Detection And Genetic Mapping Of Quantitative Trait Loci Influencing Stem Growth Efficiency In Radiata Pine

Livinus Chinenye Emebiri

Department of Forestry

Australian National University

A thesis submitted for the degree of Doctor of

Philosophy of The Australian National University

January 1997

Declaration

The work presented in this thesis is my own. Assistance and specific contributions by others are referred to in the text and acknowledgments.

L. C. Emebiri

January 1997

Department of Forestry

School of Resource and Environmental Management

Faculty of Science

The Australian National University

The following papers are based on this dissertation:

Emebiri, L. C., Devey, M. E., Matheson, A. C. and Slee, M. U. (1997) Linkage of RAPD markers to NESTUR, a stem growth index in radiata pine seedlings. *Theoretical & Applied Genetics* (in press).

Emebiri, L. C.; Devey, M. E.; Matheson, A. C. and Slee, M. U. (1995) On genetically remodelling radiata pine: identification of QTLs influencing NESTUR by use of bulked segregant analysis. Paper presented at the Conifer Biotechnology Working Group 7th International Conference, 26-30 June 1995, Surfers Paradise, Queensland, Australia.

Emebiri, L. C.; Devey, M. E.; Matheson, A. C. and Slee, M. U. (1995) Detection and mapping of QTLs influencing NESTUR in radiata pine. Paper presented at the sixth Australasian Gene Mapping Workshop & New Zealand Genetical Society Conference, Nov. 27-1 Dec., 1995, University of Otago, Dunedin, New Zealand.

Emebiri, L. C.; Devey, M. E.; Matheson, A. C. and Slee, M. U. (1996). Interval mapping of NESTUR, a stem growth efficiency index in radiata pine. Poster presentation submitted for the Scientific Regional Information Exchange Group (SRIEG) Conference, Texas A & M University, Houston, Texas, June 23-26. Abstract available at http://mslismpa.tamu.edu/staff/tom/SRIEG/srieghome.html (manuscript in preparation).

Emebiri, L. C.; Devey, M. E.; Matheson, A. C. and Slee, M. U. (1997). Age related changes in the expression of QTL for growth in radiata pine seedlings. Submitted to *Theoretical & Applied Genetics.*

Emebiri, L. C.; Devey, M. E.; Matheson, A. C. and Slee, M. U. (1997). Interval mapping of quantitative trait loci affecting NESTUR, a stem growth efficiency index of radiata pine seedlings. Submitted to *Theoretical & Applied Genetics.*

Acknowledgments

I thank my supervisory panel of Dr. M. U. Slee, Dr. M. E. Devey, and Dr. A. C. Matheson for their advice and encouragement throughout the course of my studies.

Special thanks to Dan Gardener for technical assistance with the dendrometer, and to Dane Donaldson, Nathan Caesar and Josie Morosin for assistance during NESTUR measurements. I thank David Spencer for assistance with design of field layout, planting and maintenance of the progeny evaluation.

I am especially grateful to Ms Judy Faccioni for invaluable assistance and support throughout this study. Special thanks also to Prof. Peter Kanowski for his advice and encouragement.

This study was funded by an ANU PhD scholarship from the Australian National University, Canberra, and an OPRS scholarship from the Department of Employment, Education & Training (DEET), Australia. Financial and institutional support were also provided by the Southern Tree Breeding Association (STBA), Australia, the CSIRO Forestry & Forest Products, Canberra, and by the Department of Forestry, The Australian National University, Canberra.

Abstract

Needle-to-stem unit rate (NESTUR) is a stem growth index of conifer seedling trees that measures the efficiency of stemwood production per unit of needle growth. Five experiments were carried out in this thesis using progenies of two unrelated full-sib radiata pine crosses. The initial experiment (experiment 1) applied the bulked segregant analysis technique to determine whether RAPD analysis could be successfully extended to the development of molecular markers for NESTUR in radiata pine. The NESTUR values of 174 progenies of the full-sib family 12038 x 10946 were determined. Based on the genotypic analysis of the individuals, two quantitative trait loci (QTL) controlling NESTUR were identified at ANOVA P-levels of 0.01-0.001. An absence of RAPD fragment markers generated by primers OPE-06 and OPA-10 was associated with low NESTUR values, while primer UBC-333 generated a 550 bp band that was associated with high NESTUR values. Linkage to components of NESTUR (increments in stem diameter and stem volume) was demonstrated for one of the QTL, while the other was unique to NESTUR, and not shared with the components. There was a significant interaction between the two QTLs. Presence of OPA-10_{1200} locus appeared to inhibit expression of the QTL linked to UBC-333_{550}.

To further analyse the quantitative trait loci (QTLs) controlling NESTUR, a linkage map was constructed from RAPD markers segregating in 93 haploid progeny of another full sib cross (30040 x 80121) (experiment 2). Two hundred and sixty-two (262) markers were mapped to 14 linkage groups of at least 7 markers, ranging in size from 39 to 183 cM. The 14 linkage groups covered approximately 1511 cM of genetic map distance.

In experiment 3, the linkage map was used to map QTLs controlling NESTUR, as well as increments in seedling stem diameter, volume, and height and needle volume. Altogether, five putative QTLs were detected for NESTUR, with explained variation ranging from 9 to 22%. Of the five QTLs detected, 3 were coincidental with those for stem growth in height, diameter and volume. The two QTL positions that were unique to NESTUR were flanked by QTLs for the component traits. Together, effects of the five QTLs explained 48% of the total phenotypic variation for NESTUR.

Ability of identified markers to predict the phenotype and seedlings with growth potential was assessed in the cross 30040 x 80121, using six RAPD markers associated with NESTUR at ANOVA P-levels of 0.01-0.001 (experiment 4). The correlation between observed NESTUR and predicted values was 0.70. Differences in observed vs. predicted values were not large and did not indicate serious misclassifications, such as classification of an upper ranking individual into the lower group, or vice versa.

Over a two-year growth period, the ability of NESTUR to predict stem growth was strongly affected by seedling age. In contrast, markers linked to NESTUR showed a consistent ability to predict stem growth, irrespective of seedling age. Compared with the top 1% of the original population, seedlings selected for their genotypic values showed a higher stem volume growth of 103% in the first year, and 76% in the second year.

The expression of QTLs for stem volume, stem diameter, height, number of branches, number of whorls, and branches/whorl were compared at 5, 12, and 24 months of age. Two QTLs detected for height showed contrasting expression over two years, one was gradually reduced from LOD of 2.70 to 0.43 and the other

increased from 1.12 to 2.45. Compared with the pattern observed for height, LOD scan profiles for diameter and volume showed less temporal change of peaks, suggesting that the genetic control for height growth is probably more unstable than that of diameter. QTLs controlling the phenotype at the time of measurement (ie the final phenotype) showed similar magnitude of effects on that trait's respective increments (or growth rate).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

1 Introduction

1.1 Limited progress in basic tree breeding

Radiata pine (*Pinus radiata* D. Don) is the most important plantation species in Australia and New Zealand, with a total area planted of 0.9 and 1.2 million hectares, respectively. The species is native to three localities on the central coast of California (Cambria, Monterey and Ano Nuevo), and to two small Mexican islands (Guadalupe and Cedros). The Australian plantations are mainly derived from the natural populations around Monterey, with probably some admixture of genes from the Ano Nuevo stands which occur some 50 km to the north of Monterey (Moran and Bell 1987). Breeding programs in Australia were initiated in the 1950s with the objective of improving productivity through improved stem form, increased growth rate and adaptation to sites.

Summaries of progress using basic tree breeding methods (Jett 1988; Zobel and Talbert 1984) have shown them to be effective but also slow (Tauer and Hallgren 1992; Krugman 1985). Forest tree improvement methods in the past have relied, almost exclusively, on phenotypic selection, expensive long-term field progeny testing for phenotypic traits, and generally elaborate statistical analysis of the data. The majority of characters of interest in commercial forest tree improvements are quantitative in nature with heritability (h^2) often being relatively low. Gain from phenotypic selection is limited when h^2 is small because the phenotypic value of an individual is a poor predictor of its genetic value. The genetic value of an individual is determined by progeny tests where the offspring from different families are compared in a common environment. In a breeding program, therefore, considerable efforts have to be devoted to the conduct of progeny testing procedures to identify superior parent plants.

1.2 Age-Age correlation and early selection

Gains and genetic progress are also limited both by the long life cycle of trees, and by the limited proportion of genetic variance the breeder can capture at an early stage. Many years may elapse before a tree starts to flower and yield enough seeds for testing purposes. Again after breeding, a very long time may elapse before the value of any offspring can be assessed and a selection for the next generation can be made.

Effective early assessment and selection in forest tree breeding would result in a valuable decrease of generation turnover time. Schmidt (1936) pioneered early selection in forestry, but results so far have been inconclusive. Stem growth is the cumulative result of carbon assimilation and allocation over time and stemwood production is a function of such process components as photosynthesis, net assimilation rate, and dry matter partitioning among roots, stem, branches, and foliage. These processes are in themselves complex and influenced by genetic as well as environmental variables. Growth efficiency results from favourable allocation of photosynthates to stemwood production, relative to foliage growth. The assessment of many process components for early selection, some of which may be destructive, renders the application of early selection difficult in routine breeding. A single measure of growth or other physiological parameter seems unlikely to give a generally applicable and reliable basis for early selection (Greenwood and Volkaert 1992). Derived traits, described by a function involving two or more primary traits, may provide opportunities for progress towards improved early selection (Matheson[1]).

[1] A. C. Matheson, CSIRO Forestry & Forest Products, Banks Street, Yarralumla, Canberra ACT, Australia. Postal Address: PO Box 4008, Queen Victoria Terrace, Canberra ACT 2600, Australia.

1.3 Needle-to-stem unit ratio (NESTUR)

Carbon partitioning during growth and development, rather than carbon assimilation *per se*, is highly correlated with yields of crop plants, and has proved amenable to breeding and selection (Snyder and Carlson 1984). Thus, if that same principle applies to trees, the potential for early selection in forest tree breeding would be improved with the development of a single index that is analogous, in functional terms, to the same process underlying stem growth in mature trees. Using a measure of stem:needle mass in seedlings, Matheson *et al*. (1995) obtained genetic correlations of as high as 0.98 with the relative growth rate (RGR) of stem diameter in 8-11 year-old radiata pine trees growing in several plantations in Australia. However, the methods of trait measurement were destructive, precluding further use of valuable germplasm.

Measuring systems were therefore devised for non-destructive and continuos assessment of stem:needle growth rate in individuals. Termed needle-to-stem unit rate (NESTUR), the index is derived as a ratio of unit growth in stem volume to needle volume over time. The index is similar to relative growth rate (RGR), but unlike RGR, it takes into account the influences of both photosynthetic capacity and carbon allocation on stemwood production. As a selection criterion, NESTUR identifies seedling trees, which are likely to be more efficient producers of volume increment in later-age. Although the measuring systems are non-destructive, its determination is nonetheless, tedious, time-consuming and laborious. Moreover, stem growth efficiency, like other quantitative traits, is assumed a consequence of the joint action of several genes. While the theory and techniques of quantitative genetics have proven useful in the study of quantitative traits, these characters continue to be more difficult to manipulate in breeding programs than single gene traits. If such

complex traits can be ‘tagged’ with genetic markers, it might be possible to manipulate these characters with the efficacy of single gene traits.

1.4 Molecular marker-assisted selection

The use of molecular markers to aid selection in forest species is currently a major research emphasis in many tree-improvement programs (O’Malley and McKeand 1994; Grattapaglia *et al.* 1993; Stam 1994; Williams and Neale 1992; Nance *et al*. 1992). Marker-assisted selection methods are potentially more efficient for early identification of superior individuals. The methods are based on the establishment of a linkage relationship between an easily scorable molecular marker and a character of interest.

In the past, widespread doubts were expressed about the development of genetic linkage maps of forest trees, and their utility in breeding programs (see Strauss *et al.* 1992). Markers were considered to have limitations, even in crop plants where genetic analysis was more advanced. Forest trees had high levels of linkage equilibrium, and lacked suitable pedigrees for construction of genetic linkage maps. In the absence of extended pedigrees, inbred lines or other situations where linkage disequilibrium could be readily established, it was unlikely that specific associations with important traits could be made.

The last four years has witnessed an increased catalogue of genetic linkage maps of forest tree species. Two factors have made it relatively easy to generate genetic maps of forest tree species: the high levels of heterozygosity in most forest tree populations and the advent of PCR-based genetic markers (Grattapaglia *et al.* 1993). A commonly used genetic marker system is the PCR-based system called random amplified polymorphic DNA (RAPD) (Williams *et al*. 1990; Welsh and McClelland 1990). The technique is based on the amplification of random DNA segments with single primers of arbitrary

nucleotide sequence. RAPD fragments are produced based on the frequency and location of short inverted DNA repeats. The products of RAPD amplification reactions can be viewed directly by agarose gel electrophoresis without the need for southern blots.

Unlike isozymes and RFLPs, RAPD allelic forms are generally dominant, rather than co-dominant. At an individual locus in a segregating family from a pair of cross, only two 'allelic states' are visualised, a presence or an absence of the fragment. For some applications, this might create difficulties in identification of the true genotype at a heterozygous RAPD locus. With haploid populations such as megagametophytes derived from a single conifer tree, RAPDs show 1:1 segregation, which are equivalent to a testcross. Thus, with RAPD markers, the use of the haploid tissue avoids many of the problems normally associated with quantitative trait dissection in trees. RAPDs in conjunction with haploid DNA have been used as molecular markers in linkage map construction and mapping of monogenic qualitative traits.

1.5 Quantitative trait locus (QTL) analysis

The identification of quantitative trait locus (QTL) is the first step towards developing marker-assisted selection strategies. QTL analysis makes it feasible to identify, map and measure the effects of genes underlying quantitative traits. In families where strong evidence of marker-linked QTLs have been detected, the next step is to use the information to test predictive hypotheses regarding potential of marker-assisted selection, and then to verify these potential using independent experiments.

The application of RAPDs developed from haploid DNA of conifer megagametophytes to quantitative trait loci dissection has not previously been reported. Most reported studies have only considered simply-inherited loci,

where the influence of non-genetic factors can be greatly minimised by more accurate phenotyping of the individuals (Wilcox *et al*. 1996; Devey *et al*. 1995; Lehner *et al.* 1995; Yuang and Kruger 1994; Benet *et al.* 1995; Hormaza *et al.* 1994; Cheng and Roose 1995; Mhameed *et al.* 1995). Because QTL detection using haploid megagametophytes exploits only the linkage disequilibrium between markers and QTL segregating on the maternal side of the pedigree, the paternal contribution to the phenotype may confound the power of detection. Only QTL segregating on the maternal side of the pedigree will be detectable. The extent to which these affect QTL detection and the possibility of marker-assisted selection for a quantitative trait has not previously been highlighted, and form part of the general objectives of this research.

1.6 The aim of the present study

1.6.1 Primary objectives

The primary objectives of this study were:

(i) To determine whether RAPD analysis could be successfully extended to the development of molecular markers for NESTUR in radiata pine;

(ii) Analyse the quantitative trait loci (QTL) controlling NESTUR, with the use of a RAPD-based genetic linkage map;

1.6.2 Secondary objectives

In addition to the primary objectives, the study sought to:

(iii) Compare field growth performance of seedling trees selected for NESTUR using phenotypic and marker-based methods.

(iv) Assess QTL stability for radiata pine seedlings at the early growth stages.

CHAPTER 2

Review of literature

2 Types of genetic markers

Until recently, the genetic markers used to develop linkage maps in plants have been those affecting morphological characters. The morphological markers generally correspond to qualitative traits that can be scored visually, such as seed-coat colour, dwarfism and altered leaf morphology. They occur naturally, but can be generated from mutagenesis experiments. Tanksley and Mutschler (1990) described more than 300 of such mutants occurring in tomato (*Lycopersicon* spp.). Morphological markers can be dominant or recessive, and have been found useful in plant breeding and selection. In tomato, for instance, the *Tm-2* gene for resistance to tobacco mosaic virus (TMV) is linked to an anthocyanin-less seedling marker (Robinson *et al*. 1970).

2.1 Morphological markers

Morphological traits controlled by a single locus can be used as genetic markers if their expression is reproducible over a range of environments (Staub and Serquen 1996). Morphological markers have been found useful in linkage map construction of forest trees. Such markers as yellow foliage, white cotyledons, non-white cotyledonous lethals, virescence, dwarfs, and mottled and curly needles were used in many early genetic linkage analysis studies (reviewed by Jermstad *et al*. 1994). Morphological markers are still being used (eg. Chaparro *et al*. 1994), but the majority cause such large effects on phenotype that they are undesirable in breeding programs (Tanksley *et al*. 1989). Moreover, strong epistatic interactions among morphological marker loci limits the number of segregating markers that can be unequivocally scored in the same generation (Tanksley 1993).

2.2 Molecular markers

Molecular markers reveal polymorphism at the protein or DNA level, and are often encountered as a particular isozyme or DNA fragment having a defined position when separated on an electrophoretic gel. The first linkage studies of *Pinus* were based on segregation of isozymes extracted from megagametophytes. More than 10 species were studied for about 15 loci (reviewed by Tulsieram *et al*. 1992). However, the number of genetic markers provided by isozyme assays is insufficient for many applications in plant breeding (Tanksley *et al*. 1989). As a result, even with the use of isozymes as genetic markers, genetic mapping in forest trees did not become widespread until recently.

The current increase in genetic linkage mapping of forest tree species is primarily due to the number of polymorphic DNA markers that can be generated in a single mapping population. DNA markers are quantifiable characters that distinguish, or mark, underlying genetic differences between individuals. They are inherited according to the Mendelian laws of inheritance, and through linkage to important genes, facilitate the detection of differences in the genetic information carried by individuals.

Polymorphism in the nucleotide sequence of genomic DNA usually is sufficient for it to function as a DNA marker in genetic analysis. These polymorphisms are revealed by molecular techniques such as restriction fragment length polymorphism (RFLP), random-amplified polymorphic DNA (RAPD), sequence characterised amplified regions (SCAR), amplified fragment length polymorphism (AFLP) microsatellite or simple-sequence repeat polymorphism (SSRP), cleavable amplified polymorphic sequences (CAPS) and single-strand conformational polymorphism (SSCP). The most commonly used DNA markers are restriction fragment length polymorphisms (RFLPs) and randomly amplified polymorphic DNAs (RAPDs). In the last three years, however, usage of such markers as amplified fragment length polymorphism (AFLP) and simple sequence repeats (SSR) or microsatellites has also become widespread.

2.2.1 RFLP markers

Among the various molecular markers developed, RFLPs were the first to be used in human genome mapping (Botstein *et al*. 1980). It was later adopted for plant genome (Weber and Helentjaris 1989) and forest trees (Devey *et al*. 1994). Restriction fragment length polymorphisms (RFLPs) are detected by the use of restriction enzymes that cut genomic DNA molecules at specific nucleotide sequences (restriction sites), thereby yielding variable-size DNA fragments (Staub and Serquen 1996). Identification of genomic DNA fragment is made by Southern blotting, a procedure whereby DNA fragments, separated by electrophoresis, are transferred to nitrocellulose or nylon filter (Southern 1975). Filter-immobilised DNA is then allowed to hybridise to radioactively labelled probe DNA. The filter is placed against a photographic film where radioactive disintegration from the probe results in visible bands that are often codominant.

2.2.2 AFLP

Amplified fragment length polymorphism (AFLP) is based on PCR amplification of restriction fragments generated by specific restriction enzymes and oligonucleotide adapters of few nucleotide bases (Vos *et al*. 1995). Multiple bands are generated in each amplification reaction that contains DNA markers of random origin. AFLPs are dominant, but heterozygous and homozygous genotypes can be quantitatively differentiated by the intensity of the amplified bands.

2.2.3 SSR or Microsatellites

Simple sequence repeats or microsatellites are based on tandem arrays of 2-to-5bp monomeric repeat units in a defined region of the genome (Rafalski and Tingey 1993). Polymorphisms appear because of variation in the number of tandem repeats in a given repeat motif (Staub and Serquen 1996). Most SSRs

are based on a dinucleotide repeat of the simple sequences [(AC)n, (AG)n, and (AT)n]. The number and composition of microsatellite repeats differ in plants and animals. In humans, AC or TC is a very common repeat unit, but in plants, AT is the more common followed by AG to TC (Powell *et al*. 1996) but in general, plants have 10 times less SSRs than humans. Polymorphism of such repeats is amplified by designing primers from sequenced regions flanking the repeat motifs (Staub and Serquen 1996).

Microsatellites have been developed as genetic markers for *Pinus radiata* (Smith and Devey 1994), as well as several tree species such as *Eucalyptus* (Byrne *et al*. 1996) and apple (Guilford *et al*. 1997). Two advantages of microsatellite loci are that they are highly variable codominant loci and they are PCR-based markers. However, a disadvantage, at least in forest tree breeding, is the considerable length of time taken in development of microsatellite loci (Byrne *et al*. 1996).

2.2.4 RAPDs

Random amplified polymorphic DNA markers (Williams *et al*. 1990; Welsh and McClelland 1990) are obtained by polymerase chain reaction (PCR) amplification of random DNA segments from single arbitrary primers (Saiki *et al.* 1988). The arbitrary primers used for the procedure are usually 9-10 bp in size, with a GC content of 50 to 80%, and do not contain palindromic sequences (Yu *et al.* 1993). The sequence of the primers is arbitrary, and no probe libraries, radioactivity or southern transfers are required, making the technique accessible to non-molecular biologists. All components of the reaction mixture (template DNA, primers, enzyme, nucleotides and an appropriate buffer mixture) are added together, and the temperature changes necessary for amplification to occur are automated.

Polymorphisms for RAPDs may result from single base changes, deletions or insertions in the template DNA. When two sites are separated by no more than

4000 bases and inverted in orientation, the PCR reaction amplifies the DNA between the sites. With a typical PCR cycle of 25-50, concentration of the units increases exponentially. When the reaction mixture is subjected to electrophoresis in an agarose gel, the amplified sequence is seen as a DNA band. Since a higher plant genome is very large, several amplified fragments are normally observed when a single primer is employed in the reaction. If one or more of the target sites are absent, no amplification takes place, and a null phenotype is indicated by the absence of a band.

Ethidium bromide-stained agarose gels are widely used to separate the amplification products (Williams *et al*. 1990), but in some cases silver-stained polyacrylamide gels (Bassam *et al.* 1991), capillary electrophoresis systems (Marino *et al.* 1994) and denaturing-gradient-gel electrophoresis (Dweikat *et al.* 1993; He *et al.* 1992) have been used.

Several reviews have compared RAPDs and RFLPs for detecting genetic polymorphisms in both animals and plants (Hallden *et al.* 1996; Ragot and Hoisinton 1993; Tigy and Tufo 1993; Arheim 1990; Weber 1990). There is a general agreement that RAPDs offer a number of important advantages over RFLPs, although their use in genetic studies and improvement programs of forest tree species has only recently become widespread. Like all molecular marker systems, there are potential problems with RAPDs, which have been addressed in several reports. The most important of these are related to reproducibility and nonspecific amplifications.

