this feat in temperate eucalypt species was partially thwarted by the highly variable and often poor rooting ability of cuttings from temperate species such as *E. globulus*.

Theoretically, the selection of rooting ability in *E. globulus* breeding programs could increase the potential for clonal deployment – one possible route to improved plantation productivity. Marker-assisted selection had been mooted as a possible means to speed the selection process, but it could only be applied if QTLs were detected for this trait. In chapter four of this thesis, the possibility of detecting QTLs for this trait using *in vitro* cuttings is investigated. It should also be noted that a further reason for undertaking linkage mapping and QTL detection studies, apart from the potential direct application of marker assisted selection, is that basic biological/ genetic information is likely to be gleaned from such studies.

CHAPTER TWO

THE CONSTRUCTION OF LINKAGE MAPS OF *EUCALYPTUS GLOBULUS* USING RAPD AND MICROSATELLITE MARKERS

This chapter has been published as:

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

2.1 INTRODUCTION

In temperate regions *Eucalyptus globulus* is the most significant hardwood tree species grown as a source of fibre for the manufacture of paper products, whilst in tropical regions *E. grandis, E. urophylla* and their hybrids are the most commonly planted *Eucalyptus* for pulpwood (Eldridge *et al.* 1993). Worldwide there has been considerable effort to select and breed genetically superior *Eucalyptus* trees and it has been recognised that molecular markers have potential application in both breeding and deployment (eg. Williams 1995; Kerr *et al.* 1996; Dale and Chaparro 1996; Grattapaglia 1997). Consequently linkage maps composed of DNA based markers have been constructed and published for several species of *Eucalyptus*. The crosses used for generating the segregating mapping populations were either interspecific F_1s (Grattapaglia and Sederoff 1994; Verhaegen and

Plomion 1996; Marques *et al.* 1998) or an F_2 cross using grandparents from widely disjunct populations (Byrne *et al.* 1995). This would virtually ensure that parent trees would be genetically distant from one another. Using genetically distant parent trees is likely to increase the efficiency of mapping. However since *Eucalyptus globulus* is generally grown as a pure species, intraspecific (and also intraprovenance) crosses are required in most breeding programs. An AFLP map of *E. globulus* has been published, however it was based on an interspecific F_1 cross with *E. tereticornis* (Marques *et al.* 1998). This chapter reports on the first maps published for *Eucalyptus globulus* based on a pure species cross and the first published *Eucalyptus* maps using an intraprovenance cross.

Like the previously published maps for *E. grandis* and *E. urophylla* (Grattapaglia and Sederoff 1994; Verhaegen and Plomion 1996), the *E. globulus* maps in this study are based on RAPD markers. Although RAPD markers are extremely useful for map construction, especially using an F_1 cross, they have limited transferability because they are dominant markers with only two alleles. Co-dominant markers such as microsatellite loci on the other hand, are expensive to isolate but have much broader transferability and are also potentially fully informative in crosses other than an F_1. Including microsatellite markers as a significant component of *Eucalyptus* genomic maps should increase the informativeness, transferability and reliability of these maps – factors important to their future applicability. Byrne *et al.* (1996) reported that four microsatellite loci isolated from *E. nitens* were amplifiable and polymorphic in several other species in the genus, demonstrating the potential transferability of microsatellites between species of *Eucalyptus*. Brondani *et al.* (1998) list primer sequences used to amplify twenty microsatellite loci that originate from *E. grandis* and *E. urophylla*

and also show the location of these loci on RAPD maps from a cross between these two species. These primer sequences have been used in the present study to add *E. grandis/E. urophylla* microsatellite loci to the RAPD maps of *E. globulus*. In addition microsatellite loci originating from *E. globulus*, *E. nitens* and *E. sieberi* have been mapped in this cross. This has aided the identification of linkage group homology between the two *E. globulus* parental RAPD maps, and enabled the identification of homology between the *E. globulus* linkage groups reported in this paper and the *E. grandis* and *E. urophylla* linkage groups of Brondani *et al.* (1998).

2.2 MATERIALS AND METHODS

2.2.1 Plant material and DNA extraction

The mapping population consists of a single full-sib family of 165 progeny from an *E. globulus* subsp. *globulus* intraprovenance cross, carried out by CSIRO Division of Forestry and North Forest Products (Australia) and planted at a number of field sites (Vaillancourt *et al.* 1995b). Both parents originate from King Island which is located in Bass Strait between continental Australia and Tasmania. The male parent (G164) is located at a distance of several kilometres from the mother tree of the female parent (KI2) and on this basis the two parents would not be expected to be closely related (Skabo *et al.* 1998). The female parent is an open pollinated progeny planted in a seed orchard.

Two grams of frozen leaf material was ground to a smooth powder in liquid nitrogen using a mortar and pestle. The DNA was extracted using the CTAB extraction protocol of Doyle and Doyle (1990). The CTAB extraction buffer was modified by the addition of polyvinylpyrolidone (PVP-40, Sigma) at 2% (w/v). Two volumes of ice-cold ethanol (~65% final concentration), rather than isopropanol, was found to precipitate higher quality DNA for PCR. By routinely adding a phenol, phenol-chloroform extraction step, readily PCRable DNA was almost always obtained. DNA was quantified using a Hoefer DNA Fluorometer (TKO 100) and Hoechst 33258 dye.

2.2.2 Molecular marker assays

Ten-mer RAPD primers were obtained from the University of British Columbia (UBC; Dr. J. B. Hobbs, c/o Biotechnology Laboratory, Wesbrook Building, 6174 University Boulevard, Vancouver, B.C. V6T1Z3) and Operon Technologies Inc (OP; 1000 Atlantic Ave., Alameda CA 94501 USA). Four hundred UBC RAPD primers comprising sets 1, 2, 3 and 5 and OP kit B were screened to identify those that amplified strong, reliable and polymorphic RAPD bands. DNA from the two parents and six progeny were used for the screening of the RAPD primers.

Amplification conditions are based on those of Williams *et al.* (1993). RAPD reactions (20μL) were composed of the following components: 50mM KCl, 10mM Tris-HCl pH 9.0 @ 25°C, 0.1% Triton X-100, 200μM dNTPs, 3mM $MgCl_2$, 150μg/mL Bovine Serum Albumin (BSA), 0.25μM RAPD primer, 1.6 units Taq DNA Polymerase and 20ng genomic DNA and overlaid with 30μL of mineral oil prior to amplification. An MJ Research Inc. PTC-100 programmable thermal controller was used for amplification using the following cycling profile: 94° – 2min, [94° – 1min, 35° – 1min, 72° – 2min] x 40, 72° – 5min, 10° – hold. Reactions were electrophoresed in 1.5% agarose gels at 22 volt.hours/cm in 1 x TBE buffer. Fluorescence from ethidium bromide (incorporated into the gel at 0.2μg/mL) was used to photograph the RAPD bands using Polaroid 665 film. Polymorphic bands were scored manually from the Polaroid negative. RAPD markers inherited from one parent only and segregating in an apparent 1:1 pattern were classed as originating from either the male or the female parent, creating two separate data sets depending on the parent of origin. A third data set was created

for those RAPD markers that originated from both parents and had a 3:1 segregation pattern.

All of the microsatellite loci that were used in this study were originally isolated as dinucleotide repeats. There were three sources of microsatellite primer sequences: EMCRC (*Eucalyptus* Microsatellites from Co-operative Research Centre for Sustainable Production Forestry) loci originate from *Eucalyptus globulus* DNA enriched for microsatellite sequences (Steane *et al.* 2001); CSIRO primer sequences were obtained from Dr Gavin Moran at CSIRO Forestry and Forest Products and were isolated from *E. nitens* and from *E. sieberi*; and the EMBRA microsatellites were cloned from *E. grandis* and *E. urophylla* and originate from Dr Dario Grattapaglia's lab in Brazil. The primer sequences for the EMBRA loci are those of Brondani *et al.* (1998) and the coding of loci is the same as in that publication.

The details for the primer sequence and amplification conditions for the EMCRC microsatellites are given in Steane *et al.* 2001 and for the CSIRO microsatellites in Byrne . *et al.* 1996 and the CSIRO forestry website. The reaction conditions used for the EMBRA microsatellites were essentially those of Brondani *et al.* (1998) with the following modifications: 0.1% Triton X-100, 0.1mg/mL BSA, no DMSO and annealing at 57°C. All amplifications were carried out on an MJ Research Inc. PTC-100 Thermal Cycler. All amplifiable loci were screened for their potential to be mapped on both parental maps with sizing of alleles carried out on polyacrylamide gels using an Applied Biosystems automated DNA sequencer. Reaction product fragment sizes were calculated using GENESCAN software based on an internal standard. Metaphor (FMC) agarose gels [3.5% (w/v)

in 1xTBE buffer] were used to separate microsatellite reaction products for scoring all microsatellite loci in the progeny. Depending on the size of the microsatellite and the size difference between alleles, the running of the gels varied from 20 volt.hours/cm at 4° to 16 volt.hours/cm at room temperature. Ethidium bromide (0.4μg/mL) was incorporated into the gels which were scored from Polaroid (665) photographs of the fluorescing PCR products.

A number of different enzymes were tested for their ability to be detected using frozen mature leaf tissue as a source and starch gel electrophoresis for separation with only one polymorphism reliably detectable. The methods used for detecting enzyme activity and starch gel electrophoresis were based on those of Moran and Bell (1983).

2.2.3 Linkage analysis

All loci segregating 1:1 from both parents were tested for evidence of linkage to each 3:1 segregating locus using a chi-squared goodness of fit test. Only those progeny that were found to be homozygous absent at the 3:1 locus were used in the test where a 1:1 segregation pattern would be expected for an unlinked (1:1) marker. Linkage is indicated where a significant departure from the 1:1 pattern arises, in this study where $\alpha < 1 \times 10^{-4}$.

The program MAPMAKER Version 3.0b (UNIX) was used to determine linkage groupings and ordering of markers within linkage groups for loci segregating in a 1:1 pattern (Lander *et al.* 1987; Lincoln *et al.* 1992). Data for loci segregating 1:1 was entered as F_2 backcross data and markers were scored as either present (H), absent (A) or undetermined (-) for each offspring. Because MAPMAKER

recognises linkage in coupling phase only, each marker needed to be represented by an original and also an inverse (repulsion phase) form in the data set. This allowed linkages in repulsion to be recognised by MAPMAKER as defacto linkages in coupling to inverse markers.

The LOD score threshold for declaring linkage using MAPMAKER was calculated based on the maximum number of independently segregating ('unlinked') positions expected in the genome and the required type I error. This is based on the expected number of linkage groups, an estimate of the genome size and a prior definition of linkage in map units. The expected number of linkage groups is 11 since cytological observations suggest this to be the haploid number of chromosomes in *Eucalyptus* (Potts and Wiltshire 1997). An upper estimate for the size of the *Eucalyptus* genome from other mapping studies is approx. 1500cM. Linkage can be defined arbitrarily as two markers being less than 50cM (Kosambi) apart. The number of 'independently segregating positions' is thus the maximum number of positions that are 50cM or more apart in the genome (ie. 41). There are 820 unique pairwise linkage tests that can be made between these 41 positions. For a probability of Type I error of 0.01 for the genome as a whole, the appropriate LOD score threshold is - $\log_{10}$ (820/0.01) = 4.9. Thus a LOD score threshold of 4.9 for declaring linkage between markers was used with the "group" command of MAPMAKER with a consequent estimated probability of Type I error of 0.01 for each linkage map. Since the maximum recombination fraction at which linkage is declared will vary according to the number of progeny which have been scored in common for a marker pair, the recombination fraction parameter was left as non-restrictive (ie. r = 0.49) when using the "group" command of MAPMAKER.

For ordering of markers within each linkage group a subgroup of six or less markers were chosen that were well spaced from one another and had a minimum of missing data. The most likely marker orders for the subgroup were found using the "compare" command of MAPMAKER with a LOD 3.0 threshold for alternative marker orders. The "build" command of MAPMAKER was used to identify any remaining markers in a linkage group that could be added to the order established for the initial subgroup of markers at LOD 3.0. Framework marker orders were used as the basis for data checking using the error detection system in MAPMAKER and discernible errors corrected. Error correction did not affect marker ordering as tested using the "ripple" command of MAPMAKER on the corrected data sets. The "build" command was used after error correction to add any previously unplaced framework markers to the corrected data sets at LOD 3.0 and to add accessory markers at LOD 2.0. Several markers which had missing data and that significantly inflated the length of a linkage group due to dubious double crossovers were excluded from the framework and placed as accessories. Accessory markers were located on the framework map alongside the nearest framework marker. Unplaced microsatellite loci in a linkage group were assigned to the most closely linked framework marker using the "near" command of MAPMAKER.

2.2.4 Test for segregation distortion

All loci classified as segregating 1:1 were tested for distortion from this expected ratio using a chi-squared goodness of fit test. Loci segregating from one parent with significant departure from a 1:1 pattern at $\alpha = 0.05$ were compared with expected ratios for 2:1 and 3:1 segregation. To determine the frequency of

segregation distortion in the genomes of the two parents, the number of regions (rather than markers) expected to have distorted segregation was estimated for each map (a region is defined as a group of linked markers or a lone mapped marker). Estimates are based on the number of 'independently segregating positions' which is the size of each linkage group in cM divided by 50 rounded upwards and summed for the entire map. The expected number of regions with distorted segregation is the number of these 'independently segregating positions' multiplied by the threshold value used for declaring distortion as significant (eg. 0.05, 0.01). Since distortions from a 1:1 ratio can alter the probability of linkage between two markers, pairs of framework markers with strong distortion and with weak linkages were checked for any extreme alterations to their probability of linkage. Two point LOD scores for these marker pairs were also recalculated based on an adjusted θ value for no linkage.

2.3 RESULTS

2.3.1 Screening and scoring of molecular markers

Sixty-five RAPD primers out of 400 were selected for use in genotyping, approx. one out of every six screened. The primers amplified 326 RAPD markers segregating in an apparent 1:1 pattern with 173 inherited from the male parent and 153 inherited from the female parent, resulting in five markers segregating 1:1 per primer. Twenty loci were scored as being heterozygous in both parents and segregating in an apparent 3:1 pattern.

The results of screening the microsatellite loci are summarised in Table 2.1. From the 35 amplifiable loci, 25 were heterozygous in one or both parents with the male parent heterozygous at 21 loci and the female parent at 16 loci. Of the 11 EMBRA loci scored in common with *E. grandis* and *E. urophylla*, the male *E. globulus* parent was heterozygous at 10 and the female parent heterozygous at 8 loci. This compares with the parent trees in Brondani *et al.* (1998) where the *E. grandis* parent is heterozygous at 10 and the *E. urophylla* parent at all 11 of the loci scored in common. Although seven CSIRO microsatellites were found to be potentially mappable, only the three fully informative loci (ie. segregating from both parents) were assayed for all progeny and used in the linkage analysis. All EMCRC and EMBRA microsatellites that were found to be segregating from one or both parents were used for genotyping and linkage analysis.

All of the 21 microsatellite loci that were scored for segregation in the progeny originate from species of *Eucalyptus* in the subgenus *Symphyomyrtus* except for CSIRO-03 which is from *E. sieberi* belonging to the subgenus *Monocalyptus*. The

male parent was found to be heterozygous for 18 of these loci and the female parent heterozygous at 14 loci with 11 loci heterozygous in both parents. Thus across the two parental linkage maps the 21 microsatellite loci scored in the progeny provided a total of 32 markers to be tested for linkage with only one (EMBRA 10) remaining unlinked at LOD 4.9.

Table 2.1. The source, amplifiability and informativeness of microsatellite loci used for mapping.

Source	No. tested	No. amplifiable	Heterozygosity in *E. globulus* parents		
			Neither	*One*	*Both*
EMBRA[a]	20	15	4	4	7
EMCRC[b]	12	11	4	5	2
CSIRO[c]	10	9	2	4	3
total	42	35	10	13	12

[a]EMBRA = *Eucalyptus grandis/E. urophylla* - Brondani *et al.* (1998)

[b]EMCRC = *E. globulus* – Steane *et al.* (2001)

[c]CSIRO = *E. nitens/E. sieberi* – Byrne *et al.* (1996)

Of the three enzyme systems found to have scorable activity using frozen adult leaf material (MDH, AAT and GPI) only glucosephosphate isomerase 2 (GPI-2, E.C. 5.3.1.9) was polymorphic, being heterozygous in the male parent only.

2.3.2 Linkage maps

For the male parent (G164) a total of 192 loci segregating 1:1 were used in the initial linkage analysis. Thirteen linkage groups were defined over a range of LOD score values from 4.25 to 4.95 with 15 markers (7.6%) remaining unlinked. There were 101 framework markers ordered with 19 accessory markers added to

the framework map (Fig. 2.1). The 101 framework markers map to 90 positions covering a total distance of 1013cM (Kosambi mapping function). There are 77 intervals between markers on the linkage groups with an average size of 13cM and a maximum size of 37cM. Linkage group frameworks varied in size from two positions covering 5.7cM (Group 11) to 13 positions covering 195cM (Group 8). Three groups (5, 11 and 13) consist of only two framework markers each. At LOD 4.20 twelve linkage groups form as a consequence of Groups 4 and 5 coalescing and forming a new interval of 41.7cM whilst at LOD 4.0, Groups 1 and 2 form a single linkage group with a new interval of 69.0cM (dashed lines on Fig. 2.1). Thus at LOD 4.0, eleven linkage groups form, which is the observed haploid number of chromosomes in *Eucalyptus*.

For the female parent (KI2) 167 loci segregating 1:1 were used in the analysis. These formed into 11 linkage groups from LOD score thresholds ranging from 4.35 to 6.15 with 15 markers (approx. 9%) remaining unlinked at LOD 4.9. There were 97 framework markers at 75 positions on the linkage map of 11 groups with 11 accessory markers (Fig. 2.2). The framework spans 64 intervals covering 701cM with an average interval size of 11cM and a maximum interval size of 37.1cM (Kosambi mapping function). Three groups (4, 9 and 11) are composed of two framework positions only. The largest group (Group 3) has 18 framework positions and covers 148cM with the smallest group (Group 11) being 6.9cM. Using a LOD of 4.3 with the group command of MAPMAKER, ten linkage groups form rather than 11, with Groups 4 and 5 coalescing to form a new interval of 38.4cM (dashed line Fig. 2.2).

Fig. 2.1. and 2.2 Linkage maps of *Eucalyptus globulus* using RAPD and microsatellite markers.

Framework marker orders were determined using the program MAPMAKER at a LOD threshold of 3.0. Microsatellite loci are in bold text. Microsatellites that are not framework or accessory markers are bracketed and listed next to the nearest framework marker in the linkage group. RAPD loci amplified from UBC primers are listed by the UBC number followed by the ranked size of the amplified band from large to small. RAPD loci amplified with OP primers are listed using the OP designation followed by the ranked size of the amplified band. The isozyme locus, GPI-2 is underlined. The distance between framework markers is in Kosambi centimorgans on the left hand side of each linkage group. Markers that mapped to the same position are listed separated by a comma. Accessory markers that were ordered at LOD 2.0 are located in smaller text alongside the nearest framework marker with the two point distance in centimorgans from this marker. Loci which had distorted segregation at $\alpha = 0.05$ are followed by a single asterisk, those with distortion at the $\alpha = 0.01$ level are followed by a double asterisk. A dashed line between linkage groups indicates linkage at a reduced LOD score threshold.

Next page Fig. 2.1 Map of the male (G164) parent
Following page Fig. 2.2 Map of the female (KI2) parent

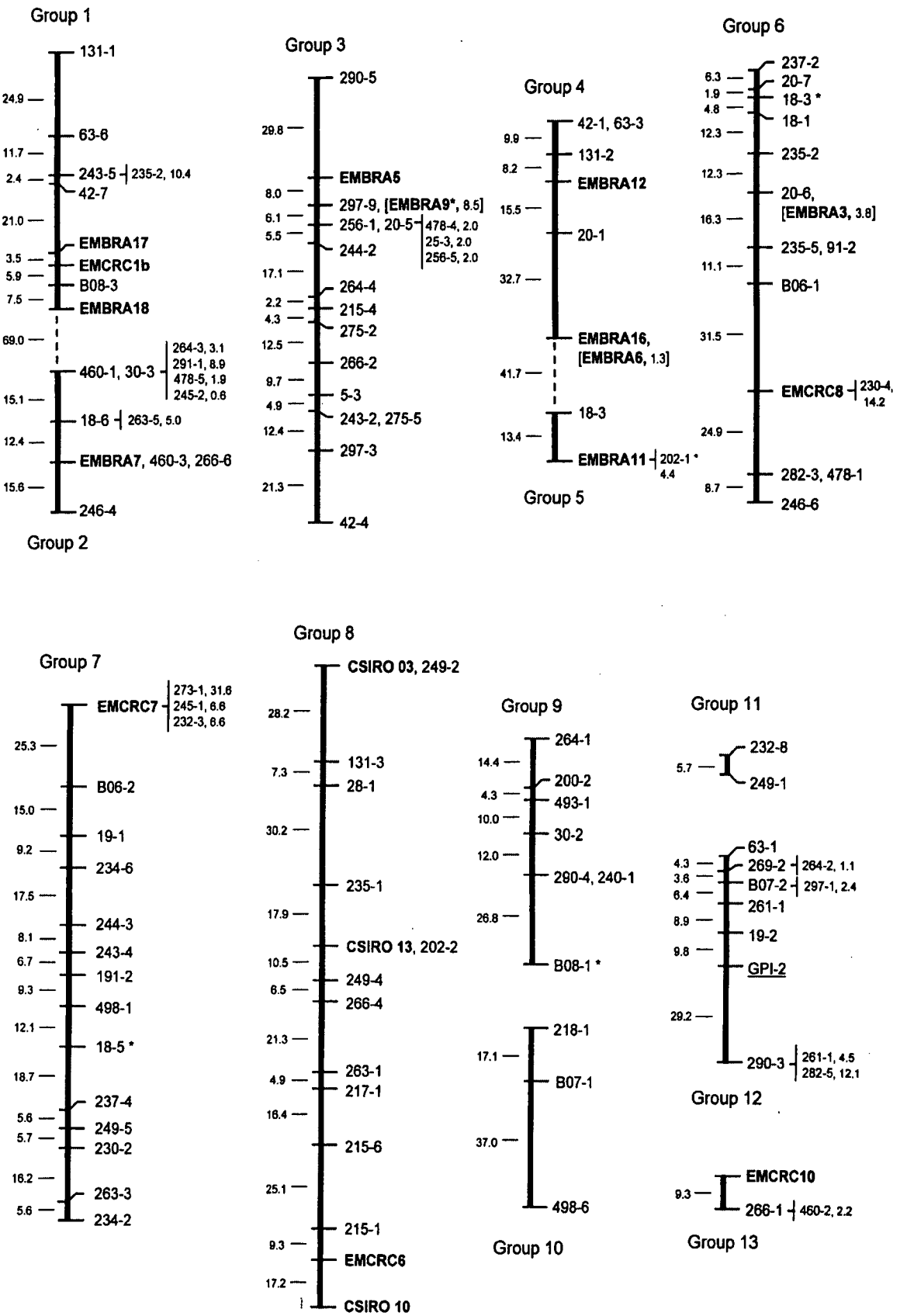


Fig. 2.1. Linkage map of the male (G164) parent

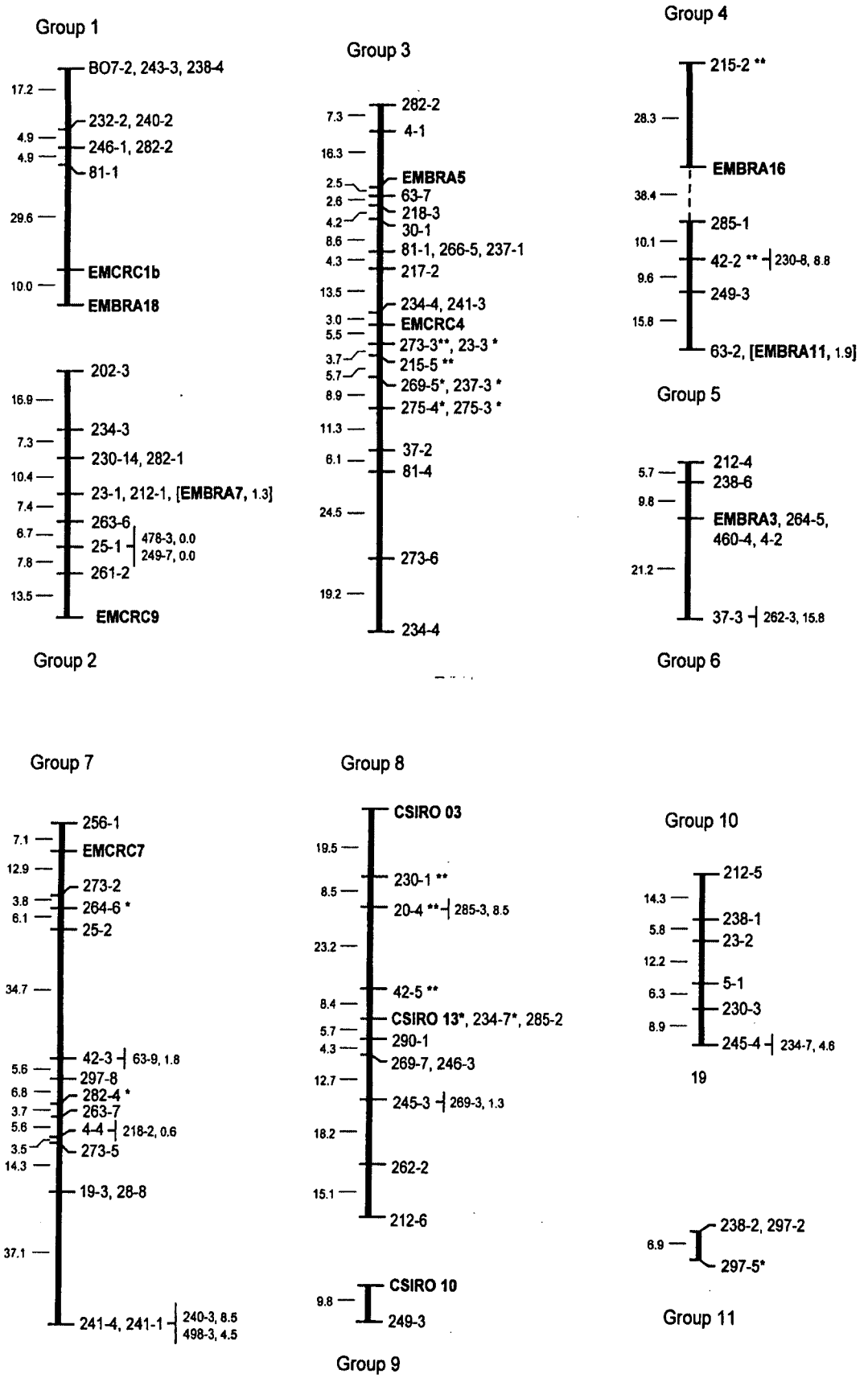


Fig. 2.2. Linkage map of the female (KI2) parent

2.3.3 Linkage group homology

The presence of microsatellite loci mapped in both parents indicates there is homology between eight linkage groups from the male parent with nine linkage groups from the female parent (Table 2.2). Groups one through to eight are proposed as homologous between parents, with group nine of the female parent homologous with one end of Group 8 of the male parent. Homology with four linkage groups of the *E. grandis/E. urophylla* maps of Brondani *et al.* (1998) is also suggested based on the sharing of EMBRA microsatellite loci (Table 2.2). In the four cases where more than one EMBRA microsatellite was found on an *E. globulus* linkage group, conservation of linkage of EMBRA loci was found between *E. globulus* and *E. grandis/E. urophylla*. Of the 20 RAPD loci segregating in a 3:1 pattern, 11 were found to link ($\alpha = 0.0001$) to framework loci segregating 1:1 in both of the parental maps. This indicated homology between five pairs of linkage groups of the two parents (Fig. 2.3). Homology is also indicated for each of these pairs of linkage groups on the basis of sharing of microsatellite loci (Fig. 2.3). There is also strong evidence for linkage of a 3:1 locus (226-3) to Group 10 of the male parent and an unlinked marker of the female parent.

Table 2.2. Homology of linkage groups between *E. globulus* parent trees and between *E. globulus* and *E. urophylla/E. grandis* based on the mapping of shared microsatellite loci.

Microsatellite locus	Male parent linkage group	Female parent linkage group	*E. urophylla/ E. grandis* group[a]
EMCRC1b	1	1	-
EMBRA17	1	-	9
EMBRA18	1	1	9
EMBRA7	2	2	9
EMBRA5	3	3	5
EMBRA9	3	-	5
EMBRA6	4	-	1
EMBRA12	4	-	1
EMBRA16	4	4	1
EMBRA11	5	5	1
EMBRA3	6	6	8
EMCRC7	7	7	-
CSIRO03	8	8	-
CSIRO13	8	8	-
CSIRO10	8	9	-

[a]*E. urophylla/E. grandis* linkage group no. from Brondani *et al.* (1998)

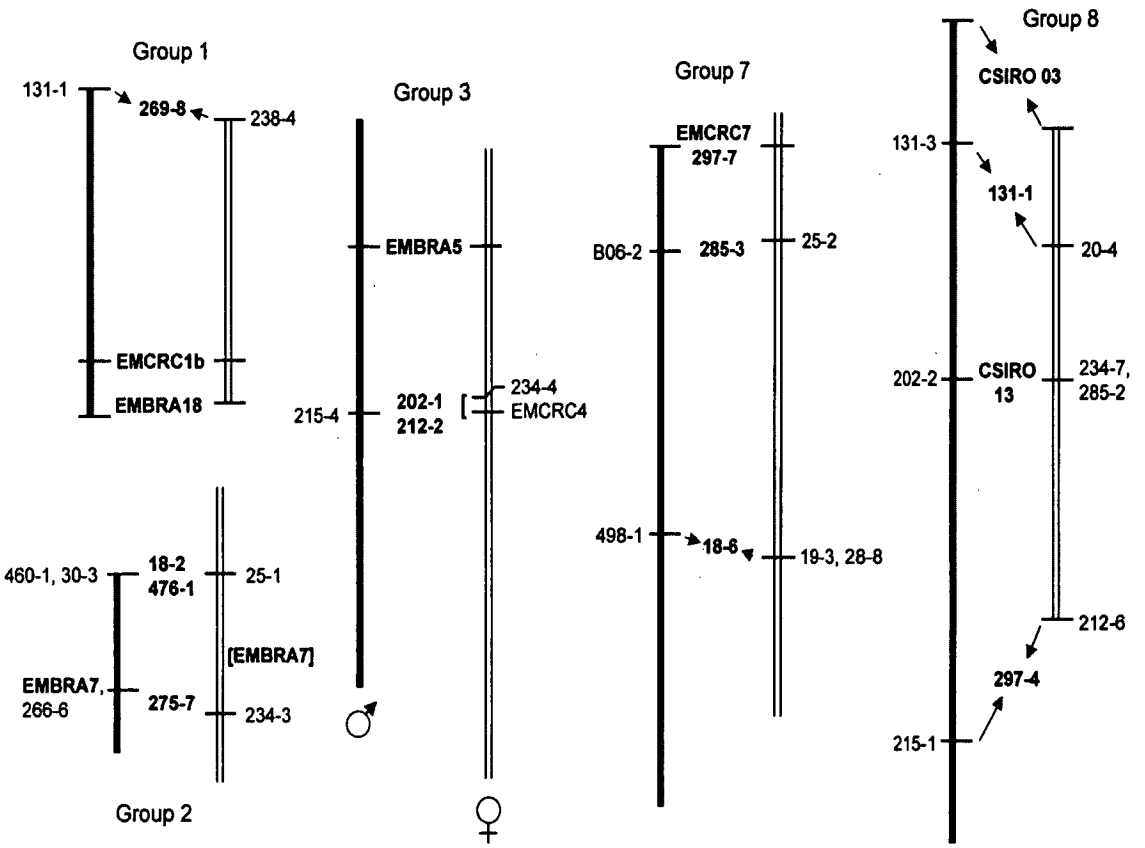


Fig. 2.3. Linkage of 3:1 segregating RAPD markers to loci on both parental maps. Linkage group homology inferred from 3:1 loci is supported in each case by the mapping of one or more fully informative microsatellite loci. RAPD loci segregating 3:1 are located between each pair of homologous linkage groups and numbered in bold text along with framework microsatellite loci. Framework RAPD loci which were most strongly linked to each 3:1 locus are shown in plain text. Otherwise coding of RAPD and microsatellite loci is as per Fig. 1.1 and 1.2.

2.3.4 Evidence for linkage based on homology

In several instances EMBRA microsatellite loci that are present on one group of the *E. grandis* and/or *E. urophylla* maps of Brondani *et al.* (1998) are present on two *E. globulus* linkage groups (Table 2.2). Linkage Group 9 of the *E. urophylla* map of Brondani *et al.* (1998) has microsatellite loci which occur on both linkage Group 1 (EMBRA 17 and 18) and linkage Group 2 (EMBRA 7) of the male *E. globulus* parent (Table 2.2 and Fig. 2.1). Similarly linkage Group 1 of both the *E. grandis* and *E. urophylla* maps has loci which occur on group 4 (EMBRA 6, 12 and 16) and Group 5 (EMBRA 11) of the male *E. globulus* parent (Table 2.2 and Fig. 2.1). It has been noted that at reduced threshold LOD scores linkage was detected between these two pairs of groups in the male parent (see Linkage Maps above). The evidence thus suggests that Groups 1 and 2 belong to a single linkage group and that Groups 4 and 5 similarly belong in a single group and that the arrangement of the EMBRA loci on these groups has been conserved between species.

Linkage Groups 1 and 2 of the female parent also share microsatellite loci (EMBRA 18 and 7) with Group 9 of *E. urophylla* and linkage Groups 4 and 5 have microsatellite loci (EMBRA 16 and 11) that map to Group 1 of both the *E. grandis* and *E. urophylla* maps (Table 2.2, Fig. 2.2). At a reduced threshold LOD score linkage is detected between Groups 4 and 5 of the female parent (see Linkage Maps above) but not, however, between Groups 1 and 2. Again the evidence supports conservation in the arrangement of EMBRA loci on Groups 4 and 5. Linkage between Groups 1 and 2 in the female parent can be inferred from

the linkage between the homologous groups in the male parent and from the arrangement of EMBRA microsatellites in both the male parent and *E. urophylla*.

The three CSIRO microsatellites map to one group in the male parent (Group 8) but to two groups (Groups 8 and 9) in the female parent (Table 2.2, Fig. 2.1 and 2.2). Since the linkages in Group 8 of the male parent are of high likelihood and assuming the distribution of microsatellites is the same between the two parents it is likely that Groups 8 and 9 of the female parent belong to one linkage group. If all the linkages inferred from homology in both the *E. globulus* parents are taken as correct then the male parent would have eleven linkage groups and the female parent eight.

2.3.5 Segregation distortion

There are six markers with segregation distortion at $\alpha = 0.05$ segregating from the male parent. These map to five regions and include three framework markers (Fig. 2.1, asterisked markers). With 28 'independently segregating positions' on the map of the male parent (see Materials and Methods), 1.4 regions are expected to have distorted segregation at $\alpha = 0.05$. There is thus approx. three and a half times the number of regions expected to have distortion at $\alpha = 0.05$. There were no markers with segregation distortion at $\alpha = 0.01$ segregating from the male parent. (If it is assumed that segregation patterns for each marker are derived from random and independent events, then the expected number of markers with segregation distortion at $\alpha = 0.05$ is 9.6 and based on this method of assessment there would be fewer markers observed to have distorted segregation than expected.)

In the overall data set of the female parent there are twenty-three markers with distorted segregation at $\alpha = 0.05$. These map to seven regions and include 17 framework markers (Fig. 2.2, asterisked markers). The female parental map of 701cM has a total of 19 'independently segregating positions' with 0.95 regions expected to have segregation distortion at $\alpha = 0.05$. There are thus seven times the expected number of regions with distorted segregation at $\alpha = 0.05$. There are eight markers overall with distorted segregation at a threshold of $\alpha = 0.01$. These map to four regions with three being framework markers on linkage Group 8 and one framework marker on each of Groups 3, 4 and 5 (Fig. 2.2, double asterisks). Since there are 0.19 regions expected to have distorted segregation at $\alpha = 0.01$ there are approx. 21 times the expected number of regions with segregation distortion at $\alpha = 0.01$. (Assuming that segregation of each marker results from random and independent events, then 8.35 markers would be expected to have segregation distortion at $\alpha = 0.05$ and 1.67 markers at $\alpha = 0.01$. Using this assumption there are 2.8 times the expected number of loci with distortion at $\alpha = 0.05$ and 4.8 times the expected number of loci with distortion at $\alpha = 0.01$.)

For the regions with segregation distortion at $\alpha = 0.01$ in the female parent the region on Group 3 and the region on Group 8 have markers which have been scored on the full array of progeny and there are linkages in repulsion with markers that also have distorted segregation. Thus it is extremely unlikely that the segregation distortion is due to 'unreliability of RAPD markers', at least for these two regions. Lone markers, both of which gave strong unambiguous banding, represent the other two regions with strong distortion in the female parent. It

would thus appear very likely that the observed distortions are not artifactual and the excess segregation distortion requires explanation.

All marker loci with significant segregation distortion were found to have segregation ratios between 1:1 and 2:1, as distinct from the 3:1 ratio expected from an unlinked duplication of a locus. None of the linkages involving framework markers with segregation distortion were found to have probabilities (of linkage) which were unduly affected by the distortion. In the most severe case on Group 8 of the female parent (Fig. 2.2), RAPD loci 42-5 and 20-4 are 23.2cM apart. For two loci with the same degree of distortion, the probability of linkage at this level or closer is 7.8×10^{-10}, which although 70 times larger than two loci segregating in a perfect 1:1 ratio, still provides very strong evidence for linkage. The two point LOD score calculated using $\theta = 0.5$ for no linkage was 11.05 and an adjusted LOD score of 10.96 was obtained for $\theta = 0.455$.

2.3.6 Estimates of genome size

Some idea of the completeness of the two maps overall can be gained from the fact that 32 out of 33 microsatellite and allozyme markers were placed into linkage groups with only one marker (EMBRA 10) remaining unlinked at LOD 4.9. For the male parent all 19 of the non-RAPD markers were placed into linkage groups and for the female parent one marker out of 14 remained unlinked at LOD 4.9. A method of estimating the overall length of the genome as detailed in Vallejos *et al.* (1992) was used. The following formula was used for estimation: $G = 2MX/K$ where G is the estimate of genome size, M is the no. of locus pairs, X is the largest estimated map distance value among the K observed no. of locus pairs that are linked with a LOD score of Z or greater ($Z = 4.9$). This method gave

an estimated genome length for the male parent of 1277cM with 79% of the genome covered by the framework map. The female parent was estimated to have a genome of size 1133cM with 62% of the genome estimated to be covered by the framework map.