2.2.4.1 Reproducibility

RAPD amplification products are very dependent upon several factors including target DNA, primer length and sequence, reaction components (especially Mg^{++}), primer concentration, template concentration, type and concentration of polymerase enzyme, cycling temperatures, transition interval, duration, number of amplification cycles, and gel scoring procedures (Davinregli *et al.* 1995; Schweder *et al.* 1995; Park and Kohel 1994; Pammi *et al.* 1994; Penner *et al.*

1993; Ellsworth *et al.* 1993; Scheierwater and Ender 1993; Wolff *et al.* 1993; Muralidharan and Wakeland 1993; Macpherson *et al.* 1993). Individual brands of thermocyclers have characteristic temperature ramping profiles and display relative to actual block temperature (Smith and Williams 1994; Macpherson *et al.* 1993; Meunier and Grimont 1993).

Initially, the RAPD technology met with scepticism in respect of achievable repeatability. The reservations have been critically examined. In the period 1993-1996, there are over a thousand references, which attest to the robustness and high repeatability of RAPDs, provided rigorous attention is paid to the establishment of suitable experimental protocols. Profiles generated by arbitrary primers can be no more problematic to obtain and database than multibanded profiles generated by other methods (Hallden *et al.* 1996; Smith and Williams 1994; Kresovich *et al.* 1994; Jagt *et al.* 1993; Penner *et al.* 1993; Schierwater and Ender 1993; Wilkie *et al.* 1993).

2.2.4.2 Nonspecific amplification

Another potential problem with RAPD is the amplification of products other than those predicted. High background and low specificity of primer binding can occur in a PCR reaction when reaction components are mixed at room temperature. Reaction setup below the optimal primer annealing temperature permits non-specific primer annealing and extension. Undesired, non-specific primer extension products formed this way may be amplified in the PCR, resulting in misprimed products. Background amplification not only confuses test results, it interferes with amplification of predicted products by consuming reaction reagents.

Background amplification can be reduced by using the 'Hot-start' (Chou *et al*. 1992) or 'Heat-soaked' (Ruano *et al*. 1992) technique. The 'Hot-start' technique involves withholding at least one of the essential reagents (DNA polymerase, magnesium chloride, and primers) from the PCR mixture until the system has reached a temperature that favours specific primer annealing (Erlich *et al*.

1991). 'Hot-start' technique has been implemented by mechanically separating a reagent, and by blocking the polymerase enzyme with an antibody. Other variations in the technique have used a heat-labile uracil glycosylase to destroy nonspecific amplification products, and a modified form of DNA polymerase.

Conditions that increase the stringency of primer hybridisation, such as higher annealing temperatures and lower MgCl2 concentrations, enhance specific amplification (Riedel *et al.* 1992; Erlich and Arnheim 1992). Other approaches that may improve specificity include the use of two primers in each reaction, instead of one (Hu *et al.* 1995), and a secondary amplification of the PCR product at a higher stringency, using the same primer (Carlson *et al.* 1991).

Heteroduplex molecules formed between allelic RAPD products are also a potential source of artifactual polymorphism that can arise during RAPD analysis. Ayliffe *et al.* (1994) amplified genomic DNA of an F_1 strain of flax rust (*Melampsora lini*) and its two parent strains. One primer out of 160 tested was unusual in that it amplified a product from F_1 DNA that was not amplified from either parental DNAs. The nonparental band was only generated from DNAs of F_2 individuals that were heterozygous for these two allelic sequences. Sequence analysis of the two RAPD alleles demonstrated greater than 99% sequence identity, although the larger allele possessed an additional 38 base pairs (bp) relative to the smaller. Mixing of the two allelic sequences followed by denaturation and annealing in the absence of polymerase activity resulted in the formation of the nonparental band. Thus nonparental band present in some RAPD reactions consists of a heteroduplex molecule formed between two allelic sequences of different size, and are potential sources of errors during RAPD analyses. However, while this class of molecular markers can confound RAPD analyses, they offer a source of co-dominant RAPD markers (Davis *et al*. 1995), which are of value in genetic relatedness estimates and as markers for studying breeding behaviour (Novy and Vorsa 1996).

2.2.4.3 Inheritance of RAPDs

A key feature of RAPD is the generation of multiple polymorphic fragments per primer. While the majority of these bands segregate as expected, a small percentage exhibit distorted segregation patterns (Boscherini *et al.* 1994; Tanhuanpaa *et al.* 1994; Carlson *et al.* 1991; Nelson *et al.* 1994; Byrne *et al.* 1995; Bucci and Menozzi 1995; Rajapakse *et al.* 1995; Tulsieram *et al.* 1992; Hunt and Page 1995; Binelli and Bucci 1994; Rowland and Levi 1994; Lanaud *et al.* 1995; Cai *et al.* 1994; Bradshaw *et al.* 1994; Kubisiak *et al.* 1995; Grattapaglia and Sederoff 1994). Mechanisms capable of producing such distributions include chromosome loss, presence of an allele for pollen lethality, conflict between genetic isolating mechanisms evolved in the parental species, expression of genetic load via a lethal recessive allele (Bradshaw and Stettler 1994), and unconscious artificial selection for vigour loci during propagation of segregating population (Woeste *et al*. 1996). Cai *et al.* (1994) suggested the occurrence of distorted segregation might be indicative of the fact that RAPD marker loci are not all neutral. Both coding and noncoding sequences can be under positive or negative selection pressure (Landry *et al.* 1992).

There is a lack of consensus among researchers as to whether markers with aberrant segregation ratios should or should not be excluded from linkage analysis. While Tulsieram *et al.* (1992) and Woeste *et al*. (1996) have had to exclude such markers, others did not preclude their use in mapping studies (Byrne *et al.* 1995; Kubisiak *et al.* 1995; Cai *et al.* 1994; Bradshaw *et al.* 1994; Grattapaglia and Sederoff 1994).

Multiple RAPDs produced by the same primers generally assort independently within a family (Byrne *et al.* 1995; Hemmat *et al.* 1994; Carlson *et al.* 1991). The independence of these loci is an important aspect of the use of RAPD loci in the identification of individuals or diversity studies (Byrne *et al.* 1995). However some fragments have been shown to demonstrate linkage or even tight clustering (Skroch and Nienhuis 1995; Vandezande and Bijlsma 1995; Hemmat *et al.* 1994; Rowland Levi 1994).

2.2.4.4 Efficiency of genetic mapping with RAPDs

Several authors (Lu *et al*. 1995; Ronning *et al*. 1995; Jermstad *et a*l. 1994; Boscherini *et al*. 1994; Heun and Helentjaris 1993; Carlson *et al*. 1991) have investigated the segregation pattern and inheritance of RAPD markers. RAPDs are dominant markers and therefore less informative at a locus than some other types of markers; however, in conifers, this problem is avoided by use of DNA from haploid megagametophyte tissue. The conifer megagametophyte is derived from the single haploid megaspore that also produces the maternal gamete. RAPD analysis of DNA from megagametophytes derived from a single tree show 1:1 segregation and are equivalent to test-cross (Lu *et al.* 1995; Boscherini *et al.* 1994; Jermstad *et al.* 1994; Bucci and Menozzi 1993; Tulsieram *et al.* 1992). Thus, RAPD markers in megagametophyte tissues avoid many of the problems normally associated with quantitative trait dissection in trees, particularly the long generation intervals and absence of suitable pedigrees for gene mapping.

The dominant nature of RAPD polymorphisms is one of the limitations to use of the technique in forest tree genetic research and improvement (Smith and Williams 1994; Clark and Lanigan 1993). Since most RAPD markers are dominant, rather than co-dominant, genetic maps are less accurate in F_2 populations and more accurate in haploid, backcross and recombinant inbred populations (Smith and Williams 1994). Reiter *et al.* (1992) observed that the use of RAPD markers in a backcross progeny, haploid megagametophytes or recombinant inbreds sufficient to allow recombination fractions of $r \leq 0.2$ renders an efficiency equivalent to the use of F_2 co-dominant markers. Williams *et al.* (1993) have further shown that in a multipoint RAPD linkage map where the true mean interval between randomly-placed markers is 20% recombination, the standard deviation of marker interval estimate is 1.4% recombination in a backcross population and 2.2% recombination in an F_2 population. Therefore, in radiata pine where haploid DNA can be extracted from

the maternal megagametophyte, RAPD analysis provides full information on heterozygote and homozygote individuals for efficient linkage mapping.

The main limitation of RAPDs for applications in breeding is its repeatability across pedigrees and ploidy levels. There is a lack of agreement among scientists as to whether RAPD markers can be easily scored across pedigrees. Studies in maize showed that half of the bands of apparently the same molecular weight amplified by the same primer mapped to different locations in two populations (Johns 1992). DNA products of similar molecular weight amplified from different species of the same genus can have different sequences (Vandezande and Bijlsma 1995; Thormann and Osborn 1992). This can occur even within a single species.

There is also a question as to whether individual RAPD markers can be easily scored across ploidy levels. The dominant character of RAPD variation limits the inference of the parental genotype when data are available for only diploid material (Roy *et al.* 1992). When a fragment is detected in both parents and in all of the full-sib progeny, it is not possible to distinguish between a dominant homozygote (AA) and a dominant heterozygote (Aa) in either of the two parents involved in the cross. The problem can be resolved by assaying the marker in haploid tissues of the individuals parents, but only if individual RAPD markers can be easily scored across ploidy levels. Lu *et al.* (1995) found in 38 comparisons that parental genotypes of *Pinus sylvestris* (L.) inferred from diploid material conformed to those inferred from analysis of haploid megagametophytes, except for three cases. Plomion *et al.* (1995b) compared RAPD maps constructed for maritime pine (*Pinus pinaster*) using DNA from selfed and open-pollinated seeds of the same individual tree. The authors found RAPD markers to be repeatable and order of markers on framework maps robust, except for 2% of the markers. However, inconsistencies were found by Devey *et al.* (1996) in amplification products in haploid and diploid DNA samples. Some RAPD loci which should have been segregating based on megagametophyte assays either were not present or were monomorphic in

diploid needle DNA. RAPD fragments segregating 1:1 were also observed in needle but not in megagametophyte DNA samples. Similar inconsistencies were also reported by Hunt and Page (1992) in the honey bee, and Pooler and Scorza (1995) in peach. Ploidy level variation in RAPD amplification products may be due to competition effects of RAPD binding sites, which are 2-fold greater in diploid than in haploid DNA (Devey *et al.* 1996; Hunt and Page 1992).

2.3 Approaches to QTL analysis in forest and tree crop species

Various methods and analytical techniques have been used to facilitate the association of markers with QTLs. In forest and tree crop species, these fall into two basic categories: bulked segregant analysis and the use of a genetic linkage map.

2.3.1 Bulked segregant analysis

Bulked segregant analysis (Michelmore *et al.* 1991) is also referred to as the trait-based analysis (Lander and Botstein 1989; Lebowitz *et al.* 1987; Stuber *et al.* 1980) and selective DNA pooling (Darvasi and Soller 1994; Wang and Paterson 1994; Tinker *et al.* 1994). The method is based on the principle of marker allele frequency difference at the extremes of phenotypic distributions as a result of linkage to the trait. In the procedure, two DNA pools are screened for marker allele frequency changes, one pool contains individuals with high phenotypic value for the trait of interest, and the other individuals with low phenotypic value. Screening of the entire population is only required when polymorphisms between the two bulks are detected, resulting in a considerable saving in time and resources, particularly when used with RAPDs.

The method has been more widely applied to annual crops than to trees. It is most useful when the analysis is aimed primarily at a single trait. In trees, the method has been used to detect markers linked to genes such as disease-resistance in sugar pine, loblolly pine, apples and Chinese elm (Wilcox *et al*. 1996; Devey *et al.* 1995; Yuang and Kruger 1994; Benet *et al.* 1995), sex determination in *Pistacia* (Hormaza *et al.* 1994), dwarfing genes in citrus

(Cheng and Roose 1995), the *pendula* gene in Norway spruce (Lehner *et al.* 1995), fruit skin colour in apple (Cheng *et al*. 1996), and genes controlling agriculturally important traits in *Persea americana* (Mhameed *et al.* 1995) and peach (Warburton *et al*. 1996).

2.3.2 Genomic mapping

A requirement for genome-wide QTL analysis is the construction of a moderate to high-density linkage map. Genetic linkage maps aid genetic study of forest trees by a greater understanding of the relationship between chromosome structure, gene function and the genetical architecture of these characteristics. Genetic linkage maps also have an important impact on tree improvement since they afford the opportunity for directed manipulation of specific genes through marker-assisted selection. Properly used, such methods have the potential to improve the efficiency of early selection, thus reducing the time necessary to complete a breeding cycle, and increasing genetic gain per time.

The first linkage map in pines was generated using isozymes extracted from megagametophytes, and contained about 12 loci (Adams and Joly 1980). Conkle (1981) later located more loci but the genome coverage was still too low for use of the linkage map in genetic studies. Using two-dimensional polyacrylamide-gel-electrophoresis (2D-PAGE), Bahrman and Damerval (1989) mapped 119 protein markers in *P. Pinaster* Ait. into 12 linkage groups, with a total genetic distance of 420 cM. Gerber *et al*. (1993) also used two-dimensional electrophoresis to generate a 65 loci linkage map of the maritime pine with a 530-cM coverage of the genome.

The development of comprehensive genetic linkage maps of forest trees has only recently become widespread. This is primarily due to research on use of various mapping populations and the large number of polymorphic DNA markers that can be obtained in a short time in a single mapping population.

2.3.3 Available mapping populations

The preferred materials for linkage map construction are individuals of an F2 generation derived from F1 hybrids and backcross populations (Young 1994), recombinant inbred (Burr and Burr 1991) and doubled haploid lines (Heun *et al*. 1991). For most forest species, populations such as these are difficult to develop due to genetic load, generation time and pollen incompatibility. Thus, researchers have had to use:

(i) A three-generation pedigree of the target species. Devey *et al.* (1994) used a three-generation pedigree consisting of four grand-parents, two parents, and 95 progeny to develop a linkage map of loblolly pine (*Pinus taeda* L.), comprising 73 RFLP, and two isozyme loci. There were 20 linkage groups in the map, with 17 RFLP and 4 isozyme loci unlinked. Byrne *et al.* (1995) used a similar 3-generation out-bred pedigree of *Eucalyptus nitens* to develop a linkage map of 12 linkage groups from 210 RFLPs, 125 RAPDs and 4 isozymes. In poplar (*Populus* spp.), Bradshaw *et al.* (1994) also used a 3-generation pedigree, with the F1 produced by interspecific hybridisation between a *P. trichocarpa* female and a *P. deltoides* male. The mapping population consisted of 26-90 F_2 trees, generated by inbreeding to the maximum degree permitted by the dioecious mating system of *Populus*.

(ii) Use of an interspecific backcross. In walnut (*Juglans* spp.), Fjellstrom and Parfitt (1994) used 63 backcross progeny of an interspecific cross of *J. hindsii* x *J. regia* to establish a linkage map with 48 RFLP loci. Twelve linkage groups were identified by 42 of the RFLP loci, spanning 302 cM of walnut genome. The linkage map was extended to 107 loci using RAPD markers (Woeste *et al*. 1996) and 15 linkage groups. Durham *et al.* (1992) and later Cai *et al.* (1994) reported the construction of a linkage map in *Citrus* using 60 BC_1 progeny from an intergeneric cross of *C. grandis* x *C. grandis* x *C. poncirus*. The map consisted of 109 RAPD markers and 51 RFLPs, and covered 70-80% of the citrus genome.

(iii) F2 families. Chaparro *et al.* (1994) used nine different F_2 families obtained by selfing F1 trees derived from various parental crosses of peach (*Prunus persica* (L.) Batsch). The genomic map comprised 15 linkage groups, made up of 83 RAPDs, one isozyme, and four morphological markers.

(iv) A pseudo testcross strategy. A pseudo-test cross strategy has been used to avoid loss of information caused by dominance of RAPD markers. In a cross between two heterozygous individuals, many single-dose RAPD markers will be heterozygous in one parent, null in the other and therefore segregate 1:1 in their F_1 progeny. Two separate data sets are obtained one for each parent. In the pseudo-test cross configuration, markers are present in one parent and absent in the other or vice versa, and are expected to segregate 1:1 in the F1 generation. Grattapaglia and Sederoff (1994) were the first to use the approach to generate genetic linkage maps of *Eucalyptus* using RAPD markers. The map for the maternal parent (*E. grandis*) had a total of 240 markers grouped into 14 linkage groups (1552 cM), while the paternal *E. urophylla* map had 251 markers in 11 linkage groups (1101 cM). A longleaf pine (*Pinus palustris* Mill.) x slash pine (*Pinus elliottii* Englm.) interspecific F1 family was used by Kubisiak *et al.* (1995) to construct similar parent-specific linkage maps. Hemmat *et al.* (1994) also reported the construction of similar parent-specific linkage maps of apple (*Malus* x *domestica* Borkh.). In *Theobroma cacao*, Lanaud *et al.* (1995) combined marker information from two heterozygous parents of a full-sib cross into a single linkage map, using the linkage software, JOINMAP (Stam 1993).

(v) Use of the haploid megagametophyte DNA of conifer species. Conifer megagametophytes have been used in conjunction with RAPDs to develop single-tree genetic linkage maps in slash pine (*Pinus elliotii*) (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; 1995b), Loblolly pine (Wilcox *et al*. 1996) and *Pinus sylvestris* (Yazdani *et al*. 1995).

2.4 Marker analysis of quantitative traits

Tree growth and development are under the control of many important genes. Over time, scientists have identified several of these genes, but many still remain undetected and poorly characterised. Early genetic studies focused on characters that could be described simply, such as dwarfs in citrus and the *pendula* character in Norway spruce. Usually this meant that only single locus genes with major qualitative effect were accessible to study. However, variation in quantitative traits, such as stem growth efficiency, is considered to be the result of cumulative, integrated physiological processes that involve many genes, each with a relatively small contribution to the phenotype. Until quite recently, biometrical genetics has been most commonly used to analyse such quantitative variation.

Payne (1918) was the first to try, using marker genes, to understand the genetic nature of quantitative variation among *Drosophila* strains accumulated as the result of artificial selection. The Mendelian analysis of QTL gained wide currency in the 1920s and 1930s, but was later superseded by biometrical genetics. However, in the 1960s, QTL analysis regained ascendancy (Spickett and Thoday 1966; Thoday 1961).

Serebrovsky (1970, 1936) was one of the first to begin applying mathematical methods in this area. The trend has continued because of the abundance of molecular markers and the availability of detailed genetic linkage maps. QTL analysis involves the association of marker allele variation with phenotypic variation for the character under study. Normally, it is not possible to determine whether the effect detected with a marker locus is due to one or more linked genes affecting the trait. For this reason, the term quantitative trait locus (or QTL) was coined to describe a region of a chromosome (usually defined by linkage to a marker) that has significant effect on a quantitative trait (Tanksley 1993).

The theoretical basis for interpreting the association of marker loci with QTL has been outlined by Mather and Jinks (1982), Tanksley *et al.* (1982), Soller and Beckman (1983), and Edwards *et al.* (1987). The theory exploits the fact that the marker locus serves to identify, or 'mark', the chromosomal region in its vicinity, and enables that region to be followed in inheritance studies (Stuber *et al.* 1987).

There are now several statistical procedures for determining whether a QTL is linked to a marker. The simplest approach is to partition the population into different genotypic classes based on genotypes at the marker locus, and then to use correlative statistics to determine whether the individuals of one genotype differ significantly compared with individuals of other genotype(s) with respect to the trait being measured. This approach is generally referred to as single-point analysis, and does not require a complete genetic linkage map. However, the further a QTL is from the marker, the less likely it is to be detected using single-point approach due to crossover events between the marker and the QTL that result in miscalculation (Tanksley 1993). Furthermore, the magnitude of the effect of any detected QTL will normally be underestimated, due to recombination between marker locus and QTL.

Both problems are resolved when a large number of segregating markers is used along with interval mapping procedures (Lander and Botstein 1989). Instead of analysing the population one marker at a time, a set of linked markers are analysed simultaneously with regard to their effects on the target trait. By using linked markers for analysis, it is possible to compensate for recombination between the markers and the QTL, increasing the probability of statistically detecting the QTL and also providing an unbiased estimate of the QTL effect on the character (Tanksley 1993).

Single-point and interval mapping methods have been successfully used for several quantitative trait mapping studies in forest trees including *Eucalyptus* (Grattapaglia *et al.* 1995; Byrne *et al.* 1997; Chaparro *et al.* 1995; Vaillancourt

et al. 1995), maritime pine (Plomion *et al*. 1996), Poplar (Bradshaw *et al.* 1995; Han *et al.* 1994) and loblolly pine (Groover *et al*. 1994).

The results now permit detailed examination of the fundamental assumptions of the quantitative model as applicable to forest tree breeding, including: (i) approximately equal effects of individual polygenes, (ii) independent assortment of polygenes, and (iii) minimal epistasis (Tanksley 1993).

2.4.1 Magnitude of QTL effects

Several QTL mapping studies with forest trees have shown that different QTLs are unequal in the magnitude of effect they exact on a character. Grattapaglia *et al.* (1995) used a pseudo testcross strategy and RAPD markers to detect QTL in *Eucalyptus* for micropropagation response (10 QTLs), stump sprouting ability (6 QTLs), and rooting ability (4 QTLs). Individual QTL explained 7 to 21% of the phenotypic variance, while multipoint estimates of the total genetic variation explained by the QTLs ranged from 63 to 89%. This indicated a large proportion of the variation in these traits is controlled by a relatively small number of major-effect QTL (Grattapaglia *et al.* 1995). Bradshaw *et al.* (1995) also found that two QTLs explained 45% of the genetic variation in *Populus* for stem volume after 2 years of growth. Han *et al.* (1994) reported adventitious root and shoot regeneration in *Populus* are controlled by major genes, with effects ranging from 35 to 51% of the genetic variance. In papaya (*Carica papaya* L.), Sondur *et al.* (1995), used a RAPD-based linkage map to locate and characterise QTLs affecting growth rate and flowering. The amount of variation explained by the respective QTLs were 22 to 77% and 14 to 58% of the phenotypic variances. In loblolly pine, five genomic regions were identified which showed strong QTL-associated effects on wood specific gravity (WSG) (Groover *et al.* 1994). Together, the effects of the five regions explained 23% of the total phenotypic variance for WSG.

These studies would suggest a tendency towards detection of QTLs with large effects on the phenotype. As demonstrated by Bradshaw *et al*. (1995), growth

and adaptive characters may not necessarily be polygenic. The generally continuos distribution of phenotypes manifested by such characters may be due to genetic control by a large number of unlinked genes each with a small effect on the character, as assumed in the quantitative 'polygenic' model. Alternatively, these traits may be controlled by a few loci with relatively large effects (the 'oligogenic' model), with non-genetic (environmental) variation masking any discrete phenotypes (Allard 1960). For practical breeding purposes these are likely to be the QTLs of greatest interest to manipulate through association with molecular markers.