2.4 DISCUSSION

2.4.1 Utility of microsatellite loci

Microsatellite loci are extremely useful for the identification of linkage group homology and for enabling the integration of linkage information (Dib *et al.* 1996; Dietrich *et al.* 1996). In this study linkage analysis of eleven fully informative microsatellite loci has enabled the identification of homology between eight (pairs of) linkage groups from the RAPD maps of the parents. In addition homology with four linkage groups of *E. grandis/E. urophylla* have been identified based on the mapping of eleven EMBRA microsatellite loci (Brondani *et al.* 1998). In all cases where it could be examined, linkages between EMBRA microsatellites that were found in *E. grandis/E. urophylla* were conserved in *E. globulus*. The close correspondence of microsatellite distributions may be a reflection of the fact that all three of these species belong to the subgenus *Symphyomyrtus*, although *E. globulus* belongs in a different section to *E. grandis* and *E. urophylla*. In total 21 microsatellite loci from five different species of *Eucalyptus* have been placed on the two parental maps detailed here.

A potential outcome of identifying linkage group homology between maps is that the arrangement of loci on one map can be used to infer linkage between groups on the second map. In this study such homology has been used to support several weak linkages found between linkage groups. The two lines of evidence lend support to one another to indicate that these linkages are likely to be correct. Importantly homology was based on a between species alignment of microsatellite loci. This illustrates the potential power of microsatellites to draw upon mapping information from both within and between species to aid in map construction.

This information could be used in strategies to search for markers to fill in gaps in a linkage map, for example using bulked segregant analysis (Michelmore *et al.* 1991).

Using homology to infer linkage has had a considerable influence on map construction in this study, impacting on the number of linkage groups and the structure of the maps. If all the linkages inferred from homology in both the *E. globulus* parents are taken as correct then the male parent would have eleven linkage groups and the female parent eight. Cytological studies indicate that n = 11 in *Eucalyptus* (Potts and Wiltshire 1997). So for the male parent, eleven linkage groups matches the number expected based on chromosome counts. However for the female parent there are three linkage groups less than expected from chromosome counts. It is likely that by scoring additional RAPD markers inherited from the female parent these three linkage groups would be defined, especially since some of the 15 unlinked markers (approx. 9%) probably belong to these three groups.

2.4.2 Mapping in *Eucalyptus* using an intraprovenance cross

The levels of heterozygosity and variation within a provenance in *E. globulus* are clearly sufficient to allow the construction of RAPD maps from an intraprovenance cross. Since this species is cultivated as a pure species, the construction of such maps from intraspecific and intraprovenance crosses will be required if QTL detection and marker assisted selection is to be carried out as part of *E. globulus* breeding programs. Since the two parents of the cross in this study originate from the same island provenance, it might have been expected that map construction would be hampered due to low genetic

divergence of the parents and a consequent paucity of RAPD loci segregating in a fully informative manner. However if there is low genetic divergence between the two parents in this study it has not proven to be a significant hindrance and has been overcome by using the easy to implement strategy of screening a large number of RAPD primers and selecting a subset which detect a maximum number of polymorphisms. Approximately one in every six primers screened was selected for genotyping, yielding 5 loci segregating 1:1 on average per selected primer. This compares with Grattapaglia and Sederoff (1994) where half of the screened primers were used for genotyping, yielding an average of 3.7 markers per primer and Verhaegen and Plomion (1996) where almost three-quarters of screened primers were used for genotyping, yielding 3.2 markers per primer. In both these cases an *E. grandis* x *E. urophylla* interspecific cross was used.

2.4.3 Segregation distortion

Commonly, the frequency of segregation distortion expected due to chance is calculated using the total number of markers that have been scored in the data set. However, since linked markers are not independent, a more meaningful method of calculating the degree of expected segregation distortion for linkage maps may be to calculate the expected number of 'regions' with segregation distortion. The number of linkage groups and their sizes determines the expected number of regions. In this study this has been shown to provide a different basis on which to judge the expected extent of segregation distortion. For example, in the male parent at $\alpha = 0.05$, the number of regions observed to have segregation distortion was five, the number expected was 1.4 (28/20) with the ratio of observed to

expected being 3.5 (somewhat more than expected) whilst for calculations based on the number of markers, the ratio of observed (six markers) to expected (192/20 = 9.6) is 0.625 (fewer than expected). The method of calculating the extent of segregation distortion can thus have a significant bearing on the assessment of whether excess distortion is occurring.

In this study it has been found that the number of regions with significantly distorted marker ratios in the map of the female *E. globulus* parent in particular is much greater than would be expected by chance alone. This greater than expected frequency/degree of skewing of Mendelian segregation ratios is not uncommon in plants (Zamir and Tadmor 1986; Bradshaw and Stettler 1994), with *Eucalyptus* being no exception (Byrne *et al.* 1994; Vaillancourt *et al.* 1995a; Verhaegen and Plomion1996; Marques *et al.* 1998).

A number of selection based genetic mechanisms have been suggested to explain these distorted marker ratios. These include incompatibility systems (Gebhardt *et al.* 1991), preferential chromosome loss (Vaillancourt and Slinkard 1992), expression of genetic load (Sorensen 1969; Bradshaw and Stettler 1994; Vaillancourt *et al.* 1995a), meiotic drive (Gillet and Gregorius 1992) and haploid expressed deleterious alleles. It is to be noted however that the four regions with strong distortion ($\alpha = 0.01$) in the female parent in this study do not align with regions of distortion in the homologous linkage groups of the male map. This excludes incompatibility as an explanation since incompatibility systems in higher plants usually operate to exclude fertilisation by the male or in the case of rare haplo-homophasic systems would operate to cause distorted segregation in both sexes (Gillet and Gregorius 1992). If genetic load is considered as an explanation

for the observed segregation distortion then a deleterious recessive allele must be segregating from the parent with the distortion. The second parent could be either heterozygous or homozygous for the deleterious recessive. However the observed non-alignment of distorted regions excludes the possibilty that both parents are heterozygous for a deleterious recessive allele at the same locus since the genetic maps of both parents would be expected to have distorted segregation in the same region. To invoke genetic load as an explanation the male parent in particular would need to be homozygous for deleterious recessive alleles at several loci which is probably unlikely as it is a naturally established tree. Genetic load would therefore appear not to be a satisfying explanation for the segregation distortion.

Chromosome loss is also an unlikely explanation for the distortion as it is usually only considered in cases where one parent is an interspecific or intersubspecific hybrid. It is to be noted however that hybridisation is common in the genus *Eucalyptus* (Potts and Wiltshire 1997), and it is not known if there are small chromosomal rearrangements between and within species. Meiotic drive and haploid expressed deleterious alleles remain as possible explanations for segregation distortion of any region since there is no evidence for selection operating at any particular stage of development. Other models based on post-fertilisation selection, for example selection for co-adapted allelic combinations, might also explain the skewed segregation ratios.

2.5 CONCLUSION

RAPD maps were readily constructed for two parents of an intraprovenance cross of *E. globulus* demonstrating the utility of RAPD markers for map construction in *Eucalyptus* from nonwide crosses. Microsatellite markers which originate from several *Eucalyptus* species have been mapped enabling the identification of homologous linkage groups between *E. globulus* and *E. grandis/E. urophylla*. The conservation of linkage of microsatellite loci and trasnsfer of loci between crosses in different species indicates the potential for the construction of a consensus map based on microsatellites. Segregation distortion in the genome was observed to a greater extent than expected by chance and a biological cause is suspected although several of the usual genetic mechanisms to account for the distortion appear to be unlikely.

CHAPTER THREE

THE DETECTION OF QUANTITATIVE TRAIT LOCI (QTL) FOR TRAITS OF COMMERCIAL IMPORTANCE IN A *EUCALYPTUS GLOBULUS* CROSS

3.1 INTRODUCTION

In common with crop plants, most of the traits of economic importance to tree breeding are quantitative in nature (Namkoong *et al.* 1988). Most of these can be considered as continuously variable traits with the variance arising from some combination of environmental, genetic or interaction effects. Until recently, information has not been available on the number of genetic factors controlling such traits, their size of effect, epistatic interaction or their distribution in the genome. Genetic theory has handled this lack of information by considering quantitative traits to be controlled by many genes of small effect – the infinitesimal or polygenic model. However with the advent of molecular markers, genetic linkage maps have been constructed and utilised to detect loci contributing to the variance of quantitative traits, so-called quantitative trait loci or QTLs. In *Eucalyptus* alone there have been reports of QTL detection for vegetative propagation traits (Grattapaglia *et al.* 1995; Marques *et al.* 1999), wood specific gravity/ density (Grattapaglia *et al.* 1996; Verhaegen *et al.* 1997), growth

(Grattapaglia *et al.* 1996; Verhaegen *et al.* 1997) and form (Verhaegen *et al.* 1997), seedling height and leaf area (Byrne *et al.* 1997a), frost tolerance (Byrne *et al.* 1997b) and foliar oil composition (Shepherd *et al.* 1999).

QTL detection studies add to the fundamental understanding about the genetic control of quantitative traits both in general and in particular. However there is great interest in the detection of quantitative trait loci (QTL) in forest trees due to the potential of marker assisted selection (MAS). It is hoped that MAS could improve the accuracy of selection for traits with low heritabilities and reduce the time to evaluate performance by partial or complete replacement of phenotypic evaluation with marker genotyping. The selection of superior parent trees is usually considered the objective in tree breeding, but it has also been suggested that marker based selection could be applied to tree seedlings before deployment in plantations, as a cost effective method to increase gains (Kerr *et al.* 1996). Recently a benefit cost analysis of MAS for growth and wood density in *Pinus radiata* has indicated the potential profitability of MAS based on propagation by cuttings of marker selected genotypes (Wilcox *et al.* 2001a).

No doubt a necessary step in the process of MAS is the detection of QTLs in relevant crosses of the species of interest. However, further information on most detected QTLs will be required before MAS can be successfully implemented. Reported QTL effects are often based on small population sizes and many require validation in other populations (Sewell and Neale 2000). An issue of major concern that requires investigation before QTL information can be utilised for MAS, is QTL stability; in different genetic backgrounds, with time and across environments (Bradshaw and Grattapaglia 1994).

Eucalyptus globulus is grown mainly for pulpwood production and the main traits identified for selection in breeding programs usually include wood density, pulp yield and volume (growth) (Greaves and Borralho 1996; Borralho 2001). However other traits have been studied in *E. globulus* with the possibility of inclusion in breeding programs, these include: frost tolerance (Volker *et al.* 1994), rooting ability (Borralho and Wilson 1994; Marques *et al.* 1999), early flowering (Chambers *et al.* 1997), survival/mortality (Chambers and Borralho 1997), susceptibility to fungal leaf disease (Dungey *et al.* 1997), wood chemistry (Rodrigues *et al.* 1999; Miranda and Pereira 2001), autumn gum moth defoliation (Jones *et al.* 2002), sawfly attack (Jordan *et al.* 2002), resistance to marsupial browsers (O'Reilly-Wapstra *et al.* 2002) and fibre morphology and pulp/paper properties (Miranda and Pereira 2002; Wimmer *et al.* 2002). In the present study measurements of growth, wood density, bark thickness and early flowering have been analysed for QTL detection. Of these growth rate has been the most thoroughly studied and most commonly incorporated into objectives for breeding programs (Borralho 2001).

Stem diameter commonly suffices as a measure of growth in lieu of volume since it is easy to measure. A recent estimate of heritability of 0.20 for stem diameter in *E. globulus* at age 4 years was significant but relatively low (600 families across 5 sites) (MacDonald *et al.* 1997) and close to an average heritability for diameter (h^2 = 0.21) calculated for *E. globulus* based on open pollinated families (Lopez *et al.* 2002). Even these relatively low estimates may be exaggerated however, as heritability estimates of growth based on open pollinated progeny can be inflated due to inbreeding depression (Potts *et al.* 1995; Hodge *et al.* 1996). Although average stem diameter of *E. globulus* has been shown to vary considerably

between sites, genetic correlations across sites are often high, indicating that family performance is relatively stable across sites and that genotype x environment interaction is small (MacDonald *et al.* 1997). Generally there is an increase in the estimated heritability of diameter with increasing age for both open pollinated (Lopez *et al.* 2002) and control pollinated populations (Volker 2002). However in general a very strong age-age genetic correlation is found for growth (Lopez *et al.* 2002). If QTL expression follows these patterns, QTLs for growth would not be expected to exhibit large QTL x environment effects and overall QTL effects would be expected to be larger at later ages compared with earlier ages.

Wood density has been linked to pulp, paper and processing characteristics and it is recognised as an important trait for selection in breeding of eucalypts for pulpwood (Greaves *et al.* 1997). Basic density is one of the more straightforward wood properties to measure and is defined as the dry weight of wood per unit of green volume (Greaves *et al.* 1995). However the direct measurement of density can be prohibitively expensive for a breeding program where many trees may need to be measured. A cheap alternative to direct density measurement is the use of a Pilodyn, which is a hand held instrument that operates by driving a steel pin into a debarked tree stem with a known force. The further the pin penetrates, the less dense the wood. Two observations of the measurement of the distance of penetration (mm) have been shown to provide a reliable indication of the phenotypic value for a tree (Greaves *et al.* 1995). The relationship between Pilodyn penetration and basic density of wood appears to be linear and relatively constant across sites (MacDonald *et al.* 1997) and between ages (Greaves *et al.* 1995). Genetic correlations between Pilodyn measurements and density in

Eucalyptus have been very high with -0.93 reported for *E. globulus* (Dean *et al.* 1990) and -0.84 for *E. nitens* (Greaves *et al.* 1995). Pilodyn penetration across sites has been shown to be highly genetically correlated, both within and between races (MacDonald *et al.* 1997; Muneri and Raymond 2000; Lopez *et al.* 2002). It has thus been concluded that genotype by environment interactions for Pilodyn penetration are small (Lopez *et al.* 2002).

Early flowering (flowering precocity) is a trait that has come under consideration for inclusion in tree breeding programs because its selection could potentially reduce the generation interval, which is a limiting factor in tree breeding (Griffin 1989). Chambers *et al.* (1997) found this trait to be under a high degree of genetic control in *E. globulus* ssp. *globulus*, with a heritability averaged across four sites of 0.59 (based on 600 families). The abundance of flower buds in three-year-old trees (bud abundance), was measured for the present study at roughly the same age as flowering precocity was measured in the study by Chambers *et al.* (1997). Bud abundance as measured in the present study, would be expected to share some of the genetic control of flowering precocity, since bud abundance is an extension of the binary trait. No studies of bud abundance have been reported for *E. globulus*, however the results for flowering precocity should be indicative. Significant environmental effects were found for flowering precocity with 1% to 25% of trees flowering by age four years depending on site (Chambers *et al.* 1997). Provenance (location) performance was consistent across trial site with no significant genotype by environment interaction (Chambers *et al.* 1997). Precocious flowering and growth appeared to be genetically, relatively independent at the assessment age of four years, however there was a slight

tendency for precocious flowering genotypes to have denser wood (Chambers *et al.* 1997).

Relative bark thickness is the proportion of the stem diameter that is due to the presence of bark. It has been hypothesised that relative bark thickness may influence traits of significance for plantation grown trees such as susceptibility to damage from pests such as sawflies (*Perga affinus*) and protection from environmental stress such as drought (Dutkowski and Potts 1999). Heritability estimates for relative bark thickness are moderate ($h^2 = 0.3$) with high genetic correlation across sites (Lopez *et al.* 2002). The relative thickness of bark has been observed to be constant across ages and sites (Kelly 1997).

The present chapter is a report of a QTL detection study of growth, wood density, relative bark thickness and early flowering, using a single full-sib family from an intraprovenance cross of *E. globulus*. The same family (G1025) has been used for the construction of parental linkage maps using RAPD and microsatellite markers (Chapter 2 of this thesis). The two-generation pedigree has been analysed using a pseudo-testcross model in which the effects contributed by each parent are analysed separately. A preliminary investigation of QTL stability with time and across environments has also been carried out, and both incremental and cumulative growth was analysed.

3.2 MATERIALS AND METHODS

3.2.1 Population used for QTL analysis

A single full sib family of *E. globulus* trees, family G1025, was included in the QTL analysis. Trait measurements were made on up to 155 trees belonging to this family. A linkage map based on 165 trees from family G1025 was constructed using RAPD and microsatellite markers (Chapter 2 of this thesis). The G1025 family was derived from a control pollinated cross of two parents both originating from the King Island provenance of *E. globulus* (Chapter 2 of this thesis). The family was planted out as part of the CSIRO/NFP hybrid trial, which included progeny from an incomplete factorial crossing design of *E. globulus* parents belonging almost exclusively to the two provenances of King Island and Taranna (Vaillancourt *et al.* 1995; Hodge *et al.* 1996). The crossing design included 26 male and 8 female parent trees. This resulted in progeny from a total of 177 control pollinated outcrossed families being planted out at various field trial sites. Trait measurements from all of these families were included in the analysis of the trait data to maximise the accuracy of the estimation of environmental effects. The G1025 family was chosen for QTL analysis because it had good growth characteristics and was the largest family present in the trial.

3.2.2 Trial sites and trial design

Trees of the G1025 family were grown at eight trial sites across southern Australia. Leaf samples were taken for DNA extraction from all surviving trees at each of these sites at years four or five. DNA extraction, marker genotyping and mapping of marker loci, were carried out as per chapter 2 of this thesis on 165 trees from the G1025 family. Quantitative trait measurements were made at seven of the eight trial sites on 155 of the trees that were genotyped. Of the seven trial sites at which quantitative trait measurements were taken, one was in Western Australia at Manjimup (W. Aust.), two were in Victoria at Flynn and Mansfield and four were in Tasmania at Boyer, Franklin, West Ridgley (Ridgley) and Parkham. There is considerable variation in the location of sites, their altitude and annual average rainfall (Table 3.1). Other important factors that are likely to vary between sites are soil type, soil fertility and temperature regimes.

With the following exceptions there were 20 trees from the G1025 family planted per trial site with 5 trees per plot in four replicates (Table 3.1). The site at Boyer was planted in plots of ten trees with four replicates. However there was a high mortality rate in the first year at this site and only 50% of trees in family G1025 survived (20 trees). At Parkham there was an additional ten G1025 trees planted as fillers either singly or in pairs, sometimes in the same replicate and incomplete block as a 5 tree plot, in which case they were included in that plot for the analysis. At the Ridgley site, there is an additional planting of 25 trees from the G1025 family in a plot (archive) alongside the trial. These were included in the analysis as if planted in the nearest replicate and incomplete block. At Manjimup (WA) there were 21 G1025 trees planted in plots of three in seven replicates.

Table 3.1. Location and design of NFP/CSIRO trial sites with plantings of the *E. globulus* family G1025.

	Boyer (TAS)	*Flynn (VIC)*	*Franklin (TAS)*	*Mansfield (VIC)*	*Parkham (TAS)*	*Ridgley (TAS)*	*Manjimup (WA)*
Latitude	42° 46'	38° 18'	43° 04'	36° 55'	41° 26'	41° 09'	34° 12'
Longitude	147° 07'	146° 40'	146° 53'	146° 14'	146° 37'	145° 46'	116°01'
Altitude	40m	170m	370m	950m	205m	185m	240m
Mean annual rainfall (mm)	532	620	865	1000-1250	1025	1200	1069
Replicates	4	4	4	4	4	4	7
Incomplete blocks /Rep	n.a.	11	9	11	13	15	11
Plots/block	30/rep	18	13	8	14	20	11
Trees/plot	10	5	5	5	5	5	3
G1025 trees	20	18	17	8	28	43	21

3.2.3 Traits

The following nine traits were measured on *E. globulus* trees representing 177 families at the given age intervals: Height age 1 year (ht1); height age 2 years (ht2); diameter of the tree stem at breast height (dbh) age 2 years (dbh2); dbh age 3 years (dbh3); dbh age 4 years (dbh4); dbh age 6 years (dbh6); mean Pilodyn penetration (mm) age 6 years (pilo6), based on the mean of two measurements; bark thickness (in mm) age 6 years (b); bud abundance (early flowering) at age 3 years measured on a scale from 1 to 6 (buds3) with 1 = no buds, 2 = 1-10 buds, 3 = 11-100, 4 = 101-1000, 5 = 1001-10,000 and 6 = 10,000 or more. Bark thickness (b) in mm, was transformed to a proportion of the stem diameter (relative bark thickness) according to the following equation: bark6 = (2 x b)/dbh6.

In addition, three traits measuring relative incremental growth (rig) were derived from the measurements of height and stem diameter. Relative incremental growth is the change in relative performance of each tree between two ages of measurement. It was calculated for each tree by subtracting the plot residual (see section 3.2.4) of an earlier growth measurement from a later measurement. The following relative incremental growth traits were used in QTL analysis: rig for height between years 1 and 2, ht2-1; rig for stem diameter at years 2 and 4, dbh4-2; and stem diameter at years 4 and 6, dbh6-4.

The number of trees in the G1025 family that were measured at each site for each of the nine measured traits is shown in Table 3.2. From Table 3.2 it can be seen that the minimum number of G1025 trees measured for a trait was 84 for dbh3. There was a maximum of 155 trees measured for any trait with a minimum of seven trees measured at any one site for a trait.

Table 3.2. The number of trees in the G1025 family measured for each trait at each site used for QTL detection.

Site	*ht1*	*ht2*	*dbh2*	*dbh3*	*dbh4*	*dbh6*	*pilo6*	*bark6*	*buds3*	*Total*
Boyer	20	20	18	20	20	20	18	18	20	20
Flynn	18	18	18	-	18	18	18	18	-	18
Franklin	17	17	17	17	17	17	17	17	17	17
Mansfield	8	8	-	-	-	8	7	7	-	8
Parkham	28	28	28	28	18	28	28	28	28	28
Ridgley	19	19	19	19	19	43	43	43	19	43
West. Aust.	21	21	21	-	-	21	21	21	21	21
All sites	131	131	121	84	92	155	152	152	105	155

3.2.4 Removal of environmental effects and standardisation of trait values

Family G1025 was grown as part of a large trial with multiple sites at disparate locations (Table 3.1) and with expected differences in environmental conditions such as soil, rainfall, humidity, temperature and biological components. The environmental heterogeneity would be expected to significantly increase the variance of the main traits of interest in this study – growth (stem diameter), wood density (Pilodyn penetration) and flowering (bud abundance). Environmental heterogeneity would also be expected between replicates within each site. If this environmental contribution is not accounted for, the power to detect segregating QTLs for these traits may be greatly reduced. The estimation and removal of the environmental contributions to trait variances would on the other hand help to maximise the power for QTL detection. Thus an analysis was undertaken to remove environmental effects from the trait data to increase the power for QTL detection.

A relatively large number of outcrossed *E. globulus* trees were planted and measured for each trait at each site (Table 3.3). These measurements were utilised in the analysis to provide as much information as possible for estimating between and within site environmental effects. Tree measurements from a total of 177 families that resulted from controlled pollinated outcrosses (but excluded self-pollinated and open pollinated families) were used in the analysis to remove environmental effects from the G1025 trait data. The number of individual tree measurements used in the analysis to remove environmental effects is given in Table 3.3 broken down by site. The number of families these represented (including G1025) is given in Table 3.4.

Table 3.3. The number of *E. globulus* control pollinated outcrossed trees at each site measured for each trait and the total number of trees included in the analysis for each trait.[a]

Site	*ht1*	*ht2*	*dbh2*	*dbh3*	*dbh4*	*dbh6*	*pilo6*	*bark6*	*buds3*	*planted*
Boyer	565	558	555	551	552	540	82	82	551	800
Flynn (Vic)	2140	2133	2133	-	2112	2066	1365	1365	-	2253
Franklin	1394	1394	1394	1386	1375	1318	83	83	1380	1466
Mansfield (Vic)	471	331	-	-	-	303	24	24	-	911
Parkham	2670	2656	2656	2633	2615	2517	1661	1662	2631	2761
Ridgley	3262	3259	3259	3231	3196	3147	2092	2092	3231	3616
West. Aust.	1906	1898	1898	-	-	1830	1827	1827	1898	1998
All sites	12408	12229	11895	7801	9850	11721	7135	7135	9691	13805
In analysis[b]	12200	12012	11682	7572	9431	11401	6961	6961	9246	n.a.

[a] The analysis of each trait was undertaken to remove environmental effects from the data (see text this section).

[b] Some trees measured for a trait were excluded from the analysis for various reasons (see text this section).

Table 3.4. The number of *E. globulus* families represented by measurements for each trait at each trial site

Site	*ht1*	*ht2*	*dbh2*	*dbh3*	*dbh4*	*dbh6*	*pilo6*	*bark6*	*buds3*	*Total*
Boyer	20	20	20	20	20	20	13	13	20	20
Flynn (Vic)	114	114	114	-	114	114	114	114	-	114
Franklin	76	76	76	76	76	76	23	23	76	76
Mansfield (Vic)	56	56	-	-	-	56	9	9	-	57
Parkham	139	139	139	139	139	139	139	139	139	139
Ridgley	173	173	173	173	173	173	173	173	173	173
West. Aust.	98	98	98	-	-	98	98	98	98	98
All sites	177	177	177	177	177	177	177	177	177	177

Individual tree measurements were excluded from the analysis to remove environmental effects where there were obvious mistakes made in recording of observations, where trees were severely damaged and generally if trees were fillers in the trial layout.

It can be seen from Table 3.3 that measurements for each trait were not always made at every site. Three traits were measured on virtually all trees at all sites: ht1, ht2 and dbh6. Dbh2 was measured on all trees at six sites, dbh3 was measured on all trees at four sites and dbh4 was measured on all trees at five sites (Table 3.3). Pilodyn penetration and bark thickness were measured at all sites but generally only on a subset of individuals (except WA where all trees were measured). The subsets measured for Pilodyn penetration and bark thickness always included each plot of family G1025 trees and at least the plots either side of each G1025plot. Early flowering (buds3) was assessed for all trees at five sites (Tables 3.3 and 3.4).

To remove site effects, the unadjusted quantitative measurements were standardised separately for each site using the SAS procedure Proc Standard (SAS Institute 1989). This set the mean value for each trait at each site to zero and the standard deviation to one. For each trait the standardisation was carried out considering all available measurements (all trees in all families measured) at each site (Table 3.3). The standardisation removed between site differences in the trait data and gave equivalent variances to the data from each site.

The model for deriving the standardised site residuals can be defined as follows:

Model for site residuals

$$\text{trait} = \text{site} + \text{standardised residual } (\mu = 0, \sigma = 1)$$

After removing the effect of site, the standardised residuals, hereafter termed site residuals, were used for QTL analysis and as the input for the estimation of the plot residuals. The estimation of the plot residuals was carried out on the full set of measurements from all families according to the following model (based on Hodge *et al.* 1996):

Model for plot residuals

$$\text{trait} = \mu + \text{site} + \text{replicate} + \text{incomplete block} + \text{female} + \text{male} + \text{family} + \text{plot} + \text{residual}$$

random effects: incomplete block, female, male, family, plot, residual

fixed effects: site, replicate

After the removal of the random effects of incomplete block, female, male, family and plot and the fixed effects of site and replicate, the residuals, hereafter termed the plot residuals, were used for QTL analysis. The estimation and removal of effects was carried out using the program ASREML, which uses an average information restricted maximum likelihood algorithm for variance component estimation (Gilmour *et al.* 1995). ASREML generated frequency histograms of the residuals, enabling any major departures from a normal distribution to be identified. As part of the model checking procedures ASREML also produced a plot of residuals

versus fitted values that were scrutinised to check for any trends in the data sets. Output from ASREML included estimates of variance components for random effects with associated standard error. The probability of each variance component being a chance deviation from zero was tested using a z-test where z = component/SE.

Preceding QTL analysis, each of the twelve traits analysed, from both datasets were checked for departure from normality by using the 'show trait' command of MAPMAKER/QTL (Patterson *et al.* 1988; Lincoln *et al.* 1992). Output from this command includes the plotting of a frequency histogram and calculations for skewness and kurtosis of the frequency distribution. For G1025 trees, Pearson Correlation coefficients were calculated for each combination of the nine measured traits, along with probabilities for each coefficient using the SAS procedure PROC CORR. The correlations were carried out between traits for the site residuals and plot residuals separately and between respective traits from the two datasets.

3.2.5 QTL detection

Single marker analysis for association between a trait and a marker locus was performed using the PROC ANOVA procedure in SAS. The following model was used in the analysis:

$$\text{trait} = \text{marker genotype} + \text{residual}$$

All markers and both site residuals and plot residuals for all nine measured traits were used in this analysis. In total there were 192 markers inherited from the male parent (G164) and 167 markers inherited from the female parent (KI2) that were tested for association using single marker analysis. These markers included all RAPD and microsatellite markers segregating in an approximate 1:1 ratio and included 30 unlinked markers (15 from each parent). In addition the trait of early flowering (buds3 – plot residuals) was tested for association with each marker using the non-parametric Wilcoxon Rank sum test (Wilcoxon 2-Sample Test - Normal Approximation) with the SAS procedure NPAR1WAY.

Markers that formed the framework of both parental linkage maps (Chapter 2 of this thesis) were used for QTL detection analysis by interval mapping, using the programs MAPMAKER/QTL Version 1.1 (Patterson *et al.* 1988; Lincoln *et al.* 1993) and QTL Cartographer Version 1.13g (Basten *et al.* 1997). The two parental genomes were included as a single entity for QTL detection purposes and interval mapping consisted of scanning a total of 1714cM over the two genomes of 24 linkage groups and 198 framework markers bracketing 141 intervals. QTL positions were identified in MAPMAKER/QTL using the 'show peaks' command with a LOD score threshold of 2.0 and a falloff of minus 1.0 (Verhaegen *et al.* 1997; Lerceteau *et al.* 2001). For the nine measured traits the QTL Cartographer programs Zmap and Eqtl were used to confirm QTL detection carried out in MAPMAKER/QTL, and to carry out permutations to establish experiment-wise threshold significance levels using 1,000 (or 5,000) permutations of the data set (Churchill and Doerge 1994).

For QTL that had LOD scores greater than 2.0, QTL positions were fixed (defining the QTL position using the MAPMAKER/QTL 'sequence' command) and the linkage group harbouring the QTL was scanned to determine if there were any peaks discernible that might indicate the presence of two QTL rather than the single detected QTL. With the detected QTL fixed, the estimated proportion of the variance explained by the QTL is excluded in the scan. Scans of all intervals were also carried out with the QTL position fixed, since this increases the possibility of detecting further QTL.

Where there were two or more QTL detected for a trait, both QTL were included in a model to determine if there was epistatic interaction. The two QTL positions were defined by the 'sequence' command, and the LOD scores for the combined model determined using the 'map' command. There is evidence for epistatic interaction if the combined LOD score is appreciably larger than the LOD scores added when the QTL were detected separately (Lincoln *et al.* 1993).

3.2.6 Detection of QTL by site interaction

Conventional QTL detection depends on the segregation of a marker allele with trait values above (or below) the average trait value. However across a number of environments, QTL with strong environmental interaction may not be detected using this method. The identification of marker genotype (QTL) by site (environment) interaction may give an indication of the presence of undetected site interactive QTL. Conversely, conventionally detected QTL can be tested for interaction with the environment. To detect

marker (genotype or class) by site interaction, an analysis of variance was carried out for each marker according to the following model:

$$\text{trait} = \text{marker genotype} + (\text{marker genotype x site}) + \text{residual}$$

This analysis was carried out using the SAS GLM (General Linear Models) procedure for each marker, for all nine measured traits, using plot residuals. Probabilities were calculated from F tests, to determine the significance of the fixed effects of marker genotype and the fixed effects of marker genotype x site.

3.2.7 Temporal stability of QTL for growth

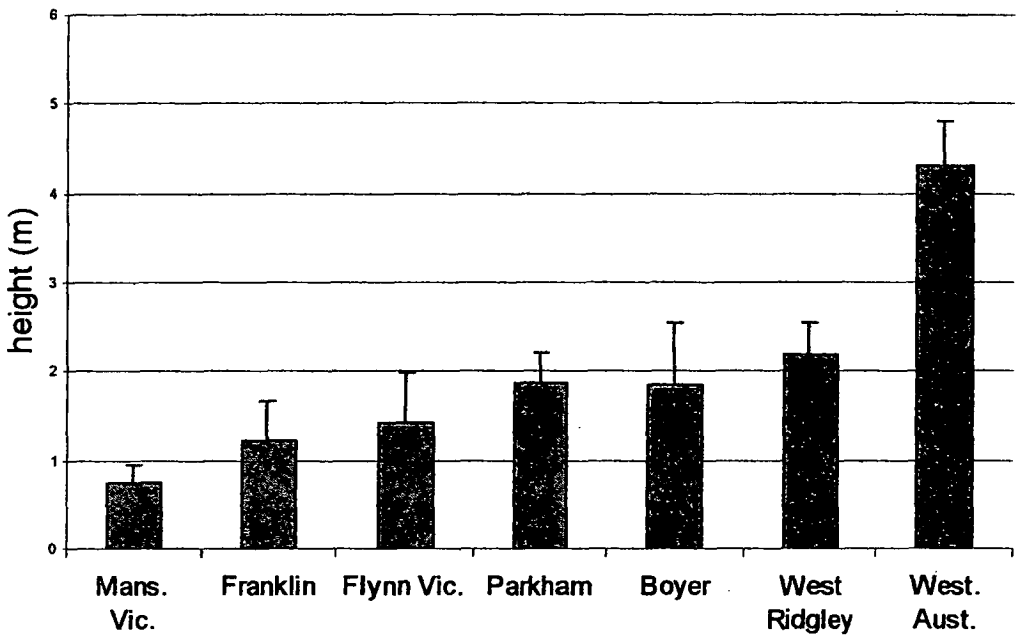
Intervals that had been identified as harbouring cumulative growth QTL were subjected to analyses to determine QTL stability with age. The initial genome scans carried out for each growth trait included all trees measured at each age. LOD score peaks and associated variances were recorded for these QTL interval/trait combinations. However, to compare these QTL intervals across the same environments/sample populations, analyses were carried out using only those trees measured in common between ages under comparison. These analyses enabled QTL effects to be compared between ages, without the confounding effects of any QTL by environment interaction or additional sampling effects.

3.3 RESULTS

3.3.1 Quantitative trait values for family G1025

Trees in the family G1025 were measured for each of the traits at the sites listed in Table 3.2. The average trait values (unstandardised) at each site for the G1025 Family are presented in Fig. 3.1(a-i) for each trait. The sites are arranged in order from the lowest to the highest average for the trait, for all trees measured at a site (including family G1025). Generally the trend in performance of the G1025 family across sites agrees with the overall trend found when all families are considered.

(a) average height year 1 (ht1)



(b) average height year 2 (ht2)

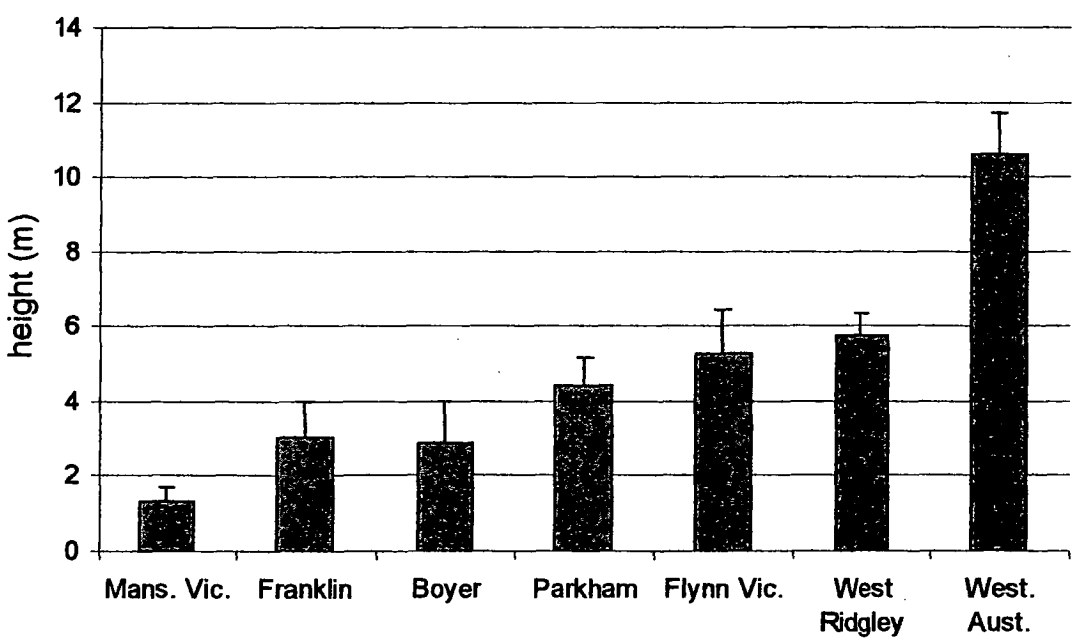
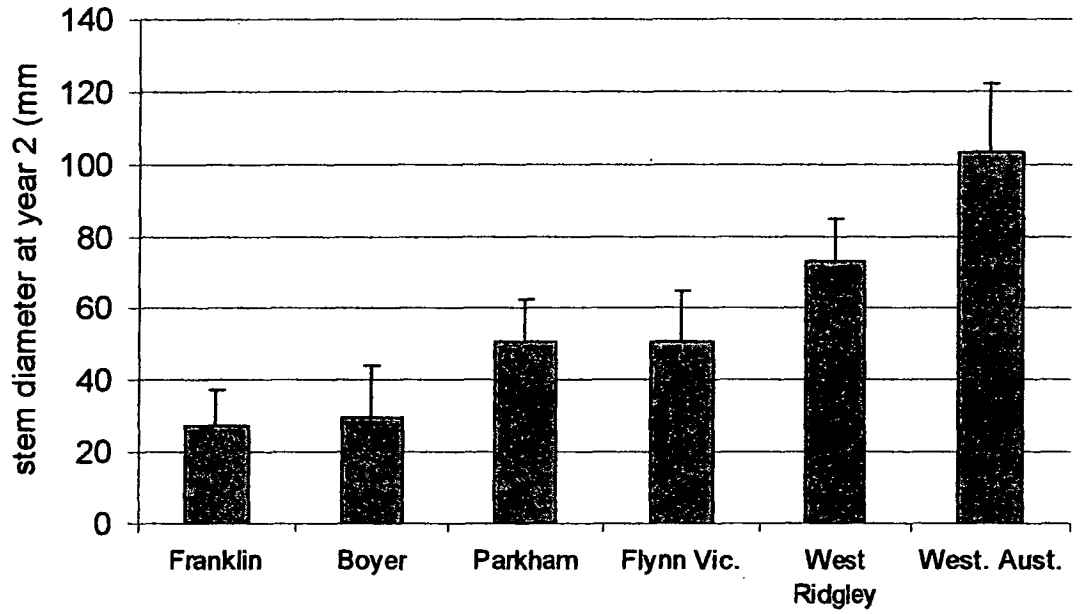


Fig. 3.1. (a-i). Trait means at each site for family G1025. Means are for unadjusted trait measurements of all nine measured traits. Error bars are one unit of standard deviation in length. For years one and two, the West. Aust. Site was measured approx. six months later than the measurements at the other sites

(c) average stem diameter at year 2 (dbh2)



(d) average stem diameter at year 3 (dbh3)

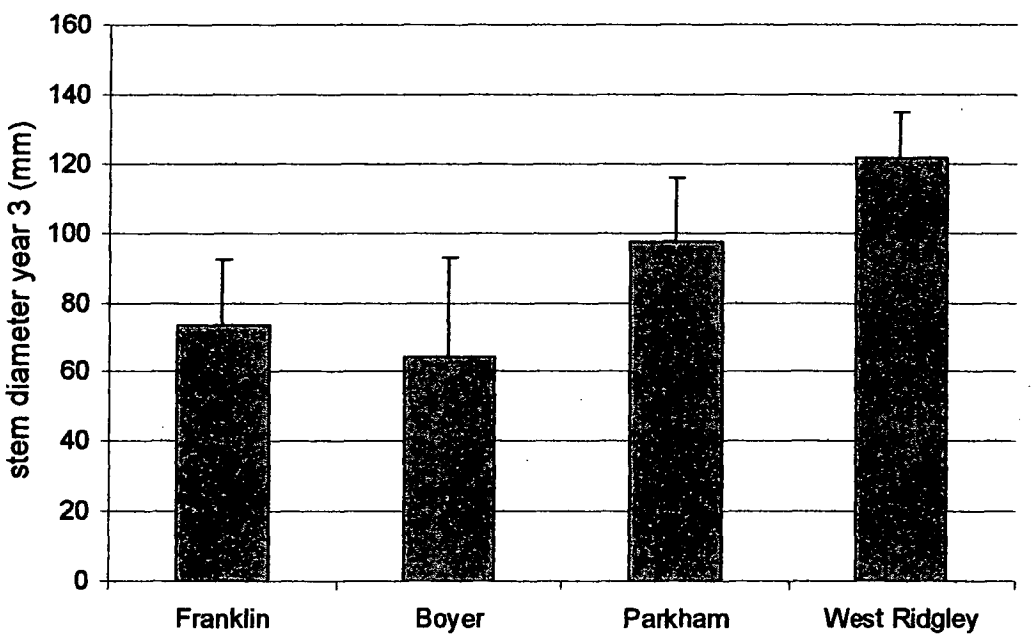
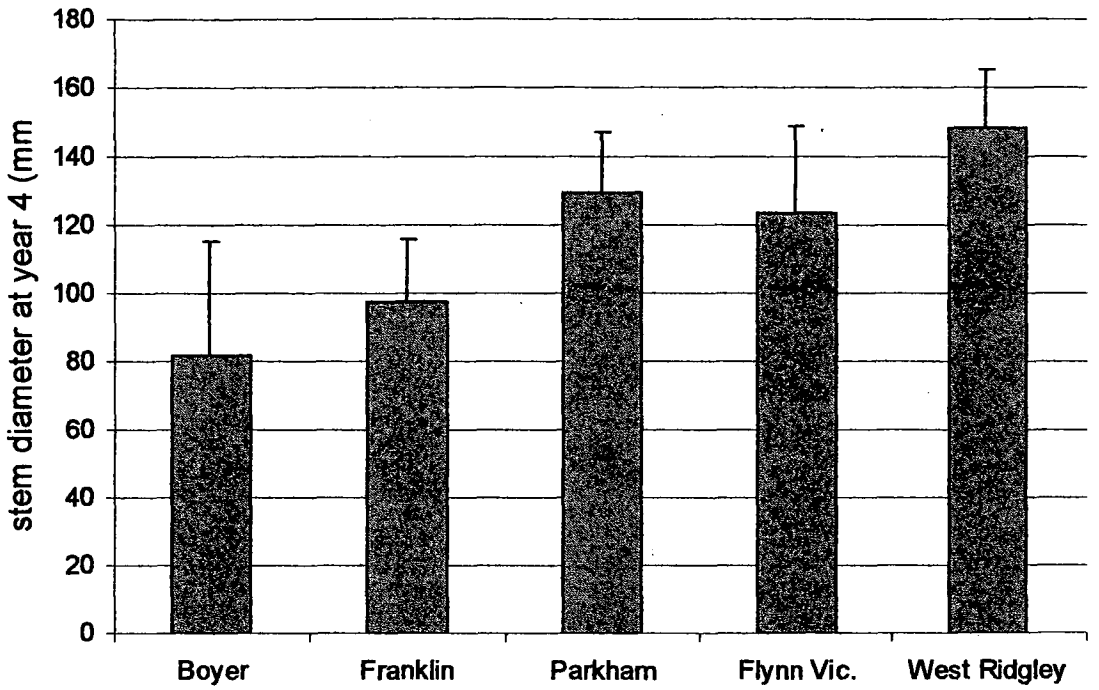


Fig. 3.1. cont.

(e) average stem diameter at year 4 (dbh4)



(f) average stem diameter at year 6 (dbh6)

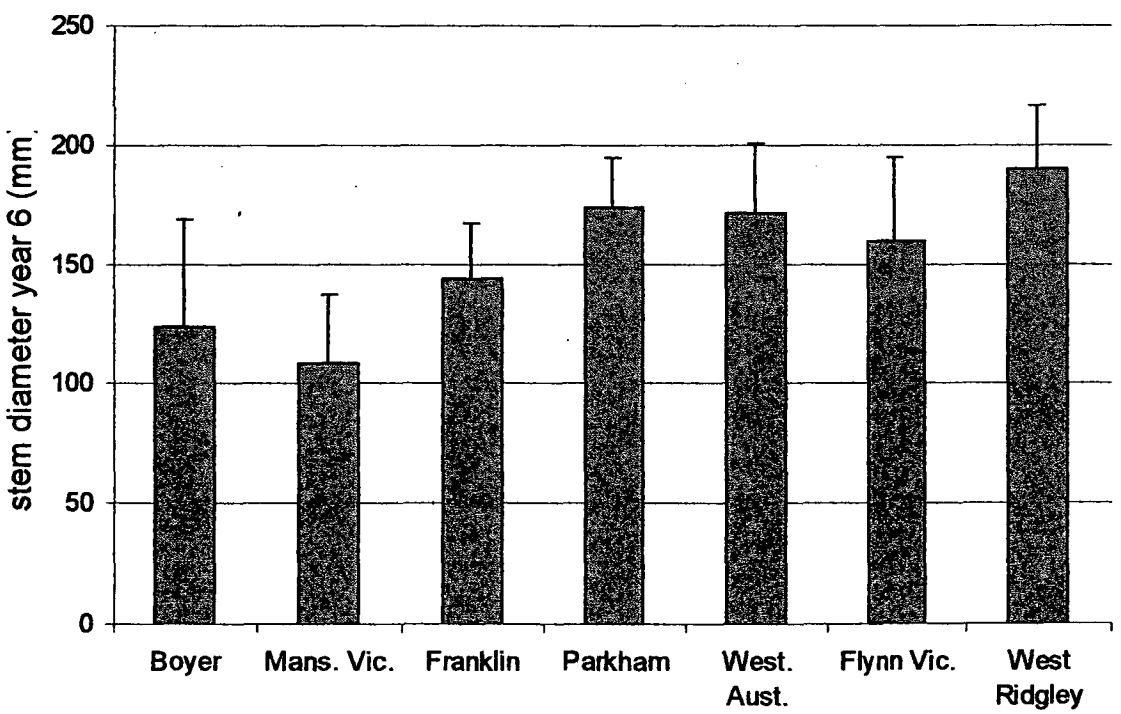
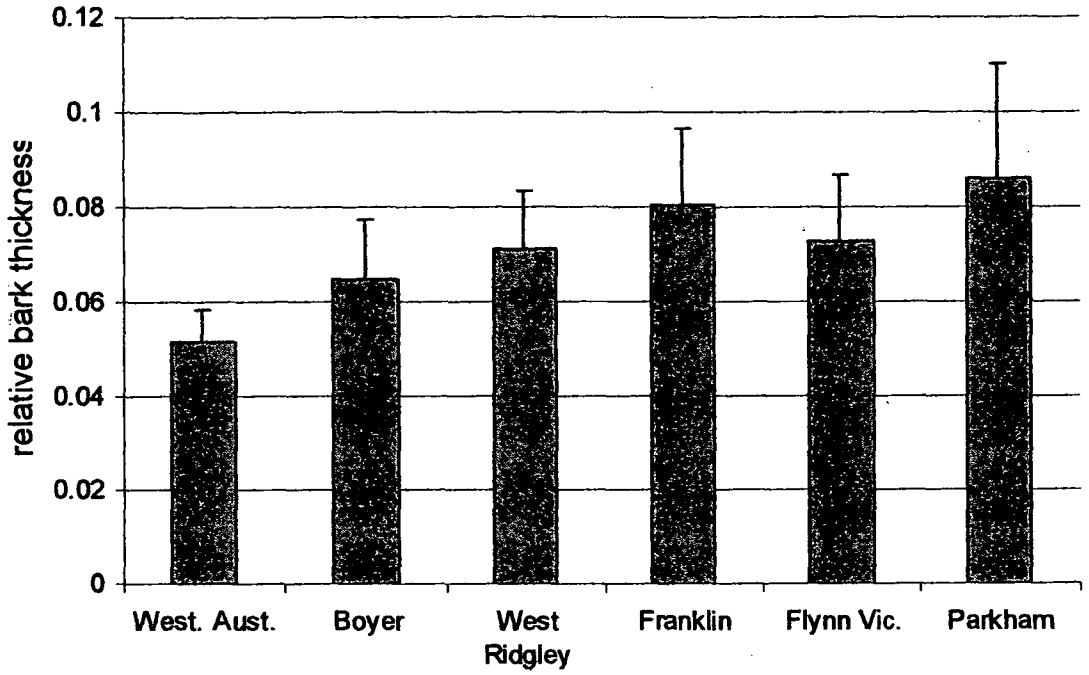


Fig. 3.1. cont.

(g) average bark thickness at year 6 (bark6)



(h) average pilodyn penetration year 6 (pilo6)

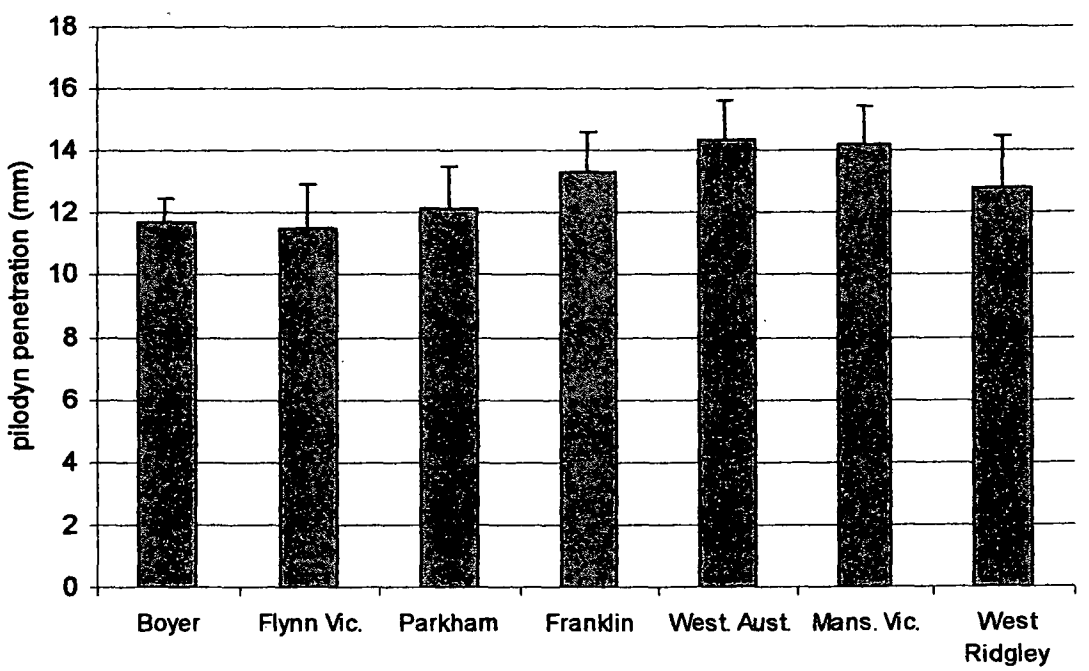


Fig. 3.1. cont.

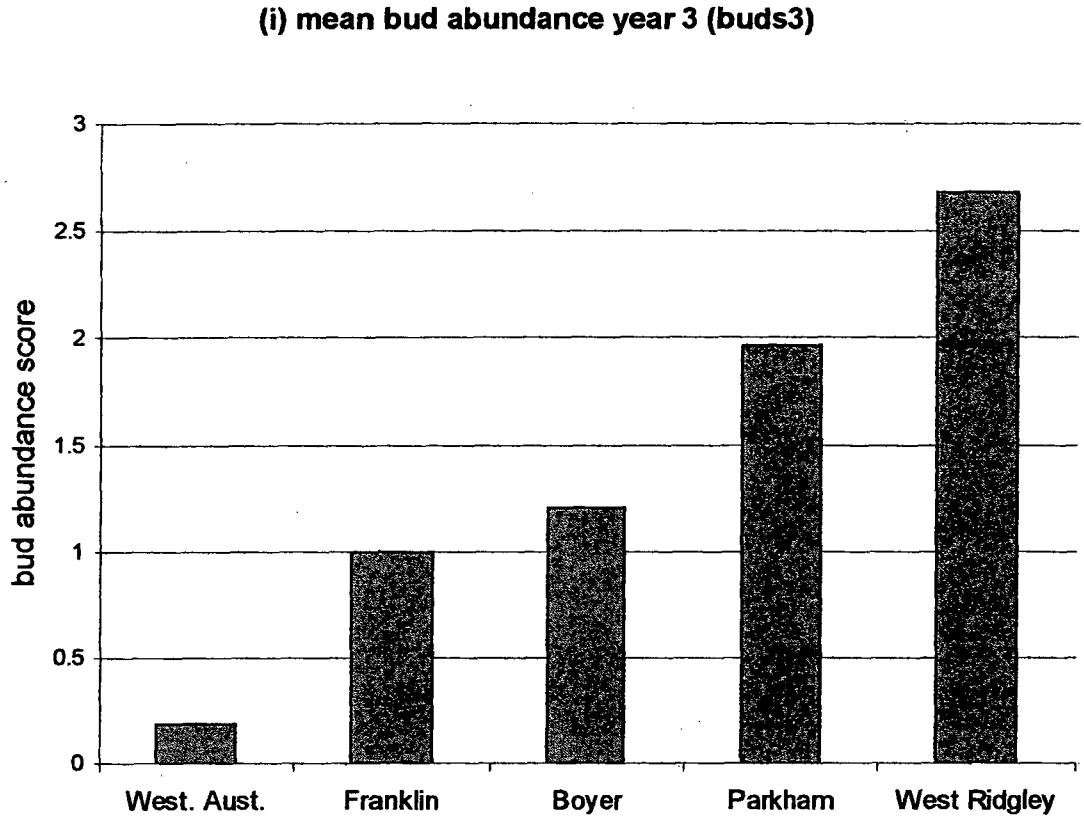

(i) mean bud abundance year 3 (buds3)
bud abundance score
3
2.5
2
1.5
1
0.5
0
West. Aust.
Franklin
Boyer
Parkham
West Ridgley

Fig. 3.1. cont.

To determine if the apparent differences in trait mean values between sites were significant, an analysis of variance was conducted for each trait for Family G1025 using unstandardised data. Site means were found to be highly significantly different from one another for all eight of the traits that were continuous variables (Table 3.5.).

Table 3.5.. Significance of differences between site means for each trait.

Trait	*p-value* [a]
ht1	8.5×10^{-26}
ht2	1.3×10^{-59}
dbh2	1.2×10^{-34}
dbh3	4.7×10^{-11}
dbh4	1.2×10^{-14}
dbh6	1.5×10^{-15}
bark6	2.0×10^{-15}
pilo6	2.8×10^{-10}

[a] From a one-way analysis of variance F test (ANOVA) (F = between site mean sq./error mean sq.)

3.3.2 Analysis to remove environmental effects from trait values

The plot residuals generated from standardised measurements of all 177 families were found to approximate a normal distribution for all traits. No trend between residuals and fitted values was found for any of the traits. Variance estimates for the components identified in the model for plot residuals, are presented in Table 3.6 for the six growth traits and in Table 3.7 for the three non-growth traits. For the growth traits the variance components 'incomplete block', 'family', 'plot' and the residual were significantly different from zero in all cases ($P < 0.001$). The components 'male' and 'female' were significant at the 0.05 level, except in two instances. Most of the variance was partitioned into the residual, the proportion of which increased with age of the trees from 67% for ht1 to 80% for dbh4 and dbh6.

For the non-growth traits four effects are significantly different from zero for all three traits ($P < 0.001$), 'incomplete block', 'male', 'plot' and the residual. The family effect was small for the non-growth traits, but comparable in size to the male effect for the growth traits. Again the residual makes up most of the total variance with 63% for Pilodyn penetration to 81% for early flowering (buds3).

Table 3.6. Variance estimates for components of growth traits[a].

Trait	*Source*[a]	*Count*	*Variance*	*S.E.*	σ^2/σ^2_T[b]	*p-value*[c,d]	*Sig.*[e]
ht1	inc. block	315	0.1156	0.0130	0.12	0	***
	female	8	0.0324	0.0187	0.03	0.042	*
	male	26	0.0108	0.0052	0.01	0.019	*
	family	177	0.0270	0.0059	0.03	0	***
	plot (family*iblk)	3041	0.1271	0.0084	0.13	0	***
	residual	12200	0.6300	0.0093	0.67	0	***
ht2	inc. block	313	0.0953	0.0110	0.11	0	***
	female	8	0.0083	0.0058	0.01	0.076	n.s.
	male	26	0.0094	0.0047	0.01	0.023	*
	family	177	0.0249	0.0057	0.03	0	***
	plot (family*iblk)	2999	0.1228	0.0081	0.14	0	***
	residual	12012	0.6104	0.0090	0.70	0	***
dbh2	inc. block	273	0.0917	0.0108	0.10	0	***
	female	8	0.0384	0.0218	0.04	0.039	*
	male	26	0.0045	0.0035	0.01	0.095	n.s.
	family	177	0.0279	0.0060	0.03	0	***
	plot (family*iblk)	2864	0.1027	0.0077	0.11	0	***
	residual	11682	0.6322	0.0094	0.70	0	***
dbh3	inc. block	152	0.0846	0.0131	0.09	0	***
	female	8	0.0349	0.0204	0.04	0.044	*
	male	26	0.0101	0.0060	0.01	0.046	*
	family	176	0.0400	0.0086	0.04	0	***
	plot (family*iblk)	1772	0.0815	0.0096	0.09	0	***
	residual	7572	0.6996	0.0129	0.74	0	***
dbh4	inc. block	196	0.0526	0.0082	0.05	0	***
	female	8	0.0416	0.0236	0.04	0.039	*
	male	26	0.0138	0.0062	0.01	0.012	*
	family	177	0.0245	0.0059	0.03	0	***
	plot (family*iblk)	2149	0.0639	0.0085	0.07	0	***
	residual	9431	0.7715	0.0127	0.80	0	***
dbh6	inc. block	313	0.0279	0.0050	0.03	0	***
	female	8	0.0600	0.0332	0.06	0.035	*
	male	26	0.0165	0.0064	0.02	0.005	**
	family	177	0.0151	0.0044	0.02	0	***
	plot (family*iblk)	2987	0.0738	0.0081	0.08	0	***
	residual	11401	0.7544	0.0115	0.80	0	***

[a] These components are identified in the site residuals in the model for the plot residuals:
trait = μ + site + rep + incomplete block + female + male + family + plot + plot residual
Random effects: incomplete block (iblk), female, male, family, plot, residual
[b] The variance of the component as a proportion of the total variance.
[c] The p-value is the prob. that the variance is a chance deviation from zero (ie. H_0: $\sigma^2 = 0$).
[d] p-values of zero are < 0.001.
[e] Significance levels are: * = 0.05, ** = 0.01 and *** = 0.001.

Table 3.7. Variance estimates for components of non-growth traits[a].

Trait	*Source*[a]	*Count*	*Variance*	*S.E.*	σ^2/σ^2_T[b]	*p-value*[c,d]	*Sig.*[e]
bark6	inc. block	244	0.0485	0.0077	0.05	0	***
	female	8	0.0499	0.0276	0.05	0.035	*
	male	26	0.0734	0.0217	0.08	0	***
	family	177	0.0065	0.0037	0.01	0.038	*
	plot (family*iblk)	2437	0.0454	0.0107	0.05	0	***
	residual	6961	0.7265	0.0152	0.76	0	***
pilo6	inc. block	244	0.0891	0.0113	0.10	0	***
	female	8	0.1088	0.0591	0.12	0.033	*
	male	26	0.0708	0.0207	0.08	0	***
	family	177	0.0061	0.0034	0.01	0.035	*
	plot (family*iblk)	2437	0.0639	0.0092	0.07	0	***
	residual	6961	0.5884	0.0122	0.63	0	***
buds3	inc. block	229	0.0253	0.0054	0.03	0	***
	female	8	0.0065	0.0047	0.01	0.082	n.s.
	male	26	0.0742	0.0219	0.08	0	***
	family	177	0.0149	0.0047	0.02	0	***
	plot (family*iblk)	2406	0.0684	0.0093	0.07	0	***
	residual	9246	0.7893	0.0134	0.81	0	***

[a, b, c, d, e.] See Table 3.6.

Summary statistics based on both site residuals and plot residuals for each of the twelve traits for family G1025 are presented in Table 3.8. Here the G1025 family trait means can be compared to the overall mean values, which for each data set are close to zero with a standard deviation close to 1.0. From the site residuals, where parental and family effects have not been removed, the performance of family G1025 can be compared to the performance of trees in all families. For the growth traits, it can be seen that the mean value for family G1025 is above the overall mean of zero in each case and that this increases with age. For diameter at breast height at year 6 (dbh6) the mean for family G1025 is 0.5 std. deviations above the overall mean with a standard deviation that is almost as large as the overall standard deviation of 1.0. Thus family G1025 has a growth performance that is considerably above average with a variance that is reasonably large. Bark thickness on the other hand is considerably lower than average by almost 0.5 standard deviations. Pilodyn penetration is above average by about one quarter of a standard deviation with variance within the family equivalent to that overall ($\sigma^2 = 1.0$). Early flowering (buds3) is also above average by about 1/5 of a standard deviation with variance within the family larger than overall.

It can be noted that once the parental, family, block and plot effects are removed (plot residuals) the mean values for family G1025, as would be expected, are much closer to the overall mean of zero with smaller standard deviations.

Table 3.8. Descriptive statistics of the twelve traits used in QTL analyses for family G1025.

Traita	No. Obs.	Mean	Std. Dev.	Minimum	Maximum
site	residuals				
ht1	131	0.11	0.90	-2.38	1.98
ht2	131	0.16	0.89	-2.12	2.21
dbh2	121	0.23	0.86	-2.12	2.04
dbh3	84	0.36	0.87	-1.84	2.71
dbh4	92	0.40	0.88	-2.60	2.44
dbh6	155	0.50	0.94	-2.82	3.00
bark6	152	-0.46	0.97	-2.57	5.35
pilo6	152	0.24	1.00	-2.09	4.31
buds3	105	0.20	1.14	-0.96	6.26
ht2-1	131	0.04	0.63	-1.91	1.60
dbh4-2	90	0.25	0.45	-0.79	1.42
dbh6-4	92	0.25	0.45	-0.79	1.42
plot	residuals				
ht1	131	0.06	0.66	-1.49	1.26
ht2	131	0.05	0.59	-1.56	1.29
dbh2	121	0.04	0.62	-1.78	1.50
dbh3	84	0.04	0.65	-1.70	1.66
dbh4	92	0.07	0.70	-2.31	1.60
dbh6	155	0.04	0.77	-2.62	1.83
bark6	152	0.00	0.91	-1.92	5.55
pilo6	152	0.01	0.85	-1.89	4.07
buds3	105	-0.04	1.03	-1.32	5.68
ht2-1	131	0.00	0.54	-1.82	1.22
dbh4-2	90	0.06	0.39	-0.88	1.00
dbh6-4	92	0.02	0.31	-0.81	0.69

[a] Traits are: ht1 - height at year 1, ht2 - height at year 2, dbh2 - diameter at breast height year 2, dbh3 - dbh yr 3, dbh4 - dbh yr4, dbh6 - dbh yr 6, bark6 - relative bark thickness yr 6, pilo6 - Pilodyn penetration yr 6, buds3 - bud abundance year 3, ht2-1 - relative incremental growth for height between years 1 and 2, dbh4-2 - rig for diameter at breast height between years 2 and 4, dbh6-4 – rig for dbh between years 6 and 4.

The similarity between the site residuals and plot residuals is illustrated by the high correlation values presented in Table 3.9. The Pearson correlation coefficients and associated probabilities for correlations between the nine measured traits are presented in Table 3.10 (for the site residuals) and Table 3.11 (plot residuals).

Table 3.9. Correlation between standarised site residuals and plot residuals for each trait[a].

ht1	*ht2*	*dbh2*	*dbh3*	*dbh4*	*dbh6*	*bark6*	*pilo6*	*buds3*
0.81	0.76	0.81	0.86	0.91	0.93	0.97	0.92	0.97

[a] Pearson correlation coefficients calculated using SAS Proc CORR. Each correlation is significant at an individual 0.0001 level. Traits as per Table 3.8.

Table 3.10. Correlations between Site residual trait values used in QTL detection analysis [a,b].

	buds3	*pilo6*	*bark6*	*dbh6*	*dbh4*	*dbh3*	*dbh2*	*ht2*
ht1	0.171	0.012	-0.097	**0.599**	**0.703**	**0.763**	**0.837**	**0.755**
	n.s.	n.s.	n.s.	0.0001	0.0001	0.0001	0.0001	0.0001
	105	128	128	131	92	84	121	131
ht2	0.306	-0.019	-0.136	**0.651**	**0.720**	**0.841**	**0.847**	
	0.0015	n.s.	n.s.	0.0001	0.0001	0.0001	0.0001	
	105	128	128	131	92	84	121	
dbh2	**0.323**	0.002	-0.080	**0.775**	**0.865**	**0.935**		
	0.0008	n.s.	n.s.	0.0001	0.0001	0.0001		
	105	121	121	121	90	82		
dbh3	0.336	-0.049	-0.095	**0.823**	**0.945**			
	0.0018	n.s.	n.s.	0.0001	0.0001			
	84	82	82	84	74			
dbh4	**0.387**	-0.039	-0.086	**0.930**				
	0.0007	n.s.	n.s.	0.0001				
	74	90	90	92				
dbh6	**0.401**	-0.033	-0.091					
	0.0001	n.s.	n.s.					
	105	152	152					
bark6	-0.044	0.033						
	n.s.	n.s.						
	103	152						
pilo6	-0.157							
	n.s.							
	103							

[a] At the top of each cell is the Pearson correlation coefficient followed by the p-value for Prob > |R| under H_0: rho = 0 and lastly the number of observations on which the calculations are based.

[b] Correlations in bold are significant at an overall level of 0.05 adjusting for 36 tests (Bonferroni correction) (ie. an individual level of $< 1.4 \times 10^{-3}$).

Table 3.11. Correlations between plot residual trait values used in QTL detection analysis [a,b].

	buds3	*pilo6*	*bark6*	*dbh6*	*dbh4*	*dbh3*	*dbh2*	*ht2*
ht1	0.139 n.s. 105	0.044 n.s. 128	-0.079 n.s. 128	**0.546** 0.0001 131	**0.655** 0.0001 92	**0.714** 0.0001 84	**0.826** 0.0001 121	**0.632** 0.0001 131
ht2	0.258 0.0079 105	0.017 n.s. 128	-0.140 n.s. 128	**0.644** 0.0001 131	**0.706** 0.0001 92	**0.782** 0.0001 84	**0.757** 0.0001 121	
dbh2	0.226 0.0207 105	0.044 n.s. 121	-0.082 n.s. 121	**0.744** 0.0001 121	**0.841** 0.0001 90	**0.915** 0.0001 82		
dbh3	0.226 0.039 84	0.021 n.s. 82	-0.137 n.s. 82	**0.820** 0.0001 84	**0.941** 0.0001 74			
dbh4	0.298 0.0099 74	0.092 n.s. 90	-0.080 n.s. 90	**0.921** 0.0001 92				
dbh6	**0.321** 0.0009 105	0.050 n.s. 152	-0.121 n.s. 152					
bark6	-0.056 n.s. 103	0.045 n.s. 152						
pilo6	-0.138 n.s. 103							

[a] At the top of each cell is the Pearson correlation coefficient followed by the p-value for Prob > |R| under H_0: rho = 0 and lastly the number of observations on which the calculations are based.

[b] Correlations in bold are significant at an overall level of 0.05 adjusting for 36 tests (Bonferroni correction) (ie. an individual level of $< 1.4 \times 10^{-3}$).

For the plot residuals, highly significant positive correlations were found between all six of the traits for growth (Table 3.11). No significant correlations were found involving the traits of bark thickness (bark6) and Pilodyn penetration (pilo6). The positive correlation found between early flowering (buds3) and growth was significant at an individual test level of 0.05 for all but the first growth measurement (ht1), although only the

correlation with growth at year six (dbh6) was significant at an overall 0.05 level.

For the site residuals, the significance of correlations is almost identical to those for the plot residuals (Tables 3.10 and 3.11). The site residuals for the growth traits are correlated with one another with highly significant p-values, as is the case with the plot residuals. There are significant correlations between early flowering and three of the growth traits (dbh2, dbh4 and dbh6) at an overall 0.05 level with the site residuals whereas with the plot residuals the only significant correlation was with dbh6. All the significant correlations between traits using the site residuals have a larger value for the Pearson correlation coefficient than the corresponding correlations with the plot residuals (Tables 3.10 and 3.11).

For family G1025 both the site residuals and the plot residuals were found to approximate a normal distribution when a frequency histogram and descriptive statistics were generated as part of the 'show trait' command of MAPMAKER/QTL. Frequency histograms using the plot residuals for family G1025 for each of the nine traits are presented in Fig. 3.2.

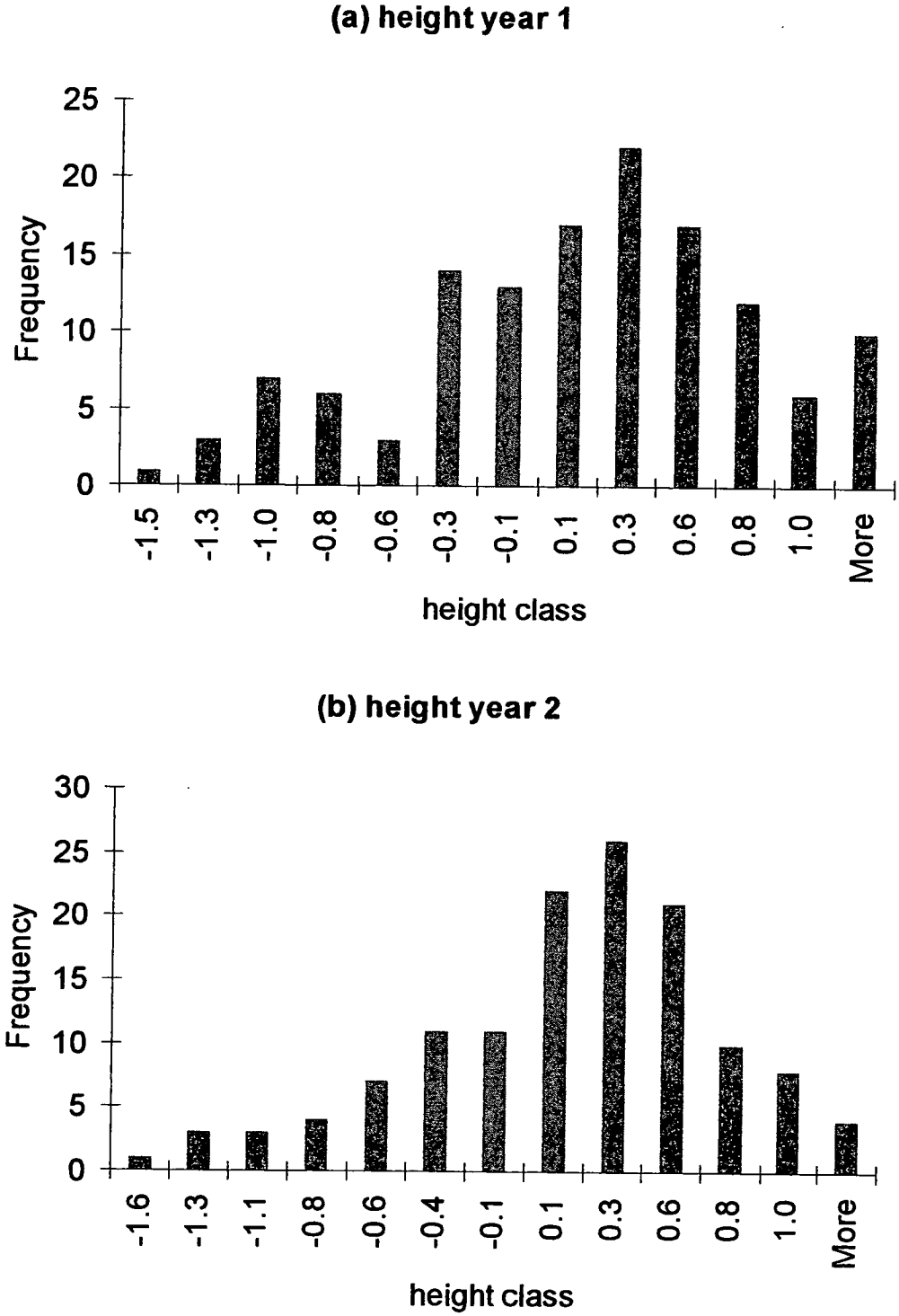
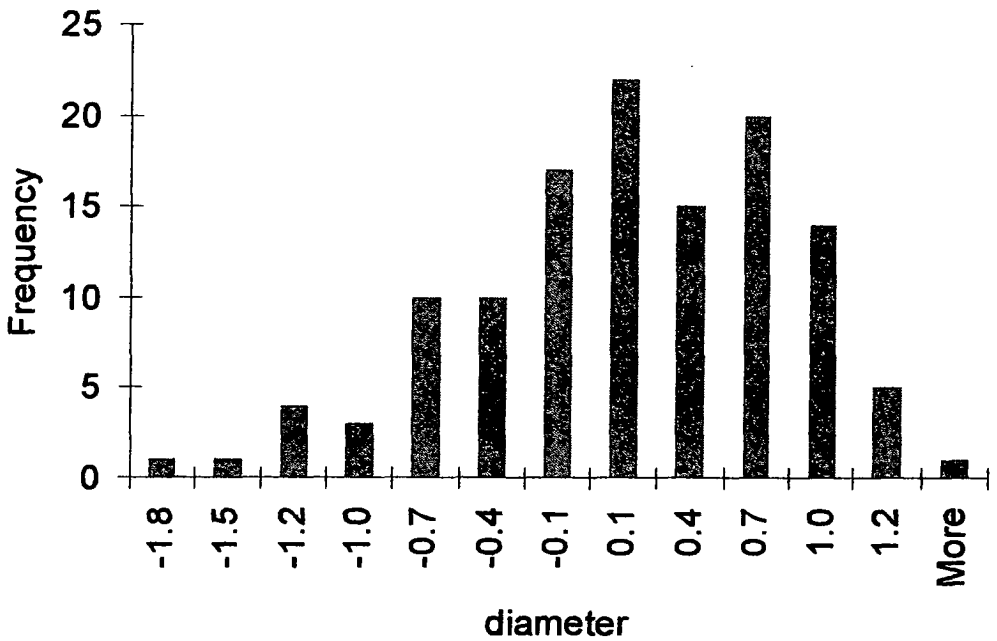
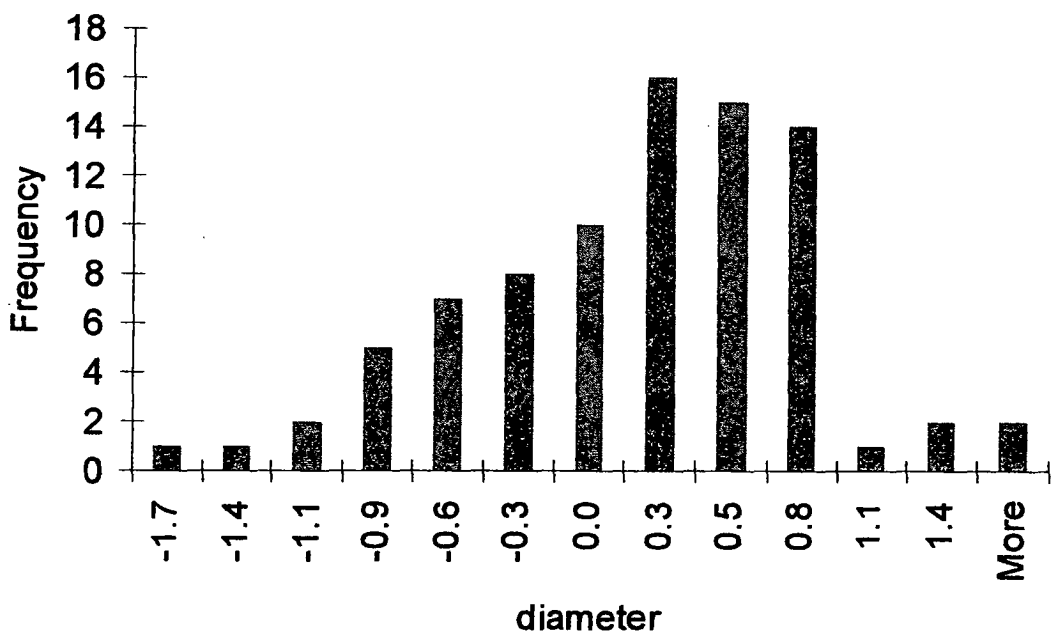


Fig. 3.2 (a-l). Frequency histograms showing the distribution of each of the twelve traits used for QTL detection for family G1025. Units are plot residuals.