2.4.2 QTL clustering and pleiotropic effects

QTL mapping results suggest that polygenes do not assort independently, but tend to cluster in short segments of the genome, suggesting either a tight linkage of different genes or pleiotropic control of the same set of genes. In *Eucalyptus*, coincidence of genomic regions associated with micropropagation response and stump sprouting ability were reported by Grattapaglia *et al.* (1994) and Bradshaw and Grattapaglia (1994). Both traits involve multiple shoot formation from dormant buds, differing in the fact that in micropropagation response, shoot formation is stimulated *in vitro* with the action of cytokinin, while in stump sprouting, it relies exclusively on the intrinsic physiological ability to break dormancy of resting juvenile buds. These traits were not phenotypically correlated in the experiment (Grattapaglia *et al*. 1995; Bradshaw and Grattapaglia 1994). In *Populus*, QTLs associated with stem basal area were found to be flanked by QTLs for sylleptic branch total area, sylleptic branch number and sylleptic branch leaf number.

Evidence of QTL clustering and pleiotropic effects can be invaluable in plant breeding programs. For instance, information regarding the frequency of coupling vs. repulsion relationships can be used to eliminate, or at least significantly mitigate the linkage of desirable genes to undesirable genes, a common problem in plant breeding (Zeven *et al*. 1983). Coincidence of QTLs also gives better indications of genetic correlations (Lebreton *et al.* 1995), and

assist in predicting the genetic consequences of selection. Furthermore, sharing of QTL regions can provide indirect biological validation for the QTLs (Lebreton *et al.* 1995; Grattapaglia *et al.* 1994).

2.4.3 Epistasis

The ability to map and characterise QTLs with molecular markers opens up possibilities for detecting and characterising interactions among specific QTLs. Epistasis is the form of interaction between nonallelic genes whereby one gene interferes with the phenotypic expression of another gene (Bateson 1907). One simple statistical method for determining two-locus interaction of QTLs is through the use of two-way analysis of variance (ANOVA) (Tanksley 1993). Once the significant QTL affecting a trait have been detected, any two significant, unlinked markers associated with the QTLs can be used as independent variable in a two-way ANOVA (Tanksley 1993). The 'interaction' factor calculated in the two-way ANOVA will be as estimate of the interaction between the two QTLs in determining the phenotype.

Despite simplicity of the statistical tests, few studies with forest trees have reported testing for epistasis, and even fewer have reported its occurrence. In loblolly pine, evidence for QTL x environment interaction was found for two of the five detected QTLs. While digenic interaction was not detected among QTLs, multiallelism was demonstrated at a QTL that showed significant intra-locus interaction.

2.5 Marker-assisted selection (MAS)

Potential applications of MAS in forest tree breeding has been reviewed in several papers (Weeden *et al.* 1994; Forrest 1994; Grattapaglia *et al.* 1993; Dudley 1993; Wagner 1992; Neale *et al.* 1992; Neale and Williams 1991; Williams and Neale 1992; Nance et al. 1992; Cheliak and Rogers 1990; Nance and Nelson 1989; Bartels *et al*. 1989; Riemenschneider *et al.* 1987).

Molecular markers could be used to increase accuracy of within family selection by choosing offspring that have favourable QTL genotypes and a superior phenotype for breeding in the next generation (O'Malley and McKeand 1994). Strauss *et al.* (1992) used Lande and Thompson's (1990) expression for efficiency to compare between- and within-family selection with and without markers in plantation species. Under a variety of within- and between-family heritabilities, MAS was found to bolster phenotypic selection by more than 50% when the heritabilities are very low, and the markers explain at least 50% of the additive variance. Dekoning and Weller (1994) used computer simulations to compare selection on known QTL with phenotypic selection indices for a single and a two-trait selection objective. The efficiency of MAS relative to phenotypic selection was higher for the two-trait selection, than was selection on a single trait. The advantage of the marker-assisted selection was found to be greater when the traits were negatively correlated.

However, the practical use of polymorphic genetic markers to facilitate selection has been studied more intensely in agronomic crops than in trees. Marker-assisted selections have been used in two ways. Firstly, where the marker allele is itself desirable and can be selected directly. Two examples of this are the avoidance of haze in beer by using proanthocyanidin-free barley mutants (von Wettstein *et al*. 1985; 1977) and the improvement of bread-making quality in wheat by selection of particular high-molecular-weight glutenin subunits (Payne 1987). Secondly, where a marker, which may itself be phenotypically neutral, is tightly linked to an allele with an important agronomic effect. This second strategy is of greatest use where the target allele is difficult or expensive to detect in single plants or requires large-scale tests. Two examples, which are being used in breeding programs, are the linkage of an endopeptidase isozyme to eyespot resistance in wheat (Worland *et al*. 1988) and the linkage of esterase isozymes to barley yellow-mosaic virus resistance in barley (Koishi *et al*. 1989). In maize (*Zea mays* L.), gains attained through marker-facilitated selection are detailed in Stuber (1994a, 1994b, 1995). Kata *et*

al. (1994) also reported the use of molecular markers to assist introgression of the opaque-2 mutation in a maize protein-quality improvement program. The opaque-2 mutation conditions are highly valuable because of a reduction in the amount of zein seed storage protein. Because the condition is recessive and endosperm-specific, conventional backcross procedures to convert elite inbreds to opaque-2 are inefficient. A restriction fragment length polymorphism (RFLP) marker assay was then utilised to accurately identify opaque-2 genotypes of individual juvenile plants in breeding populations.

In common bean (*Phaseolus vulgaris* L.), Haley *et al.* (1994) investigated the efficiency of marker-assisted selection (MAS) with RAPD markers linked in coupling and repulsion with a single pest resistance allele. The authors found repulsion phase linkages provided greater selection efficiency, even when the markers have greater linkage distances from the resistance allele. Monforte *et al*. (1996) have also reported the use of marker-assisted selection to develop salt-tolerant breeding lines in a cross of the cultivated tomato (*Lycopersicon esculentum*) with a wild relative (*L. pimpinellifolium*).

Features of forest trees, such as the long generation times, high levels of outcrossing and absence of defined inbred lines appear to be formidable barriers to genomic mapping and MAS. In addition, most conifers have large genome sizes; typically ten-fold larger than angiosperms and 200 times larger than the model plant *Arabidopsis thaliana* (Neale and Sederoff 1996). However, the high levels of heterozygosity in conifers may provide large numbers of DNA markers, and is now being exploited in various laboratories around the world.

Reports of successful use of MAS in tree breeding are now becoming available. In apple, Gianfranceschi *et al*. (1996) reported the use of markers to assist in selection for resistance to the scab disease caused by *Venturia inaequalis*. These results show conformity with theoretical expectation, and increase enthusiasm about the potential of MAS as a breeding tool.

2.5.1 QTL repeatability between populations.

Marker-linked QTLs are statistically defined entities and have the potential to be artifacts. The validity of such linkages can be verified with a different cohort of progeny within the same mapping pedigree. However, the potential benefits from MAS in forest breeding programs will depend on:

(i) Whether a marker allele accurately identified as being associated with a favourable QTL allele in one segregating population can be used to select for that favourable QTL allele in different segregating populations, and

(ii) Whether a marker allele associated with a favourable QTL allele at the seedling stage of growth will be associated with a favourable QTL allele for a mature characteristic.

The ability to repeatably detect marker-linked across segregating populations will depend on (a) the presence of segregation of the marker in the populations. A given QTL may be polymorphic in one population (ie. the two parents carry the different alleles), but monomorphic in another population (ie. the two parents carry the same allele).

It will also depend on (b) whether the linkage phase (coupling or repulsion) between the marker and the QTL alleles is the same in cross populations, and (c) genotypes of other QTLs if epistasis is present (Dudley 1993). If both the marker and the QTL are heterozygous and in the same linkage phase in the second population as in the first and epistasis is unimportant, then associations identified in one population may be used for selection in the second. Although alleles at a locus in forest trees would be expected to be heterozygous, the occurrence of multiple alleles (Groover *et al*. 1994; Byrne *et al*. 1996) suggests that very few QTLs will be common across populations (Beavis *et al*. 1991). In a study involving 10 RAPD markers closely linked to a gene for resistance to white pine blister rust in sugar pine (*Pinus lambertiana*), Devey *et al.* (1995) observed that half of the loci were easily visualised across five different families from ethidium bromide-stained agarose gels. The other five loci required more

sensitive methods of detection. However, when the gene (s) controlling the trait are present in a genomic region not subject to frequent recombination, marker-QTL linkages may be repeatable across pedigrees. Hormaza *et al.* (1994) found that a RAPD marker linked to sex-determination in *Pistacia vera* was conserved across two segregating populations, and in 14 cultivars unrelated to the crosses.

The lack of repeatability across populations is not limited to RAPD markers. Lin *et al*. (1996) found that the majority of RFLP-linked QTL for yield in maize detected in one population were undetectable in another population. Beavis *et al*. (1991) reported results on plant height for four maize populations and noted that no RFLP-linked QTL could be consistently identified in all four populations. Given the apparent lack of repeatability across pedigrees, a more practical alternative to QTL validation may be to verify marker-QTL associations using an independent, unrelated population. This would be appealing from a breeding point of view since the presence of different QTLs in different populations provide an opportunity to combine them in new breeding lines with enhanced expression of the target trait and higher performance potential.

CHAPTER 3

3 Structure of the thesis

Significant associations between molecular markers and QTLs occur when the markers are in regions of the chromosomes that carry the trait-causing genetic factors. To be detectable, the markers must be in linkage disequilibrium with the QTLs. To maximise linkage disequilibrium, researchers have commonly used F_1 backcross or F_2 populations from a cross of two inbred parents. Such populations are not available in most forest tree breeding programs, and will be difficult to obtain for many species due to high genetic load and long generation times (Bradshaw *et al*. 1995).

The use of a haploid population in conjunction with RAPD markers avoids many of the problems associated with genetic linkage mapping and QTL analysis. In haploid populations from a single coniferous tree, RAPDs show 1:1 segregation, equivalent to a testcross, and provide full information on heterozygote and homozygote individuals for linkage mapping.

The experiments reported in this thesis applied RAPD marker information derived from haploid progenies of two unrelated full-sib crosses to quantitative trait loci detection and mapping of the diploid individuals. Because QTL detection using haploid megagametophytes exploits the linkage disequilibrium between markers and QTL segregating only on the maternal side of the pedigree, the paternal contribution to the phenotype may confound the power of detection. The extent to which these limitations affect QTL detection and the possibility of marker-assisted selection for a quantitative trait has not previously been explored.

The sequence of experiments are summarised as follows:

(i) Detection of RAPD markers linked to factors underlying NESTUR by use of bulked segregant analysis (experiment 1). Bulked segregant analysis was applied to haploid megagametophyte DNA of a controlled cross (12038 x 10946) to identify linkage between RAPD markers and QTLs influencing NESTUR in radiata pine (chapter 4).

(ii) Construction of a RAPD-based linkage map (Experiment 2). A linkage map was developed using 93 haploid progeny of a full sib cross (30040 x 80121). Analysis of 267 markers resulted in 262 loci linked in 14 linkage groups covering 1511 cM (chapter 5).

(iii) Interval mapping of NESTUR (Experiment 3). The linkage map was used for QTL detection and mapping using single-point and interval mapping methods. Two methods of interval mapping were used, the conventional interval mapping of Lander and Botstein (1989), and the multiple-QTL (or composite interval) mapping method of Jansen and Stam (1994) (chapter 6).

(iv) Markers prediction of seedlings with superior growth potential (Experiment 4). Markers significantly ($P \leq 0.01$) linked to NESTUR were assessed for their ability to predict the phenotype and identification of individuals with superior growth potential. In addition, a comparison of the growth performance of seedlings with favourable and unfavourable NESTUR-linked markers was carried out using data from a two-year growth period. Seedlings were grouped on the basis of (i) high and low NESTUR values, and (ii) a high allele frequency of favourable and unfavourable RAPD markers linked to factors influencing NESTUR (chapter 7).

(v) Marker-based inferences on QTL stability (Experiment 5). The study compared QTL maps developed for the same individuals in two years of growth. The interaction of individual QTL with age was studied by comparing the frequency of identification of significant marker-linked QTLs at different growth stages. Such information can provide a basis for designing marker-assisted selection schemes that take into account significant QTLs affecting important traits that are stable across years of measurement (chapter 8).

Details of materials and methods used in the various experiments are presented in chapters 4 to 8, and a discussion of the specific results. A general discussion of the research highlights and conclusion are presented in chapter 9.

4 CHAPTER 4

Experiment 1: Detection of RAPD markers linked to NESTUR by use of bulked segregant analysis

4.1 Introduction

Identification of RAPD markers linked to quantitative trait loci requires a large number of individuals for statistical power (Darvasi *et al.* 1993) with consequent relatively high cost for marker genotyping. However, locating one or a few markers with large effects need not involve the expense and time required for a full genome mapping. If a single marker can be identified which categorically divides individuals into high and low NESTUR, the marker could be used as an indirect selection criterion to eliminate the tedious and time-consuming phenotyping. An approach that is accurate, fast, reliable, and cost-effective is bulked segregant analysis (Michelmore *et al.* 1991).

NESTUR has been observed to show high heritability between families of radiata pine (Matheson per. com.)[1]. Bulked segregant was applied to progenies of a cross between parent trees with high and low NESTUR means and within-family variances to determine whether RAPD markers could be detected associated with NESTUR.

4.2 Materials and methods

4.2.1 Plant material

In experiment 1, *Pinus radiata* seeds from a full-sib cross were obtained from CSIRO, Forestry & Forest Products. This particular cross was selected based on parental variances for NESTUR values from previous evaluations. Open pollinated (OP) seeds from the maternal parent (12038) had shown high

[1] A. C. Matheson, CSIRO Forestry & Forest Products, Banks Street, Yarralumla, Canberra ACT 2600 Australia. EMail: Colin. Matheson@cbfr.csiro.au

variance for NESTUR values, while OP seeds from the paternal parent (10946) had a low variance.

Approximately 300 seeds of the F_1 full-sibs were placed in a petri-dish lined with three layers of filter paper and stratified for 30 days at 4 ^{o}C. Only 176 seeds germinated, and these were picked out onto 150 mm diameter pots filled with a porous medium (3 parts by volume of river sand, 2 parts vermiculite, and 1 part perlite). Megagametophytic tissue (see Figure 4.1) was collected following germination without destroying the seedling and stored at -80 ^{o}C for subsequent DNA extraction, with seedling identity maintained.

The seedling trees were initially grown in a glasshouse under ambient temperature and light conditions for two weeks, before transfer to a controlled environment (Phytotron). The growth environment was maintained at 25 ^{o}C (maximum or day) and 18 ^{o}C (minimum or night) temperatures, and a relative humidity of 65%, from 20 April to 1 July 1994. Daylength was extended with artificial lighting from 5 pm to 8 am. Natural lighting in the Phytotron varied from photosynthetically active radiation (PAR) of 700 $\mu E\ s^{-1}\ m^{-2}$ in midwinter to 1500 $\mu E\ s^{-1}\ m^{-2}$ in mid summer. Nutrients (see appendix 4.1) were provided in the irrigation water through drippers (see Figure 4.2) at 12 hr intervals.

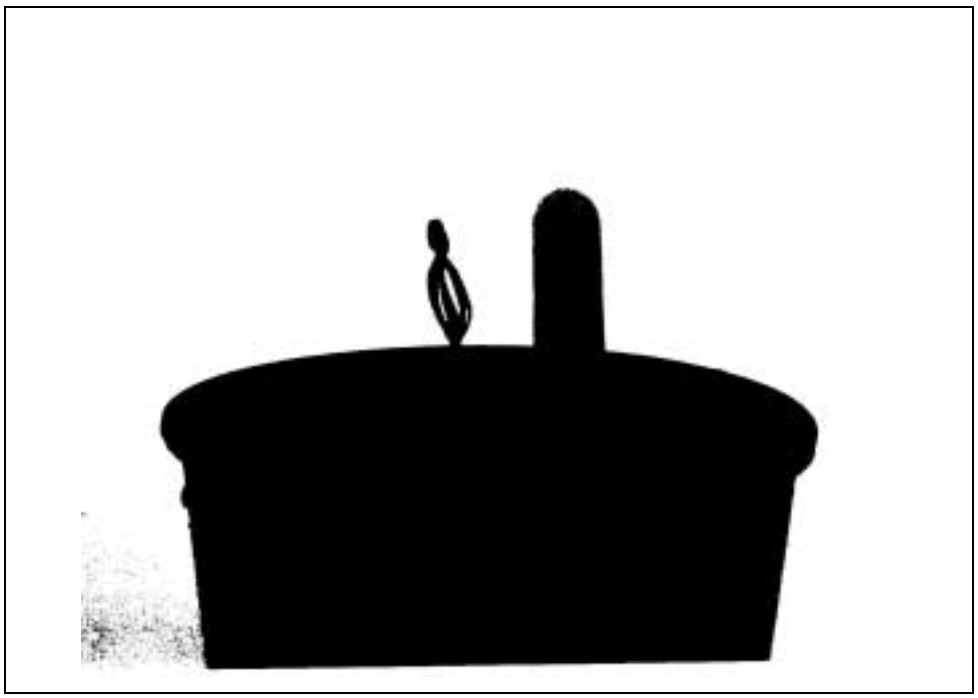
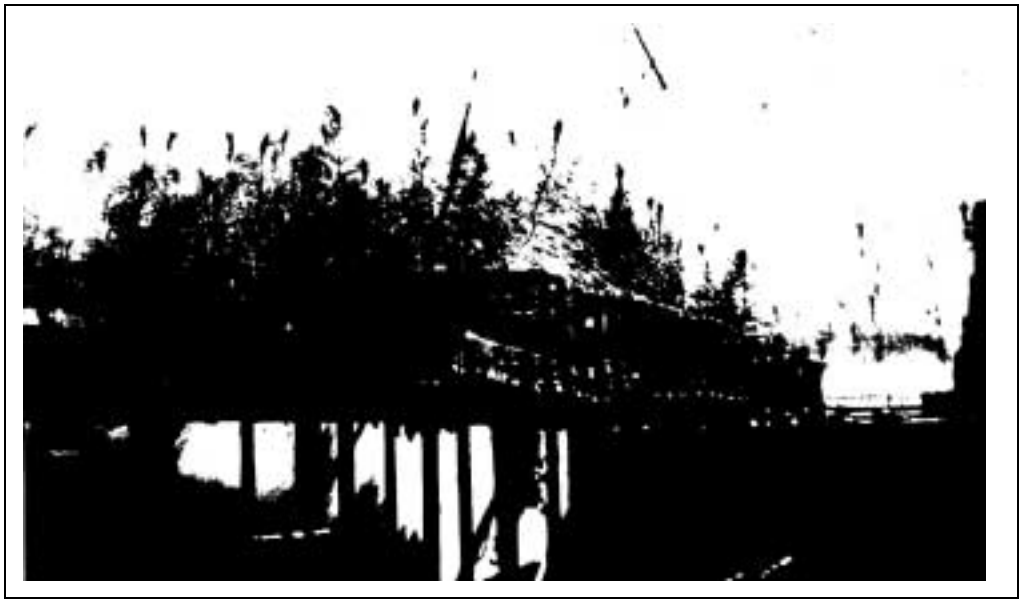


Figure 4.1 Radiata pine seedling 10 days after germination, showing the seed coat and megagametophyte tissue close to the point of being cast off. Tissues harvested at this stage provide cleaner DNA, with less contaminating polysaccharides that contribute to inconsistent PCR amplification (Grattapaglia *et al*. 1992).

Figure 4.2 Irrigation drippers used in seedling culture. Seedlings were grown under



controlled conditions in the phytotron.

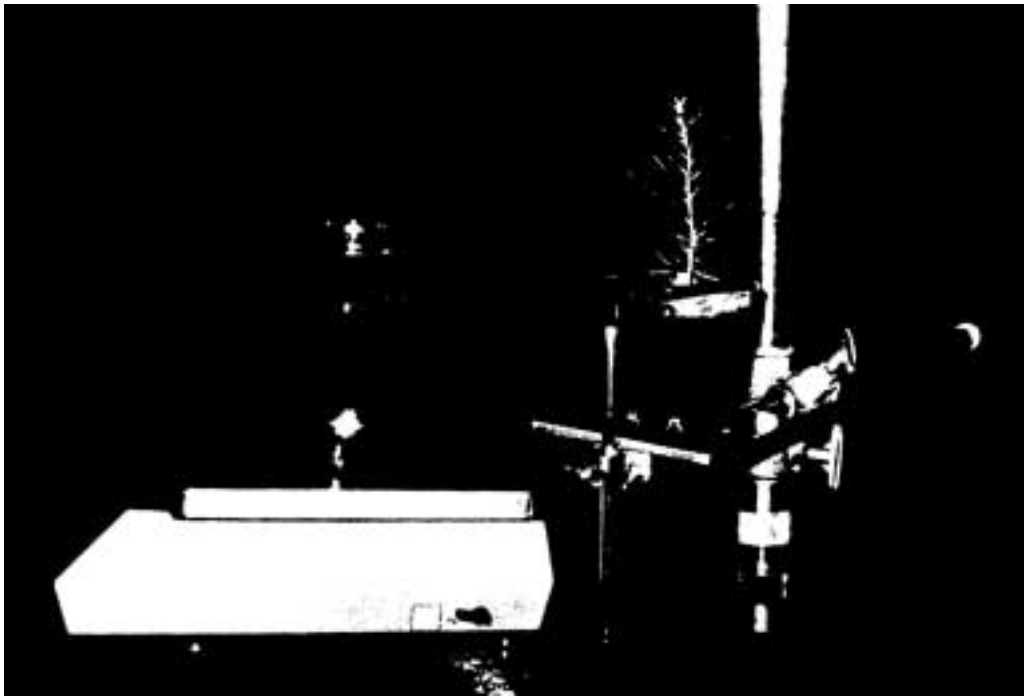
Seedlings were assessed for NESTUR after 12 and 14 weeks. Four separate measurements were made on each of the two dates: (i) total height, measured from the pot rim to the growing point, (ii) trunk height, from pot rim to the cotyledonary scar, (iii) stem diameter, using a digital dendrometer (0.0001 mm resolution) and (iv) needle volume. Crown-height was calculated as the difference between total height and trunk height. Needle volume was measured using a water displacement method. The pots were clamped firmly on a stand (Figure 4.3a), and carefully inverted with the seedling crowns positioned squarely above a cylinder of water (Figure 4.3b) resting on a top-loading balance (0.01g resolution). The crowns were carefully lowered into the water to a datum (the cotyledonary scar), and the volume of water displaced recorded. Care was taken to avoid contact of the foliage with the walls of the cylinder (Figure 4.3c). To prevent trapping of air bubbles, a wetting agent as added to the water as a surfactant at the rate of 3 mil per 15 litres.

NESTUR values were calculated for each individual using the formula:

$$\mathbf{N = [(SVol_{t2} - SVol_{t1})/(NVol_{t2} - NVol_{t1})] \times [(lnNVol_{t2} - lnNVol_{t1})/T_2 - T_1)],}$$

Where SVol = stem volume (mm^3), NVol = needle volume (mm^3),
T = time of trait measurement (15 days apart from commencement),
Stem volume = basal area x total height,
Needle volume = volume of water displaced minus the volume occupied by stemwood in crown-height.

There were two 'runt' individuals with extremely low NESTUR values, and the data for these were discarded, reducing sample size to 174.



A.

B.

C

Figure 4.3 Measurement of needle volume by water displacement. Seedlings were clamped firmly on a stand (A), and carefully inverted (B) with the seedling crowns positioned squarely above a cylinder of water + wetting agent, then the crowns were carefully lowered into the water (C) to a known point.