(c) stem diameter year 2

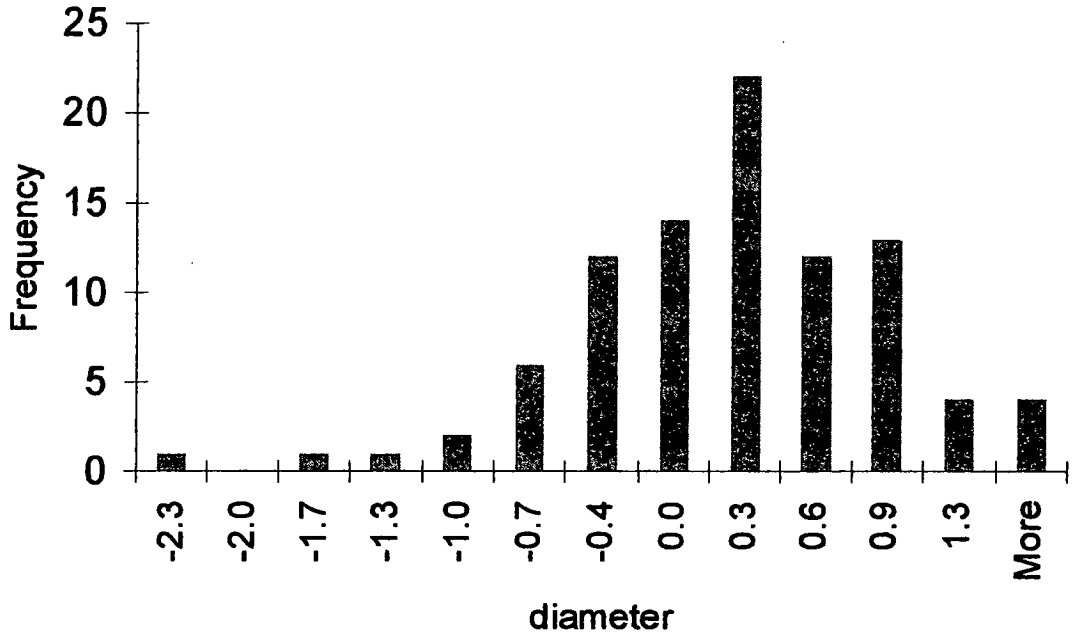


(d) stem diameter year 3

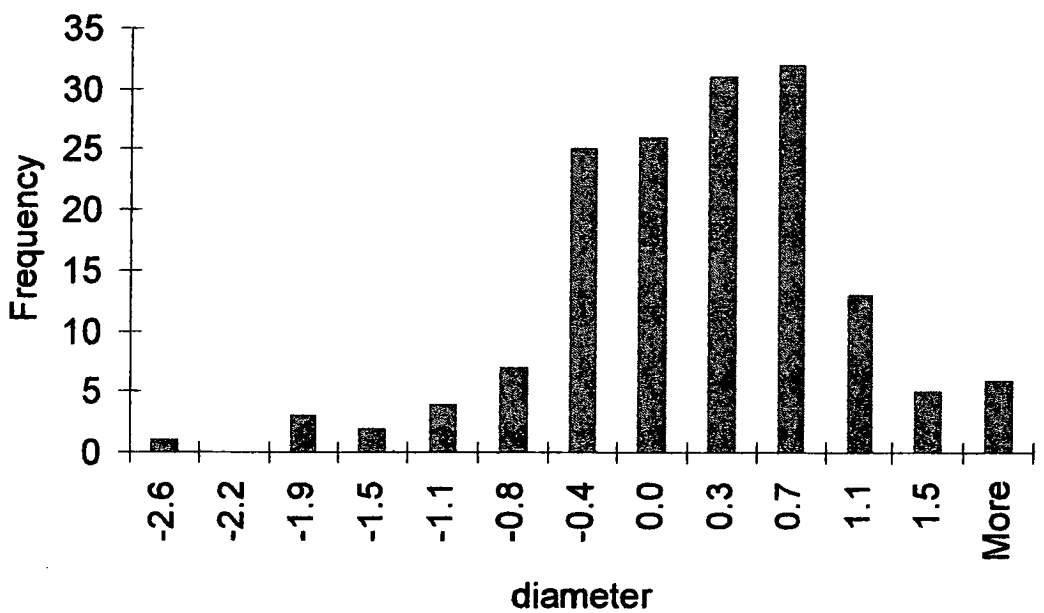


(Fig. 3.2 cont.)

(e) stem diameter year 4

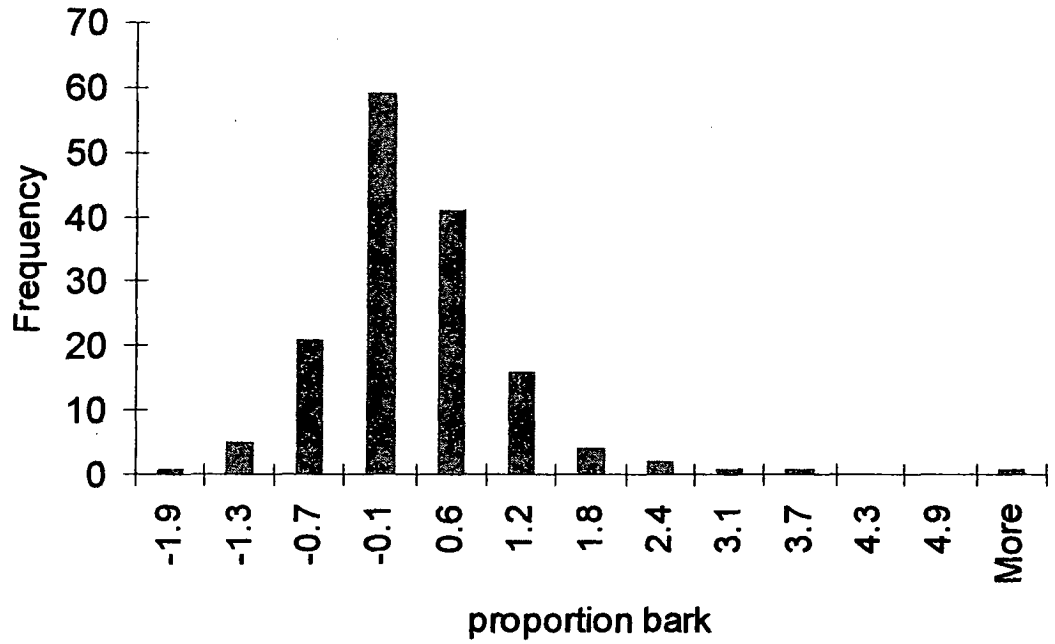


(f) stem diameter year 6

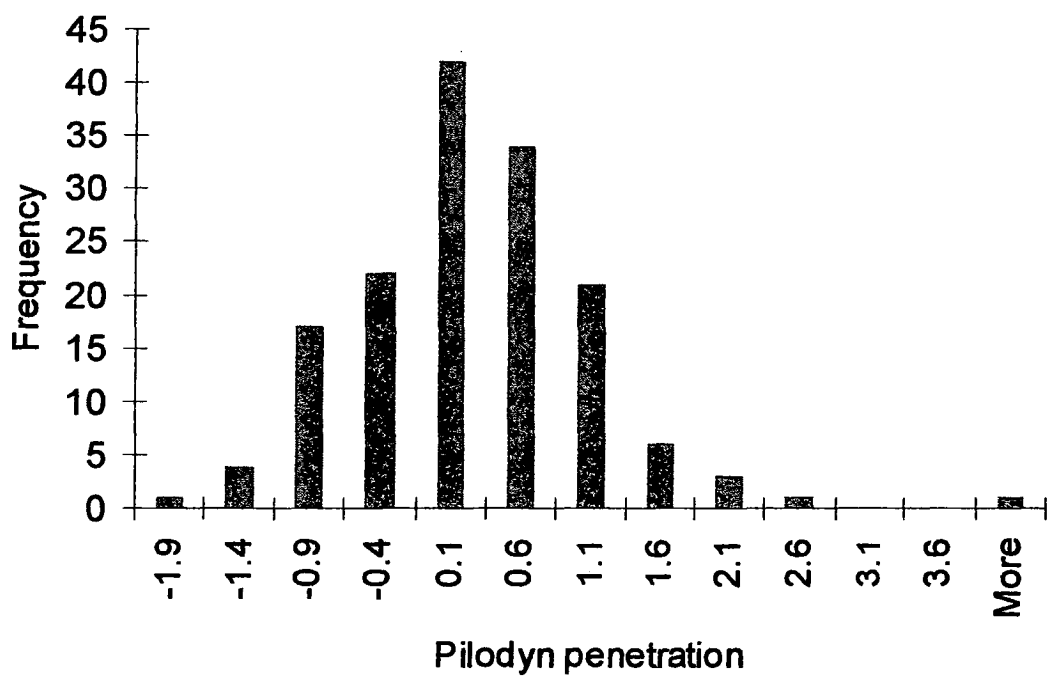


(**Fig. 3.2** cont.)

(g) relative bark thickness

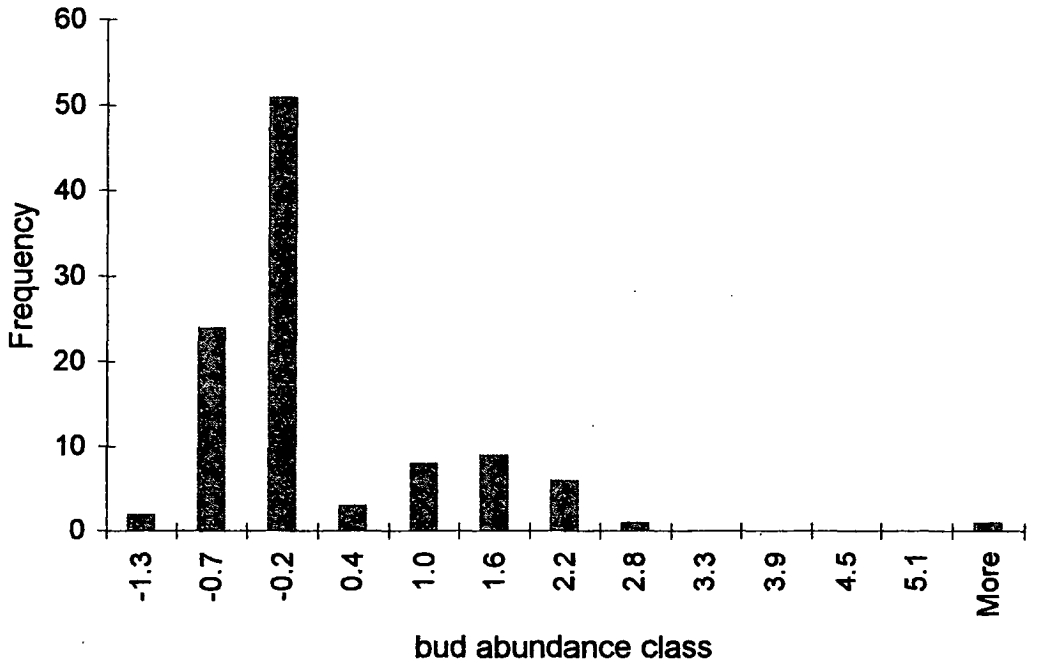


(h) wood density

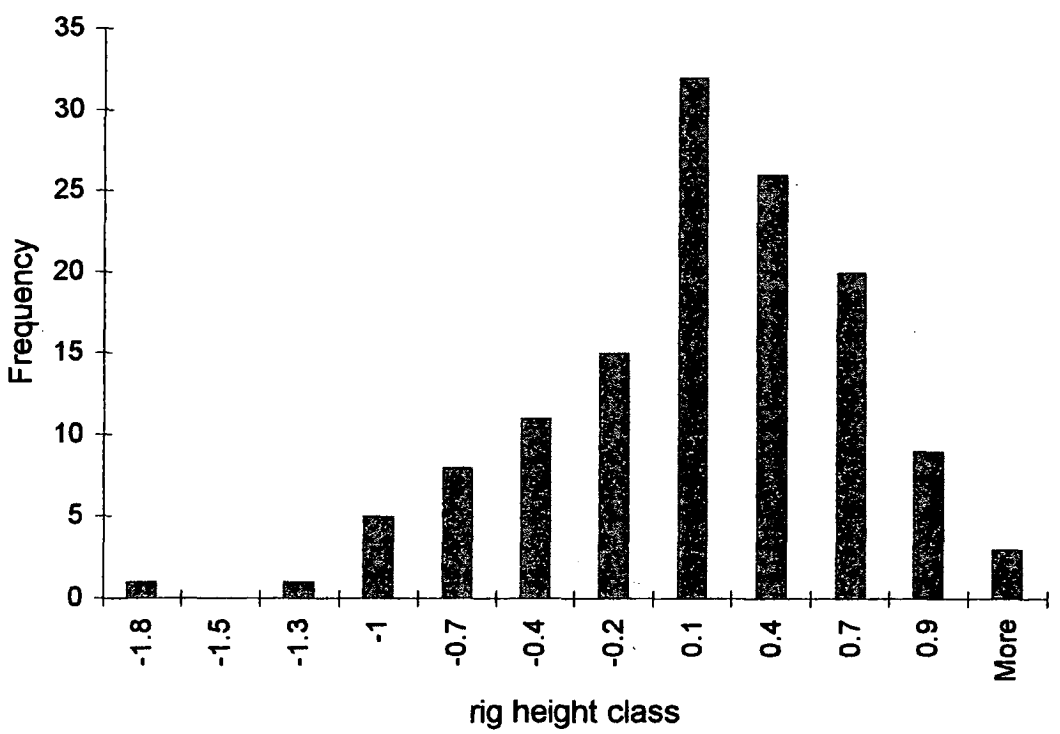


(**Fig. 3.2** cont.)

(i) early flowering

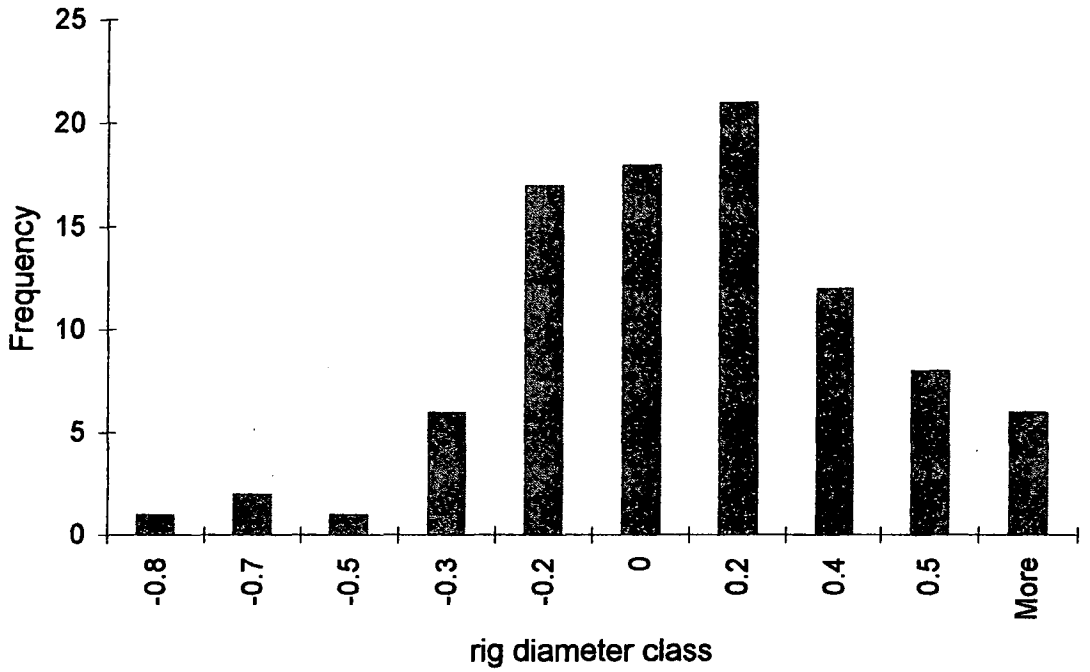


(j) relative incremental growth - height years 1 to 2

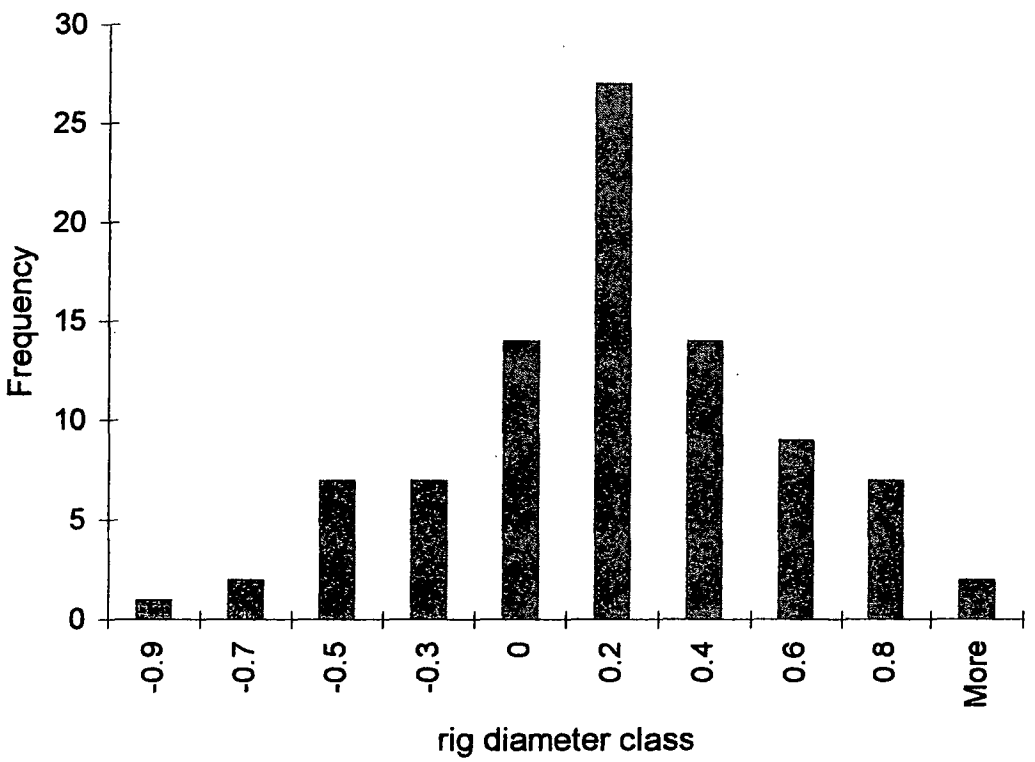


(Fig 3.2 cont.)

(k) relative incremental growth - stem diameter years 2 to 4



(l) relative incremental growth - stem diameter years 4 to 6



(Fig 3.2 cont.)

3.3.3 QTL detection

Based on the MAPMAKER/QTL analysis, map intervals that harboured LOD score peaks above 2.0, with either site residuals or plot residuals, are listed in Table 3.12 (site residuals) and Table 3.13 (plot residuals). Seven putative (LOD > 2.0) QTL were detected using the site residuals (Table 3.12), whereas eight QTL were detected using the plot residuals (Table 3.13). For seven of these eight QTL, the LOD score was larger when the plot residuals were used in the analysis, compared with the site residuals. This is expected, since the removal of environmental noise should increase the proportion of variance explained by genuine QTL.

Six intervals on six linkage groups contain putative QTL for growth traits. There are three QTLs for the non-growth traits, each on a separate linkage group (M4, M11, F6). If the results from the plot residuals only are considered, eight QTL have been detected in this study, five for growth (two for stem diameter and three for rig of stem diameter), two for Pilodyn penetration and one for bud abundance. These eight QTL correspond to seven genomic locations.

Only two linkage groups have QTL for more than one trait. There are QTL on F3 for both height increment (ht2-1) and for dbh6 (Table 3.12). However since they are approx 44cM apart, they are likely to be unrelated. The QTL on M11 for relative incremental growth (dbh4-2) is however, in the same interval as a QTL for Pilodyn penetration (pilo6 on M11, Table 3.13).

Table 3.12. List of putative QTLs for family G1025 based on analyses carried out with site residuals[a]. QTLs that have a LOD score peak > 2.0 for analyses using either site residuals or plot residuals are listed

parent/ linkage group	*trait*	*flanking framework markers*	*LOD score peak*	*Variance explained (%)*
M2[b]	*dbh3*	*246-4/266-6*	*1.53*	*8.1*
M4	buds3	42-1/131-2	2.50	10.5
M7	*dbh6-4*	*EMCRC7/B06-2*	*1.53*	*9.6*
M10	dbh6-4	218-1/B07-1	2.44	12.8
M11	dbh4-2	232-8/249-1	2.07	9.9
M11	pilo6	232-8/249-1	1.78	5.3
F3	ht2-1	273-3/215-5	2.30	8.2
F3	dbh6	217-2/234-4	2.78	8.7
F6	pilo6	212-4/238-6	2.59	8.0

[a] for an explanation of site residuals and plot residuals see section 3.2.4.
[b] QTLs in italics have LOD scores less than 2.0.

Table 3.13. List of putative QTLs for family G1025 based on analyses carried out with plot residuals[a]. QTLs that have a LOD score peak > 2.0 for analyses using either site residuals or plot residuals are listed

parent/ linkage group	*trait*	*flanking framework markers*	*LOD score peak*	*Variance explained (%)*
M2	dbh3	246-4/266-6	2.02	10.5
M4	buds3	42-1/131-2	2.38	10.0
M7	dbh6-4	EMCRC7/B06-2	2.17	13.7
M10	dbh6-4	218-1/B07-1	3.02	15.7
M11	dbh4-2	232-8/249-1	3.94	17.9
M11	pilo6	232-8/249-1	2.29	6.7
F3[b]	*ht2-1*	*EMCRC4/273-3*	*1.03*	*3.5*
F3	dbh6	81-1/217-3	2.82	8.0
F6	pilo6	212-4/238-6	4.39	13.3

[a, b] see Table 3.

Details of the eight putative QTL detected using the plot residuals are found in Table 3.14. There are two QTLs for stem diameter one each at age 3 years (dbh3) and at age 6 yrs (dbh6), three QTLs for relative incremental growth of stem diameter, two for Pilodyn penetration at age 6 yrs (pilo6/1 and pilo6/1) and one for bud abundance at age 3 yrs (buds3) (Table 3.14). Of these eight QTLs two are inherited from the female parent and six from the male parent.

Based on permutations of the trait data, three QTLs have an experiment-wise statistical significance at the generally accepted 0.05 level with two significant at the 0.01 level. Also there are three QTL that have experiment-wise significance at the 0.1 level.

A selection of LOD score plots from MAPMAKER/QTL scans is presented in Figures 3.3 to 3.8. These illustrate the occurrence of the LOD score peaks corresponding to estimated QTL effects for the traits and linkage groups indicated. By including the effect of a detected QTL in the genetic model used in the analysis and then scanning, it was possible to see if there was evidence for a second QTL elsewhere on a linkage group and also to increase the power to detect further QTL of smaller effect elsewhere in the genome.

Table 3.14. Details of QTL (LOD ≥2.0) detected in the *E. globulus* full-sib family G1025.

Trait/ QTL	*Parent/ linkage group*	*Framework markers flanking QTL*	*Dist. peak from L.H. marker (cM)*	*LOD score peak*[a]	*Variance explained (%)*	*p value from single marker analysis*[b]	*Sample size (n)*	*Marker effect* Δ_{SD}[c]	*Marker effect* $\%S^2$[d]	*exp.-wise significance level*
dbh3	M2	246-4/266-6	15	2.0	10.5	2.3×10^{-3}	84	+0.32	10.5	0.3[e]
dbh6	F3	81-1/217-3	0	2.8	8.0	3.6×10^{-4}	155	-0.28	8.0	0.1[e]
dbh4-2	M11	232-8/249-1	6	3.9	17.9	4.0×10^{-4}	90	+0.42	17.8	0.005[e]
dbh6-4	M7	B06-2/EMCRC7	25	2.2	13.7	1.8×10^{-3}	62	0.39	14.9	0.25[e]
dbh6-4	M10	218-1/B07-1	2	3.0	15.7	2.8×10^{-4}	91	+0.38	14.2	0.05[e]
pilo6/1	F6	212-4/238-6	4	4.4	13.8	9.2×10^{-6}	151	-0.32	10.5	0.002[f]
pilo6/2	M11	232-8/249-1	6	2.3	6.7	8.9×10^{-4}	152	+0.26	6.8	0.1[e]
buds3	M4	42-1/131-2	10	2.4	10.0	8.6×10^{-4}	104	+0.32	10.4	0.1[e]

[a] LOD scores and variance estimates are from analyses with the plot residuals using MAPMAKER/QTL. QTL effects at other loci were not included in the models for these estimates.

[b] The p-value was obtained from ANOVA (t-test) for the marker nearest to the QTL.

[c] The estimated difference in units of standard deviation between a population selected for a single marker nearest the QTL and the unselected population.

[d] The percentage of the phenotypic variance attributable to marker genotype for a single marker nearest the QTL. Calculated as the square of half of the difference between marker genotype means as a percentage of the total variance.

[e,f] Based on empirical estimates of experiment-wise significance from QTL analysis of either 1,000[e] or 5,000[f] permutations of the trait data (Churchill and Doerge 1994) using QTL Cartographer.

For stem diameter at year 3 (dbh3, plot residuals) a scan of M2 (linkage Group 2 of the male parent - approx. 40cM) indicates a peak in LOD score occurs close to marker 266-6 (Fig. 3.3). A second peak appears to rise in the last interval of this group. However when the QTL position is fixed at marker 266-6 the scan is almost flat (not shown) indicating that the apparent peak in the last interval can be accounted for by QTL effects within the first two intervals.

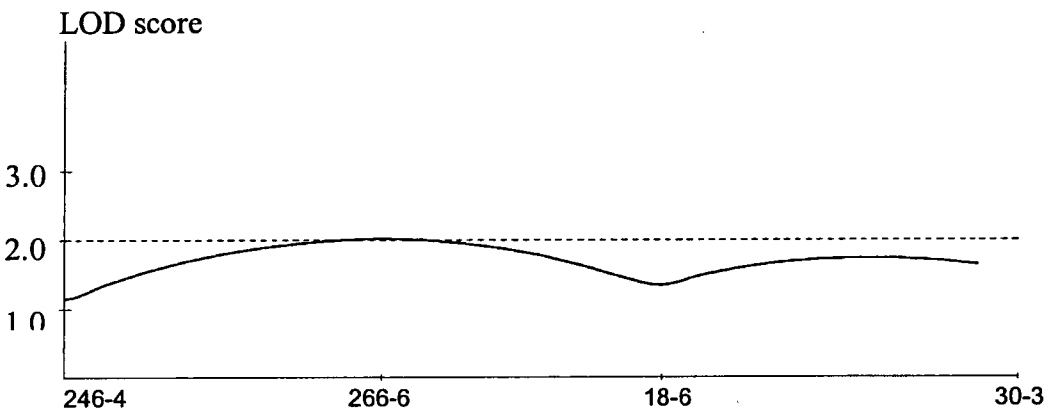


Fig. 3.3. QTL for growth at year 3 (dbh3). Scan of linkage Group 2 of the male parent (G164) using plot residuals.

A scan of the linkage group on which the QTL for growth at year 6 (dbh6) is located (F3) is presented in Fig. 3.4. In the scan, which covers approximately 150cM, there appears to be several LOD peaks that could be due to the presence of other QTLs. However from a scan in which the estimated effects of a QTL at marker 81-1 has been taken into account (Fig. 3.5), no other QTL peaks appear to be present and it can be concluded that a single QTL is likely to be responsible for the multiple peaks.

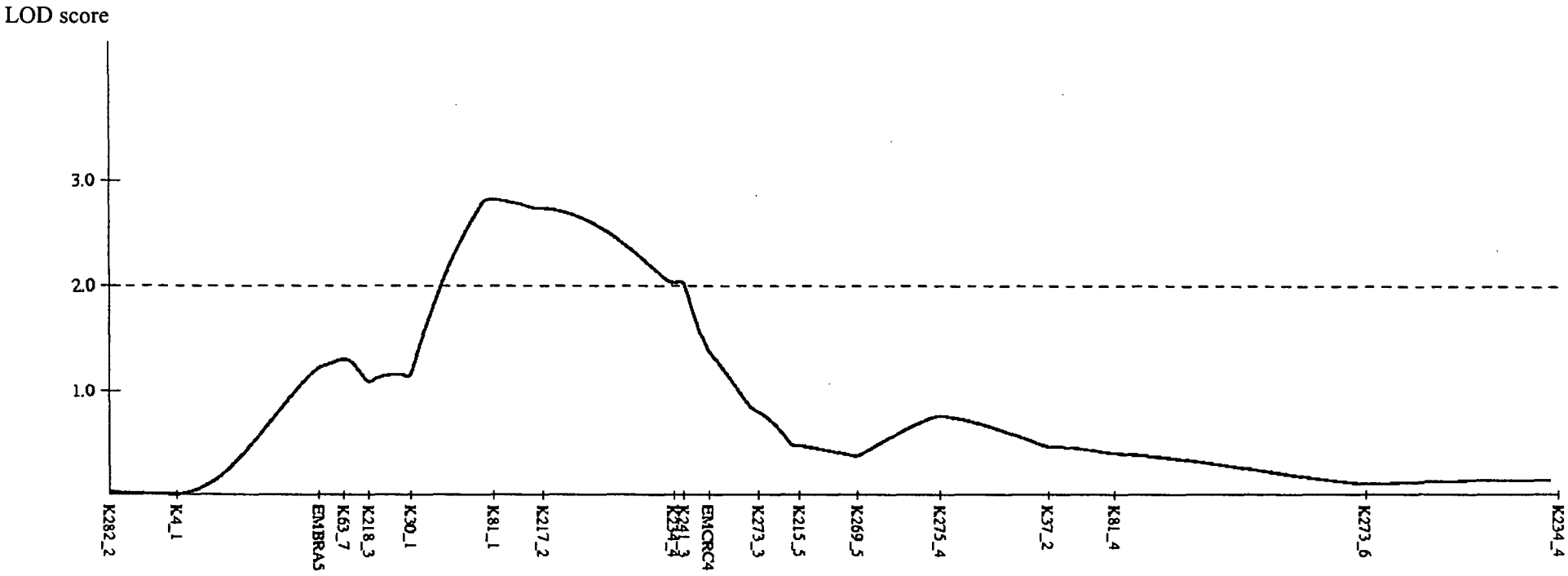


Fig. 3.4. QTL for growth at year 6 (dbh6). Scan of F3 (linkage group 3 of the female parent) using plot residuals. A LOD score peak occurs at marker K81_1.

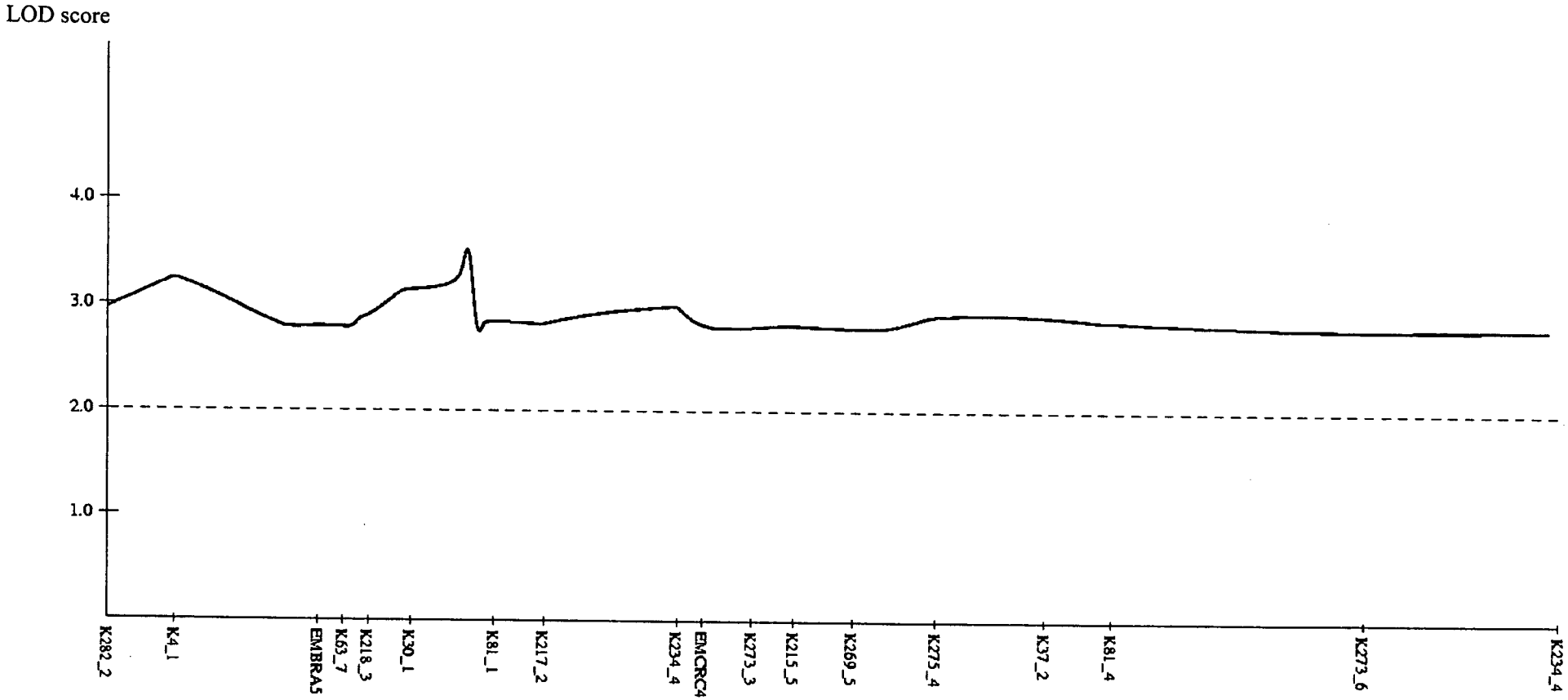


Fig. 3.5. Scan of linkage Group 3 of the female parent with effect of QTL for growth at year six removed (QTL 'fixed').

Graphs of the scans from the output of MAPMAKER/QTL for the two QTLs for Pilodyn penetration are presented in Fig. 3.6 (a-b). For QTL F6 (pilo6/1, Table 3.14) the scan of the linkage group covers 37cM and there is a single QTL peak to the right of marker 212-4. The scan of M11 (pilo6/2, Table 3.14) covers approximately 6cM and since it consists of only a single interval only one peak is resolvable. The flatness of the scan is likely due to the closeness of the bracketing markers and the lack of information from outside the interval.

The intervals containing the two QTLs for Pilodyn penetration were included together in a single model to determine if there was any discernible epistatic interaction and to increase the power to detect additional QTLs. The combined LOD score was only slightly larger than the sum when the QTL were analysed separately (Table 3.15). The estimated variance was also about the same as the sum of the individual variances indicating there is no epistatic interaction apparent between these two QTLs (Table 3.15).

Table 3.15. QTLs 1 and 2 for Pilodyn penetration (pilo6 - plot residuals) analysed separately and together. [a]

QTL	*QTL position – marker/ distance (cM)*	*LOD score*	*Variance explained (%)*
1	212-4 +4.8	4.4	13.3
2	232-8 +4.1	2.3	6.7
1 + 2	-	7.1	20.4

[a] LOD score, position and variance estimated using MAPMAKER/QTL 'map' command after defining the two QTL intervals.

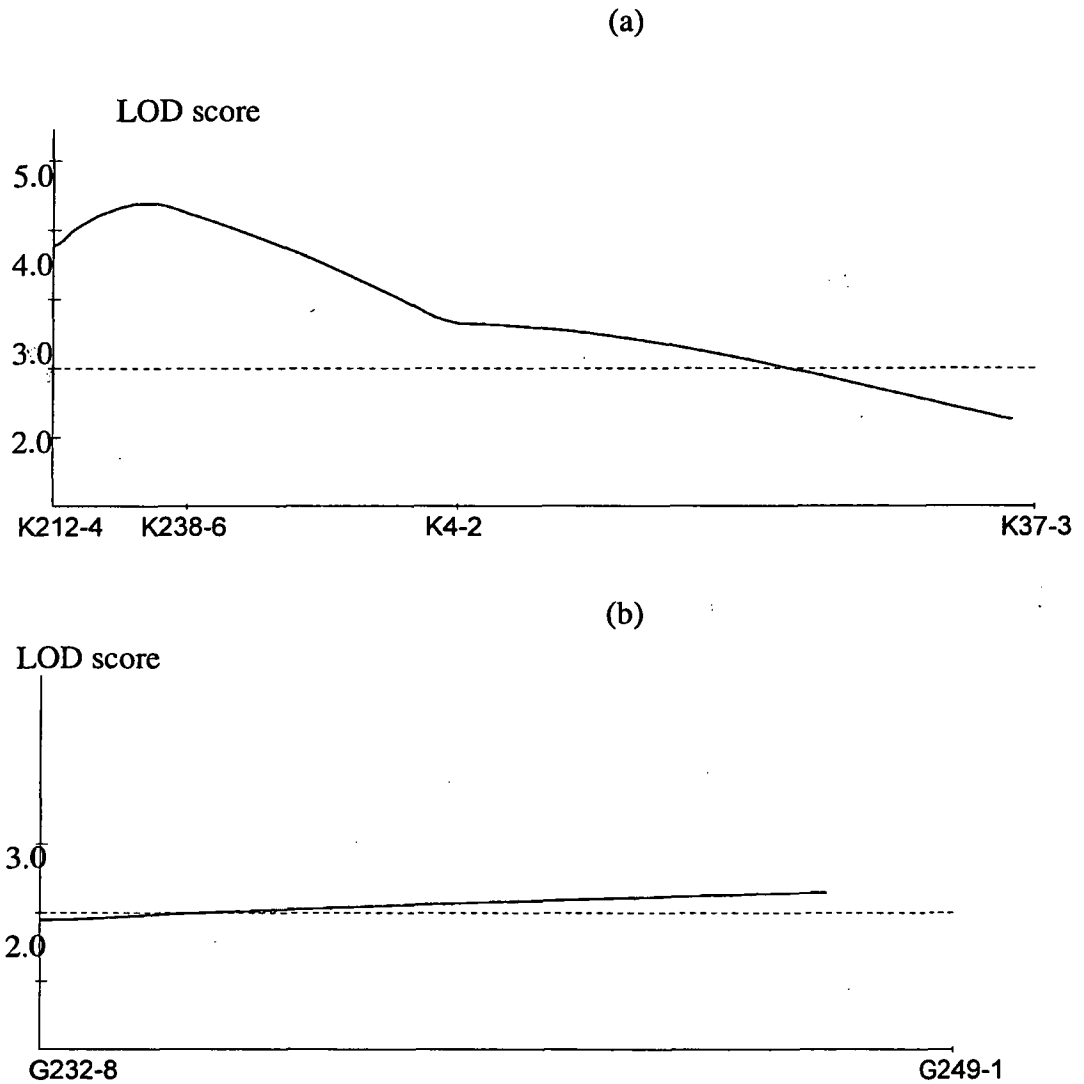


Fig. 3.6. Graphs of QTL scans for Pilodyn penetration. (a) QTL F6. Scan of linkage Group 6 of the female parent using plot residuals (pilo6). (b) QTL M11. Scan of linkage Group 11 of the male parent using plot residuals (pilo6).