4.2.2 DNA extraction and RAPD assays

Total genomic DNA was extracted from megagametophytes according to Dellaporta *et al*. (1983). Average yield of DNA was approximately 2.5 μg DNA per megagametophyte. Reaction mixtures for RAPD assays consisted of 3 ng DNA from each bulk, 2.0 mM $MgCl_2$, 200 μM each dNTP, 1U of Taq polymerase, and

0.33 μM primer in a 10x reaction buffer containing 10 mM Tris-HCL (pH 8.3 at 25 ^{o}C) and 50 mM KCL, and for a total reaction mixture of 15 μl. Amplifications were carried out in a Perkin Elmer Cetus 9600 DNA Thermal Cycler at 3 cycles of 94 ^{o}C for 2 minutes, 37 ^{o}C for 2 minutes, and 72 ^{o}C for 2 minutes, followed by 40 cycles of 94 ^{o}C for 45 seconds, 42 ^{o}C for 45 seconds, and 72 ^{o}C for 120 seconds.

4.2.3 Electrophoresis and gel documentation

After amplification runs, 5 μl of loading buffer (10x TAE, 50% glycerol, 0.25% bromophenol blue) were added to each reaction product, which were then loaded onto horizontal gels of 1.5% agarose in 0.5x Tris-borate electrophoresis (TBE) buffer (Sambrook *et al*. 1989). Gels were powered using Bio-RAD Power/PAC 300 which maintained a constant voltage of 150, and runs lasted 150 minutes in Horizon 20.25 trays. DNA fragments were visualised by UV irradiation after staining with 0.5 μg/mil ethidium bromide. Gels were photographed with Polaroid film type 667 or the NOVALINE Gel Documentation System (Novex Australia Pty Ltd).

4.2.4 DNA pooling and RAPD assay

The present study was based on haploid segregation analyses of the maternal parent of the full-sib family 12038 x 10946, although the phenotype was measured on the diploid offspring. Bulked samples consisted of 250 ng DNA from the corresponding megagametophytes of 9 High-NESTUR and 9 Low-NESTUR individuals, representing 5% of the high and low tails of the normal distribution. These bulks were screened with 933 random 10-mer oligonucleotide primers. Of these, 800 Operon primers were supplied by Operon Technologies, Alameda, CA

and 133 primers from the "conifers RAPD set" from Dr. John Carlson, University of British Columbia.

Markers were scored as 2 for presence and 1 for absence. The approximate length of RAPD markers (in base pairs) were determined by comparing each band with size markers from 100-base pair ladder (GibcoBRL). Based on the migration distance of RAPD bands and size markers, the size of each marker was determined to the nearest 50 base pairs.

The procedure for establishing a relationship between the markers and NESTUR was as follows:

(i) The presence of a RAPD band in one bulk and its absence in the other suggested a linked marker;

(ii) To exclude false positives, the RAPD assay was repeated for each polymorphism detected. The presence or absence of each marker was assayed individually in each of the 18 seedlings, which had made up the bulk samples, and on the next extreme 30 individuals.

Thus, the presence or absence of each marker was known for the 39 individuals ranked as having the highest NESTUR values and the 39 individuals with the lowest NESTUR values. To check the relationship further, a simple analysis of variance (ANOVA) was used to relate the presence or absence of each marker to the NESTUR score. If the NESTUR score for the two groups were not significantly different at the 10% level of error probability, the marker was discarded.

(iii) Finally, the procedure was repeated for each of the retained markers using the complete population. NESTUR scores for trees with the marker were compared with those for the trees without the marker. A marker was considered associated with the trait if the probability levels associated with F-values were ≤ 0.01. The use of analysis of variance enabled the difference explained by each marker to be estimated.

Epistasis was also assessed by using a simple factorial ANOVA to compare the NESTUR values of individuals with two markers present, both absent, and those with only one marker present.

4.2.5 Relationship of NESTUR with stem growth variables

Coincidence of QTL for a complex trait with one or more of the trait's components suggests an indirect biological validation for these QTL. For related traits, coincidence of at least some QTL is likely. The relationship between NESTUR and increments in height, diameter, stem volume and needle volume was assessed by standard Spearman correlations (SPSS for Windows, SPSS Inc). Interaction terms were extracted from the following statistical model: $Y_{ijk} = \mu + M_i + M_j + MM_{ij} + E_{ijk}$, where Y_{ijk} is the observed phenotype, μ is the overall mean, M_i & M_j are the main effects of the i_{th} and j_{th} RAPD markers, and E_{ijk} is the residual effect.

4.3 Results

4.3.1 Frequency Distribution of NESTUR values

The distribution of phenotypes was continuous with a variation coefficient of 28.3%, but showed a skewness of 1.16 towards higher NESTUR values.

Therefore, a $\log_{10}$ transformation was applied to reduce skewness to 0.05. The frequency distribution of these values is shown in Figure 4.4. The progeny mean NESTUR value was 8.5, while the two groups of nine extreme phenotypes had average NESTUR values of 4.8 (for the low-NESTUR group) and 13.9 (for the high-NESTUR group).

Analyses of variance were carried out on the $\log_{10}$-transformed data since ANOVA assumes the underlying variable to be normally distributed. However, results are presented in Table 4.1 for both transformed and untransformed data.

4.3.2 Identification and genetics of marker-linked QTLs

Of the 933 primers screened, 23 produced fragments that were polymorphic between the two bulks of DNA. However, only 7 were found reliable when tested on the 18 extreme individuals. All seven marker-loci showed the expected 1:1 ratio of segregation (Table 4.1). When tested on the additional 60 extreme individuals the number was further reduced to four. The three markers that did not show sufficient evidence of association with NESTUR QTLs were excluded from further analysis.

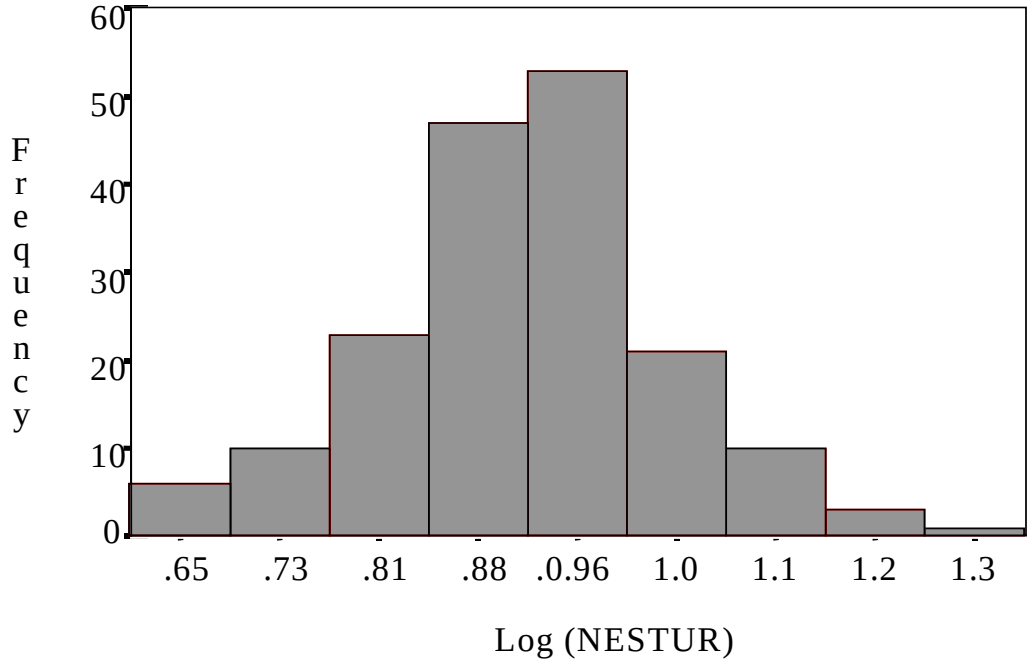


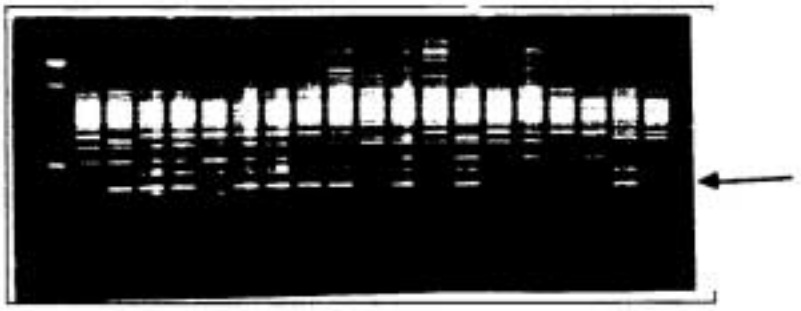
Figure 4.4 Frequency distribution of log_{10} transformed NESTUR values in 174 individuals of the full-sib cross of radiata pine (12038 x 10946).

In the second stage of the marker-trait analysis, the four markers that showed differences in genotype class means (Table 4.1) were tested on all 174 individuals. Only three were found to be significantly associated with NESTUR.

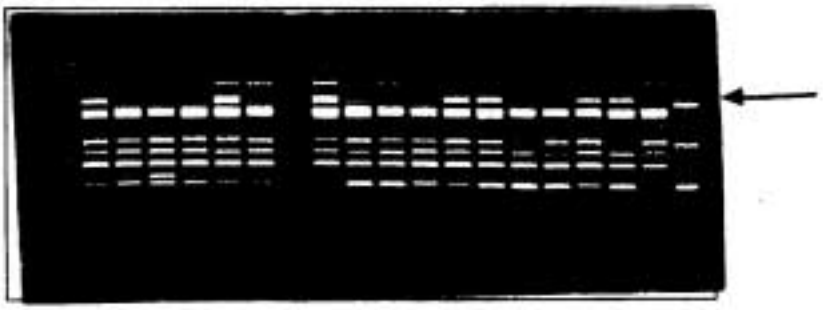
Table 4.1 Summary of RAPD loci in linkage with factors influencing NESTUR in radiata pine full-sib cross 12038 x 10946. Results are presented for log-transformed (trans.) and untransformed (untrans.) NESTUR values. R2 values were calculated as the ratio of sum of squares (markers: total).

		STAGE I Extreme phenotype progeny (n = 78)		STAGE II All progeny (n = 174)			
RAPD Locus†	χ^2 test for 1:1 ratio (P-value)	Trans. (P-value)	Untrans. (P-value)	Trans. (P-value)	Untrans. (P-value)	R^2 value (%)	Allele effect (SD units)
OPA-10_{1200}	0.100	0.004	0.005	0.001	0.001	10.56	-0.50
OPE-06_{450}	0.309	0.017	0.021	0.011	0.008	6.23	-0.41
UBC-333_{550}	0.346	0.005	0.007	0.002	0.003	9.23	0.46
OPAM-12_{600}	0.938	0.056	0.109	0.028	0.046	b	b
OPW-01_{750}	0.838	0.212	0.309	a	a	a	a
OPO-12_{650}	0.221	0.216	0.123	a	a	a	a
OPV-17_{1100}	0.307	0.350	0.319	a	a	a	a

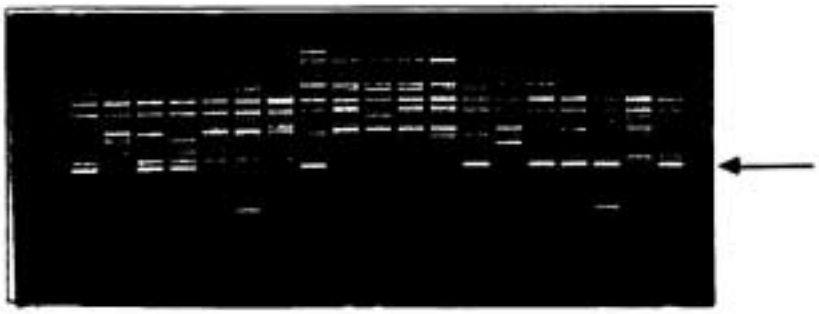
† Primer name/fragment length in base pairs.
Markers not considered further after the first stage of sampling shown by (a), and after the second stage by (b).



a. Primer: OPE-06



b: Primer: OPA-10



c: Primer: UBC-333

lucts using primers OPE-06, OPA-10 and UBC-333. Polymorphic RAPD bands linked with the quantitative trait NESTUR for each primer are indicated with an arrow. First lanes are the 100 bp ladder.

An absence of RAPD fragment markers generated by primers OPE-06 and OPA-10 (Figure 4.5) was associated with low NESTUR values, while primer UBC-333 generated a 550 bp band (Figure 4.5) that was associated with high NESTUR values.

Standardised allele substitution effects of marker-linked QTL were estimated as the slope value from simple linear regression. The marker loci, OPA-10_{1200} and OPE-06_{450}, had decreasing effects of -0.41 to -0.50 standard deviation units (SDU) from the family mean (Table 4.1), while the UBC-333_{550} locus had an increasing effect of 0.46 SDU. The amount of variation attributable to the effects of these markers ranged from 6.2 to 10.6%, while a 2-QTL model explained 20% of the phenotypic variation.

Pairwise comparisons among significant marker loci showed that the loci revealed by the Operon primers (OPE-06_{450} and OPA-10_{1200}) were linked to each other at a recombination frequency of 7% (SE_r = 0.02). The locus UBC-333_{550} was not linked to either OPE-06_{450} or OPA-$10_{1200.}$

4.3.3 QTL coincidence and pleiotropy

The relationship of NESTUR with stem growth traits is shown by the correlation coefficients (Table 4.2). The most significant of the relationships were those with increments in stem diameter (r = 0.82) and stem volume (r = 0.66).

Table 4.2 Summaries of (i) spearman rank correlations of NESTUR with stem growth traits, and (ii) associations of NESTUR-linked RAPDs with stem growth traits.

	Associations with NESTUR (r)	Associations with RAPD locus		
		OPA-10_{1200} (P-value)	OPE-06_{450} (P-value)	UBC-333_{550} (P-value)
Stem diameter growth	0.819	0.12	0.43	0.01
Stem Volume growth	0.662	0.09	0.34	0.004
Height growth	0.294	0.18	0.25	0.31
Needle Volume growth	-0.122	0.99	0.38	0.18

The two linked RAPD markers (OPE-06_{450} and OPA-10_{1200}) showed no significant ($P \leq 0.05$) linkage to any of these components, suggesting that this QTL may be specific to NESTUR, and independent of these components (Table 4.2). The RAPD marker generated by the primer UBC-333, however, was significantly linked to increments in stem diameter and stem volume, with effects similar to those observed for NESTUR; that is, a 0.46 standard deviation units (SDU) increase in NESTUR was associated with an increase of 0.46 and 0.38 SDU in stem volume and diameter growths, respectively.

4.3.4 Epistatic interactions

Analysis of variance was used to determine whether the action of one putative QTL might be dependent upon the allelic state at another locus. Using markers as treatments, it was found that allele at the locus near markers OPA-10_{1200} and OPE-06_{450} appeared to determine whether allelic variation at the QTL linked to UBC-333_{550} results in phenotypic variation (Table 4.3).

Table 4.3 P-values from 2-way analyses of digenic epistasis between putative QTLs, using markers as treatments. P-values are shown relative to NESTUR and stem growth traits.

2-Locus interaction	NESTUR	Stem diameter growth	Stem Volume growth	Height growth	Needle Volume growth
OPA-10_{1200} x OPE-06_{450}	0.4300	0.520	0.600	0.40	0.95
OPA-10_{1200} x UBC-333_{550}	0.0001	0.005	0.001	0.08	0.68
OPE-06_{450} x UBC-333_{550}	0.0001	0.002	0.001	0.12	0.70

The simplest method to graphically display such interactions is to merely display the 2-locus treatment means and demonstrate differences in slope. This is illustrated for NESTUR using all four possible 2-locus marker genotypic class means for UBC-333_{550} and OPA-10_{1200} (Figure 4.6). Presence of the UBC-333_{550} marker is associated with increased values of NESTUR. However, this was found to be dependent on the genotype at the OPA-10_{1200} locus, whose presence appeared to inhibit expression of the QTL linked to UBC-333_{550}.

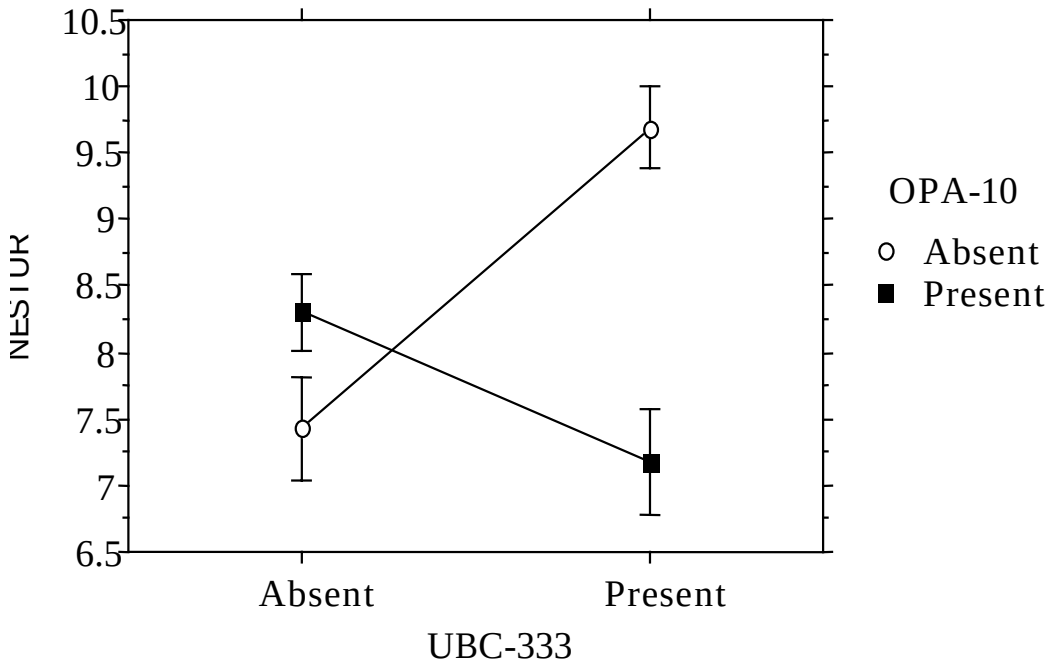


Figure 4.6 Illustration of epistasis between marker-linked QTL influencing NESTUR, using four possible genotypic class means of markers UBC-333$_{550}$ and OPA-10$_{1200}$.

4.4 Discussion

4.4.1 General considerations

Many studies have demonstrated the power of bulked segregant analysis for tagging QTLs. Most of these, however, considered simply-inherited loci altering morphology or disease reaction, where the influence of non-genetic factors can be greatly minimised by accurate phenotyping of the individuals (Devey *et al*. 1995; Lehner *et al.* 1995; Benet *et al.* 1995; Cheng and Roose 1995; Mhameed *et al.* 1995; Heinze and Geburek 1995; Yuang and Kruger 1994; Hormaza *et al.* 1994).

Using RAPD analysis in combination with bulked segregant analysis, it has been possible to identify two QTL significantly linked to factors controlling needle-to-stem unit growth rate (NESTUR) in radiata pine. The identified QTLs explained 20% of the phenotypic variation, with standardised additive effects

ranging from -0.41 to 0.50 standard deviation units (SDU). Regression analysis used in estimating standardised effects and the amount of explained variation assumes no recombination between the marker and the QTLs. Violation of this assumption would cause the QTL effects to be underestimated. Because markers are not QTL, but are linked to QTL, loose linkage between the marker and the QTL reduces the magnitude of explained variation inferred from single-point analysis.

It is pertinent to note that in the present study, only markers segregating from the maternal parent were assayed. Therefore, only some of the total number of QTLs will be potentially detectable, ie, those segregating on the female side of the pedigree. The methods used, however, have served to establish marker-QTL linkage for NESTUR, as a first step in a program aimed at mapping QTLs with respect to flanking markers on the chromosome.

The QTLs for NESTUR identified with marker UBC-333_{550} showed a significant association with stem diameter growth and stem volume growth (Table 4.2). The direction of additive gene effects of the QTL on these traits, and magnitudes of phenotypic correlations are suggestive of a pleiotropic gene action. However, the independent effects of tightly linked loci cannot be ruled out at this stage since the sampled progeny size is insufficient to distinguish between the two possibilities.

4.4.2 Limitations of bulked segregant analysis

The success of any QTL study is dependent on the influence of non-genetic factors. It is known that when a phenotype is controlled by multiple genetic loci and the heritability is low, individuals can have extreme values due to non-genetic factors. Stem growth is the cumulative result of carbon assimilation

and allocation over time. Since both assimilation and allocation are influenced by many biochemical pathways, physiological processes, and environmental variables, it is reasonable to expect that the integration of stem growth with these variables will be genetically complex.

Cross-contamination of DNA pools due to phenotyping errors blocks detection of polymorphism between pools (Wang and Paterson 1994), and thus reduces the power of bulked segregant analysis (Chaparro *et al*. 1994). The replication of phenotypic evaluations, eg, by use of clonal propagules, can mitigate the effects of non-genetic factors by increasing heritability of the trait (Bradshaw and Foster 1992), but clonal materials were not available for the present study, and bulking was based on unreplicated phenotypic information in order to avoid the confounding effects of seedling age.

A common problem in bulked segregant analysis is the large number of false positives detected from pooled DNA samples. In Norway spruce (*Picea abies* (L.) Karst. f. pendula), Lehner *et al.* (1995) detected 12 putatively linked loci among the two bulks of DNA constituted for the *pendula* gene, but only one was confirmed to be linked to the gene. For simply-inherited traits, false positives can be screened out by genotyping individuals that were used to make up the bulks. For quantitative traits, only markers tightly linked to the QTL will be detectable with this approach. In this study, a sequential sampling procedure was used, testing all reliable markers on an additional 60 extreme phenotype progeny. Markers that failed to show evidence of linkage ($P \leq 0.10$) at this stage were discarded. Subsequently, all progeny of the family were tested for markers that were retained, using a more stringent error rate of ≤ 0.01.

Potentially some QTL-linked markers, which might have been detected had all 174 progeny been genotyped, were missed by the use of bulked segregant analysis. By selecting individuals with the highest and lowest phenotypic values, some QTLs with small effects on the phenotype, (but which collectively explain a large portion of the genetic variation in a trait), probably escaped detection. However, it is unlikely that the sequential sampling reduced power of QTL detection since all markers identified from bulked DNA samples were given consideration. The sequential sampling procedure effected a major reduction in costs by allowing an early decision to discontinue scoring markers in QTL-negative chromosomal regions.

4.4.3 QTL validation

In conclusion, results from the bulked segregant analysis provide sufficient grounds for further analysis of the genetic basis of NESTUR. However, marker-linked QTLs are statistically defined entities and have the potential to be artifacts. Carson *et al*. (1995) advocated that the validity of such linkages should be verified with a different cohort of progeny within the same mapping pedigree. In the present study, limited seeds obtained from the controlled cross of 12038 x 10946 does not permit such an approach. As an alternative, additional the studies can be carried out to verify marker-QTL associations using an independent, unrelated population. Such an approach is more appealing in the present situation since the cumulative explained variation associated with detected RAPD markers was not large enough to permit a realistic judgement about the prospects of marker-assisted selection. Also, there was evidence of pleiotropy/linkage between NESTUR-linked QTL and those affecting stem volume and diameter growth. Limited genome coverage did not permit inferences on the genetic inter-dependence of NESTUR with

growth in stem volume, diameter, height and needle volume. Construction of a moderate-to-high density linkage map provides a greater probability that the entire genome is covered with molecular markers. This is important when using molecular linkage maps to detect and characterise loci underlying quantitative traits where one needs to be assured that the entire genome has been uniformly surveyed (Paterson *et al*. 1988; Lander and Botstein 1989; Tanksley *et al*. 1992). A medium density linkage map constructed from RAPD markers segregating in haploid DNA is reported in the next chapter, as first step towards QTL analysis.

5 CHAPTER 5

Experiment 2: Construction of a RAPD-based genetic linkage map

5.1 Introduction

With the use of bulked segregant analysis, it is difficult to ascertain the extent of genome coverage achieved and to be assured that the entire genome has been uniformly surveyed for marker-QTL linkages. If enough segregating markers are scattered throughout an entire genome, it is theoretically possible to detect and characterise all of the QTLs affecting a quantitatively inherited character (Tanksley 1993). On the other hand, there is no chance of detecting QTLs in chromosomal regions devoid of segregating markers.

For practical purposes, genetic mapping in forest tree species has acquired increased impetus as a means of characterising the quantitative trait loci influencing commercially important traits. Maps show the approximate chromosomal location of QTL and provide a numerical accounting of the number of such factors. Once a quantitative trait locus map is constructed, it serves as the basis for isolating and identifying genes and multigene families. The identification of QTLs may potentially feedback to the development of improved selective breeding programs, and a shortening of the interval between crossing and progeny assessment. In the shorter term, markers linked tightly to a gene of interest will allow marker-assisted selection, and permit decisions to be made well before the trait is manifested.

As a first step towards genome-wide QTL analysis, a genetic linkage map of radiata pine was constructed based solely on RAPD markers, and using DNA extracted from the megagametophyte.

5.2 Materials and methods

5.2.1 Plant materials

For this experiment, *Pinus radiata* seeds from a full-sib cross, different to that used in the previous experiment (see chapter 4) were obtained from CSIRO, Forestry & Forest Products. Parents of this cross were also selected based on parental NESTUR values from previous evaluations. Open pollinated (OP) seeds from the maternal parent (30040) had a high variance in NESTUR, while OP seeds from the paternal parent (80121) had a low variance.

Approximately 284 seeds were germinated, and then picked out onto 150-mm diameter pots filled with a porous medium. Megagametophyte collection, storage, and seedling culture were as detailed in chapter 4, section 4.2.1. DNA from ninety-three of the megagametophytes were used for linkage map construction.

5.2.2 Primer screening

To identify primers which amplified polymorphic bands, DNA from six megagametophytes were tested with an array of 432 primers, made up of 194 Operon primers supplied by Operon Technologies, Alameda, CA and 238 primers from the “conifers RAPD” set obtained from Dr. John Carlson, University of British Columbia. Based on the number of polymorphic bands produced, and on their sharpness, 132 primers were selected for genetic linkage map construction. All of the chosen primers were used to amplify DNA of the mapping population, which comprised 93 progenies of the cross, including the six already tested and used this time as a repeatability assay.

The initial primer screening was performed on a FTS-960 Thermal Sequencer (Corbett Research), while RAPD assaying of the mapping population was

carried out on GeneAmp PCR System 9600 (Perkin Elmer Cetus). These procedures provided an internal control and an estimate of repeatability for all the scored markers.

Reaction mixture for RAPD assays and amplification programs were as described in chapter 4, section 4.2.2. Amplification products were separated by gel electrophoresis, as described in section 4.2.3.

5.2.3 Nomenclature of RAPDs

Markers were assigned a two-part identity, the first part corresponding to the primer which amplified the polymorphic band (OP followed by letters and digit numbers for the Operon primers, and UBC followed by digit numbers for the University of British Columbia primers). When only a single polymorphism was amplified by a primer, it was given the letter 'a' (Kubisiak *et al*. 1995). When a primer amplified multiple polymorphisms, the bands were assigned consecutive letter designations from the largest-molecular-weight band to the smallest.

5.2.4 Linkage analysis and error detection

Linkage analyses of RAPD loci were performed using the JOINMAP ver. 2.0 suite of programs (Stam and Ooijen 1995). First, linkage groupings were made using a logarithm of odds ratio (LOD) threshold of 4. JOINMAP uses a LOD score that is derived from the probability in the chi-square test for independent segregation, which is somewhat different from normal LOD scores. The rationale behind using a test of independence rather than normal LOD scores is that distortion of segregation affects normal LOD scores, but does not affect the test of independence. The use of normal LOD scores can result in spurious linkage of markers with segregation distortion (Stam and Ooijen 1995).

Next, pair-wise data files were obtained for each of the linkage groups, which is a simple list of pair-wise recombination estimates, together with the LOD score. The linkage map and order of marker loci in each linkage group was established using a common ‘response file’ of JOINMAP parameters. The ‘response file’ specified the parameter setting for mapping as:

LOD threshold for mapping = 0.001
REC threshold (or recombination frequency) = 0.45
Triplet threshold = 8.0
Ripple value = 3
Jump threshold = 3.0
Mapping function = Kosambi.

The triplet value defined the relative log-likelihood above which the order within a triplet of markers was considered as the most likely (or best) order of the triple. The triplet with the highest log-likelihood provided the starting point for linkage map extension. During map extension, JOINMAP permutes the order of neighbouring markers within a moving window of three adjacent markers and compare the likelihood of the resulting map. The ‘Ripple’ value used specified that during map extension permutation or ripple should be performed after every addition or insertion of three markers.

5.2.4.1 Locus ordering and map construction

Locus ordering and linkage map extension with JOINMAP involved three cycles of data shuffling to find the best-fitting linear order of markers. The best-fitting order was determined by the change or ‘jump’ in goodness-of-fit of the map after a marker has been inserted, and controlled by the ‘Jump’ parameter specified in the ‘response file’. During the first cycle, only well-fitting markers were positioned. These were markers that caused a ‘jump’ or increase in the chi-square value smaller than the specified threshold of 3.0. Markers that caused a ‘jump’ of ≥ 3.0 were not discarded but temporarily kept aside because

the change in chi-square value not only depend on the marker itself but also on the markers already on the map (Stam and Ooijen 1995). During the second cycle of map construction, JOINMAP attempts to position these markers using the same fitting criterion as in the first cycle. Thus, markers ordered during the first and second cycle of map construction can be considered as 'framework' markers since they were placed with a high degree of stringency. In the third cycle of mapping, 'accessory' markers that could not be placed under the 'Jump' threshold were forced onto the map by annulling the restriction. However, when such markers caused a chi-square jump of ≥ 6.0, they were considered 'troublesome' and discarded.

Output file from JOINMAP was used as input to draw the linkage maps using the program ColorMap (Matise *et al*. 1994).

5.2.5 Genome length estimation and marker coverage

The total radiata pine genome size was estimated using the 'method of moments' procedure of Hulbert *et al*. (1988), as modified for low density genetic maps by Chakravarti *et al*. (1991):

$$G = [M * M-1] * X/K$$

Where at a LOD score of 4.0, G = genome length, M = number of markers analysed, X = maximum cM distance between linked markers, and K = number of markers at or above the LOD threshold.

5.3 Results

5.3.1 Primer screening and marker scoring on mapping population

From the 132 selected primers, 367 segregating RAPD markers were scored on the mapping population. Thirty markers could not be scored across the two

runs of amplification, ie. 90.8% repeatability of scored markers. A typical example of repeatably scored markers is shown in Figure 5.1a-d.

When six individual progeny are assessed, the probability of missing a polymorphic marker segregating 1:1 is 0.09375. It was, thus, not surprising that a RAPD band monomorphic among the six individuals was found polymorphic in the mapping population (Figure 5.1d). Markers ranged in size from 250b to 2 kb. On average, one primer produced 2.8 polymorphic RAPD markers.

The two sources of primers (UBC and OPERON) differed in the number of RAPD polymorphic markers that were found useable in genetic linkage construction (Table 5.1). There was also a difference in the average number of polymorphic bands per arbitrary primer generated by the two sources of primers.

5.3.2 Inspection of segregation ratio

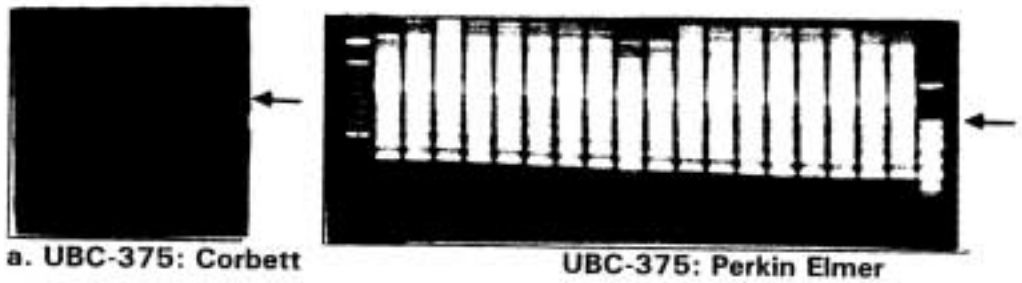
Segregation distortion was analysed using the JOINMAP module JMSLA. Of the 367 polymorphic markers scored, seven percent (7%) failed a 1:1 segregation test. Errors associated with genotyping were considered the likely main cause of the distortion. Ambiguity of genotyping RAPDs often occur due to incomplete amplification of DNA. An almost two-fold difference in the number of markers with distorted segregation was observed in the two sources of primers (Table 5.1). The affected markers were not used for linkage map construction.

Table 5.1 Scored markers, segregation distortion and mappable polymorphisms observed in the mapping population of 30040 x 80121 cross. Number of markers generated per primer were scored as present/absent, and tested for deviation from expected 1:1 Mendelian ratio.

Source of primers	Number of primers	Total RAPD markers per primer source	Mean RAPD markers per primer	Total Non-Mendelian markers	% of Total	Total Mendelian markers	% of Total
UBC primers	59	183	3.1	8.0	6.6	126	69.4
Operon primers	73	184	2.5	16.0	11.3	141	76.6
Total	132	367	-	24.0	-	267	-
Percent	-	-	-	6.5	-	73	-

5.3.3 Inspection of recombination estimates and LOD scores

Before linkage map construction, the JMREC module of JOINMAP was used to check whether the population contained very improbable genotypes. Twenty percent (20%) of the markers failed the error detection test during analysis. These were essentially markers segregating on basis of band intensity, and about which it was difficult to be categorical. Improbable genotypes were re-examined on the gel photos, and when irreconcilable data points were found, such markers were deleted.



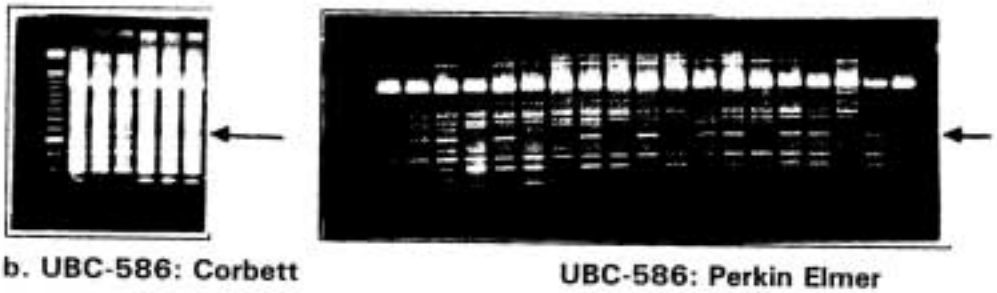


Figure 5.1 Examples of repeatably amplified RAPD bands from two separate thermocyclers, the FTS-960 Thermal Sequencer (Corbett Research) and GeneAmp PCR System 9600 (Perkin Elmer Cetus). Segregating repeatable bands are indicated by arrows; first lane on each panel is the size marker (100 bp ladder).

5.3.4 Linkage map construction

Of the 367 RAPD markers scored, 267 (or 73%) were found useable for genetic linkage mapping under the conditions of this study (Table 5.1). A plot of two-point recombination estimates obtained for the 267 markers against the LOD scores is shown in Figure 5.2. The pattern of spread of the data points is close to the ideal situation, and suggests a good data set.

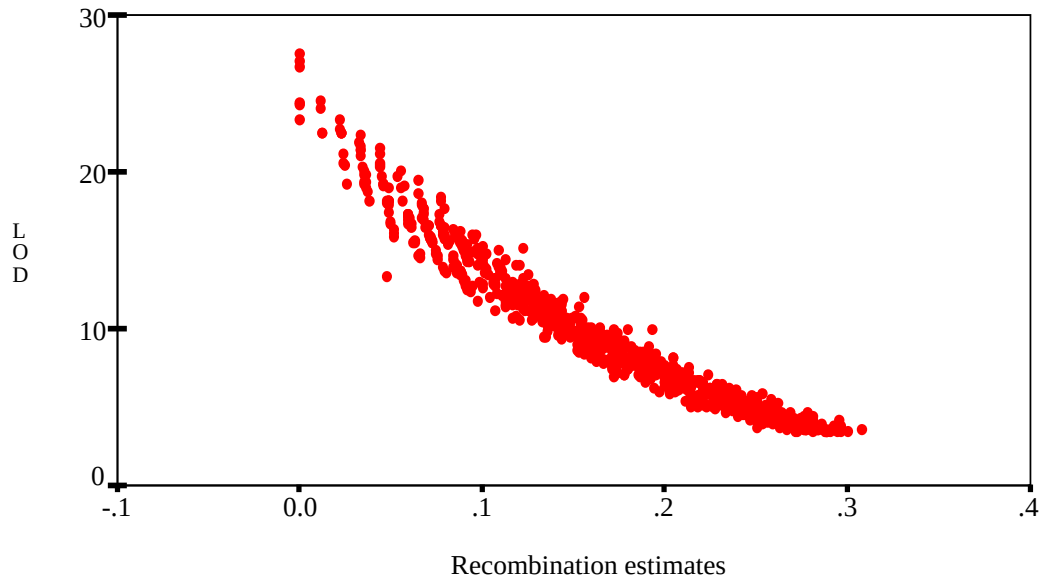


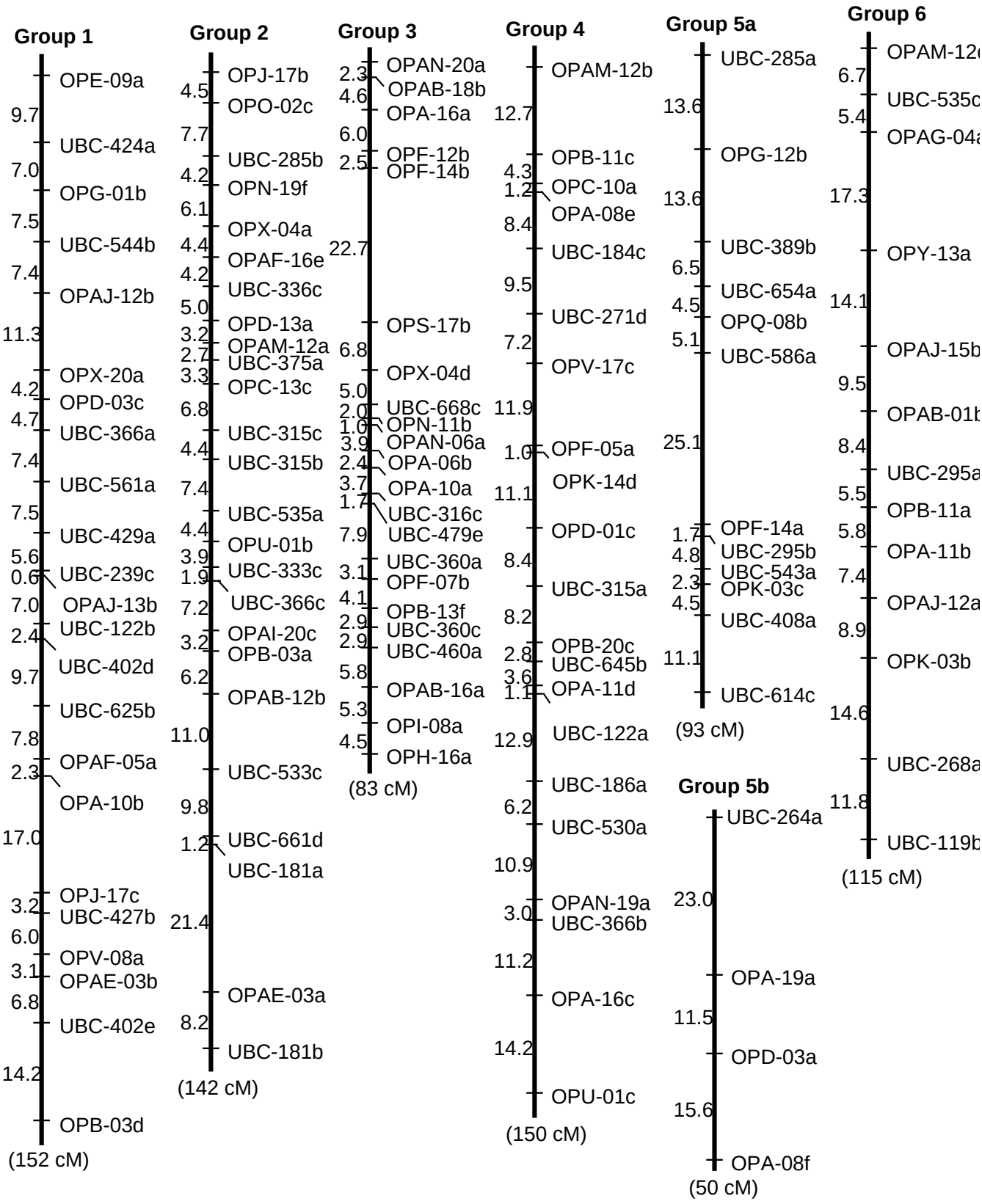
Figure 5.2 Plot of two-point recombination estimates and LOD scores obtained for 262 markers using the JOINMAP program JMREC32.

At a LOD threshold for linkage of 4.0, two hundred and sixty-two (262) markers or 98% of the scored Mendelian loci were mapped to 14 linkage groups (Table

5.2, Figure 5.3), covering an estimated 1511.1-cM of map distance. Five markers remained unlinked. The largest of the linkage groups had 37 markers, and covered 183 cM. There were two small linkage groups with seven markers each, and covering 55 and 39 cM, respectively (Figure 5.3). The average distance between adjacent markers was 6 cM (Table 5.2). Among linkage groups, the interval length was in the range of 3-5 cM (Figures 5.3 and 5.4). No interval was larger than 30 cM. The largest intervals of 21 cM were observed in linkage group 3, between markers OPF-14b and OPS-17b (Figure 5.3).

Table 5.2 Linkage groups, marker numbers, genetic length and mean interval between markers in the radiata pine RAPD-based linkage map.

Linkage group	Number of markers	Genetic length (cM)	Mean interval (cM)	Percentage of markers mapped after 1st cycle
1	24	148.0	6.2	50.0
2	25	140.7	5.6	72.0
3	23	103.0	4.5	78.3
4	37	183.4	5.0	62.2
5	30	144.5	4.8	66.7
6	13	115.3	8.9	84.6
7	14	80.0	5.7	78.6
8	20	149.1	7.5	85.0
9	8	39.4	4.9	100.0
10	7	55.1	7.9	100.0
11	17	111.2	6.5	70.6
12	22	131.4	6.0	72.7
13	15	71.0	4.7	86.7
14	7	39.0	5.6	100.0
Total	262	1511.1	-	-
Mean	19	107.9	6.0	79.1


Group 1
OPE-09a
9.7
UBC-424a
7.0
OPG-01b
7.5
UBC-544b
7.4
OPAJ-12b
11.3
OPX-20a
4.2
OPD-03c
4.7
UBC-366a
7.4
UBC-561a
7.5
UBC-429a
5.6
0.6
UBC-239c
7.0
OPAJ-13b
2.4
UBC-122b
UBC-402d
9.7
UBC-625b
7.8
OPAF-05a
2.3
OPA-10b
17.0
OPJ-17c
3.2
UBC-427b
6.0
OPV-08a
3.1
OPAE-03b
6.8
UBC-402e
14.2
OPB-03d
(152 cM)
Group 2
4.5
OPJ-17b
OPO-02c
7.7
UBC-285b
4.2
OPN-19f
6.1
OPX-04a
4.4
OPAF-16e
4.2
UBC-336c
5.0
OPD-13a
3.2
OPAM-12a
2.7
UBC-375a
3.3
OPC-13c
6.8
UBC-315c
4.4
UBC-315b
7.4
UBC-535a
4.4
OPU-01b
3.9
UBC-333c
1.9
UBC-366c
7.2
OPAI-20c
3.2
OPB-03a
6.2
OPAB-12b
11.0
UBC-533c
9.8
UBC-661d
1.2
UBC-181a
21.4
OPAE-03a
8.2
UBC-181b
(142 cM)
Group 3
2.3
OPAN-20a
OPAB-18b
4.6
OPA-16a
6.0
OPF-12b
2.5
OPF-14b
22.7
OPS-17b
6.8
OPX-04d
5.0
2.0
UBC-668c
1.0
OPN-11b
3.9
OPAN-06a
2.4
OPA-06b
3.7
OPA-10a
1.7
UBC-316c
7.9
UBC-479e
UBC-360a
3.1
OPF-07b
4.1
OPB-13f
2.9
UBC-360c
2.9
UBC-460a
5.8
OPAB-16a
5.3
OPI-08a
4.5
OPH-16a
(83 cM)
Group 4
OPAM-12b
12.7
OPB-11c
4.3
1.2
OPC-10a
OPA-08e
8.4
UBC-184c
9.5
UBC-271d
7.2
OPV-17c
11.9
1.0
OPF-05a
OPK-14d
11.1
OPD-01c
8.4
UBC-315a
8.2
OPB-20c
2.8
UBC-645b
3.6
1.1
OPA-11d
12.9
UBC-122a
UBC-186a
6.2
UBC-530a
10.9
OPAN-19a
3.0
UBC-366b
11.2
OPA-16c
14.2
OPU-01c
(150 cM)
Group 5a
UBC-285a
13.6
OPG-12b
13.6
UBC-389b
6.5
UBC-654a
4.5
OPQ-08b
5.1
UBC-586a
25.1
1.7
OPF-14a
4.8
UBC-295b
UBC-543a
2.3
OPK-03c
4.5
UBC-408a
11.1
UBC-614c
(93 cM)
Group 5b
UBC-264a
23.0
OPA-19a
11.5
OPD-03a
15.6
OPA-08f
(50 cM)
Group 6
6.7
OPAM-12
UBC-535c
5.4
OPAG-04a
17.3
OPY-13a
14.1
OPAJ-15b
9.5
OPAB-01b
8.4
UBC-295a
5.5
OPB-11a
5.8
OPA-11b
7.4
OPAJ-12a
8.9
OPK-03b
14.6
UBC-268a
11.8
UBC-119b
(115 cM)