Where effects of the QTLs with LOD score peaks greater than 2.0 were fixed in later scans, no new LOD score peaks greater than 2.0 were detected. However an additional peak of 1.9 was found when the effects of QTLs 1 & 2 for Pilodyn penetration were placed in a model together. This third potential QTL was on a separate linkage group to QTLs 1 and 2. When the effects of the three QTLs for Pilodyn penetration are considered together they are estimated to explain over 25% of the variance for this trait (plot residuals) and have a combined LOD score of 9.02 (Table 3.16).

Table 3.16. QTL 3 for Pilodyn penetration on linkage group 7 of the male (G164) parent.

QTL	*QTL position – marker + distance (cM)*	*LOD score*	*Variance explained (%)*
3 [a]	234-6 + 9.0	1.89	4.4
1 + 2 + 3 [b]	19-1 + 0.0	9.02	25.4

[a] Individual LOD score and effect are estimates for model with QTLs 1 + 2 fixed.

[b] LOD score and variance estimated using MAPMAKER/QTL 'map' command after defining the three QTL intervals.

Scans of the linkage group on which QTL 3 for Pilodyn penetration occurs are presented in Fig. 3.7 (a-b). Fig. 3.7(a) is from an initial scan of the genome, which resulted in the detection of QTLs 1 and 2. Fig. 3.7(b) is from a scan where QTLs 1 and 2 have been placed in a model together. It can be noted that the LOD peak is more pronounced in Fig. 3.7(b) where the variance due to the other two QTLs has been removed from the total variance.

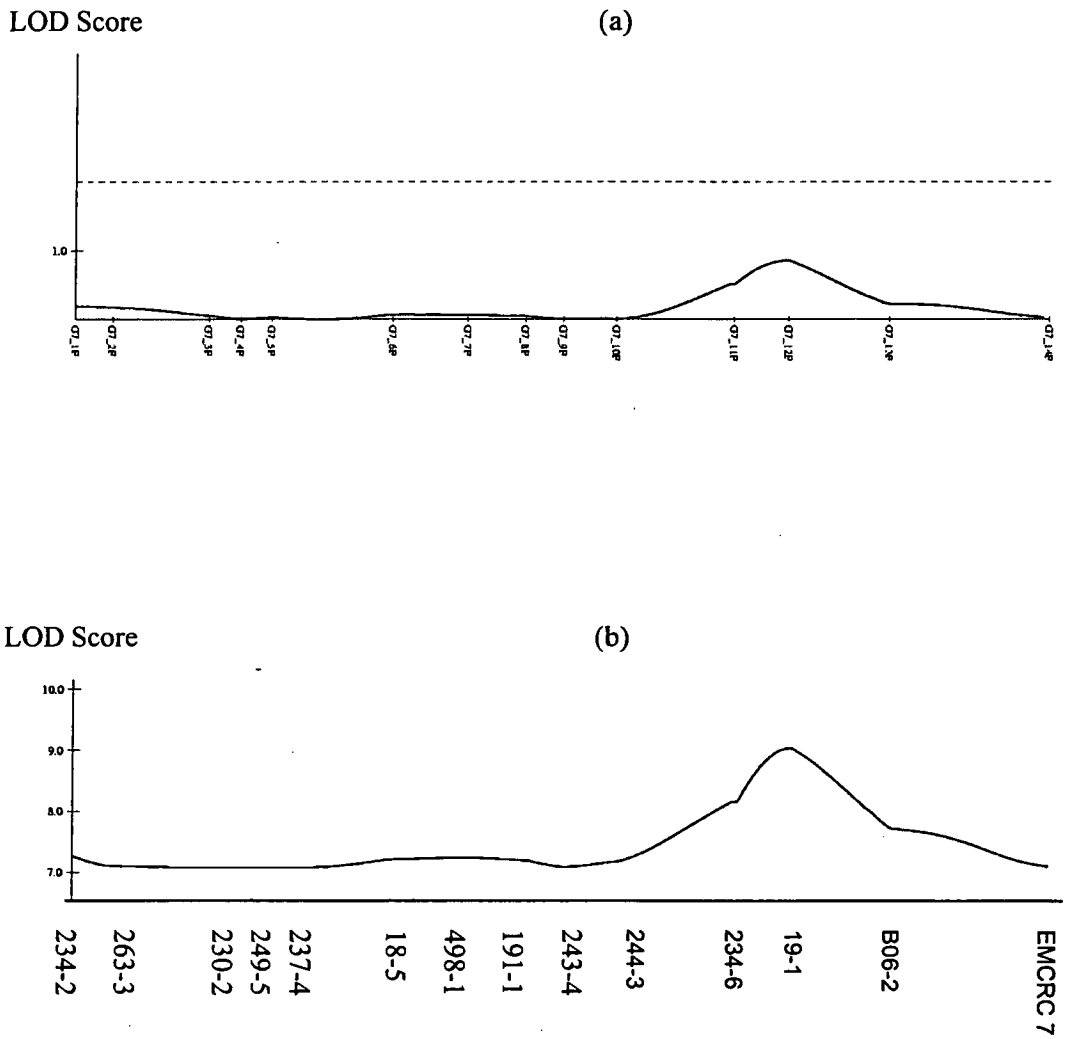


Fig. 3.7. LOD Score peak on M7 for pilodyn penetration (pilo6). (a) Initial scan – no fixed QTL effects. (b) With the effects of QTLs 1 & 2 included in the model (fixed). [Note: baseline is increased to LOD 7.1 in figure (b) due to inclusion of the two QTLs in the model].

The QTL for bud abundance gave a fairly sharp peak at marker 131-2 (Fig. 3.8) with no other peaks on the scan of the linkage group covering 66cM. Bud abundance was scored on a scale from 0-6 and was thus not a continuous variable. The frequency histogram pertaining to bud abundance using plot residuals (Fig. 3.2(i)) indicated considerable deviation from a normal distribution. Thus statistical tests that do not assume a normal distribution were carried out to detect any marker/trait associations. For all 359 markers tested the most extreme p-value using the Wilcoxon rank sum test ($P = 1.1 \times 10^{-4}$) was for marker 131-2 on M4 (linkage group 4 of the male parent) using plot residuals (Table 3.17). This is the same marker identified from interval mapping as occurring at the QTL for bud abundance. In the ANOVA for association between individual markers and bud abundance, marker 131-2 also gave the smallest p-value ($P = 8.6 \times 10^{-4}$). The results of the Wilcoxon rank sum test for the trait of early flowering are presented in Table 3.17 for the two datasets.

Table 3.17. Nonparametric test of the association of marker 131-2 with early flowering.[a]

	site residuals	*plot residuals*
p-value	6.7×10^{-5}	1.1×10^{-4}

[a] The nonparametric Wilcoxon rank sum test (normal approx.) was carried out.

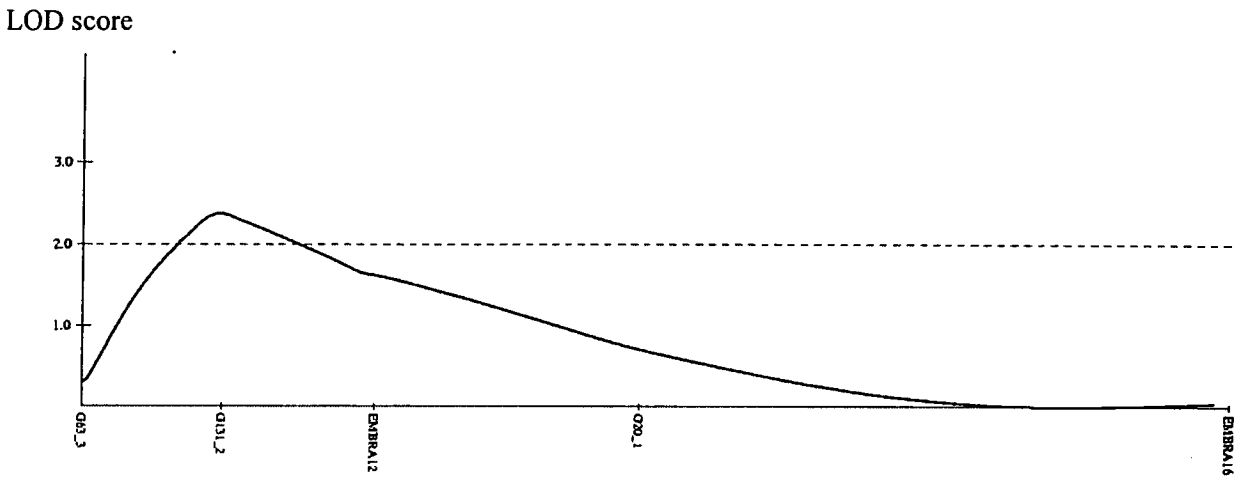


Fig. 3.8. QTL for early flowering (buds3). Scan of M4 using plot residuals. There is a LOD score peak of 2.4 at marker 131-2.

3.3.4 Markers (QTL) with site interaction

From the nine traits by 359 markers tested (approximately 3,200 tests), a total of 53 marker/trait combinations were found to have significant marker genotype by site interaction at the 0.01 level (Table 3.18). However none of these 53 marker/trait combinations that had significant site interaction were significant at the 0.01 level for the effect of marker genotype, with only one having significant marker genotype effect at the 0.05 level. The number of significant associations for each trait is given in Table 3.18.

Table 3.18. The number of significant interactions (P < 0.01) between marker genotype and site for each trait.

Trait	*No. of significant[a] marker x site associations*	*No. of independent[b] locations*
ht1	0	0
ht2	4	1
dbh2	2	2
dbh3	1	1
dbh4	2	2
dbh6	34	9
pilo6	0	0
bark6	6	4
buds3	4	3
All traits	53	23

[a] Based on ANOVA F tests with the plot residuals using the model: trait = mkr class + (mkr class x site) + residual

[b] Independent in this case means unlinked. Markers significant at 0.01 that are in linkage are counted as one.

There were approximately 3,200 tests undertaken for marker genotype by site interaction. From this number of tests, assuming independence, 32 significant associations would be expected at the 0.01 level due to chance. Thus with 53 significant associations there would appear to be an excess requiring an explanation other than chance. It should be noted however, that the assumption of independent tests is violated due to linkage between markers. However if those markers with significant interaction that are in linkage, are counted as representing one independently segregating location, there are 23 significant marker genotype by site interactions at the 0.01 level for the nine traits (Table 3.18).

To calculate the expected number of significant associations, the reasoning used in chapter 2 of this thesis, where segregation distortion of markers was considered, has been adopted. In chapter 2 the genome of the two parents was calculated to cover 50 independently segregating locations and there are in addition 30 unlinked markers. On this basis, there are 80 independently segregating locations included in the testing of marker genotype by site interaction with nine traits (720 independent tests). For significance at the 0.01 level, seven significant associations are expected to be due to chance. Thus with 23 independent significant associations, it would appear likely that factors other than chance are causing the marker genotype by site interaction.

It is notable from Table 3.18 that there are no significant marker by site associations for the trait of Pilodyn penetration, whereas for example there are nine for growth at age 6 years (dbh6).

The nine markers with significant ($p < 0.01$) marker x site associations for diameter at year six are listed in Table 3.19. Among this group of markers is marker 266-6, which was estimated to be located at the QTL on M2 for diameter at age three years (dbh3) (Table 3.14). Clearly there is a significant marker genotype by site interaction for this marker when considered across all seven sites at year six ($P = 0.0037$), however marker 266-6 did not have significant marker genotype by site interaction at year three when only the four Tasmanian sites were measured. It is also to be noted that marker 266-6 had significant marker genotype effects for stem diameter at year three (Table 3.14), but not at year six (Table 3.19).

Table 3.19. Markers with significant ($p < 0.01$) marker genotype by site interaction for growth at year six (dbh6).

linkage group/ marker	*no. progeny scored*	*p-value for marker genotype*	*p-value for mkr class x site*
M6 20-7	153	0.87	0.0011
F8 CSIRO 03	153	0.12	0.0014
M3 256-1	152	0.78	0.0035
M2 266-6	155	0.33	0.0037
M7 [a]	71	0.82	0.0041
F4 215-2	154	0.89	0.0044
F7 218-2	153	0.15	0.0076
M12 [a]	85	0.72	0.0081
M9 493-1	150	0.82	0.0087

[a] These markers were dropped during the ordering process and do not appear on the linkage map.

3.3.5 Markers with significant QTL effects and marker genotype by site interaction

Markers that were found to have a significant marker genotype (QTL) effect (the first effect specified in the model used to detect site interaction) were examined for significant marker genotype by site interaction. From the 3,200 tests carried out there were 51 marker trait associations that were found to be significant at the 0.01 level. Among these 51 markers, a group of eight (mapping to five positions) were found to be significant at the 0.05 level for marker genotype by site interaction. All eight markers were from linkage group three of the female parent (F3). Six of these eight markers were framework markers that had associations with dbh6. The framework markers 81-1, 266-5 and 237-1, which map to the same position on F3, had the most significant marker genotype effect as well as the most significant marker by site interaction amongst this group of eight markers (Table 3.20). These three markers, represented by marker 81-1, were estimated to be co-segregating with QTL F3, the QTL for growth at year 6 (dbh6) (Table 3.14).

Thus the significant marker genotype by site interaction for marker 81-1 indicates that the F3 QTL has a strongly site dependent effect on growth. This is illustrated by looking at the difference between marker genotype means for marker 81-1 (QTL F3) at each of the seven sites where measurements for growth at year 6 (dbh6) were made (Fig. 3.9). The QTL effect is strongest at the mainland sites (Mansfield, Flynn and West. Aust.), with two of the Tasmanian sites producing neutral QTL effect (Parkham and Boyer) and the other two (Franklin and West Ridgley) giving small effects in the same direction as the mainland sites (Fig. 3.9).

Table 3.20. Markers with significant (P < 0.01) marker genotype (QTL) effects that also had significant (P < 0.05) marker genotype by site interaction. [a]

Trait	*Parent/ Link. group*	*Marker*	*p-value mkr class*	*p-value mkr class x site*
dbh6	F3	81-1	0.00019	0.012
	F3	266-5	0.00019	0.012
	F3	237-1	0.00024	0.013
	F3	234-4	0.00158	0.014
	F3	241-3	0.00158	0.014
	F3	EMCRC4	0.00739	0.013
ht2	F3	N87 [b]	0.00950	0.040
	F3	N237 [b]	0.00953	0.047

[a] All p-values are based on an analysis using the plot residuals. Marker 81-1 is estimated to be 0 cM from the QTL for growth at year six (Table 3.14).

[b] These markers link to F3 but were dropped during the ordering process.

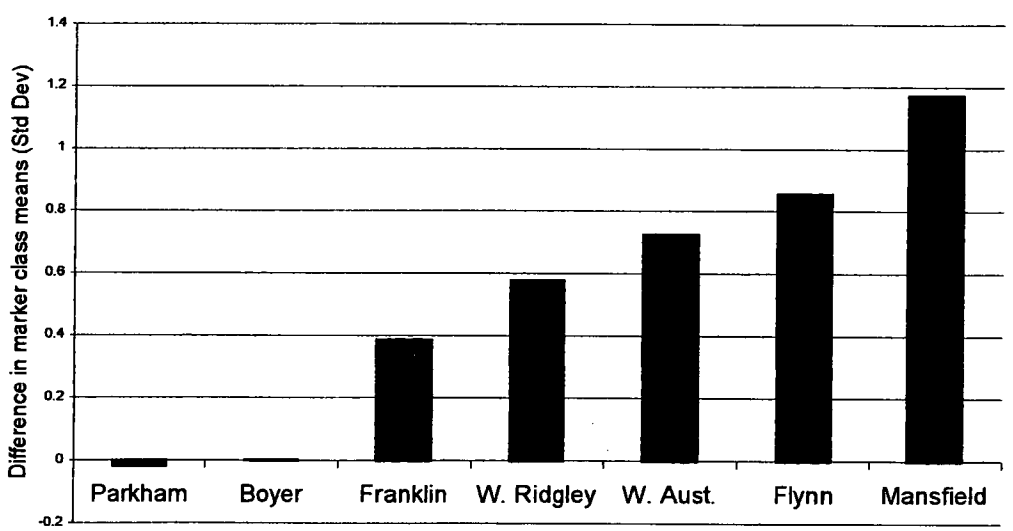


Fig 3.9. Difference in marker class means at each site for marker 81-1 (QTL F3) for the trait of stem diameter at year 6 using plot residuals.

3.3.6 Temporal stability of QTL for growth

There were two cumulative growth QTL that had LOD scores greater than 2.0 (M2 and F3, Table 3.14). Both of these QTL were examined (independently) for their stability with time. To avoid confounding QTL x site interaction with age differences, QTL analyses were carried out using trees measured in common between ages under comparison (see section 3.2.6).

3.3.6.1 Temporal stability of QTL M2 for stem diameter at year 3

To compare the size of effect of the QTL for diameter at year 3 (M2) with other years, analyses were carried out for the other growth traits using the 84 trees measured for dbh3 (Table 3.21, numbers not in parentheses). Based on these 84 trees, for the trait of stem diameter at year 2, the M2 QTL had a LOD score (LOD 1.9) close to the threshold value of 2.0 (Table 3.21). This level of significance was maintained across the four measurements of stem diameter from year two (dbh2), to the final measurement of diameter at year 6 (LOD range 1.8-2.0; Table 3.21). For height at year 1, the M2 QTL was in fact more significant (LOD 2.4) than for diameter at later ages (Table 3.21). However for height measured a year later (ht2), there was a large drop in significance from LOD 2.4 to LOD 1.3. The effect of this QTL on incremental height between year 1 and 2 (ht2-1) was of opposite direction to the effect on the cumulative growth traits, that is the QTL represented a region that was weakly correlated with a lower mean height increment. This combined with the larger LOD for dbh2 than ht2 (Table 3.21) and the greater correlation between ht1 and dbh2 than between ht1 and ht2 (Table 3.11 section 3.3.2), suggests that trees that were taller at year 1 invested more in

diameter growth than height growth in the following year, compared with shorter trees.

Table 3.21. Stability with age for QTL M2 for growth (dbh3).

Trait	*LOD score*[a]		*Variance explained (%)*		*Sample size*
ht1	2.4	(1.1)	12.6	(3.8)	84 (131)
ht2	1.3	(0.0)	6.7	(0.3)	84 (131)
dbh2	1.9	(0.7)	9.9	(2.7)	82 (121)
dbh3	2.0	(2.0)	10.5	(10.5)	84 (84)
dbh4	1.9	(0.6)	11.2	(2.9)	74 (92)
dbh6	1.8	(0.2)	9.2	(0.6)	84 (155)
ht2-1	0.5	(1.0)	2.7	(3.5)	84 (131)
dbh4-2	0.3	(0.0)	1.7	(0.0)	72 (90)
dbh6-4	0.0	(0.0)	0.0	(0.2)	74 (92)

[a] Numbers not in parentheses are the LOD score, variance and sample size for the trait based on trees measured in common with dbh3. Numbers in parentheses are the LOD score, variance and sample size when all trees measured for the trait are included in the QTL analysis.

It is possible that the effect of M2 on diameter may result solely from its association with height at year 1 and have maintained significance for diameter growth due to this initial effect. The M2 QTL effect can be seen to be site (or sample) dependent since including trees from mainland sites in analyses of cumulative growth traits resulted in reduced QTL significance in all cases (LOD scores in parentheses, Table 3.21). This effect was most apparent for dbh6 where the LOD score decreased from 1.8, when the 84 trees measured in common with dbh3 were included in an analysis, to 0.2 when all 155 trees were included.

Using the marker estimated to be closest to QTL M2 (266-6), the marker genotype differences, or QTL effect, across sites and with time is illustrated in Fig. 3.10 (a-c) for (a) height at year 1 (ht1), (b) diameter at year 3 (dbh3) and (c) diameter at year six (dbh6). For QTL M2 for the four Tasmanian sites, the QTL effect remained relatively stable with time. This was also reflected in the stable LOD scores for analyses of these 84 trees (Table 3.21). The QTL effect at the mainland sites in Western Australia and Flynn (Vic.) was of opposite orientation to the Tasmanian sites but remained relatively stable with age. The instability of the Mansfield site is likely due to the small sample size of eight trees.

(a) Effect by site of QTL M2 on ht1

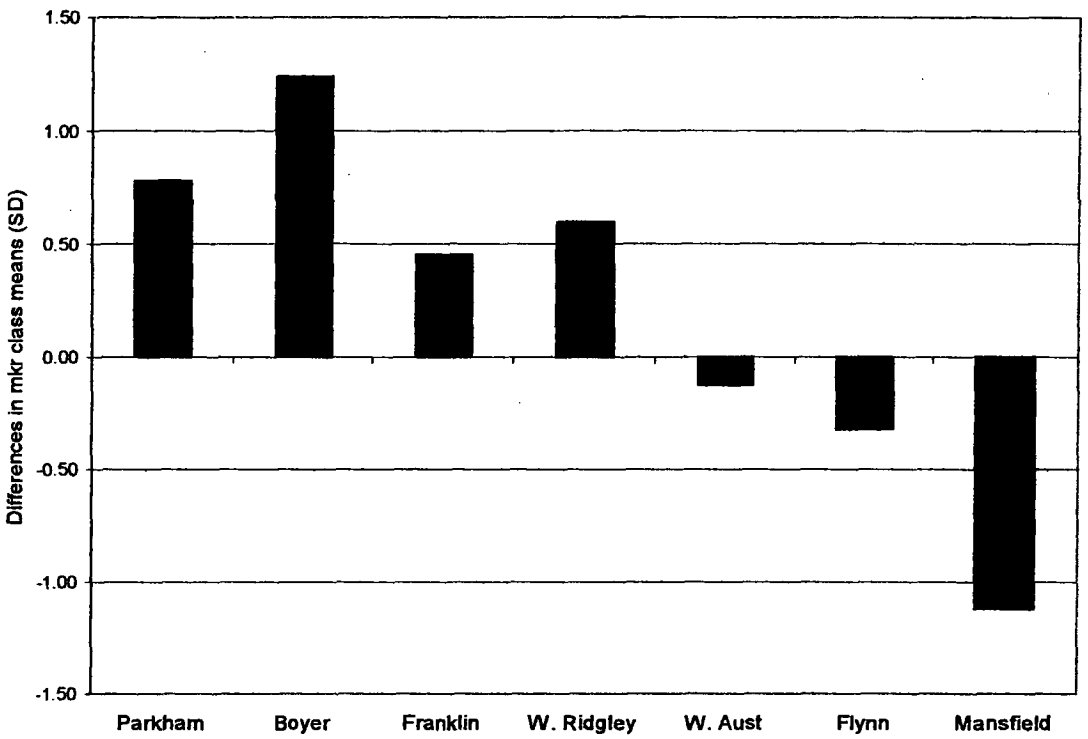
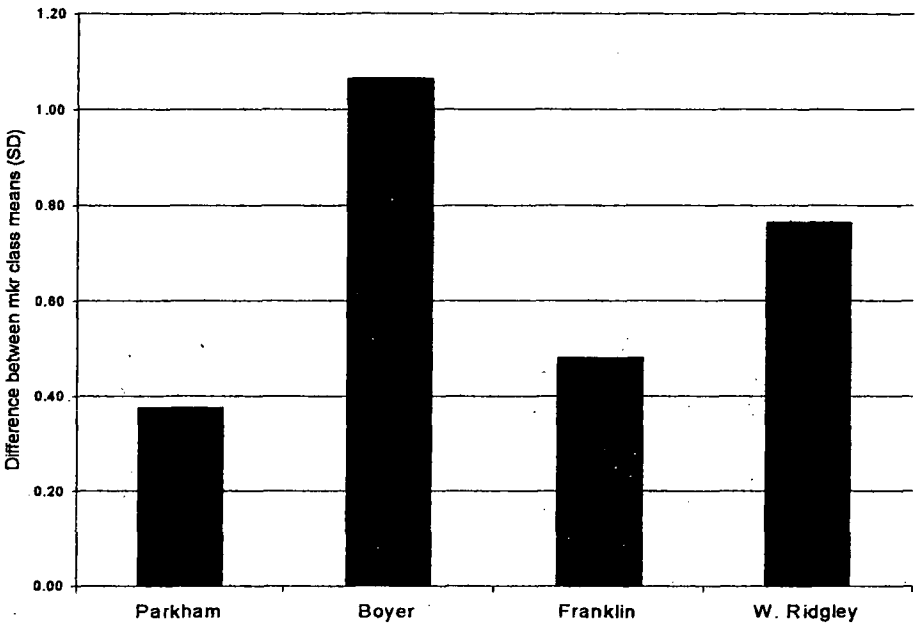
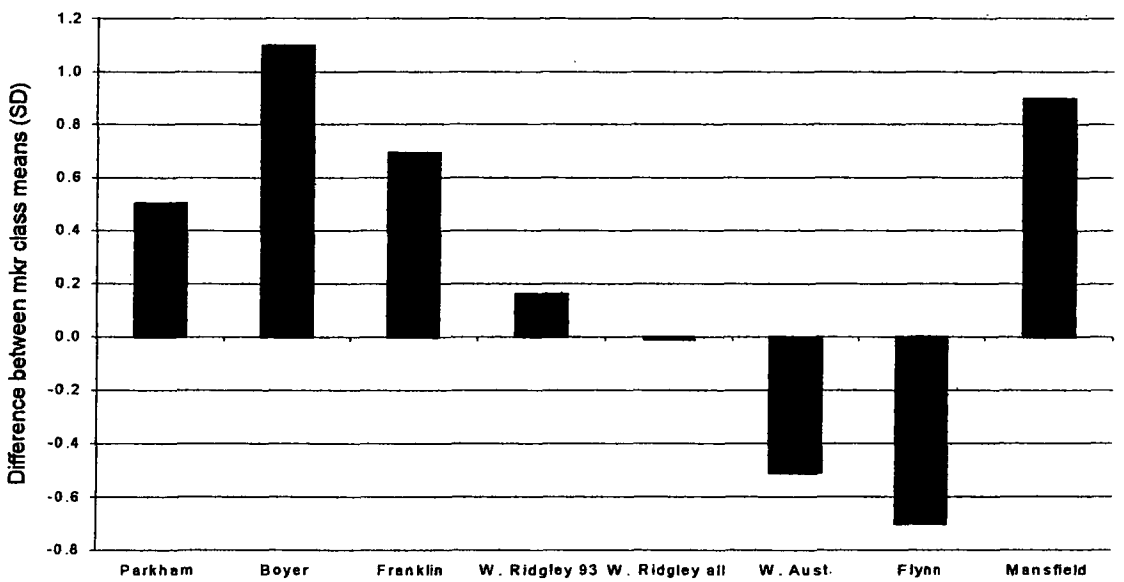


Fig. 3.10. Effect of site and age on QTL M2 for growth. The QTL effect is measured here as a difference between marker class means (in standard deviations). This has been calculated at each site for the marker (266-6) closest to QTL M2, at three ages of measurement of growth (a) height year 1 (ht1), (b) diameter at year 3 (dbh3) and (c) diameter at year 6 (dbh6).

(b) Effect by site of QTL M2 for dbh3



(c) Effect by site of QTL M2 for dbh6



(Fig. 3.10. cont.)

3.3.6.2 Temporal stability of QTL F3 for stem diameter at year 6

The F3 QTL for growth, which was estimated to be located near marker 81-1, was most significant in year six (dbh6) when 155 G1025 trees were measured for stem diameter (Table 3.22). However F3 was associated with growth, albeit below threshold QTL levels, from the first measurements of both height (ht1) and diameter (dbh2) (Table 3.22). As was the case with the QTL on M2, the association of F3 with ht1 and dbh2 was stronger than the association with ht2. These observations support the suggestion made earlier, that taller trees at year 1 invested preferentially in diameter growth rather than height growth in the second year compared with shorter trees.

Table 3.22. Stability with age for QTL F3 for growth (dbh6).

Trait	*LOD score for growth traits*[a]	*Variance explained (%)*[a]	*Equivalent LOD score for dbh6*[b]	*Variance explained dbh6 (%)*[b]	*Sample size*[a, b]
ht1	1.6	6.0	2.5	8.3	131
ht2	0.7	2.6	2.5	8.3	131
dbh2	1.6	5.9	1.9	6.9	121
dbh3	0.36	2.1	0.44	2.5	84
dbh4	0.7	4.0	0.9	4.3	92
dbh6	2.8	8.0	2.8	8.0	155
ht2-1	0.3	1.1	2.5	8.3	131
dbh4-2	0.0	0.1	0.9	4.3	90
dbh6-4	0.5	2.6	0.9	4.3	92

[a] LOD score, variance and sample size for the trait analysed at QTL F3.

[b] LOD score, variance and sample size for the trait of stem diameter at year six (dbh6) at QTL F3, using only the subset of trees measured for the growth trait indicated.

The low LOD scores observed for QTL F3 when the growth traits dbh3 and dbh4 are analysed, are however close to the scores for dbh6 when measured in the same set of trees (Table 3.22). This indicates that the low LODs for these traits are mostly due to the sites/trees measured at years 3 and 4 rather than an age effect (Table 3.22). However for each trait, the LOD score is larger when the data from dbh6 are analysed for the same set of trees indicating that the influence of the growth QTL became stronger with time.

The early association of QTL F3 with growth is illustrated by the QTL scans of the linkage group for the traits of ht1 and dbh2 compared with dbh6 (Fig. 3.11a-c). As might be the case for QTL M2, the association of QTL F3 with growth at later ages may have resulted from an association in the first year that remained detectable. However in the case of the F3 QTL, its significance increased with age, suggesting possible expression of the QTL at later ages.

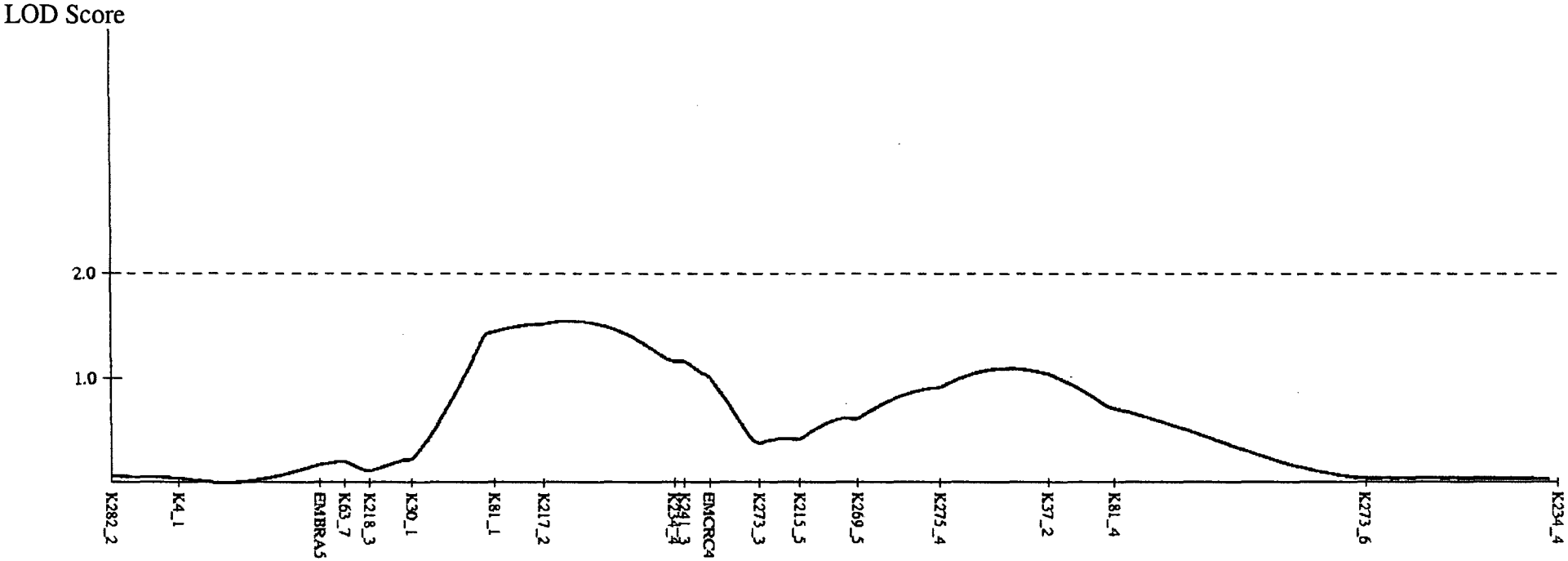


Fig. 3.11. (a) Expression of QTL F3 over time. Growth at year 1 (ht1). Scan of linkage Group 3 of the female parent using plot residuals. A LOD score peak for growth at year 6 (dbh6) occurs at marker K81_1.

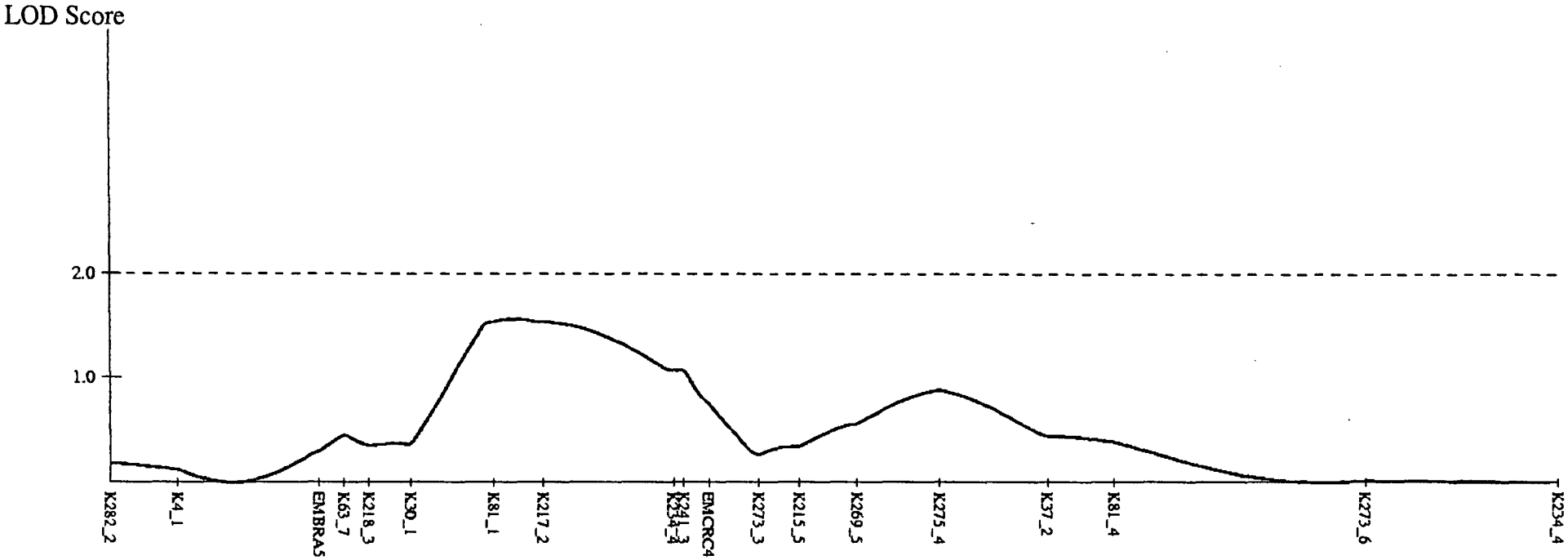


Fig. 3.11. (b) Expression of QTL F3 over time. Growth at year 2 (dbh2). Scan of linkage Group 3 of female parent using plot residuals. A LOD score peak for growth at year 6 (dbh6) occurs at marker K81_1.

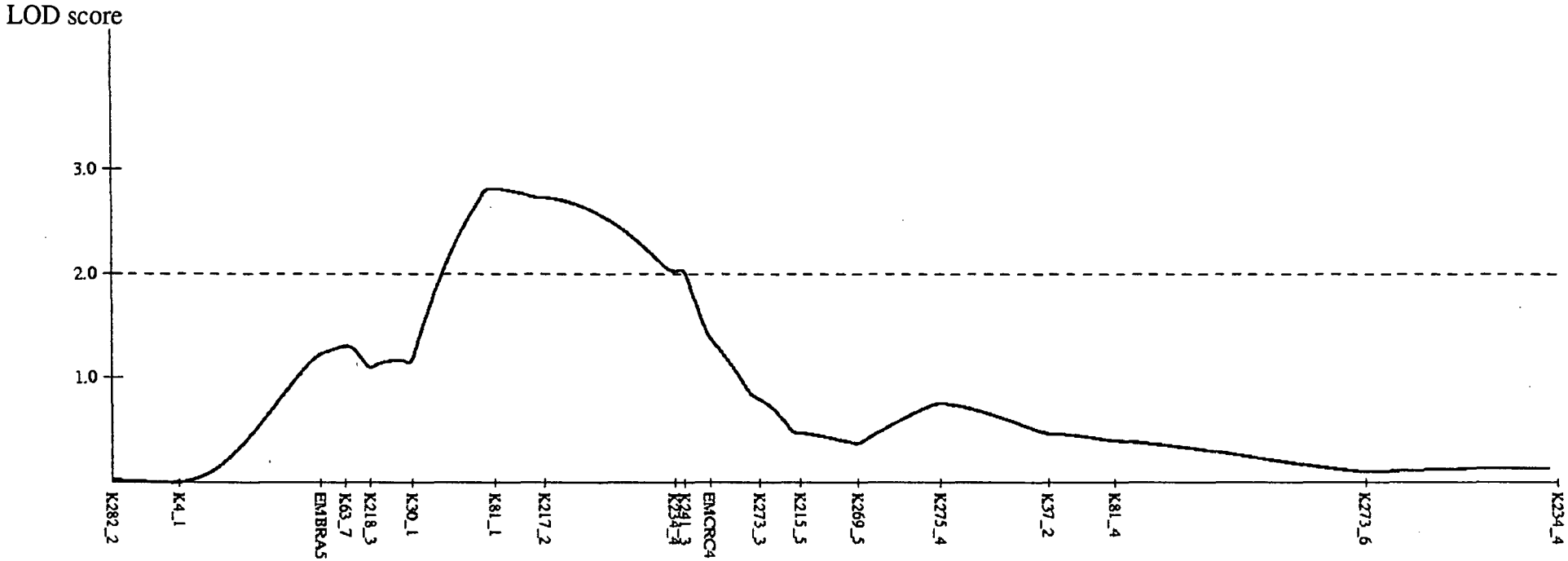


Fig. 3.11. (c) Expression of QTL F3 over time. QTL for growth at year 6 (dbh6). Scan of linkage Group 3 of the female parent using plot residuals. A LOD score peak occurs at marker K81_1.