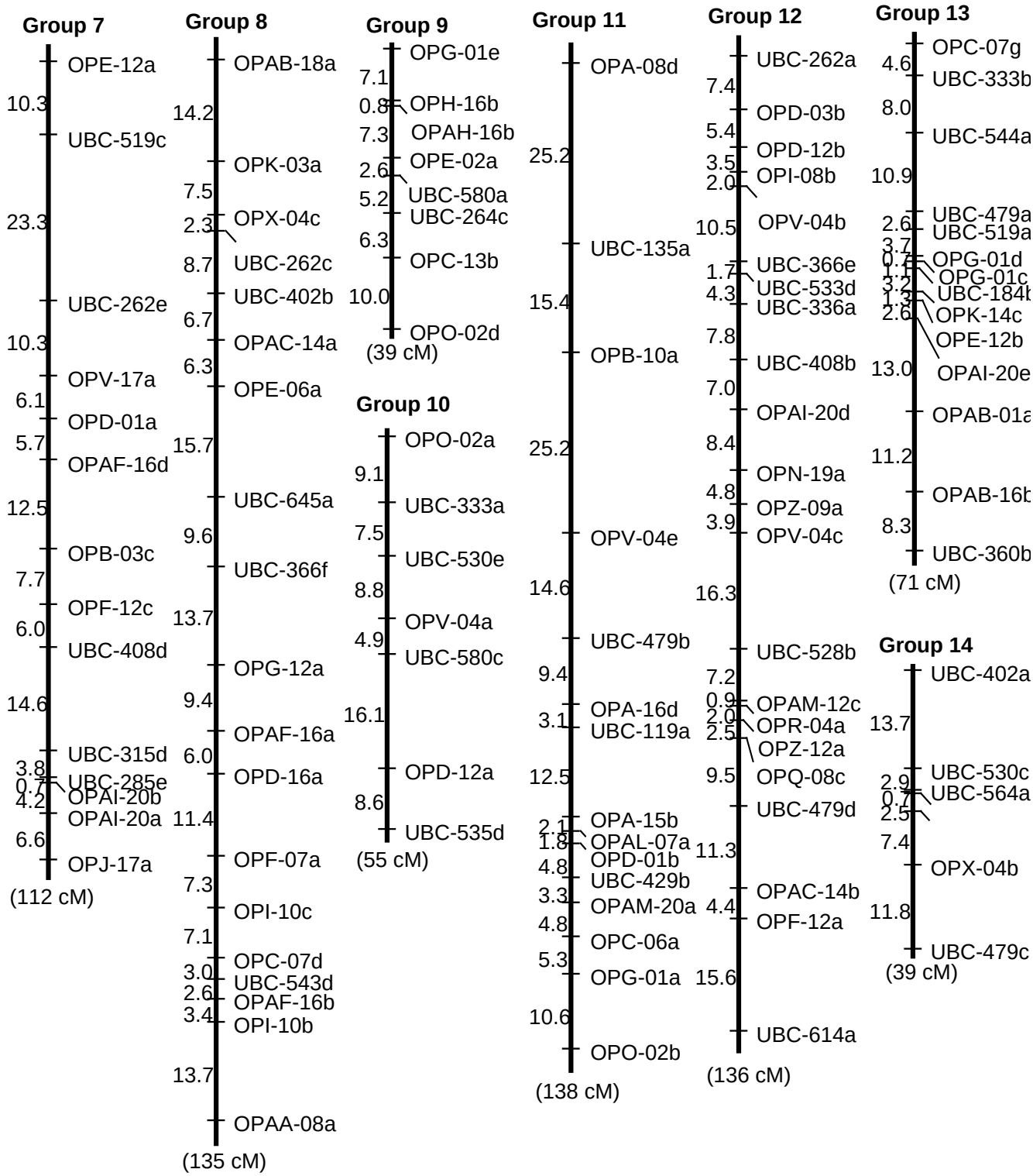


Figure 5.3 RAPD-based genetic linkage maps of radiata pine, constructed from haploid DNA of 93 megagametophytes. Marker identification (see materials and methods for nomenclature) is given on the right of the bars; cumulative distances (in centiMorgans, cM) is indicated to the left of the bars.

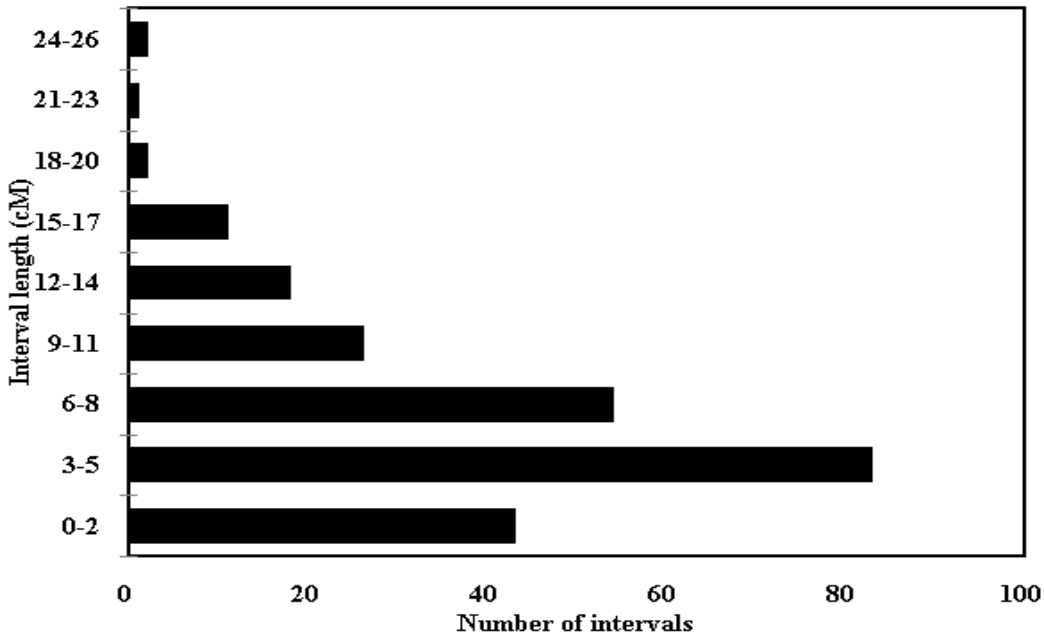


Figure 5.4 Distribution of interval sizes (in centiMorgans, cM) between adjacent markers on RAPD-based radiata pine linkage map.

5.3.5 Map goodness-of-fit

The percentage of markers that were placed in the first cycle of mapping under the jump restriction of 3.0 ranged from 50% in linkage groups 1 to 100% linkage groups 9, 10, and 14 (Table 5.2). These are markers positioned with a high degree of statistical confidence. The rest of the markers on each linkage map that failed the jump restriction were later placed after the restriction was removed.

Assuming the order of markers is correct, the assembled linkage maps were examined for discrepancies between calculated (or joint best estimate) and observed (or direct) map distances. The direct map distance is the calculated recombination frequency translated into centiMorgans (Stam and Ooijen 1995). Non-linearity can result from mistyping errors, which inflate estimates of recombination frequencies. The mean observed (or direct) distance plotted

against the calculated (or map) distance for each linkage group is presented in Figure 5.5.

The linkage groups showed good agreement with expected linear distribution of data-points along a line of unit slope.

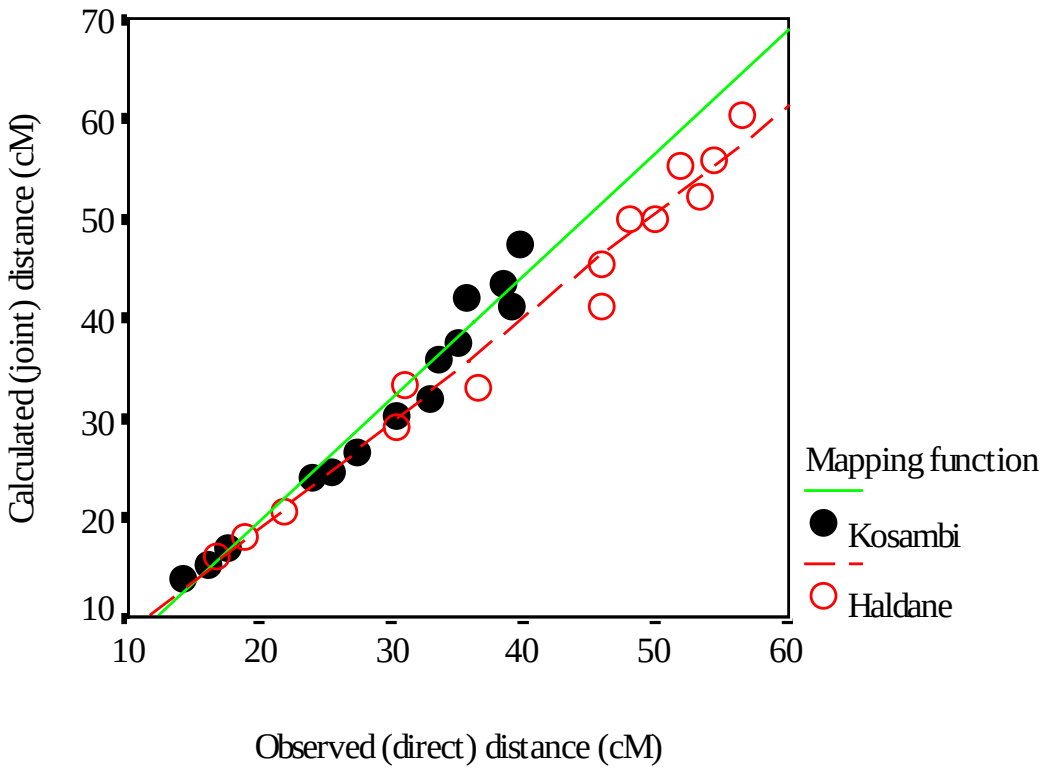


Figure 5.5 Plot of mean observed vs. calculated map distances for the 14 linkage groups assembled using 262 RAPD markers. For comparison, linkage analyses were performed using the Kosambi and Haldane mapping functions. Reference lines indicate fitted linear regression between observed and calculated distances using both functions.

However, the best fit was in linkage mapping performed with the Kosambi mapping function. In 11 out of the 14 linkage groups, distances calculated with the Haldane mapping function were longer than the observed two-point distances.

5.3.6 Estimate of genome size

JOINMAP calculated the maximum likelihood estimates of Kosambi map distances for all unique pairs of the 262 RAPD markers. For genome size estimation, data for 258 markers were used because four pairs generated by the same primer showed no recombination, and may represent codominant markers.

The largest observed map distance (or recombination values) between linked markers at a LOD score of ≥ 4.0 was 31 cM. The observed number of locus pairs with a LOD score 4.0 or greater (K) was 1017. These values were substituted in method 3 of Chakravarti *et al*. (1991) to estimate the size of the radiata pine genome as 2021 cM. Thus, the current set of 262 loci is apparently distributed over approximately three-quarters of the genome.

5.4 Discussion

5.4.1 Reliability of RAPDs

An important consideration of RAPDs as genetic markers is reproducibility. Because individual brands of thermocyclers have different characteristic temperature ramping profiles and display relative to actual block temperature differences (Smith and Williams 1994), reproducibility was assessed in the present study by re-running reaction amplifications using two different thermocyclers (see section 5.2.2). A repeatability of 91% of scored markers was observed, which agrees well with an estimate of 92.4% for RAPDs in *Eucalyptus* obtained by Grattapaglia and Sederoff (1994). However, there was variation in the intensity of bands generated by the two thermocyclers, as is evident from Figure 5.1a-d. Due to faint or inconsistent amplifications, 20% of

scored markers (as mentioned in section 5.3.3) were found unreliable for genetic mapping under the conditions of the present study.

5.4.2 Segregation distortion

Seven percent (7%) of the scored markers in this study exhibited distorted segregation patterns. Segregation distortion has also been reported by other researchers (Hunt and Page 1995; Lanaud *et al*. 1995; Kubisiak *et al*. 1995; Rajapakse *et al*. 1995; Byrne *et al*. 1995; Bucci and Menozzi 1995; Boscherini *et al*. 1994; Tanhuanpaa *et al*. 1994; Nelson *et al*. 1994; Binelli and Bucci 1994; Rowland and Levi 1994; Cai *et al*. 1994; Bradshaw *et al*. 1994; Grattapaglia and Sederoff 1994; Tulsieram *et al*. 1992; Carlson *et al*. 1991). However, there appears to be a lack of consensus as to whether markers with aberrant segregation ratios should be excluded from linkage analysis.

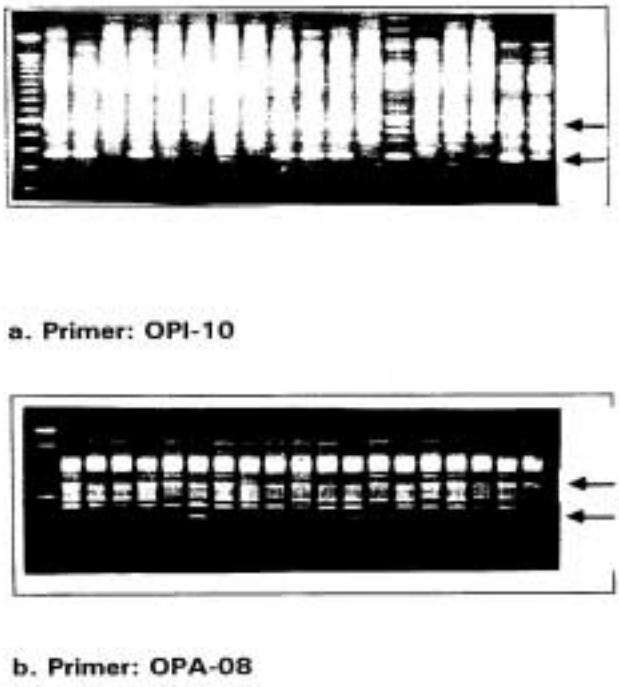
Distorted segregation can arise by chance, the probability of, which is the significance level used. Selection for any of the genotypes at a locus can also cause a skewed segregation ratio (Hallden *et al*. 1996). Other mechanisms capable of producing a skewed distribution of genotypes include chromosome loss, presence of an allele for pollen lethality, conflict between genetic isolating mechanisms evolved in the parental species, and expression of genetic load via a lethal recessive allele (Bradshaw and Stettler 1994). The effect of these sources of segregation distortion on recombination estimates is relatively small (Säll and Nilsson 1994), and bias to marker order negligible (Lin and Ritland 1996). Thus, such markers have commonly been used for linkage map construction (Byrne *et al*. 1995; Kubisiak *et al*. 1995; Cai *et al*. 1994; Bradshaw *et al*. 1994; Grattapaglia and Sederoff 1994; Kesseli *et al*. 1994).

However, errors associated with mistyping can cause segregation distortion, and even low rates of such errors can have serious effects on recombination values (Hallden *et al*. 1996; Kesseli *et al*. 1994; Säll and Nilsson 1994). With RAPDs, distorted segregation can result from mistyping errors due to inadequate amplification or competition for binding sites (Heun and Helentjaris 1993; Hallden *et al*. 1996) and also from incidental co-migration of unrelated fragments characterised by the same molecular weight (Burke et al. 1992). These would reduce reproducibility of such markers. Since it is highly impractical to establish reproducibility of every marker used in linkage map construction, it was considered more prudent to exclude markers exhibiting significantly skewed distribution (with α-value $\leq$ 0.05) in genetic map construction. Yazdani *et al*. (1995) also excluded RAPD markers with distorted segregation from linkage map construction.

5.4.3 Levels of polymorphism

An essential feature of RAPD markers is the generation of multiple polymorphic fragments per primer. Multiple RAPD bands produced by the same primers are reported to assort independently within a family (Byrne *et al*. 1995; Hemmat *et al*. 1994; Carlson *et al*. 1991). The independence of these loci is an important aspect of the use of RAPD loci identified by the same primers in the identification of individuals or diversity studies (Byrne *et al*. 1995). However, in some mapping studies, such fragments have been shown to demonstrate linkage or even tight clustering (Hemmat *et al*. 1994; Rowland and Levi 1994). Five percent (5%) of the primers used in the present study generated at least two fragments that showed linkage within the same group. The 350- and 600 bp bands generated by primer OPI-10 (Figure 5.6a), for instance showed a tight

linkage of 10-cM interval in group 8 (Figure 5.3; OPI-10b and OPI-10c). These markers were found to be distributed across the linkage groups.



a. Primer: OPI-10

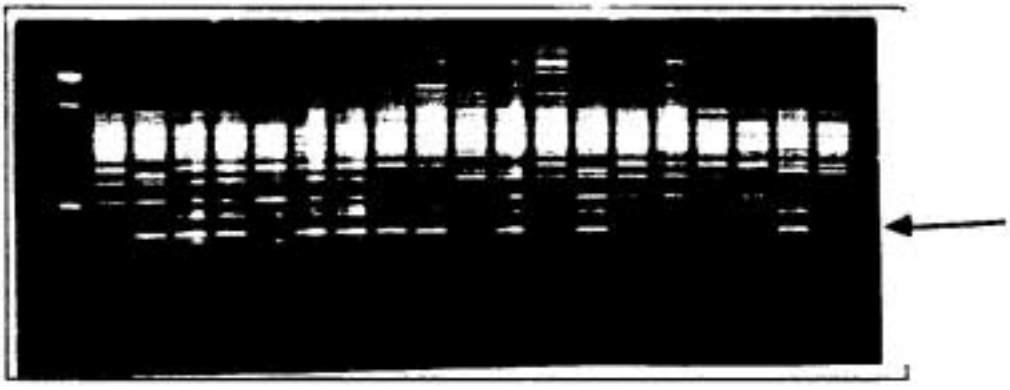
b. Primer: OPA-08

Figure 5.6 Examples of RAPD bands amplified from the same primer, and which showed linkage to the same group. Arrows indicate segregating RAPD bands; in both panels, first lane represent the size markers (100 bp ladder).

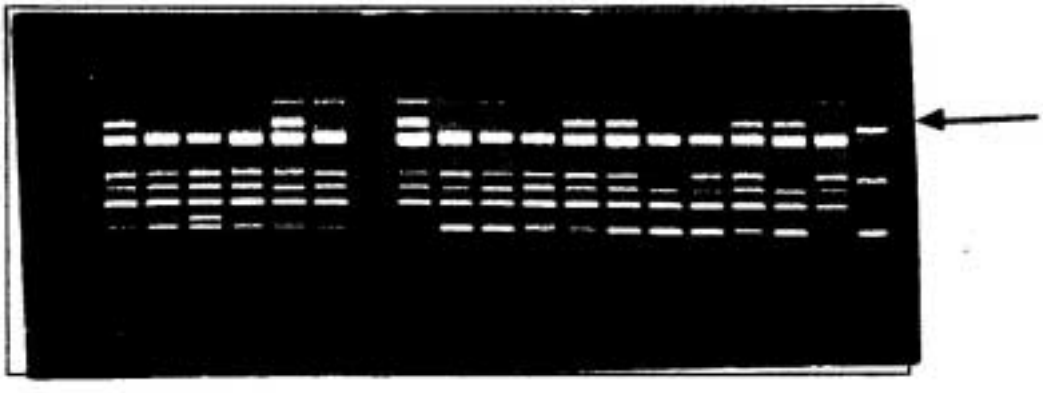
Note that at $\alpha = 0.01$, and considering a total of 34,191 markers in pair-wise combinations, about 342 pair-wise combinations are expected to display this behaviour due to chance alone. In all, three of such markers had an interval of $\geq$ 30 cM between them, and could be considered unlinked on a 30-cM map scale. Therefore, such occurrences may not have a biological basis, and no effort was made to establish sequence homology of such RAPD bands. In linkage group 5, for instance, the 400-and 750 bp fragments generated by primer OPA-08 (Figure 5.6b) mapped to $\geq$ 30 cM of each other (Figure 5.3; OPA-08b and OPA-08f).

5.4.4 Co-dominant RAPD markers

Co-dominant RAPD markers have been reported in essentially all linkage maps constructed from RAPD markers, irrespective of the species: in *Eucalyptus* (Grattapaglia and Sederoff 1994; Byrne *et al*. 1995), slash pine (Nelson *et al*. 1993), maritime pine (Plomion *et al*. 1995), the honey bee (Hunt and Page 1994), the fungus *Neurospora crassa* (Williams *et al*. 1990), and in celery (Yang *et al*. 1995). Co-dominant RAPDs are generally characterised by (i) size or fragment-length variations produced by the same primer, (ii) F_1 individuals receive either one or the other allele, ie, the two RAPD fragments are in repulsion, and (iii) no recombinants are observed in the F_1, ie, no individuals with both fragments or null for both fragments. In the present study, five pairs of markers were observed to fit these criteria (Figure 5.7a-d), representing a 4% occurrence of co-dominant markers.



a. Primer: OPE-06



b. Primer: OPA-10

Figure 5.7 Examples of RAPD bands showing co-dominant segregation patterns. Arrows indicate size or fragment-length variations produced by the same primer. The first lane on each panel represents the molecular-size markers (100-bp ladder).

The frequency of occurrence is similar to that reported by Grattapaglia and Sederoff (1994) for *Eucalyptus*. One of the fragment-length polymorphisms had a 200 bp size difference between alternate alleles from the same primer (Figure 5.7d).

5.4.5 Genome size estimation and map coverage

The radiata pine, like other gymnosperm species, is characterised by its antiquity (the first conifers appeared 140 million years before the first angiosperm) and the very high and consistent amount of DNA per nucleus (Ohri and Khoshoo 1986; Wakamiya *et al*. 1993). Due to limited development of genetic mapping in gymnosperms, previous estimates of the genome size have been speculative. Plomion *et al*. (1995b) and Neale and Sederoff (1996) estimated size of the pine genome at approximately 1500 cM by from the number of chiasmata per bivalent (2.5) observed in pine pollen mother cells (Saylor and Smith 1966). Gerber and Rodolphe (1994), however, estimate the size at 2000 cM, based on the amount of DNA per diploid cell. An alternative

estimate of pine genome size can be obtained by using the parameters outlined by Meagher *et al*. (1988) who concluded that the number of map units per chromosome would range from 50 to 300 cM regardless of the physical size of the chromosome. Since most pines have 12 metacentric chromosomes with little variation in physical length, the size of the genome would be estimated to be between 1200 and 1800 cM. This rule holds true for tree species as diverse as *Eucalyptus* (1620 cM, n = 11, Grattapaglia *et al*. 1994), Poplar (2400-2800 cM, n = 19, Bradshaw *et al*. 1994), Theobroma cacao (1200 cM, n = 10, Crouzillat *et al*. 1996), and Walnut (1660 cM, n = 16, Fjellstrom and Parfitt 1994), where 100 to 150 cM per linkage groups is a consistent estimate.

However, even with the availability of genetic linkage maps, the pine genome sizes calculated from recombinational rates have been variable, and often higher than the estimated length. For instance, Neale and Williams (1991) estimated the size of pine genome to be approximately 2500 cM. In maritime pine, Gerber and Rodolphe (1994) used protein data to calculate its genomic length at approximately 2114-2279 cM. Plomion *et al*. (1995a) used 463 RAPD and protein markers to calculate the size of the maritime pine genome at about 1860 cM. Among other conifer species, estimates of genome length from RAPDs have ranged from 2638 cM for *Pinus sylvestris* (Yazdani *et al*. 1995) to 2656 cM for Longleaf pine (Nelson *et al*. 1994). The estimate of 2021 cM from the current study is closer to the upper limit of 2000 cM predicted by Gerber and Rodolphe (1994).