Stability of the F3 QTL across sites and with age can be seen from the differences in the trait means between marker genotypes at each site, across three times of growth measurement (Fig. 3.12 a-c). The QTL effect across sites has not changed markedly with time, except for the Mansfield site where the sample size is small. By year six the mainland sites were contributing most to the observed QTL effect (Fig. 3.12 c).

The interaction of the F3 QTL with site at year 6 is highlighted, when the 71 trees not measured at year three (for dbh3) and grown mostly at mainland sites, are utilised in a QTL analysis for stem diameter at year six (dbh6). The F3 QTL in this analysis (using dbh6) has a LOD score of 2.6 (data not shown) compared with the LOD of 0.4 in a corresponding analysis (also using dbh6) using the 84 trees that were measured in year 3 (Table 3.22). The difference between LOD scores being due to site (or sampling) effects.

In summary the M2 QTL had a stable positive QTL effect when measured across the Tasmanian sites, for the four years of diameter measurements. The effect of the M2 QTL on height at year one across the Tasmanian sites was even stronger. The effect of the F3 QTL was observable in the earliest growth measurements, but was greatest at year six. The effect of the F3 QTL probably increased steadily with age, however missing data makes it impossible to determine the extent of the QTL effect in the intervening years. Effects of both QTLs were found to be strongly site dependent, with the M2 QTL effect dependent on Tasmanian sites (Fig. 3.10) and the F3 QTL more dependent on mainland sites (Fig. 3.12). For both QTLs much of the QTL effect observed at later ages might have resulted from the association with growth evident at the first year of measurement, rather than a continuous expression of these QTL.

(a) Effect by site of QTL F3 on ht1

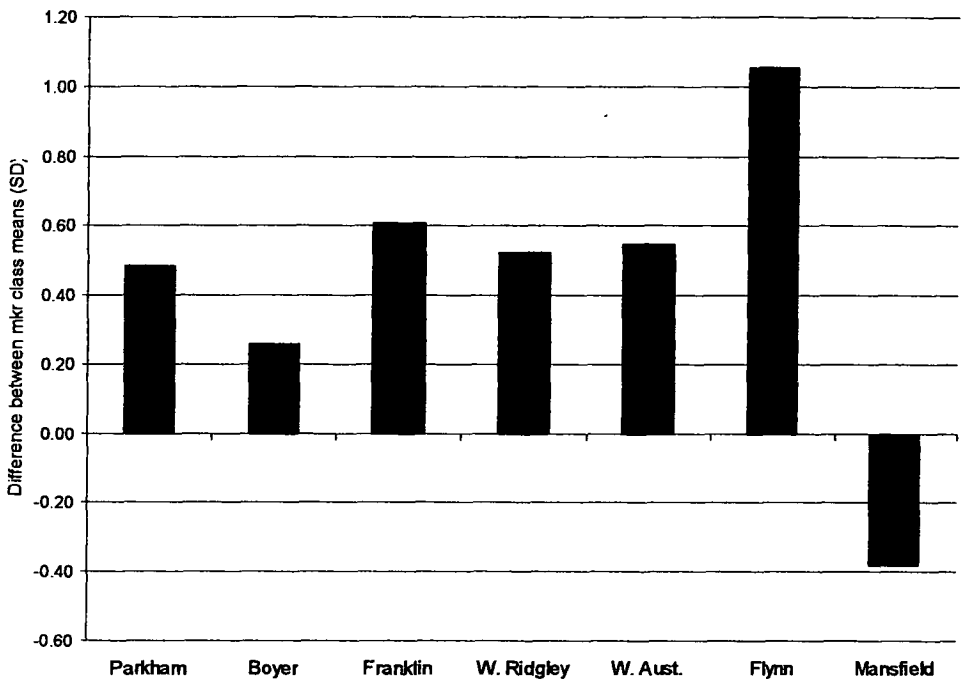
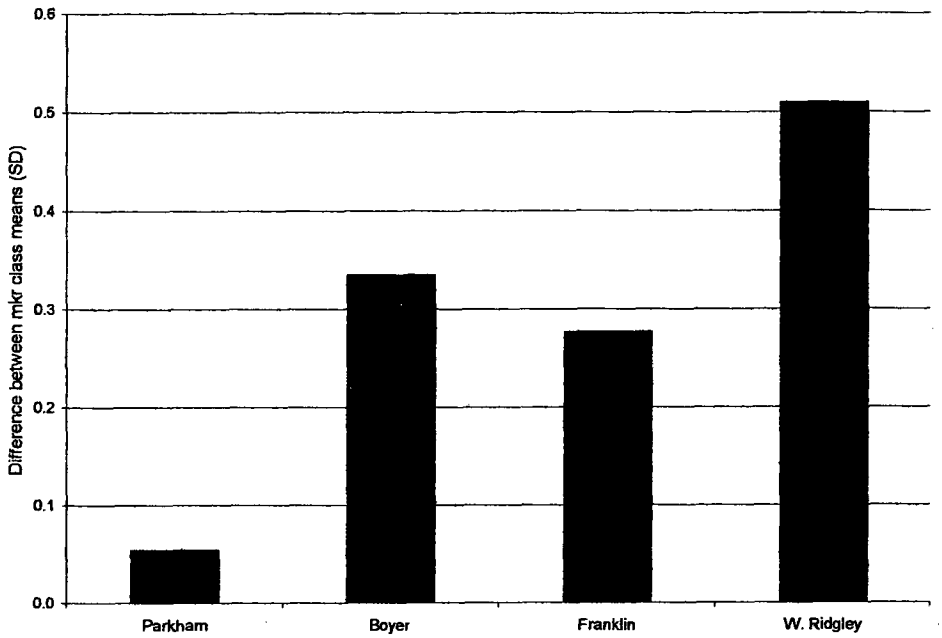
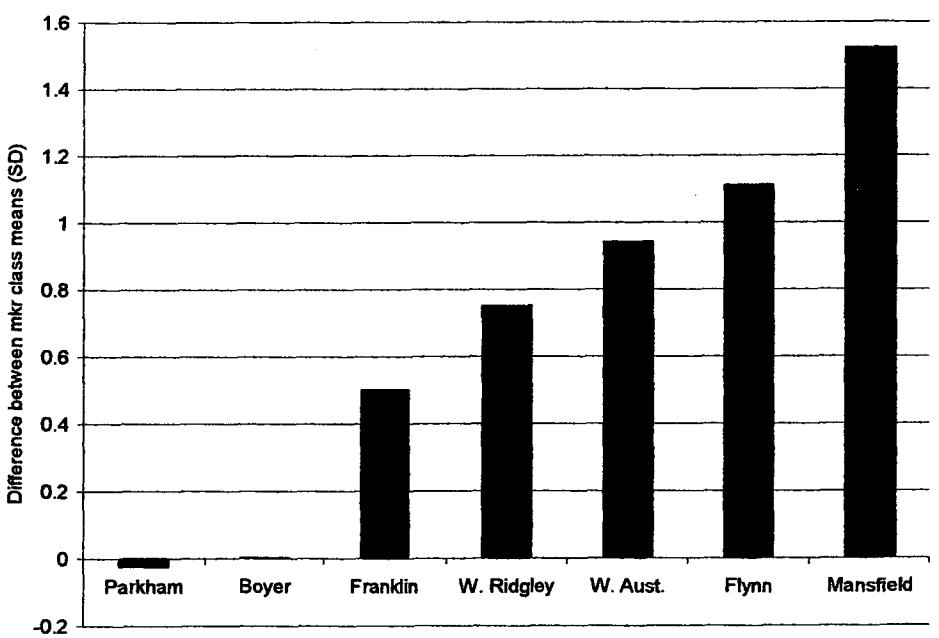


Fig. 3.12. Effect of site and age on QTL F3 for growth. The QTL effect is measured here as a difference between marker class means (in standard deviations). This has been calculated at each site for the marker (81-1) closest to QTL F3, at three ages of measurement of growth (a) height year 1 (ht1), (b) diameter at year 3 (dbh3) and (c) diameter at year 6 (dbh6).

(b) Effect by site of QTL F3 on dbh3



(c) Effect by site of QTL F3 on dbh6



3.4 DISCUSSION

3.4.1 Number and size of effect of detected QTLs

In a review of twenty QTL detection studies in forest trees, Sewell and Neale (2000) found that the number of QTLs identified for all traits reported ranged from 0 to 7 per trait with a mean of 2.7. However the probability of QTL detection is likely to be dependent on the trait studied and be greater for traits with higher heritability. Also, comparison of results between studies is not straightforward, due to the different methods of QTL analysis and varying QTL detection power as a result of different population sizes.

There have been a number of studies identifying QTLs for growth and wood density in forest trees, though none published for these traits in *Eucalyptus globulus*. In a *Eucalyptus grandis* x *urophylla* hybrid F1 family, Grattapaglia *et al.* (1996) found one putative QTL for growth (circumference at breast height) and four for wood density (measured by specific gravity) with LOD scores greater than 2.0. The QTL were inherited from the maternal parent and detected using a half-sib family of 300 progeny measured at age 6.5 years. In a second *Eucalyptus* hybrid study Verhaegen *et al.* (1997) found four putative QTLs (LOD > 2.0) for vigour (equivalent to growth) and five for Pilodyn penetration at age 2 years (26 months) based on 142 progeny. It would thus appear that in comparison the detection of two putative QTLs for cumulative growth (from 6 measurements) and two for wood density in the present study is a low yield. This is not really surprising however, since the presence of detectable QTL effects, especially in populations that have not been constructed for the purpose, will be strongly dependent on chance (Beavis 1998). In addition phenotypic measurements in the

two cited studies were from single sites compared with the present study where trees were grown across seven sites. Depending on such factors as the number of trees at each site, the number of sites, the degree of difference between sites and the extent of QTL by site interaction, the probability of QTL detection could be increased or decreased in multiple site experiments relative to single site experiments.

In the present study, with a relatively small number of trees at each site and a relatively large number of diverse sites, the likelihood would be for a reduced number of QTLs detected compared to a single site study, given there is some degree of QTL by site interaction. Also, the likelihood of QTL detection may have been reduced due to site effects, since even though these were estimated and removed in the present study, this can only be done imperfectly and some between site effects would remain in the residuals reducing QTL detection power. It is interesting to note that in the present study at age six years and the two cited studies, there has been a larger number of QTLs detected for the trait with the higher heritability – wood density, than for cumulative growth.

Most other examples of QTL detection studies in forest trees for growth traits, have been carried out on single site plantings. Lerceteau *et al.* (2001) in a study of *Pinus sylvestris*, found three QTLs (LOD > 2.0) for height (each evident at four ages) and four for stem diameter (at two ages) based on 94 progeny from a full-sib family. Emebiri *et al.* (1998) found numerous QTL for growth at four ages using 93 progeny, and for stem diameter alone at 2 years of age six QTLs were detected. In *E. nitens* in a study of seedling height at 55 days of age, three QTLs were detected ($P < 0.01$) (Byrne *et al.* 1997a).

There are no reports in the literature of QTL studies for the traits of bark thickness and only a single study of flowering abundance in which flowering was induced by gibberellin leaf spray in the conifer sugi (*Cryptomeria japonica*) with three QTL detected for this trait (Yoshimaru *et al.* 1998).

Thus the number of QTLs detected in the present study is at the lower bounds for comparable published studies in forest trees. It is suggested that one reason for the small number of QTLs detected may be the multiple sites inflating the unexplained variance due to between site effects and with only those QTL that are effective at a number of sites being detected. Chance also may play a part along with the likelihood that studies that fail to yield QTL go unreported.

There is evidence from a number of studies in forest trees, that within a family a significant proportion of the phenotypic variance for traits of commercial importance can be controlled by a few QTLs (Groover *et al.* 1994; Bradshaw and Stettler 1995; Grattapaglia *et al.* 1995; Grattapaglia *et al.* 1996; Verhaegen *et al.* 1997). This is the situation in which MAS would be most attractive, since by selecting for a few markers there would be a considerable genetic gain for the trait of interest. In the present study it was estimated that 20% of the phenotypic variance for wood density in a full-sib family was explained by the segregation of the two putative QTL, thus adding to the list of examples where a reasonable proportion of the phenotypic variance can be explained by the segregation of a small number of QTL. The other QTL detected here also explain a reasonable proportion of the phenotypic variance (8.0% - 10.5%). This would seem to indicate that there is significant and potentially useful genetic diversity for these traits within populations, since the QTL were detected in an intraprovenance

cross. However for some traits no QTL were detected and also, based on simulation studies reported by Beavis (1998), it has been claimed that the generally small sample size used for QTL detection studies in forest trees can lead to the over estimation of the size of QTL effects (Sewell *et al.* 2000; Lerceteau *et al.* 2001). Thus the reported large effects may be overestimated. Beavis (1998) has also suggested that sampling error will lead to QTLs being detected that have a broad range of effect sizes even when the real effects are of equal magnitude (van Buijtenen 2001). Clearly QTLs require validation in other populations, preferably of large size.

3.4.2 Genomic distribution of QTLs

Linkage group 11 of the male parent (M11) was the only genomic location that had a significant QTL effect (LOD>2.0) for more than one trait. M11 was significant for relative incremental growth (dbh4-2) and also for wood density (Pilodyn penetration). The influence of M11 on both incremental growth and wood density is significant because growth and wood density have been found to have slight negative genetic correlation in *E. globulus* (MacDonald *et al.* 1997). The direction of QTL effect is opposed for the two traits and a QTL acting in this manner supports the observed genetic correlation. It is tempting to hypothesise that the QTL effects are the result of the pleiotropic effects of a single gene. The association in this case would not be a chance association of closely linked loci within the male parent, but is likely to be found in other individuals with the direction of effects opposed as in the present case.

Support for this hypothesis can be obtained by testing this genomic region in other populations to determine if QTL influencing both wood density and stem

diameter are segregating. Unfortunately the linkage group of two markers (5.7cM appart) does not correspond to a homologous group on the female genome and there is no linkage to microsatellite markers. Thus to test this hypothesis directly will probably require that closely linked microsatellite markers are found for the M11 region. However QTLs for growth and wood density have also been co-localised in other QTL studies in *Eucalyptus* (Grattapaglia *et al.* 1996; Verhaegen *et al.* 1997). Grattapaglia *et al.* (1996) found two QTLs for stem circumference, both of which co-localised with QTL for wood density (specific gravity) with both QTL having opposing effects on the two traits. Verhaegen *et al.* (1997) found that four QTL for wood density (Pilodyn penetration) were co-localised with growth (vigour), three inherited from the *E. grandis* parent and one from the *E. urophylla* parent. In three of these cases the allele that increased growth also increased Pilodyn penetration (ie. growth and wood density were in opposition) as is the case observed for the QTL in this study. Taken together these results indicate that pleiotropy is likely to be a good explanation for these observations. Alternatively genes affecting growth and wood density may be clustered together in a number of regions in the *Eucalyptus* genome.

In contrast the QTL for Pilodyn penetration on F6 was not associated with QTL for growth. Wood density and growth have been found to have only a partial negative genetic correlation (MacDonald *et al.* 1997) and thus some genetic influences on wood density would be expected to be independent of genetic effects on growth. The F6 QTL potentially represents a gene or genes that have an influence on wood density without influencing the growth rate.

3.4.3 QTL by environment interaction

QTL by environment interaction has been investigated in a number of crop species. In tomato Patterson *et al.* (1991) found only four environmentally stable QTLs influencing fruit characteristics from a total of 29 that were significant across the three environments tested. Ten QTLs had partial environmental stability and were detected in two environments, with 15 QTLs being environmentally specific. Both stable and environment dependent QTLs have been found in maize (Stuber *et al.* 1992) and rice (Zhuang *et al.* 1997). However in tree species few studies have investigated QTL by environment interaction. Groover *et al.* (1994) found evidence for QTL by environment effects for two wood density QTLs in loblolly pine. Weng *et al.* (2002) found two marker intervals that had QTL by environment interaction in a longleaf pine x slash pine backcross. One interval was significant for both incremental and cumulative stem diameter between 7 months and 16 months of age, the other was significant for incremental stem diameter only, where two sites were contrasted.

In the present study a preliminary investigation of QTL stability across sites has been carried out. QTL F3 for stem diameter at year six (dbh6) was found to have significant QTL by site interaction at the 0.05 level ($p = 0.012$). Expression of QTL F3 appears to have been stronger at mainland sites than at Tasmanian sites with no correlation between size of QTL effect and average site performance for the trait (compare Fig. 3.1(f) and Fig. 3.10). The QTL interval for QTL M2 for stem diameter at year 3, was found to have significant QTL by site interaction at the 0.01 level, when tested at year six. The QTL effect for this interval was however not significant (at 0.05) when all sites were considered at year six, but there was a reasonable QTL effect (LOD 1.8) when only the sites measured at

year 3 are tested at year six. Both the observed interaction and the QTL effect require investigation in other populations before conclusions could be made on the relevance of these QTL across sites.

Some regions of the genome that did not have significant QTL effects were found to have strong site interaction. Since there was more interaction than expected by chance (23 locations found versus seven expected), it is suggested that some of these regions represent QTL that are strongly interactive with the environment. Strong environmental interaction reduces the possibility of detection of these QTLs based on average marker genotype effect used in the other QTL detection analyses where interaction was ignored. In agreement with results suggesting that Pilodyn penetration has little genotype by environment interaction (MacDonald *et al.* 1997) no marker by site associations were found for this trait. This is in contrast to most of the other traits, where 1-9 locations in the genome had marker by site interactions. Regions of the genome that were interactive with site were particularly common for stem diameter at year six (nine regions).

The multiple site nature of this study has enabled an examination of marker genotype by site interaction, which may aid the identification of QTL that are likely to have relatively stable performance across sites. It has also enabled the identification of genomic regions that warrant further examination as harbouring QTLs that are strongly interactive with site.

3.4.4 Temporal stability of QTL for growth

A number of studies of forest trees have addressed the issue of QTL stability particularly with respect to time. Verhaegen *et al.* (1997) investigated stability of

growth and wood density QTLs over a three year period in a single full-sib family of hybrid F1s in an *E. grandis* x *E. urophylla* cross. Stability across time was mixed, no QTLs were significant for all three times of measurement but 68% were significant at two ages. In other forest tree genera the stability of QTLs for growth have been investigated in a *Populus* hybrid F2 (Bradshaw and Stettler 1995) where no QTL were significant for the two measurement times. In *Pinus radiata* Emebiri *et al.* (1998) found QTLs for growth followed one of three trends with time: a linear increase (4 times), a linear decrease (12 times) and a curvilinear trend with QTL effect peaking at an intermediate age (15 times). In a study of *Pinus sylvestris* that included measurement of height at four age intervals (years 9, 10, 11 and 12) three QTL were detected with each one being significant (ie. > LOD 2.0) at each of the four ages of measurement and two significant for height increment as well (Lerceteau *et al.* 2001). In *Pinus taeda*, Kaya *et al.* (1999) found that height and diameter increment QTLs were not detected for consecutive growing seasons in two pedigrees grown on multiple sites.

In the present study the two QTL identified for growth (M2 and F3), were found to be detectable from the first measurements of height growth at year one and diameter at year two. However only one of these QTL effects had a significant LOD score peak, and this effect was found to be site dependent. Nevertheless, for a subset of sites, the M2 QTL (for dbh3) was found to be significant for height at year one, and to have a stable effect on stem diameter over the four years of measurement. The stability with time of the other cumulative growth QTL, QTL F3 (for dbh6), was difficult to evaluate due to measurements being unavailable for some sites in some years combined with significant interaction with site. However it was possible to show that the effect of this QTL was evident at year

one for height and year two for diameter (LOD 1.6), with an increase in QTL effect at year six.

The effect of both QTL M2 and QTL F3 on stem diameter through time may result solely from their initial association with height at year 1, although there is slight evidence (an increase in relative LOD score) for later expression of QTL F3. This is however exactly the type of QTL effect that would be most convenient for future evaluation, because if the QTL effects are detectable at year one, validation and further testing would be much quicker than for a QTL that is only detectable close to rotation age.

3.4.5 Statistical significance and relevance of detected QTLs

In the present study two QTLs for wood density (Pilodyn penetration) have been detected at age six years, two QTLs for cumulative growth (one each for diameter at age 3 years and diameter at age 6 years), three for relative incremental growth and one for bud abundance at three years have also been detected. This has been based on a threshold LOD score of 2.0. Only one of the five QTLs for the measured traits (QTL F6 for Pilodyn penetration) had an experiment-wise significance level below the standard 0.05 level. The experiment-wise significance level for three of the others was close, being significant at the 0.1 level. However studies such as the present study, being based on a relatively small number of progeny (<500) and with traits of relatively low heritability are best regarded as exploratory – useful to identify candidate (but not necessarily statistically significant) QTLs for further study. This is because for relatively small population sizes, sampling error is prone to lead to inaccurate estimations

of the size of QTL effects (Beavis 1998) and power for QTL detection is low, especially when the whole genome is scanned for multiple traits.

Beavis (1998) argues that for exploratory studies the usual type I error threshold of 0.05 is not entirely sensible since the aim is really to identify candidate QTL for further study, and a stringent type I error threshold will increase the frequency of type II errors, resulting in genuine QTLs being overlooked. Candidate QTLs, some of which will not be statistically significant, can be tested in subsequent studies for the appropriate trait associations, where it is likely that experiment-wise significance thresholds (type I error thresholds) would be at least an order of magnitude larger than the initial study (assuming similar numbers of progeny).

The appropriate level of type I error for exploratory studies cannot be calculated exactly and must be judged by the investigator. In the present study QTL loci below the 0.05 level of experiment-wise significance have been reported with the justification that it is an exploratory study. In fact for correct statistical treatment the threshold experiment-wise levels of significance (type I error) calculated here for each trait need to be reduced as a result of the testing of more than one trait, a fact that is rarely considered in QTL detection studies.

For marker assisted selection of growth and wood density, only those QTLs that influence performance as measured at rotation age are likely to be worthy of selection. Thus QTL for cumulative growth at rotation age would be the most useful growth QTL for MAS purposes. Incremental growth QTLs are of interest from the perspective of gene expression patterns and interactions influencing growth but also may become significant as cumulative growth QTLs at later ages. The main intention in this study for examining incremental growth traits was to

determine how cumulative growth QTL behaved over incremental periods and to see if they were significant as incremental growth QTL. There was however no correspondence between the location of cumulative growth QTL and incremental growth QTL. To determine if the incremental growth QTLs detected here become significant as cumulative growth QTLs obviously requires a follow-up study.

It would be expected that at least some of the QTL identified in this study result from actual genetic effects as opposed to chance associations. QTL F6 (pilo6/1) for Pilodyn penetration is obviously the QTL most likely to result from genetic effects since it has the most statistically significant effect (0.002 experiment-wise significance). However verification in further experiments is required to determine which QTLs are likely to be important. Although it is uncertain as to precisely how marker/QTL information would be implemented in marker-assisted selection in the breeding of *E. globulus*, it is likely that information on QTL stability at least across environments and with different genetic backgrounds would be critical (Bradshaw and Grattapaglia 1994).

This study has addressed, in a preliminary way, the issue of QTL stability across environments and also with time. For MAS it would also be likely that testing of the same QTL locations in other pedigrees would be carried out and also the use of larger population sizes will be required to accurately determine the size of the QTL effect. To enable the collection of the foregoing information markers that are likely to be polymorphic in many crosses will probably be required (eg. microsatellites).

3.5 CONCLUSION

The present study indicates that putative QTL for traits of commercial importance, with low to moderate heritability, are detectable in a *Eucalyptus globulus* cross using populations of moderate size. Although few of these QTL have experiment-wise statistical significance at the generally accepted 0.05 level, the study can be considered exploratory. For wood density, two QTL explained 20% of the variance for the trait, indicating that a small number of QTL might explain a reasonable proportion of the trait variance. One of these QTL was found to be independent of QTL for growth whereas the second QTL co-segregated with a QTL for relative incremental growth, with the presence of the marker nearest to this QTL being linked to faster growth but lower wood density. It was also observed that there was more marker genotype by site interaction than expected by chance for growth traits. This is likely to be due to the presence of QTL that are strongly site interactive. In addition, growth QTLs were observed to have temporal stability extending back to an early age.

CHAPTER FOUR

AN ANALYSIS OF THE ROOTING ABILITY OF MICROPROPAGATED SHOOTS FROM TWO FAMILIES OF *EUCALYPTUS GLOBULUS*

4.1 INTRODUCTION

Transferring genetic gains to plantations through the deployment of clones of superior genotypes is a strategy that has been used successfully with tropical and subtropical *Eucalyptus* species such as *E. grandis, E. urophylla* and their hybrids (Zobel 1993). Propagation from stem cuttings taken from coppiced mother trees has generally been found to be a successful, convenient and economic method for clonal deployment of these eucalypts.

However in some temperate species of *Eucalyptus* such as *E. globulus* cloning ability has been found to vary greatly between individuals, with the cloning ability of many genotypes too low to be economically viable for plantation deployment (Wilson 1992). A crucial factor determining clonability from cuttings is their capacity to produce roots. This capacity, known as rooting ability, has been found to be highly variable within many forest tree species with some of this variation being attributed to genetic differences (Foster 1990 and references therein).

Rooting ability of stem cuttings in *E. globulus* in particular has been found to be very variable with many clones proving to be poor rooters (Wilson 1992). Several strategies to overcome this variable and generally poor rooting ability have been tried. One strategy which has been employed by Stora Celbi in Portugal has been to screen selected plus trees for genotypes that have economically viable levels of rooting ability and use only these individuals for clonal deployment (MacRae and Cotterill 1997). However by deploying only these genotypes the number of genets that can be deployed clonally is reduced, potentially reducing genetic gains for other traits.

A possible alternative strategy is to include selection for rooting ability as part of the breeding program. Selection and breeding for rooting ability has the potential to increase and extend clonability, enabling the clonal deployment of a much broader range of genotypes than does simple screening for rooting ability, potentially capturing more genetic gain in deployed populations. Prediction of gains from selecting for rooting ability can be made based on estimates of the heritability (h^2) of this trait. Calculations of the heritability of rooting ability in *E. globulus* have been made for both stem cuttings in the field (Borralho and Wilson 1994, England and Borralho 1995, Lemos *et al.* 1997) and *in vitro* propagated stem cuttings (Ruaud *et al.* 1999). In the studies with stem cuttings, estimates of heritability for rooting ability were moderate to high (0.36, 0.22, and 0.54) indicating the potential for improvement of this trait using selection. Heritability estimates for the *in vitro* cuttings (Ruaud *et al.* 1999) were however generally lower (0.16 and 0.27).

Micropropagation of *E. globulus* has been carried out in a number of laboratories with the hope of overcoming the problem of poor rooting ability. As a means of multiplying clones of *E. globulus* it is uncertain if micropropagation has any advantage over stem cuttings produced in the field (macropropagation). One of the apparent advantages is the potentially higher rates of multiplication that can be achieved with *in vitro* methods (MacRae and Cotterill 1997), however the time to produce plantable plants has been shown to be about the same for the two propagation systems (Wilson 1995). One possible advantage of micropropagation arises from the observation that micropropagated plantlets have a root structure that resembles seedlings more closely than do stem cuttings. This may be an important factor since Stora Celbi in Portugal have found that *E. globulus* stem cuttings have mediocre field growth performance compared with seedlings (Cotterill and Brindbergs 1997) and the root structure of cuttings has been found to differ significantly from seedlings (Sasse and Sands 1995). Clones from *E. globulus* stem cuttings have also been noted to have an earlier change of phase from juvenile to adult leaves than seedlings (MacRae and Cotterill 1997). To date trials aimed at comparing micropropagation with macropropagation are too young to provide a clear comparison (MacRae and Cotterill 1997, Watt *et al.* 1995, Barbour and Butcher 1995). Even if the apparent rapid multiplication rates and growth performance of micropropagated plants does not provide a significant advantage over macropropagation systems, the shorter lag time in assessing rooting ability from seed using micropropagation makes it an attractive experimental system.

Given a high heritability for rooting ability in either a micropropagation or macropropagation system, a major problem in selecting for this trait is the high

cost and difficulty associated with its measurement. The effect of this is likely to be that only a relatively small subset of the breeding population could be assessed for rooting ability. However if a significant proportion of the variance for rooting ability could be predicted based on genotyping using molecular markers, then a much higher proportion of the breeding population could be assessed for this trait. Effective marker assisted selection would need to be based on a sound knowledge of the effect of selected QTLs and their stability across variable genetic backgrounds. Although building up this knowledge base would initially be expensive it could pay off in terms of the gains transferred to plantations.

To date two studies have been published in which QTLs for vegetative propagation traits have been found in *Eucalyptus* (Grattapaglia *et al.* 1995, Marques *et al.* 1999). Grattapaglia *et al.* (1995) found four QTLs for rooting ability in an *E. grandis* x *E. urophylla* cross using macropropagated cuttings and a family size of 96. The four QTLs were estimated to collectively explain 33% of the phenotypic variation and 63% of the genetic variation for the rooting ability trait. There were two samplings of cuttings from each genotype (two blocks) enabling an estimation of clonal repeatability, which was calculated to be 0.52 (Grattapaglia *et al.* 1995). In the study by Marques *et al.* (1999) where selective genotyping was compared with random genotyping, nine QTLs for adventitious rooting in macropropagated cuttings were hypothesised in a cross between *E. tereticornis* and *E. globulus*. Rooting was assessed over two consecutive years, which were treated as separate traits in the analysis. Five QTLs from the *E. tereticornis* parent explained an estimated total of 24% of the phenotypic variance in the 1995 assessment (%rooted/surviving) whilst QTLs at two of these locations were found in 1996 explaining 12% of phenotypic variance. Two QTLs from the

E. globulus parent explained an estimated 9% of the phenotypic variance in 1995 with one new QTL found and one lost in 1996, the resultant two QTLs explaining 12% of the phenotypic variance.

In the present study micropropagated shoots from two families of *E. globulus* were tested for their ability to form roots in agar root inducing medium, originally with the aim of carrying out a QTL analysis for this trait. The shoot cuttings were tested for rooting ability at 5 sampling dates and this enabled estimation of the variances due to between genotype effects and within genotype effects and calculation of the clonal repeatability for this trait. These estimates were used to calculate the power of detecting quantitative trait loci for rooting ability in these two families.

4.2 MATERIALS AND METHODS

4.2.1 Plant material and tissue culture

The rooting ability of *in vitro* micropropagated shoots was tested in two full sib families of *E. globulus* ssp. *globulus*. The pedigrees of the two families are shown in Fig. 4.1 and are as follows: Family 1 is derived from a cross between two full-sibs (pseudo F_2 cross) from the G1025 family. The two grandparents are both from the King Island provenance and linkage maps of these individuals have been constructed using microsatellite and RAPD markers (chapter 2 of this thesis). Family 2, supplied by Gunn's Ltd (formerly North Forest Products), also has parents that originate from the King Island provenance of *E. globulus*. The male parent 7100250 is a Forest Resources clone which has been propagated from stem cuttings and planted in clonal stands at several trial sites in Tasmania. This parent had proven to be readily clonable with good rooting ability. The clonability of the other parents/grandparents is unknown.

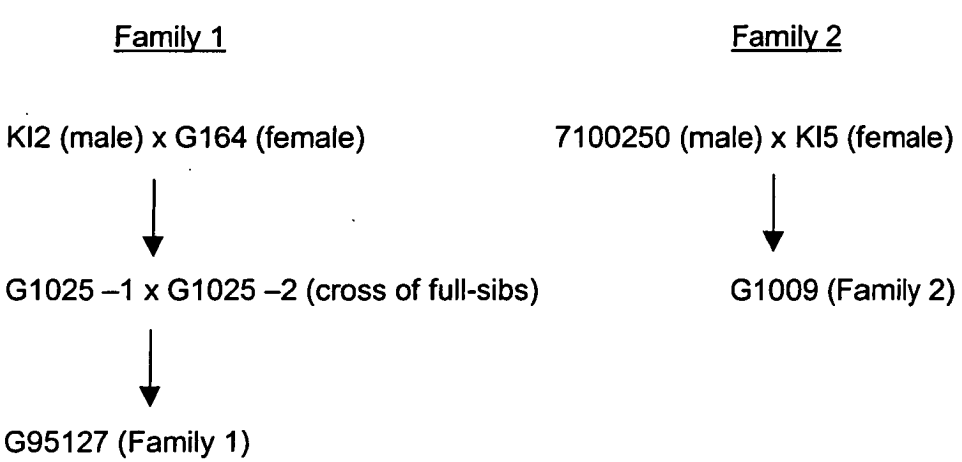


Fig. 4.1. The pedigrees of the two *E. globulus* families used in the rooting ability study

Seeds were surface sterilized in a weak solution of sodium hypochlorite for 35 mins and rinsed twice in sterile dH_2O for 10mins each before being placed onto the surface of solid agar (0.6%) medium containing 1/2 strength MS salts and 20g/L sucrose (pH 6.0) in tissue culture vials. After germination the vials were placed under fluorescent lighting in a temperature regulated culture room at 23°C and a 12 hr daylength at North Eucalypt Technologies tissue culture laboratory.

The vials were numbered and randomly allocated into trays to grow to a sufficient size for shoot excision. Excised shoots were placed on a shoot inducing medium to proliferate. After a number of rounds of subcloning as many shoot cuttings as possible were excised from shoot clumps and placed on a root inducing medium to assess rooting ability. Rooting ability was scored as a binary trait with each cutting being assessed for the presence or absence of roots from the base of the cutting. The number of cuttings tested and the proportion that produced roots was recorded for each genotype for each date of sampling.

The rooting ability trial was carried out at the tissue culture laboratory of North Eucalypt Technologies at Ridgley (Gunn's Ltd) in Tasmania. Cuttings were taken on the following five dates: 27/3/96 (Date 1), 8/5/96 (Date 2), 20/6/96 (Date 3), 29/7/96 (Date 4) and 9/9/96 (Date 5). Not all genotypes were sampled at each date and the number of cuttings sampled varied from date to date and genotype to genotype. Only those genotypes that had seven or more cuttings tested overall were included for analysis. The first round of cuttings (27/3/96) were scored for the presence/absence of roots from the stem base after three weeks on root inducing medium. However by three weeks it was found that the proliferation of roots from leaf fragments and the development of large callus clumps at the base

of shoot cuttings made assessment of rooting ability difficult. Thus for all dates thereafter the interval on root inducing medium was two weeks. The media used for the proliferation of shoots and root induction are proprietary formulations of Gunn's Ltd and under agreement cannot be divulged.

4.2.2 Statistical analysis – comparison of families

To determine if the difference in the rooting abilities of the two families was statistically significant an analysis was undertaken in which genotypes from both families were included in a statistical model. A generalised linear model (GLM) to binomial responses was used for the analysis with a probit link function and weighted according to the number of cuttings on which each proportion was calculated (Goldstein 1995). Second order penalised quasi-likelihood estimates were obtained for the fixed effects of family, date and family by date interaction (and the random effects of genotype and error) according to the following model:

$$y_{ijk} = \mu + F_k + D_i + (F.D)_{ik} + g_j + e_{ijk}$$

where y_{ijk} is the proportion of cuttings that rooted on the ith date for the jth genotype in the kth family and F_k is the fixed effect of the kth family, D_i is the fixed effect of the ith date, $(F.D)_{ik}$ is the fixed effect of the interaction between the kth family and the ith date, g_j is the random effect of the jth genotype and e_{ijk} is the residual. Extra binomial variation was observed and fitted. The program MLwin was used to carry out the analysis and estimates of the difference in rooting ability between the two families at each date were tested for significance using a chi-squared statistic (1df). A joint 95% confidence interval for the estimated difference at each date was also calculated along with a joint chi

squared test over all five dates (5df). The significance of the effect of date and the date by family interaction was also tested.

4.2.3 Statistical analysis – estimation of variance components and calculation of clonal repeatability

A separate analysis was undertaken for each family. A binomial distribution was assumed for the proportional data and the following generalised linear model using a probit link function was used in the analysis (Ruaud *et al.* 1999):

$$y_{ij} = \mu + D_i + g_j + e_{ij}$$

where y_{ij} is the observed proportion of shoots that rooted at the ith date for the jth genotype, μ is the overall mean, D_i is the fixed effect of the ith date, g_j is the random effect of the jth genotype and e_{ij} is the residual effect of the jth genotype on the ith date. Since the number of cuttings sampled varied across dates, the data was weighted using an iterative weighted Restricted Maximum Likelihood analysis where $w_{ij} = n_{ij}/[p_{ij}(1-p_{ij})]$, where w_{ij} is the weight, p_{ij} is the predicted proportion and n_{ij} is the number of cuttings sampled on the ith date for the jth genotype (Ruaud *et al.* 1999).

Estimates of the variance components were made using the program ASREML (Gilmour *et al.* 1995). In each of the ASREML analyses the relative size of the variances for the two random effects, the between genotypes variance (σ^2_b) and the within genotypes or residual variance (σ^2_w) were estimated (along with the standard error for σ^2_b). The hypothesis that the between genotype variance was a chance deviation from a variance of zero was tested using the statistic σ^2_b/SE with corresponding probability obtained from the standard normal probability density

function ($z = \sigma^2_b/SE$). Model checking included plotting fitted values against residuals to determine any apparent trend.

The clonal repeatability or repeatability is here defined as the ratio of the between genotypes variance to the sum of the between and within genotypes variances as calculated from the five measurements. That is $r = \sigma^2_b/(\sigma^2_b + \sigma^2_w)$. It is thus an estimate of the broad sense heritability of the mean phenotypic values for the genotypes.

4.2.4 QTL detection power calculations

Estimates of the power of a t-test to detect a QTL were undertaken using the following formula based on that in Soller (1991):

$$N = [2 (Z_{\alpha/2} + Z_\beta)^2 \sigma^2_m]/\delta^2$$

Where N is the number of genotyped progeny per marker class, $Z_{\alpha/2}$ is the ordinate of the standard normal distribution and α is the per marker Type I error, Z_β is the ordinate of the standard normal distribution and β is the Type II error, δ is the expected difference between marker class means and σ_m is the residual (within marker class) standard deviation.