5.4.6 Efficiency of mapping function and algorithm

Contrary to the report of Lin and Ritland (1996), a good agreement was observed between map distances calculated directly from recombination

frequencies and those from multipoint analyses obtained with JOINMAP. Map distances calculated directly from recombination frequencies do not coincide with distances from multipoint analyses due to the influence of multiple crossover events upon one another (interference). Interference is ubiquitous in eucaryotes (Korol *et al*. 1994), but even more so in the pines. Among plant species, pines possess one of the largest amounts of DNA per diploid cell, about 40 picograms (Ohri and Khoshoo 1986). Although there is no linear relationship between recombinational rates and DNA quantity, chiasma frequency tend to decrease as the DNA amount increases (Rees and Durrant 1986). There is no clear explanation for this observation, but Kaback *et al*. (1992) suggested that this could be due to an increased amount of chiasma interference with chromosome size. In yeast, the authors found that the DNA sequences of the smallest chromosome undergo recombination at a two-fold higher average than DNA sequences on the largest chromosome.

A mapping function such as Kosambi (1944) is used to transform observed recombination frequencies into map distances. Unlike the Kosambi mapping function, the Haldane (1919) function assumes no interference. The net effect, as shown in Figure 5.5 is longer multipoint distances, compared with the observed (direct) distance, over a given region of the genome.

Map distances calculated with the JOINMAP suite of programs are usually lower than those with MAPMAKER/EXP (Lander *et al*. 1987). The discrepancy results from the different methods of calculating map lengths. MAPMAKER/EXP calculates the map length as the sum of adjacent distances, ie, using adjacent marker pairs only (Qi *et al*. 1996). JOINMAP, on the other hand, uses all pair-wise estimates (above a predetermined LOD threshold

value) for calculating the total map length (Stam 1993). The likelihood method applied in MAPMAKER/EXP assumes an absence of interference, and recombination frequencies are simply translated into centiMorgans, according to the chosen map function. The JOINMAP program, however, does consider interference, and whenever the assumed level of interference does not exactly reflect the true interference, the two methods will produce different total map length (Qi *et al*. 1996). Thus, where there is interference, JOINMAP will produce shorter maps than MAPMAKER/EXP, even when both programs use the same Kosambi mapping function.

5.4.7 Genetic mapping of QTLs

Although bulked segregant analysis had previously been successfully used to detect RAPD markers linked to NESTUR (Chapter 4), complete genome coverage with molecular markers could not be assured. This is especially important when searching for molecular markers linked to quantitative traits to ensure a greater probability of QTL detection (Tanksley *et al*. 1992). The current map, with an average of 6 cM between loci, should be adequate to begin genome-wide search for QTL affecting NESTUR and related stem growth traits.

6 CHAPTER 6

Experiment 3: Interval mapping of QTL underlying NESTUR and associated seedling growth traits in radiata pine

6.1 Introduction

Lander and Botstein (1989) introduced interval mapping technique that localises the effect of a QTL to a genomic segment located between pairs of markers rather than associating the effect with an individual marker locus. Interval mapping uses phenotypic and genetic marker information to estimate the probable genotype and the most likely QTL effect at every point in the genome, by means of a maximum likelihood linkage analysis.

The basic method introduced by Lander and Botstein (1989) has been generalised for a variety of settings. In interval mapping, QTL are mapped one at a time, ignoring the effects of other QTL lying elsewhere in the genome, and which can influence the residual variance. Problems arise when several QTLs are located on the same chromosome, because estimates of the position and effects of a QTL can be biased by adjacent QTLs (Ooijen 1992; Martinez and Curnow 1992; Haley and Knott 1992; Stam 1991). An alternative approach is to use markers as co-factors to absorb most of the genetic variance caused by QTL other than the target (Zeng 1994; Jansen and Stam 1994; Jansen 1993). The two methods were used for QTL analysis in this study to determine whether they lead to different results.

6.2 Materials and methods

6.2.1 Plant materials and phenotying

Plant materials used for experiment 3 were progenies of the single controlled cross used for linkage map construction in chapter 5 (section 5.2.1). The

methods used for growing radiata pine seedlings, and for the determinations of NESTUR were as described in chapter 4, sections 4.2.1. Seedling growth (or increment) in height (Ht growth), diameter (SD growth), volume (SVol growth) and needles (NVol growth) were determined as the difference between two measurements taken at an interval of two weeks.

6.2.2 Genotyping, linkage map construction, and QTL analysis

The methods used for DNA extraction, RAPD reaction mixture and assay conditions are described in chapter 4, sections 4.2.2 and 4.2.3.

6.2.3 Linkage map re-construction

The linkage map that was reported in chapter 5 (Figure 5.3) was developed for a haploid progeny without knowledge of the linkage phase of the RAPD marker loci. In interval and multiple-QTL mapping using a HAP population model, however, the linkage phase of locus alleles must be known. The assignment of linkage phase can be made by analysis of re-coded data (ie. presence re-coded to absence and vice versa) along with the original data set (Plomion *et al*. 1995b). The analysis yields twice the expected number of linkage groups, corresponding to the two homologues for each chromosome. Inferences on linkage phase is made following two-point linkage analysis of the combined data with MAPMAKER/EXP 3.0 (Lander *et al*. 1987). In the present study, genetic linkage maps were then re-calculated *de novo* from the data file with inferred linkage phase, and checked for consistency with the linkage map described in chapter 5, Figure 5.3.

6.2.4 QTL detection methods

QTL analyses were carried out using single-point and interval mapping methods.

6.2.4.1 Single-factor analysis

The association between markers and measured quantitative trait was analysed by simple linear regression. The regression considered, as the dependent variable, the value of the measured character in each of the 93 full-sib progenies. As the regressor, the allelic composition of each RAPD locus was used, which assumed the numerical value of 1 or 2 according to the absence or presence of the RAPD band in the individual. All the analyses were carried out with pair-wise deletion of missing data and the 'include constant in equation' option. Significant associations at ≤ 0.01 are reported in the single-point analysis.

6.2.4.2 Interval mapping methods

Interval mapping with MAPQTL was performed using two methods, the conventional interval mapping of Lander and Botstein (1989), and the multiple-QTL (or composite interval) mapping method of Jansen and Stam (1994), Tinker (1994), Utz and Melchinger (1994). The multiple-QTL mapping (MQM) was carried out with the "restricted MQM mapping" option of MAPQTL, in which for a given interval all selected markers are used as cofactors, except ones on the linkage group where the QTL is fit (Ooijen and Maliepaard 1996).

6.2.4.2.1 Preselection of cofactor markers

Markers used as cofactors were selected by means of a stepwise regression with the software SPSS for Windows (SPSS, Inc.). The stepwise regression was carried out with an F-to-enter value of 5.0. To avoid using tightly linked markers as cofactors, correlation matrices were established for selected markers and only uncorrelated loci included. The number of cofactors was kept to a maximum of 15, with files generated separately for each of the measured traits.

6.2.4.2.2 Estimation of empirical LOD threshold value

The probability of the presence of a QTL at a particular location was expressed as a logarithm of odds ratio (LOD score), which is the log of the odds that a QTL is present at a location to the odds that there is no QTL at that location. Experiment-wise empirical threshold values were estimated for the HAP population as suggested by Churchill and Doerge (1994). The original data set for each of the 5 traits was repeatedly generated as random numbers for the 93 individuals, with mean set at that of the original data. The generated data sets, along with the original trait values, were then analysed by conventional interval mapping. The analyses yielded 1700 LOD scores per generated data set. The top 1% of the LOD scores from each of 100 analyses were combined and the 95th and 99th percentile used to estimate experiment-wise Type 1 error rate at 5% and 1% error probability, respectively. The values obtained were 1.6 and 2.4, but to control Type II, a threshold value of 2.0 was used as the experiment-wise LOD score.

When multiple LOD peaks were detected within a 30-cM interval, the presence of two distinct QTL was inferred if the closest markers are linked in repulsion. In summaries of the biometrical information, the interval length bracketing the putative QTL is reported, as well as the peak LOD score and its recombinational distance from the closest flanking marker. The interval length defines the distance of the map position supported by 2-unit drop of the LOD score on either sides of the peak LOD position. Such an interval is supposed to show with its length the precision with which the putative QTL has been localised. The recombination distance measures the estimated distance of the peak LOD score from the closest marker.

The percentage of the population variance explained by individual QTL was calculated as a mean of all positions ≥ 2.0 LOD, using the formula (Ooijen and Maliepaard 1996):

$$100\ [(\sigma^2_{residual}-\sigma^2_{max})/(\sigma^2_{population})]$$

where,
$\sigma^2_{residual}$ = the residual variance under the null hypothesis,
σ^2_{max} = the variance associated with the maximum likelihood QTL position, and
$\sigma^2_{populatio}$= the population variance.

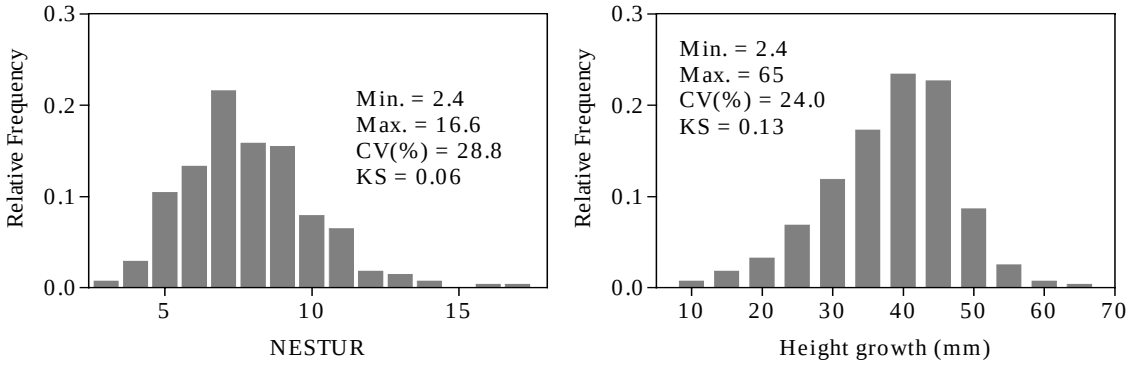
6.2.4.3 Analysis of epistasis

After apparent QTLs were located, digenic interactions were assessed between pairs of markers closest to the QTL. The analyses were performed using a simple factorial ANOVA model on SPSS for Windows, with the UNIQUE sum of squares, which evaluates all main and interaction effects simultaneously.

6.3 Results

6.3.1 Frequency of phenotypes

Frequency distribution patterns for NESTUR, and seedling growth or increment in height, diameter, stem volume, and needle volume are presented in Figure 6.1.



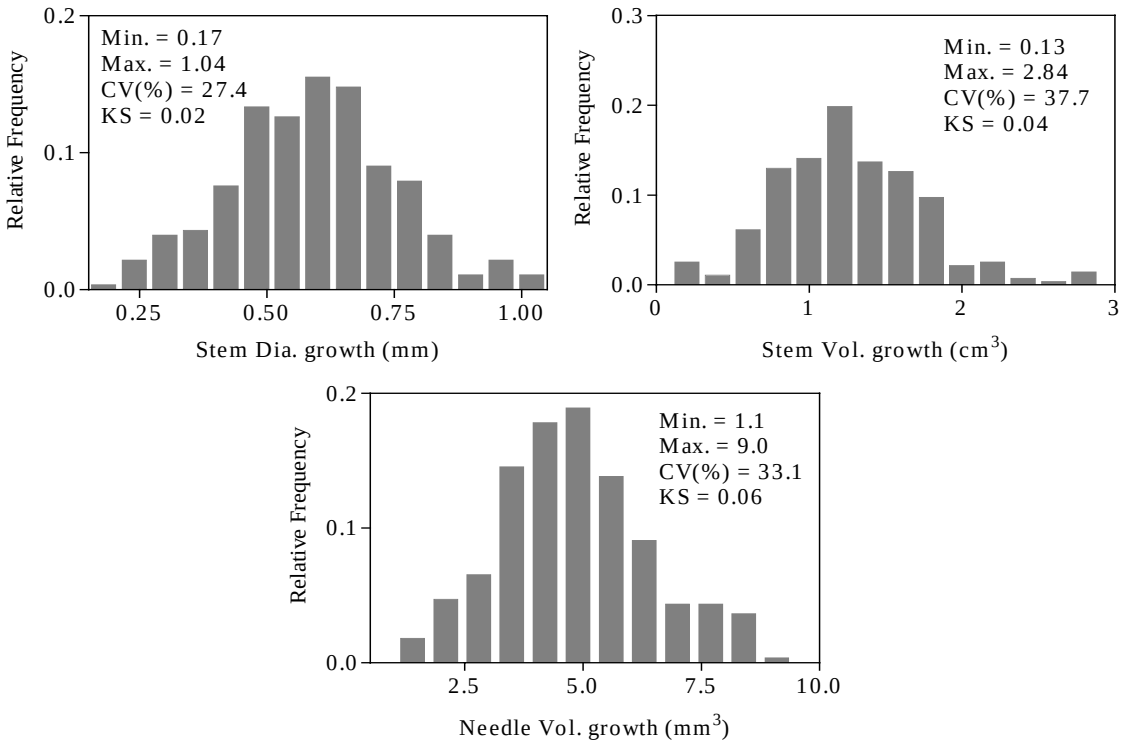


Figure 6.1 Frequency distributions of NESTUR and associated seedling growth traits in the full-sib family 30040 x 80121. Minimum (Min.), maximum (Max.), coefficient of variation (CV%) and Kolmogorov-Smirnov test for Gaussian distribution (KS distance) are indicated besides the histograms.

The five traits showed continuous variation, exhibiting a nearly normal distribution (as measured by the Kolmogorov-Smirnov test) without transformation. Height growth showed a significant (α = 0.01) KS distance statistic, but data transformation did not significantly improve normality. All analyses were therefore carried out with untransformed data.

NESTUR had a mean value of 7.8, with a minimum of 2.4 and a maximum value of 16.6. The phenotypic coefficient of variation (CV) was 28.8%. For stem growth traits, the CV ranged from 24% for height growth (Ht growth), to 38% for stem volume growth (SVol growth).

6.3.2 Linkage map re-construction

In order to use MAPQTL v. 3.0 (Ooijen and Maliepaard 1996) for interval mapping, linkage phase of marker loci on the map reported in Chapter 5 must be deduced. Inferences on linkage phase were made using the method described in section 6.2.3 and the linkage map re-constructed. Pair-wise estimates of recombination among 21 markers in the new linkage calculations exceeded 0.5. The markers were in two groups of the previous linkage map, groups 4 and 5, which had the largest number of markers (see chapter 5, Table 5.2).

The "suspect" estimates of recombination may have arisen due to random variation, which happens in particular at larger distances (see JOINMAP manual). It is also possible that these markers contained incorrectly scored bands, giving rise to an excess of double recombinants. As shown in chapter 5 (section 5.3.5; Figure 5.5), the observed (or direct) two-point map distance in some linkage groups deviated from that calculated from multipoint analysis. The markers involved, however, were deleted, and the linkage maps re-calculated *de novo*. In all, 14 markers in linkage group 4 and 7 in group 5 of the previous linkage map were deleted.

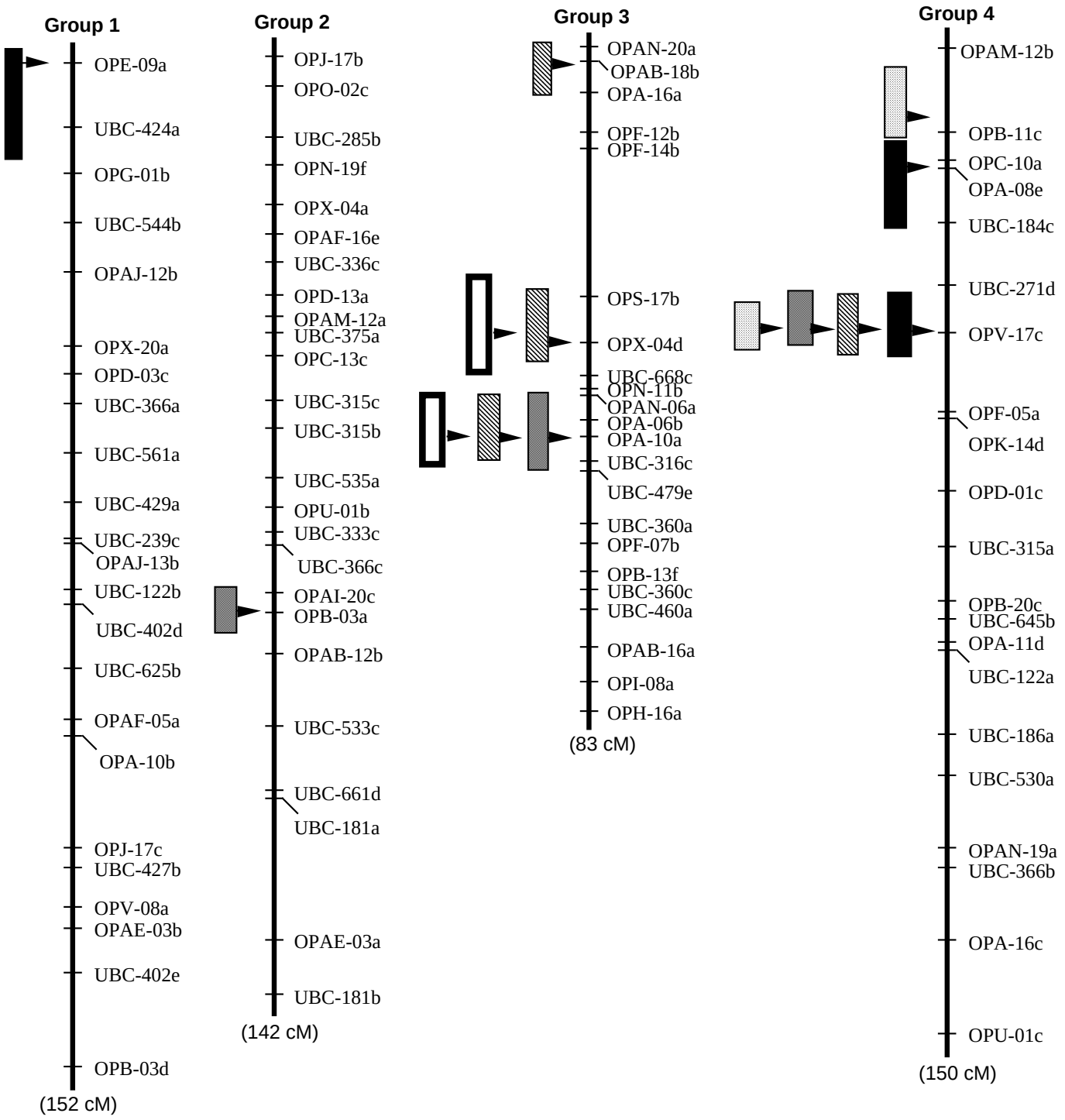
The current map showed an increase of only 4% in total map distance, compared with the previous linkage map constructed from markers with unknown phase (Table 6.1). The total map distance covered was 1567 cM, with an average of 7 cM per interval. The linkage map contained 238 markers, with three unlinked markers. Marker linkage groupings remained the same. However, linkage group 5 of the previous linkage map was split in two (now designated 5a and 5b; Figure 6.2). In addition, the three unlinked markers in the current linkage calculations previously mapped to linkage group 5.

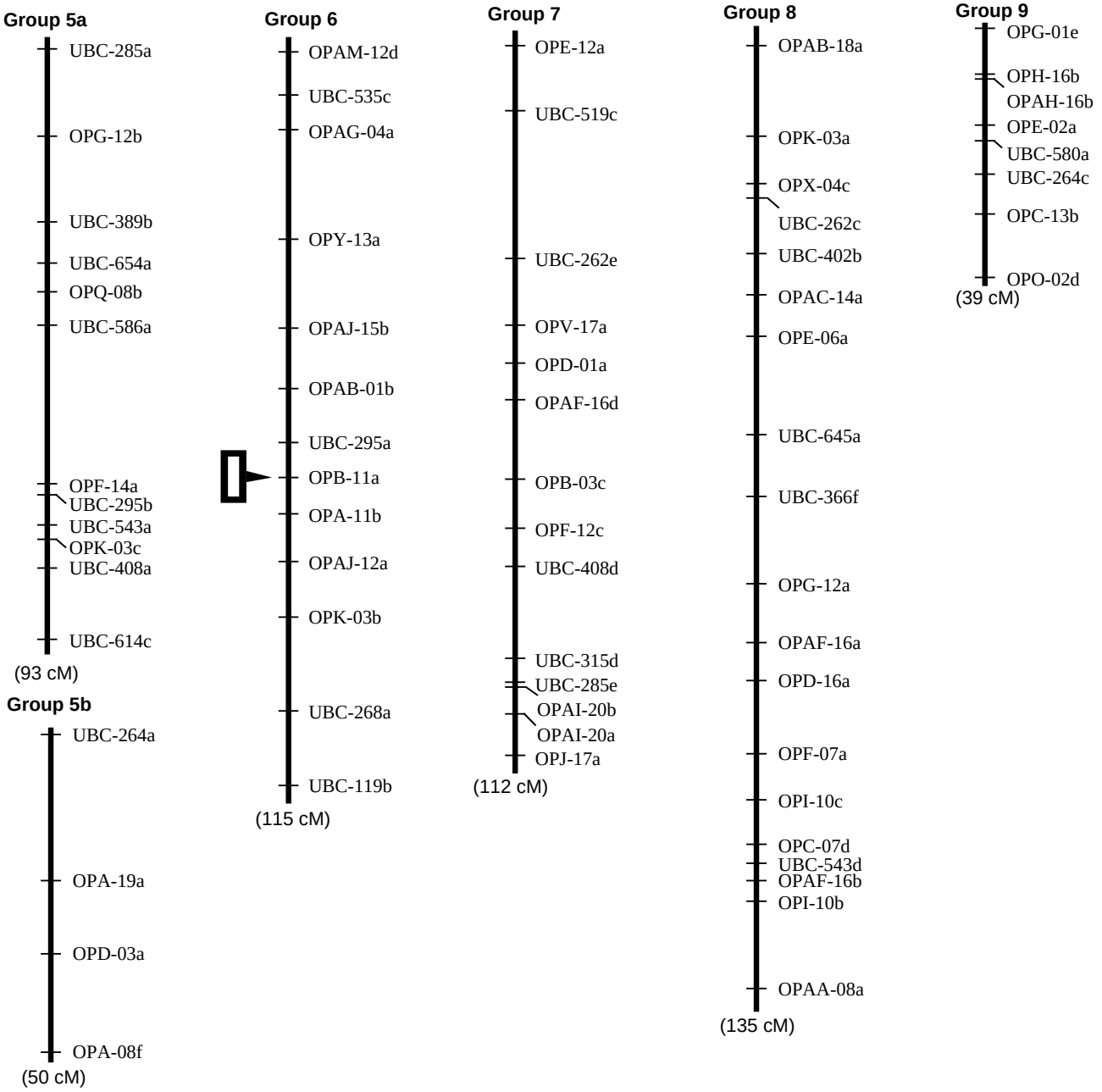
Homogeneity of marker order was tested graphically by plotting the linear order of markers in the previous map (chapter 5; Figure 5.3) and current linkage map. Markers in linkage groups 4 and 5 were not included in the analysis. Deviation from a line of unit slope would indicate heterogeneity of marker order in the present and previous maps. As shown in Figure 6.3, marker order was not significantly altered when locus phase was inferred, and linkage map re-calculated. Only 4% of the markers switched order when phase was inferred using MAPMAKER/EXP.