A rearrangement of the formula to solve for Z_β gives:

$$Z_\beta = - SQRT[(N \times \delta^2)/(2\sigma^2_m)] - Z_{\alpha/2}$$

Values for β (the area to the left of the Z_β ordinate) were obtained from the standard normal probability density function and power of the test given by $1 - \beta$.

The type I error was set at 0.05 overall covering the two parental genomes for a single trait analysis.

For Family 1, the maps of the grandparents (chapter 1 of this thesis) could be used as the basis for map construction since it would be expected that 1/2 of the markers segregating 1:1 in the G1025 family would also segregate 1:1 in Family 1 (1/4 would be expected to segregate 3:1 and 1/4 would be lost). The required type I error for testing each marker can thus be calculated using the G1025 mapping information. The calculations are based on the following: there are approximately 50 independently segregating locations on the maps of the two grandparents (see chapter 1 of this thesis: male 28, female 19) plus a total of 30 unlinked markers, 15 from each grandparent. There is thus approximately 80 independently segregating locations across the two genomes and therefore the type I error rate α, was set to $\alpha = 0.05/80 = 0.000625$ and thus $\alpha/2 = 0.0003125$ and $Z_{\alpha/2} = -3.42$. The same assumptions have been made for Family 2. For the number of individuals per marker class (N), half the number of genotypes was substituted.

When determining the power of detecting a QTL of given effect occurring at a marker the between marker-class difference (δ) is equivalent to twice the deviance (in the family overall) due to the QTL ($2 \times \sigma_{QTL}$). If the marker under test is located near but not at the QTL, the recombination fraction (θ) between the marker and the QTL and the deviance due to the segregation of the QTL (σ_{QTL}) have the following relationship to the expected difference between marker-classes (δ) (Soller 1991):

$$\delta = 2\sigma_{QTL}(1-2\theta) \text{ and thus } \delta^2 = 4\sigma^2_{QTL}(1-2\theta)^2.$$

The within marker class variance (σ^2_m) is the remaining variance due to environmental influences and remaining genetic effects and is thus obtained from:

$$\sigma^2_m = (\sigma^2_b + \sigma^2_w) - \sigma^2_{QTL}.$$

With σ^2_b and σ^2_w being estimated in the individual family analyses.

Values used in power calculations for the recombination fraction (θ) were based on the maps of the two grandparents of Family 1, which would form the basis of marker genotyping for this family. The chance of a QTL being perfectly linked to a marker (ie. $\theta = 0$) could be hypothesised to be the same as two markers being perfectly linked which occurred in 17% of the framework markers from the two grandparental maps. The average interval size from the two maps is approximately 12cM and the average distance from a marker within this interval would be 3cM or a recombination fraction θ, of 0.03. This 'average' distance of 3cM from each marker would cover 46% of the framework maps of the two parents. The proportion of genetic variance explained by a QTL (σ^2_{QTL}) was based on estimates from Grattapaglia *et al.* (1995) for the upper-limit proportion of the genetic variation that is explained by detected QTLs. From their study out of 20 QTLs for three traits the largest QTL effect was equivalent to 29% of the genetic variance ($\sigma^2_{QTL}/\sigma^2_b$). The average of the largest QTL for the three traits was 18%. Based on this information a large QTL could be expected to explain 30% of the genetic variance rather than the optimal 100% and a reasonably large QTL could be expected to explain 20%.

For example a QTL which explained 30% of the genetic variance linked to a marker 3cM away within the region of the genome which had been mapped, the

following calculations indicate the power of a t-test to detect this QTL in Family 1.

$$\sigma^2_{QTL} = 0.3 \times \sigma^2_b = 0.3 \times 0.2 = 0.06.$$

$$\sigma^2_m = \sigma^2_w + \sigma^2_b - \sigma^2_{QTL} = 1.2 - 0.06 = 1.14$$

$\delta = 2\sigma_{QTL}(1-2\theta)$, substituting $\theta = 0.03$

$\delta = 1.88\sigma_{QTL}$ and thus $\delta^2 = 3.53\sigma^2_{QTL} = 0.21$

N = 77 (half the population size of Family 1)

$$Z_\beta = -\text{SQRT}[(N \times \delta^2)/(2\sigma^2_m)] - Z_{\alpha/2}$$
$$= -\text{SQRT}[(77 \times 3.53 \times 0.06)/(2 \times 1.14)] + 3.42$$
$$= 0.7455$$
$$\beta = 0.77$$
$$\text{Power} = 1 - \beta = 0.23 \text{ or } 23\ \%$$

It would be expected that for some traits several large QTLs may be segregating in the population under study and in QTL detection experiments in plants it is common for several to be detected (Kearsey and Farquhar, 1998). Thus the probability of detecting one or more of several hypothetical QTLs has been calculated. This probability has been calculated as:

Pr = 1 - (the product of the probabilities of not detecting each QTL individually).

4.2.5 Estimated size of phenotypic effect of hypothesised QTLs

The size of the phenotypic effect has been calculated from the data collected over five measurements so that $\sigma^2_P = \sigma^2_b + \sigma^2_w$ and thus $\sigma^2_{QTL}/\sigma^2_P = \sigma^2_{QTL}/\sigma^2_b * \sigma^2_b/(\sigma^2_b + \sigma^2_w)$. The proportion of the variance attributable to the QTL was converted to a standard deviation and added to the probit function value corresponding to a proportion of 0.16 (ie. -0.997) which was the average rooting ability for Family 1 over the five dates tested. The proportion corresponding to the new probit function value was recovered and the difference between the proportion of 0.16 and the final proportion was determined.

4.2.6 Change in variance and QTL detection power resulting from doubling the number of tests

When there are multiple measurements of the phenotype of an individual (genotype), then the overall phenotypic variance for a group of individuals decreases as more measurements are made. The component of variance that reduces is that due to special environment, V_{Es} (within genotypes variance) (Falconer 1989). The phenotypic variance for multiple measurements can be calculated according to the following relationship (Falconer 1989):

$$V_{P(n)} = V_G + V_{Eg} + 1/n\ V_{Es}$$

where $V_{P(n)}$ is the phenotypic variance after n measurements, V_G is the genetic variance, V_{Eg} is the between individual environmental variance arising from permanent or non-localised circumstances and V_{Es} is the within-individual variance arising from temporary or localised circumstances. Here σ^2_b is equivalent to $V_G + V_{Eg}$ and V_{Es} is equivalent to $5\sigma^2_w$ since σ^2_w was derived from five

measurements. Doubling the number of measurements from 5 to 10 would be expected to approximately increase the ratio of between genotypes variance to within genotypes from σ^2_b/σ^2_w to $2\sigma^2_b/\sigma^2_w$ (ie. double the ratio). This ratio has been substituted into the power estimations for ten measurements.

4.3 RESULTS

4.3.1 Rooting ability of families

The scale of the experiment to assess rooting ability in genotypes from the two *E. globulus* families is presented in Table 4.1. Overall more than 5,200 micropropagated cuttings were tested for their ability to produce roots.

Table 4.1. Scale of the rooting ability assessment experiment.

	Family 1	*Family 2*
No. genotypes	154	65
No. of cuttings taken	3562	1705
Mean no. cuttings/genotype	23	26
Mean no. cuttings/genotype/date	5.7	6.0

From Table 4.1 it can be seen that although there were more genotypes tested for Family 1 than for Family 2 the mean number of cuttings tested per genotype and the mean number of cuttings tested per genotype per date was similar for the two families. The number of genotypes and cuttings sampled at each date for each family is presented in Table 4.2 and Table 4.3 for Family 1 and Family 2 respectively.

Table 4.2. The number of genotypes and the number of cuttings tested for a genotype at each date for Family 1.

	No.	*Number of cuttings tested for a genotype*				
Date	*Genotypes*	*Min.*	*Max.*	*Mean*	*Mode*	*Median*
1	146	1	8	3.4	3	3
2	153	2	20	7.8	10	8
3	130	2	13	5.4	4	5
4	138	1	14	6.3	5	6
5	59	1	10	5.2	5	5
Overall	154	8	49	23.1	22	22

Table 4.3. The number of genotypes and the number of cuttings tested for a genotype at each date for Family 2.

	No.	*Number of cuttings tested for a genotype*				
Date	*Genotypes*	*Min.*	*Max.*	*Mean*	*Mode*	*Median*
1	62	1	8	3.4	3	3
2	64	3	19	7.6	10	8
3	64	1	11	6.9	10	7
4	62	2	12	6.5	5	6
5	30	5	8	5.3	5	5
Overall	65	13	38	26.2	23	26

From Tables 4.2 and 4.3 it can be seen that the maximum, mean, mode and median number of cuttings tested per genotype varies between dates but is similar at any given date between families.

The mean rooting ability for both families at each date is presented in Fig. 4.2. Rooting ability here is scored as the average of the proportion of cuttings that produced roots for all genotypes in the family.

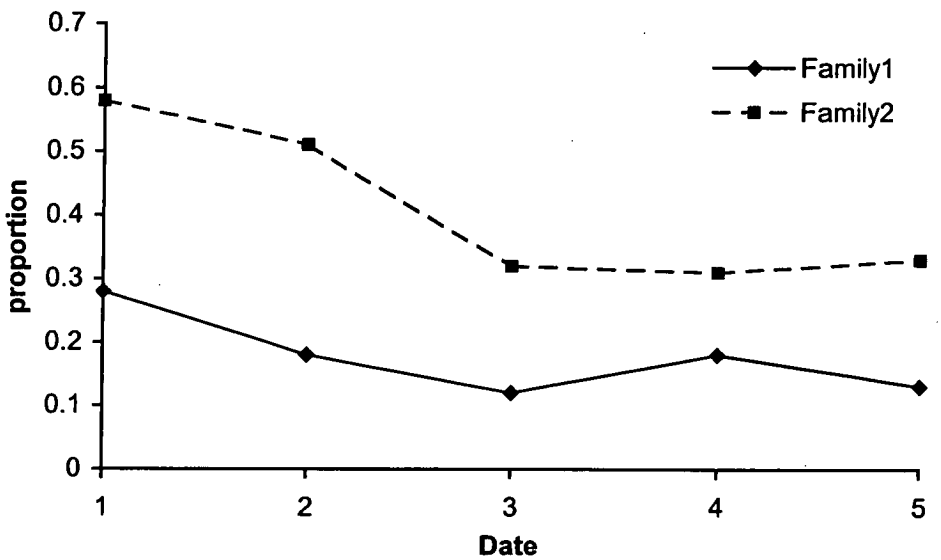


Fig. 4.2. Rooting ability of the two *E. globulus* families at each date.

From Fig. 4.2 it can be seen that the mean proportion of cuttings that rooted for a genotype was higher for Family 2 than for Family 1 at each of the five dates of sampling. The rooting ability of both families declined from the initial sampling of cuttings (date1) till the third sampling after which it tended to level out.

The estimated average rooting ability for genotypes in both families at each date are presented in Table 4.4. These estimates result from the analysis that included data from both families in a GLM. Alongside each pair of estimates is a p-value resulting from a test of the hypothesis that there is no difference between the

means of the two families. The p-value is derived from a separate chi-squared test (1d.f.) of the estimated differences between the two families at each date (ie. H_0: $\mu_1 - \mu_2 = 0$).

Table 4.4. A comparison of the estimated rooting ability of the two *E. globulus* families.

	mean proportion of cuttings producing roots		*p-value*[a]
Date	*Family 1*	*Family 2*	H_0: $\mu_1 = \mu_2$
1	0.27	0.62	5.1×10^{-10}
2	0.17	0.50	8.0×10^{-20}
3	0.10	0.32	1.0×10^{-10}
4	0.16	0.31	1.4×10^{-5}
5	0.10	0.31	4.9×10^{-5}

[a] Obtained from a chi-squared test of the estimated difference between families, 1df.

The p-values from the test at each date that the family means are equal (Table 4.4) are all highly significant indicating that it is extremely unlikely that the two families have equal rooting ability. The estimates for the mean proportion of cuttings that rooted at each date were very close to the observed values (Fig. 4.2) supporting the appropriateness of the model used in the analysis. A joint chi-squared test of the estimated differences at all of the five dates gave a very highly significant p-value of 3.4×10^{-25} indicating that overall it is extremely unlikely that the difference in rooting ability between the two families is due to chance. This is very strong evidence that the two families have different rooting abilities with Family 2 having higher rooting ability than Family 1.

The significance of the effect of date on rooting ability for both families for each pair of dates is shown in Table 4.5.

Table 4.5. Significance of the effect of date on rooting ability for the two families.

	Family 1				*Family 2*			
Date	1	2	3	4	1	2	3	4
2	**				*			
3	**	**			**	**		
4	**	n.s.	*		**	**	n.s.	
5	**	*	n.s.	n.s.	**	**	n.s.	n.s.

* = significant at 0.05 level for a separate chi-squared test
** = significant at 0.05 overall correcting for 20 tests (0.0025 level each test)

The results in Table 4.5 indicate that the rooting ability for Family 1 at date 1 is significantly different to the four subsequent dates at an overall 0.05 level (0.0025). The only other significant difference in rooting ability between dates for Family 1 was between dates 2 and 3. For Family 2 there was a significant difference in rooting ability between dates 1 and 3 and subsequent dates. There was also a significant difference in rooting ability between dates 2 and 3 and subsequent dates. All other pairs of date comparisons were not significant at the more stringent level. The trend observed in Fig.3 1 with a decline in rooting ability over time until the third date is supported in both families by the significantly different rooting abilities between initial dates and later dates.

The interaction between date and family was tested for significance for each of the ten possible comparisons. Generally the relative difference in rooting ability between families did not change significantly with the date, however when date 2 is compared with date 4 a significant interaction between date and family was found at an overall 0.05 level (allowing for 10 tests ie. 0.005). This agrees with Fig. 4.2 where it can be seen that the difference in rooting ability between families is largest at date 2 and smallest at date 4 and thus the finding of a significant interaction when comparing these dates is not surprising.

The analyses of the two families considered individually provided estimates for the between-genotype variance (σ^2_b) relative to the within-genotype (residual) variance (σ^2_w) that are presented in Table 4.6 along with the associated standard error (SE) for the estimate of between-genotype variance.

Table 4.6. Estimates for the between genotype variance (σ^2_b) relative to the within genotype variance (σ^2_w) for the rooting ability of two *E. globulus* families.

Family	σ^2_w	σ^2_b	*SE for* σ^2_b	*p value for* σ^2_b [a]	*clonal repeatability*[b]
1	1	0.201	0.036	1.28×10^{-8}	0.17
2	1	0.165	0.042	4.08×10^{-5}	0.14

[a] The hypothesis under test is: H_0: $\sigma^2_b = 0$.

[b] Clonal repeatability defined as $\sigma^2_b/(\sigma^2_b + \sigma^2_w)$.

Because a probit link function was used in the analyses the within-genotype variance (or residual variance - σ^2_w) is set to 1.0 with the between-genotype variance (σ^2_b) scaled accordingly. The p-values from a statistical test of the hypothesis that the variance due to genotype effect is due to chance (H_0: $\sigma^2_b = 0$)

was very highly significant for both families indicating that this effect was extremely unlikely to be due to chance alone. The size of the between genotype variance relative to the within-genotype variance is however small for both families.

From the analysis of each family it was found that the estimated mean proportion of cuttings that rooted for a genotype at each date was very close to the observed values (data not shown). A plot of fitted values against residuals did not reveal a trend in the data from either family. Frequency histograms of the residuals generated from the analysis of both families approximated a normal distribution. This indicated the model was appropriate for the experimental data.

The proportion of cuttings that rooted at each date for three selected genotypes from each family, are plotted in Figures 4.3 and 4.4 for Family 1 and Family 2 respectively. Genotypes that had a large number of cuttings sampled at each date were selected.

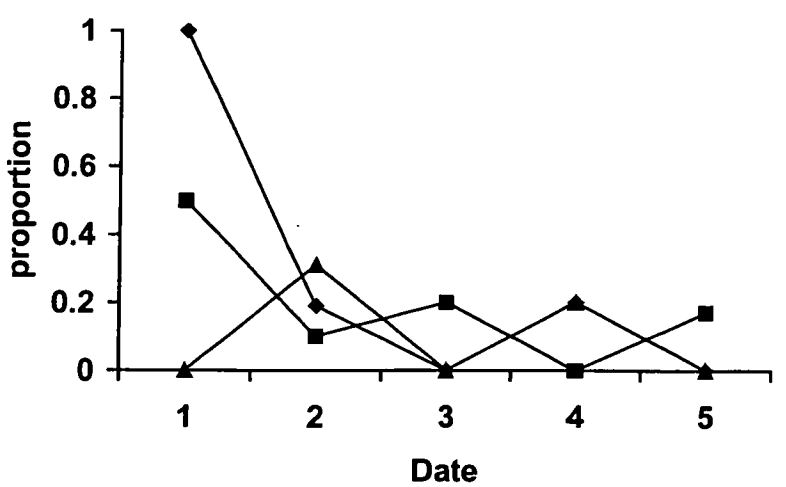


Fig. 4.3. Rooting ability of three selected genotypes from Family 1.

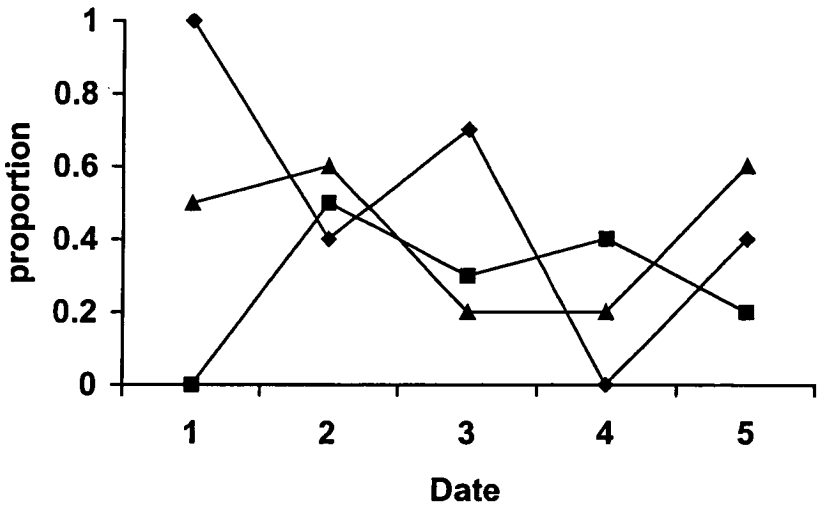


Fig. 4.4. Rooting ability of three selected genotypes from Family 2.

The plots in Figs. 4.3 and 4.4 illustrate the fluctuation in the proportion of cuttings that rooted for each of the genotypes over the five dates. Across all genotypes this fluctuation represents the within genotypes variance which was large relative to the between genotypes variance resulting in the low repeatabilities reported in Table 4.6.

4.3.2 Power of t-test to detect QTLs

The optimal power for detecting a given QTL would occur when the genotype variance (σ^2_b) is due solely to the segregation of this QTL in the population under study (ie. $\sigma^2_b = \sigma^2_{QTL}$) and the QTL is perfectly linked to a marker used for genotyping (ie. $\theta = 0$). Thus with regard to the formula for determining the power of a t-test, the between marker class difference (δ) would be equivalent to twice the deviance (in the family as a whole) due to genetic differences ie. $\delta = 2\sigma_{QTL}$ and thus $\delta^2 = 4\sigma^2_{QTL}$. The within marker class variance (σ^2_m) would be equivalent

in this case to the residual variance (σ^2_w) which would be due to environmental effects. Carrying out these substitutions into the rearranged power estimation formula we have:

$$Z_\beta = -\text{SQRT}\,[2N\sigma^2_b / \sigma^2_w] - Z_{\alpha/2}$$

and substituting the values and estimates obtained for Family 1:

$$Z_\beta = -\text{SQRT}(2 \times 77 \times 0.201) - Z_{\alpha/2}$$

$$= -5.56 + 3.42$$

$$= -2.14$$

$$\beta = 0.016$$

$$\text{Power} = 1 - \beta = 0.98 \text{ or } 98\ \%$$

Substituting the values and estimates obtained for Family 2 gives:

$$Z_\beta = -\text{SQRT}(32 \times 2 \times .165) + 3.42$$

$$= 0.17$$

$$\beta = 0.57$$

$$\text{Power} = 1 - \beta = 0.43 \text{ or } 43\ \%$$

For Family 1 the probability of detecting a QTL of this magnitude under optimised conditions is high (98%) whereas for Family 2 the probability is less than 50%. It would appear that the relatively low clonal repeatability (or small ratio of genotypic variance to environmental variance) and the small number of individuals phenotyped make the data collected for Family 2 unsuitable for QTL detection purposes. The results of calculating the power for QTL detection for a single QTL with varying size of effect and varying recombination between marker and QTL is presented in Table 4.7 using data from Family 1 only.

Table 4.7. QTL detection power of hypothesised QTLs for Family 1 using single marker analysis (t-test): influence of size of QTL effect ($\sigma^2_{QTL}/\sigma^2_b$) and recombination fraction (θ) between marker and QTL.

θ	$\sigma^2_{QTL}/\sigma^2_b$	$\sigma^2_{QTL}/\sigma^2_P$	Power
0	1	0.17	0.98
0	0.30	0.05	0.29
0	0.20	0.03	0.13
0.03	1	0.17	0.97
0.03	0.30	0.05	0.23
0.03	0.20	0.03	0.11

From Table 4.7 it can be seen that for a rooting ability QTL of large effect ($0.3\sigma^2_b$) an average distance (see Materials and Methods for definition) from a marker (3cM) there would be an expected chance of detection of 23%.

As already indicated it is extremely likely that more than one segregating QTL would be responsible for the observed genotypic variance. The power to detect one or more of several QTLs of given size effect is presented in Table 4.8.

Table 4.8. Power for detecting one or more of several hypothesised major QTLs using data for Family 1.

θ	$\sigma^2_{QTL}/\sigma^2_b$	Power
0	0.30 + 0.20 + 0.20	0.46
0.03	0.30 + 0.20 + 0.20	0.39
0	0.20 + 0.20 +0.20 + 0.20	0.43
0.03	0.20 + 0.20 +0.20 + 0.20	0.37

θ = recombination fraction between marker and QTL
$\sigma^2_{QTL}/\sigma^2_b$ = size of QTL effect relative to total genotype effect

From Table 4.8 it can be seen that even where there is no recombination and there are three unlinked QTLs one responsible for 30% the other two responsible for 20% of the genotype variance the probability of detection would be 46%, [1 – (0.71 x 0.87 x 0.87) = 0.46]. Even though these are relatively large QTL effects and there are three of them there is less than 50% probability of detection. The probability of all three QTLs being perfectly linked to markers is highly unlikely and if the average recombination distance is considered (3cM) the power decreases to 39%. For four QTLs each explaining 20% of genotype variance the calculated power of detecting one or more was lower than for the three QTLs.

The expected size of the phenotypic effect for hypothesized QTLs for Family 1 is presented in Table 4.9.

Table 4.9. Expected effect on average rooting ability of selection[a] for hypothesised QTLs in Family 1.

$\sigma^2_{QTL}/\sigma^2_b$	$\sigma^2_{QTL}/\sigma^2_P$	σ_{QTL}	$\mu + \sigma_{QTL}$	*average rooting ability*	*Δ proportion*
1.0	0.17	0.41	-0.587	0.28	0.12
0.3	0.05	0.22	-0.772	0.22	0.06
0.2	0.03	0.18	-0.813	0.21	0.05

[a] Selection based on a marker located at the QTL

From Table 4.9 it can be seen that the expected effect on rooting ability of the QTL that explains 100% of the genotype variance is relatively small at 0.12. This does however virtually double the probability from the mean of 0.16 without selection to 0.28 with selection. For QTLs with a more realistic size of effect the increase is more modest eg. an increase from 0.16 to 0.22 for a QTL explaining 30% of the genetic variance. It can be noted from the change in rooting ability that the effect of small QTLs is large relative to the proportion of variance explained.

The effect on the power for QTL detection of doubling either the number of dates of testing or the number of progeny is presented in Table 4.10 for Family 1 and Table 4.11 for Family 2.

Table 4.10. Effect on QTL detection power of doubling the number of testings of rooting ability from 5 to 10 and doubling the number of progeny from 154 to 308 for Family 1.

$\sigma^2_{QTL}/\sigma^2_b$	*5 tests*	*10 tests*	*5 tests, 308 progeny*
1.0	0.98	1.00	1.00
0.3	0.29	0.65	0.73
0.2	0.13	0.36	0.44

Here recombination fraction between marker and QTL is zero (ie.$\theta = 0$).

Table 4.11. Effect on QTL detection power of doubling the number of testings of rooting ability from 5 to 10 and doubling the number of progeny from 65 to 130 for Family 2.

		Power	
$\sigma^2_{QTL}/\sigma^2_b$	*5 tests*	*10 tests*	*5 tests, 130 progeny*
1.0	0.43	0.88	0.89
0.30	0.04	0.12	0.15
0.20	0.02	0.06	0.07

Here recombination fraction between marker and QTL is zero (ie.$\theta = 0$).

From Tables 4.10 and 4.11 it can be seen that the increase in QTL detection power from doubling the number of samplings is quite considerable for QTL of reasonable size. An even greater increase in power results from doubling the number of progeny tested.

4.4 DISCUSSION

4.4.1 Between family genetic variation for rooting ability

In the present study it has been observed that there is a considerable difference in the rooting abilities of two families of *Eucalyptus globulus* and that this difference is statistically highly significant. The most likely explanation for this observation is that genetic differences between the two families are responsible for the differences in rooting ability. This supports findings from previous studies on rooting ability in *E. globulus* (Willyams *et al.* 1992, Borralho and Wilson 1994, England and Borralho 1995, Lemos *et al.* 1997, Ruaud *et al.* 1999) which have concluded that there is genetic variation for this trait at the between family level.

The existence of genetic variation for rooting ability is of interest in this study as it was originally designed as a QTL detection experiment and the presence of genetic variation predicates the existence of QTLs. However genetic variation for rooting ability between unrelated families cannot be directly exploited for QTL detection purposes. This is because the identification of QTL/marker associations depends on the existence of linkage disequilibrium that may not be present and cannot be relied upon in a pool of individuals from unrelated families. For QTL detection purposes genetic variation for the trait of interest is required within the family or families under study rather than between families. However the observed existence of genetic variation for rooting ability between families suggests that there is potential to generate or discover large within family genetic variation for this trait (given an appropriate cross).

4.4.2 Within family genetic variation for rooting ability

It is preferable in QTL detection studies if the genetic variation within a family is large because this is likely to increase the probability of QTL detection. It is not always possible however to estimate the genetic contribution to phenotypic variance. Due to the repeated measurement (five dates of testing) of each genotype in this study it has been possible to estimate the between genotypes contribution (σ^2_b) to phenotypic variance relative to the within genotypes variance (σ^2_w). The between genotypes variance (σ^2_b) may include a component of environmental variance termed V_{Eg} by Falconer (1989) that arises from permanent or non-localised circumstances. Thus σ^2_b as a measure of genetic variance may be an overestimate. Lynch and Walsh (1998) make the point that the within individuals component of variance (ie. within genotypes or special environmental variance - σ^2_w) can be inflated due to the inclusion of measurement error. There may have been measurement error in the assessment of rooting ability since it was assessed by visual inspection of the stems of cuttings after two weeks to determine if roots were present. From a finite number of samplings it could be argued that there is an uncertain amount of measurement error present in every estimate of rooting ability (but as more cuttings are sampled the estimate is likely to increase in accuracy). Neither of these two possibilities for overestimation is resolvable from the present data but are noted here to indicate that there is uncertainty in the variance estimates. Even so from the data the best estimate for genetic variance is σ^2_b and the best estimate for environmental variance is σ^2_w. Ideally, for QTL detection purposes, σ^2_b will be large relative to σ^2_w. When measuring the vegetative propagation characteristics of an *E. grandis* x *E. urophylla* cross, Grattapaglia *et al.* (1995) calculated the repeatability for three

traits, fresh weight of shoot clumps, number of stump sprout cuttings and percentage rooting of cuttings as 0.59, 0.42, and 0.52 respectively (with n = 118, 97, 96). The genotypic variance for these three traits thus constituted a large proportion of the phenotypic variance and genotyping of individuals led to the discovery of QTLs for all three traits. Since the clonal repeatability was quite large any QTL that explained a reasonable proportion of the genetic variance also explained a considerable proportion of the phenotypic variance. In the present study however the estimated clonal repeatability ($\sigma^2_b/(\sigma^2_b + \sigma^2_w)$) was relatively small for both families with 0.17 for Family 1 and 0.14 for Family 2. With low clonal repeatabilities such as these the proportion of phenotypic variance explained by a QTL is likely to be relatively small making the probability of QTL detection (power) low.

Small clonal repeatabilities would be expected where the genotypic variation for rooting ability is small or where the environmental variation influencing each measurement of phenotype is large or a combination of the two. In this study it would appear that the generally large fluctuations in the phenotypic measurements as illustrated indicate that the environmental variation is large. Measurement error has been mentioned previously as contributing to this. No data has been obtained to determine what other factors might contribute to this variation and this is open to speculation. It is interesting to note that if there were no variation in the cuttings from a genotype and no variation in the propagation environment then it would be expected that for any number of cuttings from a genotype rooting ability would be either 0% or 100%.

4.4.3 QTL detection power

The estimates of the variance components (between and within genotypes) for rooting ability in this study were used in calculations of the power of detection for hypothesised QTLs that explain specified proportions of the genetic effect. Even where it was assumed that the linkage between a marker and a QTL was perfect the combination of low genetic variance relative to environmental variance and relatively small number of individuals in the study made the power for QTL detection very low. This was especially true for Family 2 which had the smaller ratio of genetic to environmental variance combined with a much smaller number of phenotyped individuals. The power to detect one or more of several QTLs was also computed since this is considerably greater than the probability of detection of a QTL considered on its own. Even so in all cases where the QTL was of reasonable size the power for detection was less than 50%. The lack of power for detecting QTLs for rooting ability led to the decision not to carry out marker genotyping of either of the two families.

The calculations of QTL detection power are of course only applicable to the environmental conditions under which this experiment was conducted and power would increase if the variance attributable to environmental effects was smaller (assuming the genetic variance remained the same). It should be noted that the conditions that were used for the trial had been optimised for producing rooted cuttings from *E. nitens* and not *E. globulus*. Possibly under different conditions of media formulation there would be less environmental noise and greater variation in rooting ability between 'good' and 'poor' rooting genotypes within each family.

According to the equation in Soller (1991), the power to detect a given QTL (as it relates to this study) depends upon the following: (1) the relative size of the QTL effect to the residual effect (caused by environmental noise and the segregation of other QTLs), (2) the number of progeny (that have been phenotyped and genotyped) in the family in which the QTL is segregating, and (3) the distance of the QTL from a scored marker locus. There are possibilities for improving all three of these influences in the current experiment. Continuing the assessment for rooting ability over more samplings will theoretically increase the ratio of between genotype variance to within genotype variance. Practically this would mean that the phenotypic values used for QTL detection would have a larger genetic component. With these values the proportion of the phenotypic variance explained by a given QTL would be larger increasing power for detection. The expected increase in QTL detection power from doubling the number of testings has been explored for both families and is not inconsiderable. However the cost of doubling the amount of phenotyping would be large. Certainly it would only improve power sufficiently to be considered as an option for Family 1. If cost was not a problem and the aim was solely to detect QTL for rooting ability this would be a satisfactory option.

The second method of improving power would be to increase the number of progeny assessed. Unfortunately all seed that was available for these two crosses was used in the rooting ability assessment. For Family 2 the actual number of seeds available was unfortunately much lower than estimated. The family was still included in the trial however since one of the parents was known to have high rooting ability. Larger numbers in both families would have been helpful for increasing QTL detection power and doubling the number of progeny was found

to be slightly more effective in increasing power than doubling the number of tests of rooting ability. Potentially however the cost of doubling the number of progeny would be more since as well as doubling the amount of phenotyping the genotyping costs would also potentially double. In any case both these methods of increasing the experimental power would have involved considerable time and cost and any QTL detected is still likely to be of small effect.

One approach to reducing the environmental variance that would be justified if the *in vitro* system was to be used on a commercial scale would be to examine factors that might influence within genotype variation. Factors such as the position of the sourced cutting on the shoot clump, the age of the shoot, the length of the cutting, the number of internodes and the presence/absence of an apex might be examined for effect on rooting ability. In field grown cuttings of *E. globulus* factors such as these have been found to influence rooting ability (Wilson 1993).

The third factor that could be altered to increase the power for QTL detection would be to increase the density of the linkage map to improve the chances of having a marker close by any QTL that might be discovered. This is not really an issue in this study since even with zero recombination between marker and QTL the power for QTL detection appears to be low. The cost and time required to increase map density would only be warranted if the power for QTL detection was already high. The issue of map coverage as it impinges on power for QTL detection is treated below.

4.4.4 Factors that would reduce QTL detection power

In this study the influence of distance between a marker and QTL and the size of the QTL effect have been taken into account when carrying out calculations for QTL detection power. However three factors that could reduce the power for QTL detection have not been taken into account - these are map coverage, segregation distortion and an inflated estimate of genetic variance. The power to detect a given QTL in the genome is entirely dependent on the coverage of that genome with markers. Even without statistical considerations of the power of marker contrasts, a genomic map covering less than 100% of the genome will have reduced QTL detection power relative to a map with complete coverage. This reduction in power is not likely to be a direct function of the proportion of the genome that is covered but may depend on the way in which the genome is covered (or not covered) with markers ie. the number of linkage groups covered the size of the genome and the distance between markers. The framework maps of the grandparents of Family 1 were estimated to cover 79% and 62% for the male and female parents respectively (Chapter 1 of this thesis) an average coverage of 70%. These estimates did not include the information from unlinked markers of which there were 15 for each parent. There are also 11 regions on the two maps with close to or more than 30cM intervals. It is possible that there are QTLs for numerous traits in regions of the two parental genomes either not covered by markers or at least sparsely covered. Since these marker maps would form the basis for mapping in Family 1, there is some probability of otherwise detectable QTLs for rooting ability being outside the detection boundary of the maps.

The second factor that was not taken into account when calculating the power for QTL detection is segregation distortion. The power calculations were carried out

assuming that the population would fall into two equal halves for marker genotyping purposes. However it is fairly unlikely that the two marker genotypes would consist of 77 individuals each and the probability of this given a 0.5 probability of either genotype is approximately 0.064 and it is thus more likely (0.936) that some distortion from the exact 1:1 ratio will be found. It is also to be noted, that there were more regions of segregation distortion in the genomes of the grandparents than would be expected by chance, lowering the probability of a QTL occurring at a marker segregating perfectly 1:1. Certainly the power to detect a given QTL would decrease with increasing segregation distortion at the QTL/ marker. The question remains however as to what extent detection power will decrease as a QTL/marker varies from the 1:1 segregation ratio?

If it is assumed that the number in the smaller marker class is substituted into the power equation, then using Family 1 as an example:

for a QTL explaining 30% of genetic variance that is 3cM away from a marker segregating 72:82, the power (based on 72 individuals) would be 20% rather than 23% for a 1:1 (77:77) segregating locus. Since map coverage and marker segregation are ignored in the calculations of power for QTL detection presented in the results section, these calculations are likely to overestimate QTL detection power.

A third factor that may have overestimated QTL detection power calculations is a potentially inflated genetic variance estimate. The estimate of genetic variance used in the QTL detection power calculations was the between genotypes variance. The genetic variance is potentially overestimated since the between genotypes variance may include environmental effects specific to some genotypes

(V_{Eg}). An example of this would be a cryptic bacterial infection that effects growth and is present in the *in vitro* cultures of some genotypes but not others. There is however no evidence for the existence of any effects and the estimate may in fact include very little environmental variance.

4.4.5 Size of QTL effects and usefulness for breeding

An important consideration for QTL detection with regard to breeding is the size of the QTL effect and its impact on the trait if marker assisted selection was implemented. It would appear that the impact on rooting ability of QTLs of realistic size from Family 1 under similar conditions would be relatively small. The increase expected on the mean rooting ability of Family 1 for QTLs of reasonable size is modest indeed with a QTL explaining 30% of genetic variance increasing the rooting ability from 0.16 to 0.22 an increase of 0.06. It would thus be likely that any QTLs detected would have insufficient impact on rooting ability to warrant their inclusion in a marker assisted breeding program. This would be especially true for a family such as Family 1 where in any case the rooting ability is low. The position of any QTLs could be checked in other families however to determine if there is segregation that influences the trait in other crosses. It would still also be useful to detect QTL for rooting ability in poor rooting families if their performance with regard to other traits was of merit.

4.4.6 Further investigations

As already mentioned an examination of the potential contributions to environmental variance could identify sources of environmental variation. If these sources of environmental variation could be easily reduced or removed then this

would increase the genetic contribution to rooting ability increasing both the power for QTL detection and the importance of any QTLs detected. Preferably a reduction in a source of environmental variation would improve the average rooting ability as well as improving the predictability of rooting ability.

A useful approach to QTL detection for rooting ability would be to examine a larger number of families in smaller numbers to identify which families had the largest genetic variation for rooting ability. QTL detection experiments could then be undertaken with larger progeny numbers for the families with the highest genetic variation for the trait. Any QTL detected could be tested within the smaller families already studied to determine if their is segregation occurring within these families. Genotyping in the small families would only involve the testing of a few markers.

4.5 CONCLUSION

It can be concluded that although there is considerable difference in rooting ability between the two families under study, the estimated ratio of genetic variance to environmental variance within each family was too small to give a reasonable power for detecting any segregating QTLs of realistic size. Even if a QTL of large genetic effect were to be detected, the influence of this QTL on rooting ability under similar conditions would be expected to be relatively small (an increase on average of 0.05 to 0.06 for QTLs explaining 20% and 30% of genetic variance for Family 1). Since their was an apparently large environmental variance found in this experiment further investigations into the source of the variation could be useful for *in vitro* rooting ability studies.

CHAPTER FIVE

GENERAL DISCUSSION

The aims of the PhD project reported in this thesis were:

(i) to produce a linkage map of *Eucalyptus globulus* suitable for the detection and mapping of quantitative trait loci (QTL), with the linkage map to form a basis for further mapping and marker studies in this species

(ii) to utilise the linkage map to detect and map QTLs for traits of commercial importance measured in the mapping population and carry out a preliminary study of environmental and temporal QTL stability

(iii) to investigate the trait of *in vitro* rooting ability in large families of *Eucalyptus globulus* to determine its suitability for QTL detection.

In chapter 2 of this thesis, the construction of linkage maps for two *Eucalyptus globulus* trees is presented. The trees are parents of an intraprovenance cross and the resultant linkage maps were the first reported for this cross type in *Eucalyptus*. They are also the first linkage maps constructed for a pure species cross in *Eucalyptus globulus*.

The linkage maps give good coverage of the genome, estimated to be 79% for the male framework map and 62% for the female framework map with markers ordered at a high stringency LOD score of 3.0. The inclusion of microsatellite

markers on the maps has enabled both homologous linkage groups (between parents) and homoeologous (between *Eucalyptus* species) linkage groups to be identified. It is also a useful start to converting the maps into microsatellite maps for future studies. This is important to enable map integration and transfer of QTL information both within and between species. The transfer of QTL information between *Eucalyptus* maps from different species has now been demonstrated by Marques *et al.* (2002) for QTL for vegetative propagation, based on microsatellite loci.

The linkage maps constructed for this study have also provided useful information for the use of microsatellite markers, in particular the previously unmapped EMCRC markers, for population and gene flow studies in *Eucalyptus globulus* (and potentially other *Eucalyptus* species), since they enable the selection of a subset of markers for study that are not in linkage and can thus be depended upon to provide independent information (Steane *et al.* 2001; Jones *et al.* 2002).

Based on independently segregating regions (rather than number of markers), segregation distortion was observed at a frequency greater than expected by chance, suggesting a biological cause.

In chapter 3, the linkage maps have been successfully utilised for the detection of quantitative trait loci (QTL) for traits of commercial importance. This is the first report of QTL detection for these traits in *Eucalyptus globulus*. QTL with LOD scores greater than 2.0 were detected for growth traits – two for cumulative growth and three for relative incremental growth, wood density (two QTL) and early flowering (one QTL).

Even though this is an exploratory study using a population of moderate size, one of the QTL for wood density was found to be highly significant at an empirically determined experiment-wise threshold. It is probable that this represents a QTL of reasonable effect, with an estimated 14% of the variance for the trait attributable to its segregation. This adds to the list of traits that have been found to have putative QTL that explain a reasonable proportion of the variance for the trait.

A study of QTL stability was undertaken. QTL stability is an important issue, since marker-assisted selection is likely to depend on the availability of information on the stability of detected QTLs. Two putative QTL for cumulative growth were examined for their stability across environments, their stability with time and their co-segregation with QTL for relative incremental growth. The results indicated that the two QTL for cumulative growth were both strongly interactive with the environment. In addition, evidence was found for the presence of QTL that were detectable solely on the basis of their interaction with the environment. These are important findings since most QTL detection studies in forest trees have been carried out on single sites, and it is unknown to what extent the detected QTL effects are environment dependent.

The temporal stability of QTL is also an important issue, since QTL for early growth may have no detectable effect on volume at rotation age, which is the key growth trait used for selection by breeders. The results from other studies of temporal QTL stability have been mixed, some studies indicating poor correlation between QTLs detected at different ages (Bradshaw and Stettler 1995; Plomion *et al.* 1996; Kaya *et al.* 1999), some intermediate (Verhaegen *et al.* 1997), some with good correspondence (Lerceteau *et al.* 2001) and some being a mixture (Emebiri

et al. 1998). The results from this study indicate that both of the cumulative growth QTLs were active by the first measurements of growth at year one, with QTL effects stable or having increased by year six. This should enable the future validation and further study of these QTL early in growth.

It is likely that due to the early expression of the cumulative growth QTL, there was found to be no co-segregation with QTL for relative incremental growth. There was however co-segregation between an incremental growth QTL and a QTL for wood density. The QTL effects were in opposite directions for the two traits, that is, an increase in relative growth performance occurs with a decrease in wood density. It has been discussed that this is similar to results found in other *Eucalyptus* species by Verhaegen *et al.* (1997) and Grattapaglia *et al.* (1996), and the possibility was mooted for the existence of QTLs with pleiotropic effects, influencing growth and wood density but in opposing directions. It was also noted that the second QTL for wood density was independent of any QTL for growth and this QTL may represent part of the genetic variation for wood density that is genetically independent of growth.

In chapter 4 an analysis of the trait of rooting ability of micropropagated cuttings was reported. The two *E. globulus* families studied were found to have significantly different rooting abilities, a finding that bodes well for the possibility of finding or constructing a family in which there is segregation of QTLs for this trait. Large family sizes were utilised with the aim of undertaking marker genotyping and QTL detection, if sufficient genetic variance for the trait was found in one or both families. The level of genetic variance was however, small compared to the large environmental variance found for both families. The

population size for the family with the greatest genetic variance was small and estimations of the power to detect segregating QTL of reasonable size indicated that marker genotyping for QTL detection was not worthwhile in either family.

The aims outlined at the beginning of this discussion have been fulfilled by the experimental work and a number of significant findings highlighted. There is certainly scope for further experimental work relating to the studies described here, especially in the areas of QTL validation, exploration of site interaction and temporal stability of QTLs. The fact that the population used for mapping and QTL detection was part of a half diallel crossing design, allows for the possibility of studying QTL stability across genetic backgrounds, another critical issue concerning QTL stability in forest trees. In addition a reference linkage map has been established for mapping in *E. globulus* and the trait of rooting ability explored for its potential for QTL detection.

CHAPTER SIX

REFERENCES

Adams WT, Joly RJ (1980) Linkage relationships among twelve allozyme loci in loblolly pine. *Journal of Heredity* **71**, 199-202.

Arcade A, Anselin F, Rampant PF, Lesage MC, Paques LE, Prat D (2000) Application of AFLP, RAPD and ISSR markers to genetic mapping of European and Japanese larch. *Theoretical & Applied Genetics* **100**, 299-307.

Barbour EL, Butcher T (1995) Field testing vegetative propagation techniques of *Eucalyptus globulus*. In 'Eucalypt plantations: Improving Fibre Yield and Quality, Proc. CRCTHF-IUFRO Conf.' Hobart, Tasmania. (Eds BM Potts, NMG Borralho, JB Reid, RN Cromer, WN Tibbits and CA Raymond) pp. 313-314. (CRC for Temperate Hardwood Forestry)

Barone A, Ritter E, Schachtschabel U, Debener T, Salamini F, Gebhardt C (1990) Localization by restriction fragment length polymorphism mapping in potato of a major dominant gene conferring resistance to the potato cyst nematode. *Molecular & General Genetics* **224**, 177-182.

Barreneche T, Bodenes C, Lexer C, Trontin JF, Fluch S, Streiff R, Plomion C, Roussel G, Steinkellner H, Burg K, *et al.* (1998) A genetic linkage map of *Quercus robur* L. (pedunculate oak) based on RAPD, SCAR, microsatellite, minisatellite, isozyme and 5S rDNA markers. *Theoretical & Applied Genetics* **97**, 1090-1103.

Basten CJ, Weir BS, Zeng ZB ('1997) 'QTL Cartographer: A reference manual and tutorial for QTL mapping.' (Department of statistics, North Carolina State University: Raleigh, NC USA)

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

Binelli G, Bucci G (1994) A genetic linkage map of *Picea abies* Karst., based on RAPD markers, as a tool in population genetics. *Theoretical & Applied Genetics* **88**, 283-288.

Bonierbale MW, Plaisted RL, Tanksley SD (1988) RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. *Genetics* **120**, 1095-1103.

Borralho NMG (2001) The purpose of breeding is breeding for a purpose. In 'IUFRO International Symposium on developing the Eucalypt of the future'. Valdivia, Chile. (Ed. S Barros). (INFOR)

Borralho NMG, Wilson PJ (1994) Inheritance of initial survival and rooting ability in *Eucalyptus globulus* Labill. stem cuttings. *Silvae Genetica* **43**, 238-242.

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* **32**, 314-331.

Bradshaw HD, Grattapaglia D (1994) QTL mapping in interspecific hybrids of forest trees. *Forest Genetics* **1**, 191-196.

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*. III. A genetic linkage map of a hybrid poplar composed of RFLP, STS, and RAPD markers. *Theoretical & Applied Genetics* **89**, 167-178.

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 & Genomics: MGG* **267**, 338-347.

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

Brondani RPV, Campinhos EN, Grattapaglia D (1997) Mapped RAPD markers are transferable among *Eucalyptus* trees from the same population. In 'Proceedings of the IUFRO Conference on Silviculture and Improvement of Eucalypts'. Salvador, Brazil pp. 111-115. (EMBRAPA, Colombo)

Byrne M, Marquezgarcia 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, Murrell JC, Tibbits WN (1994) Detection and inheritance of RFLPs in *Eucalyptus nitens*. *Theoretical & Applied Genetics* **89**, 397-402.

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

Byrne M, Murrell JC, Owen JV, Kriedemann P, Williams ER, Moran GF (1997a) Identification and mode of action of quantitative trait loci affecting seedling height and leaf area in *Eucalyptus nitens*. *Theoretical & Applied Genetics* **94**, 674-681.

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

Cato SA, Gardner RC, Kent J, Richardson TE (2001) A rapid PCR-based method for genetically mapping ESTs. *Theoretical & Applied Genetics* **102**, 296-306.

Cervera MT, Storme V, Ivens B, Gusmao J, Liu BH, Hostyn V, Van Slycken J, Van Montagu M, 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.

Chambers PGS, Borralho NMG (1997) Importance of survival in short-rotation tree breeding programs. *Canadian Journal Of Forest Research* **27**, 911-917.

Chambers PGS, Potts BM, Tilyard PA (1997) The genetic control of flowering precocity in *Eucalyptus globulus* ssp. *globulus*. *Silvae Genetica* **46**, 207-214.

Chase M, Kesseli R, Bawa K (1996) Microsatellite markers for population and conservation genetics of tropical trees. *American Journal of Botany* **83**, 51-57.

Cheliak WM, Pitel JA (1985) Inheritance and linkage of allozymes in *Larix larcina*. *Silvae Genetica* **34**, 142-148.

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

Costa P, Pot D, Dubos C, Frigerio JM, Pionneau C, Bodenes C, Bertocchi E, Cervera MT, Remington DL, Plomion C (2000) A genetic map of Maritime pine based on AFLP, RAPD and protein markers. *Theoretical & Applied Genetics* **100**, 39-48.

Cotterill PP, Brindbergs ML (1997) Growth of first and second generation Eucalyptus globulus clonal cuttings and seedlings. In 'Proceedings of the IUFRO conference on silviculture and improvement of Eucalypts'. Salvador, Brazil pp. 233-238. (EMBRAPA)

Dale G, Chaparro J (1996) Integration of molecular markers into tree breeding and improvement programs. In 'Tree improvement for sustainable tropical forestry Proceedings QFRI-IUFRO Conference'. Caloundra, Qld, Australia. (Eds MJ Dieters, AC Matheson, DG Nikles, CE Harwood and SM Walker) pp. 472-477. (QFRI)

Darvasi A, Weinreb A, Minke V, Weller JI, Soller M (1993) Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. *Genetics* **134**, 943-951.

Dean HG, French J, Tibbits WN (1990) Variation in pulp making characteristics in a field trial of *Eucalyptus globulus*. In '44th Annual Appita General Conference'. Rotorua, New Zealand

Devey ME, Fiddler TA, Liu BH, Knapp SJ, Neale DB (1994) An RFLP linkage map for loblolly pine based on a three-generation outbred pedigree. *Theoretical & Applied Genetics* **88**, 273-278.

Devey ME, Jermstad KD, Tauer CG, Neale DB (1991) Inheritance of RFLP loci in a loblolly pine three-generation pedigree. *Theoretical & Applied Genetics* **83**, 238-242.

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

Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, *et al.* (1996) A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* **380**, 152-154.

Dietrich WF, Miller J, Steen R, Merchant MA, Damronboles D, Husain Z, Dredge R, Daly MJ, Ingalls KA, Oconnor TJ, *et al.* (1996) A comprehensive genetic map of the mouse genome. *Nature* **380**, 149-152.

Dow BD, Ashley MV (1996) Microsatellite analysis of seed dispersal and parentage of saplings in bur oak, *Quercus macrocarpa*. *Molecular Ecology* **5**, 615-627.

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

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.

Echt CS, Vendramin GG, Nelson CD, Marquardt P (1999) Microsatellite DNA as shared genetic markers among conifer species. *Canadian Journal of Forest Research* **29**, 365-371.

Edwards MD, Stuber CW, Wendel JF (1987) Molecular marker facilitated investigations of quantitative trait loci in maize. I. Numbers, genomic distribution, and types of gene action. *Genetics* **116**, 113-125.

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

El-Kassaby YA, Sziklai O, Yeh FC (1982) Linkage relationships among 19 polymorphic allozyme loci in coastal Douglas-Fir (*Pseudotsuga menziesii* var. *menziesii*). *Canadian Journal of Genetics and Cytology* **24**, 101-108.

Elsik CG, Williams CG (2001) Low-copy microsatellite recovery from a conifer genome. *Theoretical & Applied Genetics* **103**, 1189-1195.

Emebiri LC, Devey ME, Matheson AC, Slee MU (1998) Age-related changes in the expression of QTLs for growth in radiata pine seedlings. *Theoretical & Applied Genetics* **97**, 1053-1061.

England NF, Borralho NMG (1995) Heritability of rooting success in *Eucalyptus globulus* stem cuttings. In 'Eucalypt plantations: Improving Fibre Yield and Quality, Proc. CRCTHF-IUFRO Conf.' Hobart, Tasmania. (Eds BM Potts, NMG Borralho, JB Reid, RN Cromer, WN Tibbits and CA Raymond) pp. 237-238. (CRC for Temperate Hardwood Forestry)

Falconer DS (1989) 'Introduction to quantitative genetics, 3rd ed.' (Longman Scientific and Technical: Harlow, Essex, England)

Fisher PJ, Richardson TE, Gardner RC (1998) Characteristics of single- and multi-copy microsatellites from *Pinus radiata*. *Theoretical & Applied Genetics* **96**, 969-979.

Foster GS (1990) Genetic control of rooting ability of stem cuttings from loblolly pine. *Canadian Journal of Forest Research* **20**, 1361-1368.

Fripp YJ, Griffin AR, Moran GF (1987) Variation in allele frequencies in the outcross pollen pool of *Eucalyptus regnans* F Muell. throughout a flowering season. *Heredity* **59**, 161-171.

Gebhardt C, Ritter E, Barone A, Debener T, Walkemeier B, Schachtschabel U, Kaufmann H, Thompson RD, Bonierbale MW, Ganal MW, *et al.* (1991) RFLP maps of potato and their alignment with the homoeologous tomato genome. *Theoretical & Applied Genetics* **83**, 49-57.

Gebhardt C, Ritter E, Debener T, Schachtschabel U, Walkemeier B, Uhrig H, Salamini F (1989) RFLP analysis and linkage mapping in *Solanum tuberosum*. *Theoretical & Applied Genetics* **78**, 65-75.

Gelderman H (1975) Investigations on inheritance of quantitative characters in animals by gene markers. I. Methods. *Theoretical & Applied Genetics* **46**, 319-326.

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

Gilmour AR, Thompson R, Cullis BR (1995) Average information REML, an efficient algorithm for variance parameter estimation in linear mixed models. *Biometrics* **51**, 1440-1450.

Gion JM, Boudet C, Grima-Pettenati J, Pichavant FH, Plomion C, Baillères H, Verhaegen D (2001) A candidate genes approach identifies CCR, PAL and C4H as loci for syringyl/guaiacyl ratio in an interspecific hybrid between *E. urophylla* and *E. grandis*. In 'Developing the Eucalypt of the Future, Proceedings of IUFRO International Symposium'. Valdivia, Chile. (Ed. Sea Barros)

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.

Goldstein H (1995) 'Multilevel statistical models, 2nd edn.' (Edward Arnold: London)

Gonzalez-Martinez SC, Gerber S, Cervera MT, Martinez-Zapater JM, Gil L, Alia R (2002) Seed gene flow and fine-scale structure in a Mediterranean pine (*Pinus pinaster* Ait.) using nuclear microsatellite markers. *Theoretical & Applied Genetics* **104**, 1290-1297.

Graner A, Jahoor A, Schondelmaier J, Siedler H, Pillen K, Fishbeck G, Wenzel G, Herrmann RG (1991) Construction of an RFLP map of barley. *Theoretical & Applied Genetics* **83**, 250-256.

Grattapaglia D (1997) Opportunities and challenges for the incorporation of genomic analysis in Eucalyptus breeding. In 'Proc. IUFRO Conf. On Silviculture and Improvement of Eucalypts'. Salvador, Brazil, pp. 129-136. (EMBRAPA)

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 & Applied Genetics* **90**, 933-947.

Grattapaglia D, Bertolucci FLG, 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, 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.

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

Greaves BL, Borralho NMG, Raymond CA (1995) Use of a Pilodyn for indirect selection of basic density in *Eucalyptus nitens*. In 'Eucalypt plantations: Improving Fibre Yield and Quality, Proc. CRCTHF-IUFRO Conf.' Hobart, Tasmania. (Ed. CA Raymond) pp. 106-109. (CRC for Temperate Hardwood Forestry)

Greaves BL, Borralho NMG, Raymond CA (1997) Breeding objective for plantation eucalypts grown for production of kraft pulp. *Forest Science* **43**, 465-472.

Griffin AR (1989) Sexual reproduction and tree improvement strategy - with particular reference to *Eucalyptus*. In 'Breeding tropical trees: Population structure and genetic improvement strategies in clonal and seedling forestry. Proc. IUFRO Conf.' Pattaya, Thailand. (Ed. AC Matheson). (Oxford Forestry Institute, Oxford, UK and Winrock International, Virginia, USA)

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

Haley CS, Knott SA, Elsen JM (1994) Mapping quantitative trait loci in crosses between outbred lines using least squares. *Genetics* **136**, 1195-1207.

Helentjaris T, Slocum M, Wright S, Schaefer A, Nienhuis J (1986) Construction of genetic linkage maps in maize and tomato using restriction fragment polymorphisms. *Theoretical & Applied Genetics* **72**, 761-769.

Hodge GR, Volker PW, Potts BM, Owen JV (1996) A comparison of genetic information from open-pollinated and control-pollinated progeny tests in two eucalypt species. *Theoretical & Applied Genetics* **92**, 53-63.

Isagi Y, Suhandono S (1997) PCR primers amplifying microsatellite loci of *Quercus myrsinifolia* Blume and their conservation between oak species. *Molecular Ecology* **6**, 897-899.

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.

Jones CJ, Edwards KJ, Castaglione S, Winfield MO, Sala F, van de Wiel C, Bredemeijer G, Vosman B, Matthes M (1997) Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories *Molecular Breeding* **3**, 381-390.

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 Forest 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 foliar waxes. *Canadian Journal of Forest Research* (in press).

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 & Management* **160**, 189-199.

Karhu A, Dieterich JH, Savolainen O (2000) Rapid expansion of microsatellite sequences in pines. *Molecular Biology & Evolution* **17**, 259-265.

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 & Applied Genetics* **98**, 586-592.

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

Kelly CM (1997) Natural variation and genetic control of relative bark thickness in *Eucalyptus globulus* ssp. *globulus*. Unpublished B.Sc. Honours thesis, University of Tasmania.

Kerr RJ, Jarvis SF, Goddard ME (1996) The use of genetic markers in tree breeding programs. In 'Tree improvement for sustainable tropical forestry. Proc. QFRI-IUFRO Conference'. Caloundra, QLD, Australia. (Eds MJ Dieters, AC Matheson, DG Nikles, CE Harwood and SM Walker) pp. 498-505. (Queensland Forestry Research Institute, Gympie, Australia)

Kesseli RV, Paran I, Michelmore RW (1992) Efficient mapping of specifically targeted genomic regions and the tagging of these regions with reliable PCR-based genetic markers. In 'Joint Plant Breeding Symposia Series - Applications of RAPD technology to plant breeding'. Minneapolis, Minnesota, USA pp. 31-36. (Crop Science Society of America, American Society for Horticulture, American Genetic Association)

Kesseli RV, 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.

Knott SA, Neale DB, Sewell MM, Haley CS (1997) Multiple marker mapping of quantitative trait loci in an outbred pedigree of loblolly pine. *Theoretical & Applied Genetics* **94**, 810-820.

Kubisiak TL, Nelson CD, Nance WL, Stine M (1995) RAPD linkage mapping in a longleaf pine x slash pine F-1 family. *Theoretical & Applied Genetics* **90**, 1119-1127.

Lagercrantz U, Ellegren H, Andersson L (1993) The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. *Nucleic Acids Research* **21**, 1111-1115.

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.

Lemos L, Carvalho A, Araujo JA, Borralho NMG (1997) Importance of additive genetic and specific combining ability effects for rooting ability of stem cuttings in eucalyptus globulus. *Silvae Genetica* **46**, 307-308.

Lerceteau E, Plomion C, Andersson B (2000) AFLP mapping and detection of quantitative trait loci (QTLs) for economically important traits in *Pinus sylvestris*: a preliminary study. *Molecular Breeding* **6**, 451-458.

Lerceteau E, Szmidt AE, Andersson B (2001) Detection of quantitative trait loci in *Pinus sylvestris* L. across years. *Euphytica* **121**, 117-122.

Lian CL, Miwa M, Hogetsu T (2001) Outcrossing and paternity analysis of *Pinus densiflora* (Japanese red pine) by microsatellite polymorphism. *Heredity* **87**, 88-98.

Lincoln S, Daly M, Lander E (1992) 'Constructing genetic maps with MAPMAKER/EXP 3.0.' Whitehead Institute Technical Report, 3rd edition, Cambridge, Mass., USA.

Lincoln S, Daly M, Lander E (1993) 'Mapping genes controlling quantitative traits using MAPMAKER/QTL Version 1.1: A tutorial and reference manual' Whitehead Institute Technical Report, 2nd edition, Cambridge, Mass., USA.

Liu Z, Furnier GR (1993) Inheritance and linkage of allozymes and restriction fragment length polymorphisms in trembling aspen. *Journal of Heredity* **84**, 419-424.

Lopez GA, Potts BM, Dutkowski GW, Apiolaza LA, Gelid PE (2002) Genetic variation and inter-trait correlations in *Eucalyptus globulus* base population trials in Argentina. *submitted.*

Lynch M, Walsh B (1998) 'Genetics and analysis of quantitative traits.' (Sinauer Associates Inc: Sunderland, Massachusetts, USA)

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

Macrae S, Cotterill PP (1997) Macropropagation and micropropagation of *Eucalyptus globulus*: means of capturing genetic gain. In 'Proceedings of the IUFRO conference on silviculture and improvement of Eucalypts'. Salvador, Brazil pp. 102-110. (EMBRAPA)

Mariette S, Chagne D, Decroocq S, Vendramin GG, Lalanne C, Madur D, Plomion C (2001) Microsatellite markers for *Pinus pinaster* Ait. *Annals of Forest Science* **58**, 203-206.

Marques CM, Araujo JA, Ferreira JG, Whetten R, Omalley DM, Liu BH, Sederoff R (1998) AFLP genetic maps of *Eucalyptus globulus* and *E. tereticornis*. *Theoretical & 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 & Applied Genetics* **105**, 474-478.

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

McCouch SR, Kochert G, Yu ZH, Wang ZY, Khush GS, Coffman WR, Tanksley SD (1988) Molecular mapping of rice chromosomes. *Theoretical & Applied Genetics* **1988**, 815-829.

Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 9828-9832.

Miranda I, Pereira H (2001) Provenance effect on wood chemical composition and pulp yield for *Eucalyptus globulus* Labill. *Appita Journal* **54**, 347-351.

Miranda I, Pereira H (2002) Variation of pulpwood quality with provenances and site in *Eucalyptus globulus*. *Annals of Forest Science* **59**, 283-291.

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

Moran GF, Bell JC (1983) *Eucalyptus*. In 'Isozymes in Plant Genetics and Breeding'. (Eds SD Tanksley and TJ Orton) pp. 423-441. (Elsevier: Amsterdam)

Morgante M, Olivieri AM (1993) PCR-amplified microsatellites as markers in plant genetics. *Plant Journal* **3**, 175-182.

Mukai Y, Suyama Y, Tsumura Y, Kawahara T, Yoshimaru H, Kondo T, Tomaru N, Kuramoto N, Murai M (1995) A linkage map for sugi (*Cryptomeria japonica*) based on RFLP, RAPD, and isozyme loci. *Theoretical & Applied Genetics* **90**, 835-840.

Muneri A, Raymond CA (2000) Genetic parameters and genotype-by-environment interactions for basic density, pilodyn penetration and stem diameter in *Eucalyptus globulus*. *Forest Genetics* **1**, 81-95.

Namkoong G, Kang HC, Brouard JS (1988) 'Tree breeding: Principles and strategies.' (Springer-Verlag: New York, NY)

Nelson CD, Kubisiak TL, Stine M, Nance WL (1994) A genetic linkage map of longleaf pine (*Pinus palustris* Mill) based on random amplified polymorphic DNAs. *Journal of Heredity* **85**, 433-439.

Nelson CD, Nance WL, Doudrick RL (1993) A partial genetic linkage map of slash pine (*Pinus-elliottii* Engelm. var. *elliottii*) based on random amplified polymorphic DNAs. *Theoretical & Applied Genetics* **87**, 145-151.

O'Brien SJ (Ed.) (1987) 'Genetic maps (5th edn).' (Cold Spring Harbor Laboratory: Cold Spring Harbor, New York)

O'Malley DM, McKeand SE (1994) Marker assisted selection for breeding value in forest trees. *Forest Genetics* **1**, 207-218.

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

Paglia GP, Olivieri AM, Morgante M (1998) Towards second-generation STS (sequence-tagged sites) linkage maps in conifers - a genetic map of norway spruce (*Picea abies* K). *Molecular & General Genetics* **258**, 466-478.

Patterson AH (1998) Of blending, beans, and bristles: the foundations of QTL mapping. In 'Molecular dissection of complex traits'. (Ed. AH Patterson) pp. 1-10. (CRC Press: Boca Raton, Fla.)

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

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

Plomion C, Durel CE, Omalley DM (1996) Genetic dissection of height in maritime pine seedlings raised under accelerated growth conditions. *Theoretical & Applied Genetics* **93**, 849-858.

Plomion C, Hurme P, Frigerio JM, Ridolfi M, Pot D, Pionneau C, Avila C, Gallardo F, David H, Neutelings G, *et al.* (1999) Developing SSCP markers in two *Pinus* species. *Molecular Breeding* **5**, 21-31.

Plomion C, Omalley DM, Durel CE (1995b) Genomic analysis in maritime pine (*Pinus pinaster*) - comparison of two RAPD maps using selfed and open-pollinated seeds of the same individual. *Theoretical & Applied Genetics* **90**, 1028-1034.

Potts BM, Volker PW, Hodge GR, Borralho NMG, Hardner CM, Owen JV (1995) Genetic limitations to the exploitation of base populations of *Eucalyptus globulus* ssp. *globulus*. In 'Eucalypt plantations: Improving Fibre Yield and Quality, Proc. CRCTHF-IUFRO Conf.' Hobart, Tasmania. (Ed. CA Raymond) pp. 217-221. (CRC for Temperate Hardwood Forestry)

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

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey SV, Rafalski JA (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular Breeding* **2**, 225-238.

Rafalski JA, Vogel JM, Morgante M, Powell W, Andre C, Tingey SV (1996) Generating and using DNA markers in plants. In 'Nonmammalian genomic analysis, a practical guide'. (Eds E Birren and E Lai) pp. 75-134. (Academic Press)

Rajora OP, Rahman MH (2001) Microsatellite DNA markers and their usefulness in poplars, and conservation of microsatellite DNA loci in Salicaceae. *Genetic Response Of Forest Systems To Changing Environmental Conditions* **70**, 105-115.

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 & Applied Genetics* **98**, 1279-1292.

Rodrigues J, Meier D, Faix O, Pereira H (1999) Determination of tree to tree variation in syringyl/guaiacyl ratio of *Eucalyptus globulus* wood lignin by analytical pyrolysis. *Journal of Analytical and Applied Pyrolysis* **48**, 121-128.

Ruaud JN, Lawrence N, Pepper S, Potts BM, Borralho NMG (1999) Genetic variation of in vitro rooting ability with time in Eucalyptus globulus. *Silvae Genetica* **48**, 4-7.

Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Mullikin JC, Mortimore BJ, Willey DL, Hunt SE, Cole CG, *et al.* (2001) A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* 409, 928-933

Sarfatti M, Katan J, Fluhr R, Zamir D (1989) An RFLP marker in tomato linked to the *Fusarium oxysporum* resistance gene *I2*. *Theoretical & Applied Genetics* **78**, 755-759.

SAS Institute (1989) 'SAS/STAT Users Guide, version 6, 4th ed.' (SAS Institute Incorporated: Cary, NC USA)

Sasse J, Sands R (1995) Root system development in cuttings of *Eucalyptus globulus*. In 'Eucalypt plantations: Improving Fibre Yield and Quality, Proc. CRCTHF-IUFRO Conf.' Hobart, Tasmania. (Eds BM Potts, NMG Borralho, JB Reid, RN Cromer, WN Tibbits and CA Raymond) pp. 299-303. (CRC for Temperate Hardwood Forestry)

Sewell MM, Bassoni DL, Megraw RA, Wheeler NC, Neale DB (2000) Identification of QTLs influencing wood property traits in loblolly pine (*Pinus taeda* L.). I. Physical wood properties. *Theoretical & Applied Genetics* **101**, 1273-1281.

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

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

Shepherd M, Cross M, Maguire TL, Dieters MJ, Williams CG, Henry RJ (2002) Transpecific microsatellites for hard pines. *Theoretical & Applied Genetics* **104**, 819-827.

Skabo S, Vaillancourt RE, Potts BM (1998) Fine-scale genetic structure of Eucalyptus globulus ssp. globulus forest revealed by RAPDs. *Australian Journal of Botany* **46**, 583-594.

Soller M (1991) Mapping quantitative trait loci affecting traits of economic importance in animal populations using .nolecular markers. In 'Gene-mapping techniques and applications'. (Eds LB Shook, HA Lewin and DG McLaren) pp. 21-49. (Marcel Dekker Inc.: New York)

Soller M, Brody T, Genezi A (1976) On the power of experimental designs for the detection of linkage between marker loci and quantitative trait loci in crosses between inbred lines. *Theoretical & Applied Genetics* **47**, 35-39.

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.

Streiff R, Ducousso A, Lexer C, Steinkellner H, Gloessl J, Kremer A (1999) Pollen dispersal inferred from paternity analysis in a mixed oak stand of *Quercus robur* L- and *Q. petraea* (Matt.) Liebl. *Molecular Ecology* **8**, 831-841.

Stuber CW, Lincoln SE, Wolff DW, Helentjaris T, Lander ES (1992) Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. *Genetics* **132**, 823-839.

Tanksley SD, Orton TJ (Eds) (1983) 'Isozymes in plant genetics and breeding parts A and B.' (Elsevier Science Publications: Amsterdam)

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 & Applied Genetics* **104**, 379-387.

Thormann CE, Osborn TC (1992) Use of RAPD and RFLP markers for germplasm evaluation. In 'Joint Plant Breeding Symposia Series - Applications of RAPD technology to plant breeding'. Minneapolis, Minnesota, USA pp. 9-11. (Crop Science Society of America, American Society for Horticulture, American Genetic Association)

Travis SE, Ritland K, Whitham TG, Keim P (1998) A genetic linkage map of pinyon pine (*Pinus edulis*) based on amplified fragment length polymorphisms. *Theoretical & Applied Genetics* **97**, 871-880.

Tsarouhas V, Gullberg U, Lagercrantz U (2002) An AFLP and RFLP linkage map and quantitative trait locus (QTL) analysis of growth traits in *Salix*. *Theoretical & Applied Genetics* **105**, 277-288.

Tulsieram LK, Glaubitz JC, Kiss G, Carlson JE (1992) Single tree genetic linkage mapping in conifers using haploid DNA from megagametophytes. *Bio/Technology* **10**, 686-690.

Vaillancourt RE, Potts BM, Manson A, Eldridge T, Reid JB (1995a) Using RAPDs to detect QTLs in an interspecific F2 hybrid of *Eucalyptus*. In 'Eucalypt plantations: Improving Fibre Yield and Quality, Proc. CRCTHF-IUFRO Conference'. Hobart, Tasmania. (Eds BM Potts, NMG Borralho, JB Reid, RN Cromer, WN Tibbits and CA Raymond) pp. 430-433. (CRC for Temperate Hardwood Forestry)

Vaillancourt RE, Potts BM, Watson M, Volker PW, Hodge GR, Reid JB, West AK (1995b) Detection and prediction of heterosis in *Eucalyptus globulus*. *Forest Genetics* **2**, 11-19.

Vaillancourt RE, Slinkard AE (1992) Inheritance of new genetic markers in lentil (Lens Miller). *Euphytica* **64**, 227-236.

Vallejos CE, Sakiyama NS, Chase CD (1992) A molecular marker-based linkage map of *Phaseolus vulgaris* L. *Genetics* **131**, 733-740.

van Buijtenen JP (2001) Genomics and quantitative genetics. *Canadian Journal of Forest Research* **31**, 617-622.

van de Ven WTG, McNicol RJ (1996) Microsatellites as DNA markers in Sitka spruce. *Theoretical & Applied Genetics* **93**, 613-617.

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, Kremer A (1997) Quantitative trait dissection analysis in *Eucalyptus* using RAPD markers .1. Detection of QTL in interspecific hybrid progeny, stability of QTL expression across different ages. *Theoretical & Applied Genetics* **95**, 597-608.

Vogl C, Karhu A, Moran G, Savolainen O (2002) High resolution analysis of mating systems: inbreeding in natural populations of *Pinus radiata*. *Journal of Evolutionary Biology* **15**, 433-439.

Volker PW (2002) Unpublished PhD thesis, University of Tasmania.

Volker PW, Owen JV, Borralho NMG (1994) Genetic variances and covariances for frost tolerance in *Eucalyptus globulus* and *E. nitens*. *Silvae Genetica* **43**, 366-372.

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

Watt MP, Duncan EA, Ing M, Blackway FC, Herman B (1995) Field performance of micropropagated and macropropagated Eucalyptus hybrids. *South African Forestry Journal* **173**, 17-21.

Welsh J, McClelland M (1990) Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research* **21**, 7213-7218.

Weng C, Kubisiak TL, Nelson CD, Stine M (2002) Mapping quantitative trait loci controlling early growth in a (longleaf pine x slash pine) x slash pine BC1 family. *Theoretical & Applied Genetics* **104**, 852-859.

Wilcox PL, Carson SD, Richardson TE, Ball RD, Horgan GP and Carter P (2001a) Benefit-cost analysis of DNA marker-based selection of Pinus radiata seed orchard parents. *Canadian Journal of Forest Research* **31**, 2213-2224.

Wilcox P, Richardson T, Corbett G, Ball R, Lee J, Djorovic A, Carson S (2001b) Framework linkage maps of *Pinus radiata* D.Don. based on pseudotestcross markers. *Forest Genetics* **8**, 109-117

Williams CG (1995) Beyond marker-assisted selection: Use of genomic mapping in breeding strategy. In 'Eucalypt plantations: Improving Fibre Yield and Quality, Proc. CRCTHF-IUFRO Conf.' Hobart, Tasmania. (Eds BM Potts, NMG Borralho, JB Reid, RN Cromer, WN Tibbits and CA Raymond) pp. 438-442. (CRC for Temperate Hardwood Forestry)

Williams CG, Neale DB (1992) Conifer wood quality and marker-aided selection: a case study. *Canadian Journal of Forest Research* **22**, 1009-1017.

Williams JGK, Hanafey MK, Rafalski JA, Tingey SV (1993) Genetic analysis using random amplified polymorphic DNA markers. *Methods in Enzymology* **218**, 704-740.

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

Willyams D, Whiteman P, Cameron J, Chandler S (1992) Inter- and intra-family variation in rooting capacity in micropropagated *Eucalyptus globulus* and *Eucalyptus nitens*. In 'Mass production technology for genetically improved fast growing tree species, AFOCEL-IUFRO Symposium 1992'. Nangis, France pp. 177-181. (Association Foret Cellulose)

Wilson PJ (1992) The development of new clones of *Eucalyptus globulus* and *E. globulus* hybrids by stem cuttings propagation. In 'Mass production technology for genetically improved fast growing tree species, AFOCEL-IUFRO Symposium 1992'. Nangis, France pp. 379-386. (Association Foret Cellulose)

Wilson PJ (1993) Propagation characteristics of *Eucalyptus globulus* Labill. ssp. *globulus* stem cuttings in relation to their original position in the parent shoot. *Journal of Horticultural Science* **68**, 715-724.

Wilson PJ (1995) Multiplication rates in vitro and by stem cuttings propagation, and the development of seedling-origin clones of *Eucalyptus globulus*. In 'Eucalypt plantations: Improving Fibre Yield and Quality, Proc. CRCTHF-IUFRO Conf.' Hobart, Tasmania. (Eds BM Potts, NMG Borralho, JB Reid, RN Cromer, WN Tibbits and CA Raymond) pp. 304-307. (CRC for Temperate Hardwood Forestry)

Wilson PJ (1999) The growth and form of potted mother plants of *Eucalyptus globulus* Labill. ssp. *globulus* in relation to the rooting ability of stem cuttings. *Journal of Horticultural Science & Biotechnology* **74**, 645-650.

Wimmer R, Downes GM, Evans R, Rasmussen G, French J (2002) Direct effects of wood characteristics on pulp and handsheet properties of *Eucalyptus globulus*. *Holzforschung* **56**, 244-252.

Wu RL, Han YF, Hu JJ, Fang JJ, Li L, Li ML, Zeng ZB (2000) An integrated genetic map of *Populus deltoides* based on amplified fragment length polymorphisms. *Theoretical & Applied Genetics* **100**, 1249-1256.

Yoshimaru H, Ohba K, Tsurumi K, Tomaru N, Murai M, Mukai Y, Suyama Y, Tsumura Y, Kawahara T, Sakamaki Y (1998) Detection of quantitative trait loci for juvenile growth, flower bearing and rooting ability based on a linkage map of sugi (*Cryptomeria japonica* D. Don). *Theoretical & Applied Genetics* **97**, 45-50.

Zamir D, Tadmor Y (1986) Unequal segregation of nuclear genes in plants. *Botanical Gazette* **147**, 355-358.

Zeng ZB (1993) Theoretical basis of precision mapping of quantitative trait loci. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 10972-10976.

Zhou Y, Bui T, Auckland LD, Williams CG (2002) Undermethylated DNA as a source of microsatellites from a conifer genome. *Genome* **45**, 91-99.

Zhuang JY, Lin HX, Lu J, Qian HR, Hittalmani S, Huang N, Zheng KL (1997) Analysis of QTL x environment interaction for yield componenets and plant height in rice. *Theoretical & Applied Genetics* **95**, 799-808.

Zobel BJ (1993) Clonal forestry in the *Eucalyptus*. In 'Clonal Forestry II, Conservation and Application'. (Eds MR Ahuja and WJ Libby) pp. 140-148. (Springer-Verlag: Berlin, Heidelberg)