Table 6.1 Number of linkage groups and genetic length of the previous linkage map constructed from markers with unknown phase (chapter 5) , compared with the current map constructed with markers whose phases were inferred using MAPMAKER/EXP.

	Number of markers		Genetic length (cM)	
Linkage group	Phase-unknown	Phase-inferred	Phase-unknown	Phase-inferred
1	24	24	148.0	152.1
2	25	25	140.7	142.4
3	23	23	103.0	106.6
4	37	23	183.4	169.3
5	30	-	144.5	-
5a	-	16	-	93.4
5b	-	4	-	50.1
6	13	13	115.3	115.5
7	14	14	80.0	111.8

	Number of markers		Genetic length (cM)	
Linkage group	Phase-unknown	Phase-inferred	Phase-unknown	Phase-inferred
8	20	20	149.1	148.7
9	8	8	39.4	39.4
10	7	7	55.1	55.1
11	17	17	111.2	136.8
12	22	22	131.4	136.6
13	15	15	71.0	70.0
14	7	7	39.0	39.0
Total	262	238	1511.1	1566.8
Mean	19	17	107.9	111.9


Group 1
OPE-09a
UBC-424a
OPG-01b
UBC-544b
OPAJ-12b
OPX-20a
OPD-03c
UBC-366a
UBC-561a
UBC-429a
UBC-239c
OPAJ-13b
UBC-122b
UBC-402d
UBC-625b
OPAF-05a
OPA-10b
OPJ-17c
UBC-427b
OPV-08a
OPAE-03b
UBC-402e
OPB-03d
(152 cM)
Group 2
OPJ-17b
OPO-02c
UBC-285b
OPN-19f
OPX-04a
OPAF-16e
UBC-336c
OPD-13a
OPAM-12a
UBC-375a
OPC-13c
UBC-315c
UBC-315b
UBC-535a
OPU-01b
UBC-333c
UBC-366c
OPAI-20c
OPB-03a
OPAB-12b
UBC-533c
UBC-661d
UBC-181a
OPAE-03a
UBC-181b
(142 cM)
Group 3
OPAN-20a
OPAB-18b
OPA-16a
OPF-12b
OPF-14b
OPS-17b
OPX-04d
UBC-668c
OPN-11b
OPAN-06a
OPA-06b
OPA-10a
UBC-316c
UBC-479e
UBC-360a
OPF-07b
OPB-13f
UBC-360c
UBC-460a
OPAB-16a
OPI-08a
OPH-16a
(83 cM)
Group 4
OPAM-12b
OPB-11c
OPC-10a
OPA-08e
UBC-184c
UBC-271d
OPV-17c
OPF-05a
OPK-14d
OPD-01c
UBC-315a
OPB-20c
UBC-645b
OPA-11d
UBC-122a
UBC-186a
UBC-530a
OPAN-19a
UBC-366b
OPA-16c
OPU-01c
(150 cM)


Group 5a
UBC-285a
OPG-12b
UBC-389b
UBC-654a
OPQ-08b
UBC-586a
OPF-14a
UBC-295b
UBC-543a
OPK-03c
UBC-408a
UBC-614c
(93 cM)
Group 5b
UBC-264a
OPA-19a
OPD-03a
OPA-08f
(50 cM)
Group 6
OPAM-12d
UBC-535c
OPAG-04a
OPY-13a
OPAJ-15b
OPAB-01b
UBC-295a
OPB-11a
OPA-11b
OPAJ-12a
OPK-03b
UBC-268a
UBC-119b
(115 cM)
Group 7
OPE-12a
UBC-519c
UBC-262e
OPV-17a
OPD-01a
OPAF-16d
OPB-03c
OPF-12c
UBC-408d
UBC-315d
UBC-285e
OPAI-20b
OPAI-20a
OPJ-17a
(112 cM)
Group 8
OPAB-18a
OPK-03a
OPX-04c
UBC-262c
UBC-402b
OPAC-14a
OPE-06a
UBC-645a
UBC-366f
OPG-12a
OPAF-16a
OPD-16a
OPF-07a
OPI-10c
OPC-07d
UBC-543d
OPAF-16b
OPI-10b
OPAA-08a
(135 cM)
Group 9
OPG-01e
OPH-16b
OPAH-16b
OPE-02a
UBC-580a
UBC-264c
OPC-13b
OPO-02d
(39 cM)

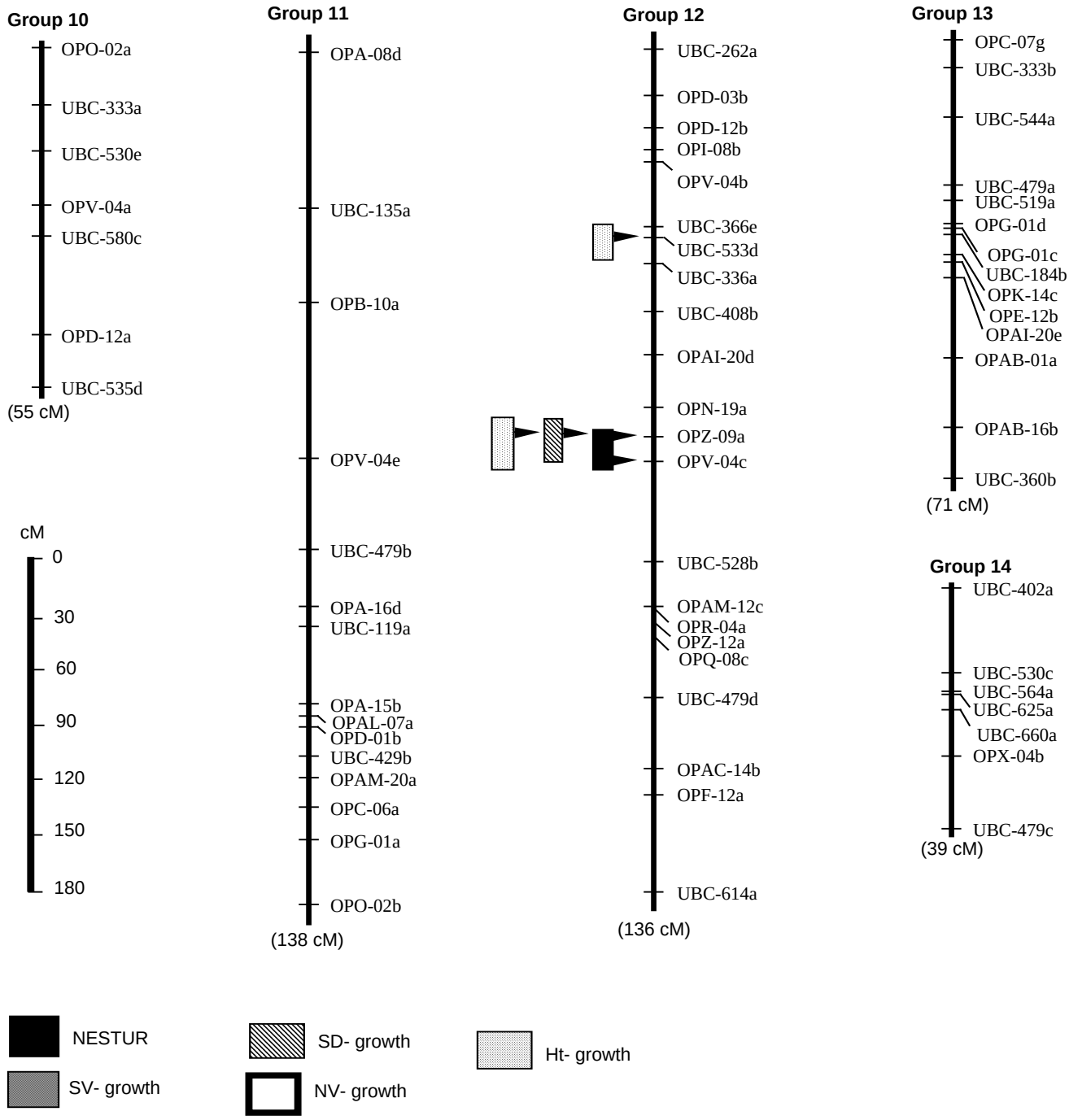


Figure 6.2 Quantitative trait locus (QTL) map of NESTUR and associated stem growth variables in radiata pine seedlings. Linkage map of RAPD markers was constructed using a haploid population progeny of a heterozygous diploid individual. QTL detection was by MAPQTL with linkage phase of marker loci inferred using MAPMAKER/EXP. Bars to the left of linkage groups correspond to intervals over which QTLs are most likely to be located.

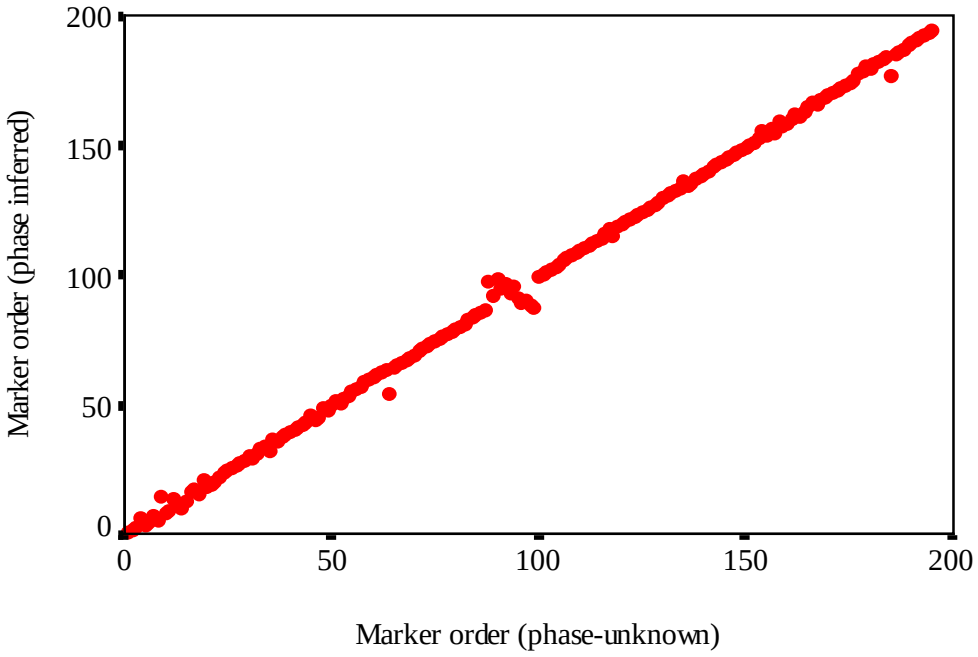


Figure 6.3 Scatter plot of marker order in linkage maps constructed with markers of known and inferred phases. Homogeneity of marker order is indicated by the linear distribution of common markers along a line of unit slope.

6.3.3 QTL detection by single-point analyses

Single-point analyses detected six markers associated with NESTUR at $F_p \leq 0.01$(Table 6.2). Those with the highest degree of association ($F_p \leq 0.001$) were OPA-08e and OPV-17c, both on linkage group 4, with phenotypic effects of different directions (Table 6.2). Other significant linkages observed were those of OPE-09a on linkage group 1, OPZ-09a and OPV-04c, both on linkage group 12, and OPX-04a on linkage group 2.

Ten (10) RAPD loci showed significant linkage to stem diameter increment (SD growth) in the present study (Table 6.2). These were distributed on 3 linkage groups, with an additional marker that could not be placed on any linkage group during linkage map construction. The RAPD markers OPE-09a on linkage group 1, and OPX-04a on linkage group 2, which were significantly linked to SD growth, had earlier been found to be linked to NESTUR.

Table 6.2 Linkage group, marker loci, and single-locus effects of RAPD markers significantly associated with NESTUR and stem growth traits.

Trait	Linkage group	Marker locus	F_p-value	Phenotypic effect (SD units)	Explained variation (%)
NESTUR	1	OPE-09a	0.0024	-0.63	12.4
	2	OPX-04a	0.0063	0.58	9.1
	4	OPA-08e	0.0007	-0.70	13.5
	4	OPV-17c	0.0006	0.73	15.5
	12	OPV-04c	0.0078	-0.55	8.1
	12	OPZ-09a	0.0069	0.57	8.6
Total					40.1
SD growth	3	UBC-316c	0.0013	-0.66	12.6
	3	OPA-06b	0.0010	-0.71	14.4
	3	OPAN-06a	0.0038	-0.60	9.8
	3	OPN-11b	0.0053	0.59	9.2
	3	UBC-668c	0.0090	-0.56	8.2
	3	OPX-04d	0.0010	0.68	12.9
	3	OPF-12b	0.0061	-0.57	9.8
	3	OPAB-18b	0.0025	0.61	11.1
	4	OPV-17c	0.0010	0.71	13.9
	6	OPB-11a	0.0049	-0.58	9.3
	unlinked	UBC-167a	0.0110	0.54	7.9
Total					50.2
SVol growth	3	OPA-06b	0.0032	-0.63	11.4
	3	OPX-04d	0.0045	0.59	9.5
	3	OPAB-18b	0.0074	0.56	8.7
	4	OPV-17c	0.0005	0.74	15.4
	6	OPB-11a	0.0042	-0.59	9.7
	unlinked	UBC-530d	0.0119	0.52	7.6
	unlinked	UBC-167a	0.0110	0.57	9.0
Total					42.0
Ht growth	4	OPV-17c	0.0062	0.59	9.4
	4	OPB-11c	0.0065	-0.58	8.6
	12	UBC-533d	0.0023	0.65	11.8
	12	OPZ-09a	0.0069	0.60	10.1
	12	OPN-19a	0.0137	0.51	7.0
	unlinked	UBC-530d	0.0034	0.62	10.4
Total					26.3
NVol growth	3	UBC-316c	0.0094	-0.55	8.1
	3	OPA-06b	0.0071	-0.59	9.4
	3	UBC-668c	0.0106	-0.54	7.8
	3	OPX-04d	0.0042	0.59	9.5
	3	OPS-17b	0.0068	0.60	9.3
	5a	UBC-668a	0.0050	-0.62	10.5
	6	UBC-119b	0.0119	-0.52	7.2
	6	OPA-11b	0.0055	0.57	8.9
	6	OPB-11a	0.0009	-0.70	13.3
Total					33.9

For SVol growth, seven (7) RAPD markers were detected with significant associations at Fp ≤ 0.01 (Table 6.2). These were essentially the same markers found to be linked to QTLs for SD growth, except for UBC-530d that was not

placed on any of the linkage groups. The standardised effect of allele substitution (see chapter 4, Figure 4.6) ranged from -0.58 standard deviation units to 0.65.

Regression analyses of needle volume increment (NVol growth) with RAPD markers identified nine loci that were significantly ($\alpha \leq 0.01$) associated with the trait. These loci were distributed on 3 linkage groups, with the highest concentration on linkage groups 3 and 6.

6.3.4 QTL detection by interval mapping

Interval mapping was performed using the conventional method and the multiple-QTL (or composite) method, which uses cofactor markers. Results from both methods of interval mapping were not generally in agreement, as shown in the likelihood maps (Figures 6.4, 6.5, 6.6) and in Table 6.3. For NESTUR, conventional interval mapping detected 3 putative QTL positions on linkage groups 1 and 4, but multiple-QTL mapping detected 5 QTL distributed on linkage groups 1, 4, and 12 (Table 6.3). Both interval mapping methods detected a QTL on linkage groups 1 and 4. The single QTL on linkage group 1 had the LOD peak exactly at the marker OPE-09a, and explained 11% of the population variance (Table 6.3).

Two putative QTL were detected on linkage group 4, linked in repulsion, and separated by 26 cM from each other (Table 6.3; Fig 6.4). Both QTL had positive additive effects. The QTL position closest to marker OPA-08e had a major effect on NESTUR, explaining 21% of the population variance, as estimated from both methods of interval mapping (Table 6.3). About 0.2 cM from the marker OPV-17c, a putative QTL position was detected by both methods, but

the amount of variation associated with the QTL was higher with the conventional, relative to the multiple-QTL (MQM) mapping method.

A case of discrepancy in results from the two methods was found in linkage group 12 (Table 6.3; Figure 6.5). The MQM mapping detected two QTL linked in repulsion, but the LOD score from conventional mapping was below the threshold value. A similar case of disagreement occurred for a QTL detected for NVol growth in linkage group 6 (Table 6.3). The LOD score from MQM mapping was 2.18, but conventional mapping gave a lower LOD score of 1.48 (Table 6.3). One case was observed where the reverse situation occurred, ie, a QTL detected by conventional interval mapping failed to be significant under the MQM mapping. In linkage group 12, a putative QTL was detected for Ht growth by conventional interval mapping, but the LOD score from MQM mapping was below the threshold value (Table 6.3).

Table 6.3 Biometrical parameters of detected QTL affecting NESTUR, SVol growth, Ht growth and NVol growth in the radiata pine family 3040 x 80121.

	Linkage group	QTL interval length (cM)	Peak LOD		Recomb. Dist. (cM)		% var. expl.		Closest Marker	[a]Phase type	Dir. of add. effects
			CQTL[b]	MQM[c]	CQTL	MQM	CQTL	MQM			
NESTUR	1	16.4	2.00	2.75	0.0	0.0	11.0	10.9	OPE-09a	{0}	+
	4	17.9	2.94	5.16	2.0	1.0	21.4	21.5	OPA-08e	{0}	+
	4	13.1	3.47	2.32	0.2	0.2	18.2	9.0	OPV-17c	{1}	+
	12	4.0	1.53	3.10	3.0	2.0	7.7	10.6	OPV-04c	{1}	-
	12	4.3	1.25	3.08	2.0	2.0	10.2	18.1	UBC-336a	{0}	-
SD growth	3	12.3	3.02	2.31	0.0	0.0	13.9	9.0	OPA-10a	{1}	-
	3	14.8	2.23	2.35	1.0	0.0	11.8	9.1	OPX-04d	{0}	-
	3	7.7	2.42	2.50	0.0	0.0	11.7	9.9	OPAB-18b	{0}	-
	4	13.1	3.24	2.60	0.2	0.0	17.2	9.4	OPV-17c	{1}	+
	12	8.8	2.00	2.06	0.0	0.0	9.3	7.3	OPZ-09a	{0}	-
SV growth	2	5.8	2.49	2.64	3.0	3.0	16.2	15.0	OPB-03a	{0}	+
	3	12.3	2.18	2.28	0.0	0.0	10.2	10.5	OPA-10a	{1}	-
	4	6.8	3.43	2.36	0.2	0.2	18.1	9.9	OPV-17c	{1)	+
Ht growth	4	15.6	2.04	2.08	0.8	1.8	10.8	12.4	OPB-11c	{0}	+
	4	1.0	2.28	2.34	1.2	1.2	18.9	19.3	OPV-17c	{1}	+
	12	4.0	2.05	1.68	0.0	0.0	9.7	6.4	OPZ-09a	{0}	-
	12	4.3	2.49	2.49	1.3	0.3	16.4	10.7	UBC-533d	{0}	-
NV growth	3	14.8	2.00	2.29	3.8	2.8	11.6	11.4	OPX-04d	{0}	-
	5a	10.8	1.48	2.18	0.0	0.0	7.1	7.5	UBC-614c	{0}	+
	6	5.8	2.40	2.53	1.3	2.0	13.3	10.4	OPB-11a	{1}	-

[a] Locus pairs with the same phase code are assumed to be in coupling in the parent, and in repulsion otherwise.
[b] CQTL = Conventional interval mapping method; [c] MQM = Multiple-QTL interval mapping method.

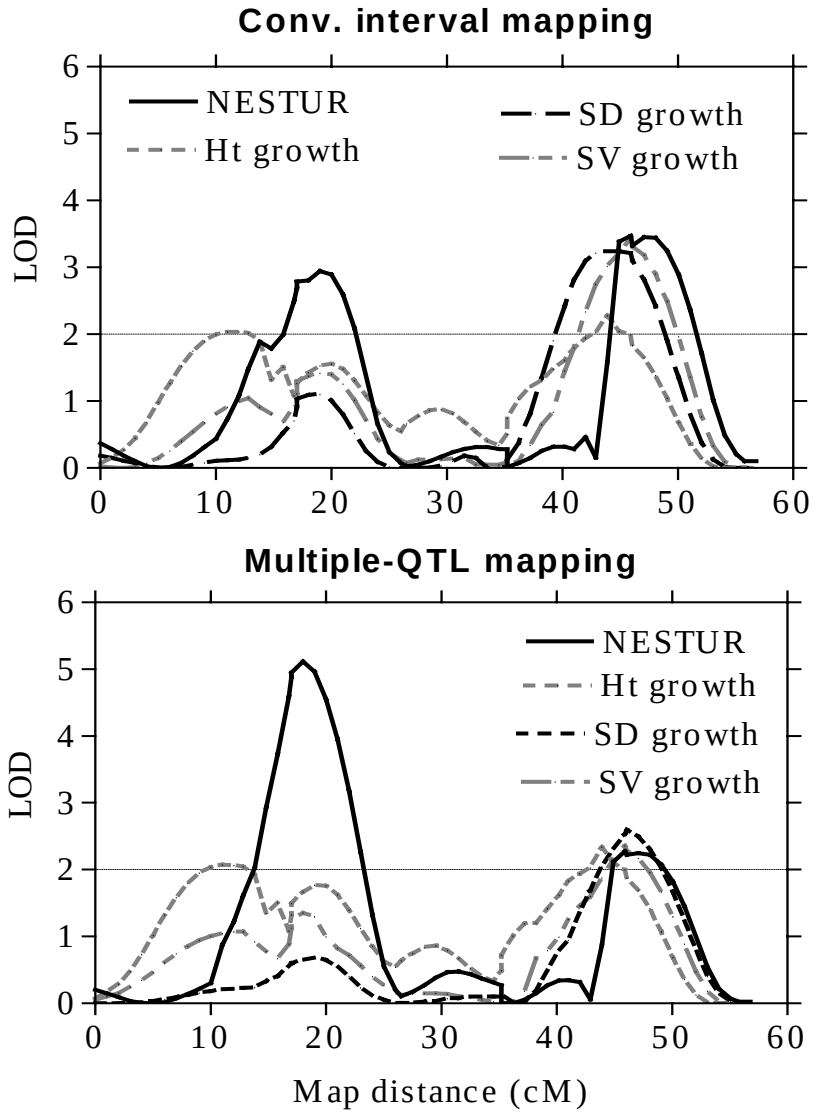


Figure 6.4 Logarithm of odds ratio (LOD) plot of significant genomic region in group 4, showing scan profiles for NESTUR, height growth (Ht growth), volume growth (SVol growth), and diameter growth (SD growth). The LOD scores were calculated using MAPQTL programs for conventional and multiple-QTL interval mapping methods. Putative QTLs were inferred within an interval with a LOD score of ≥ 2.0.

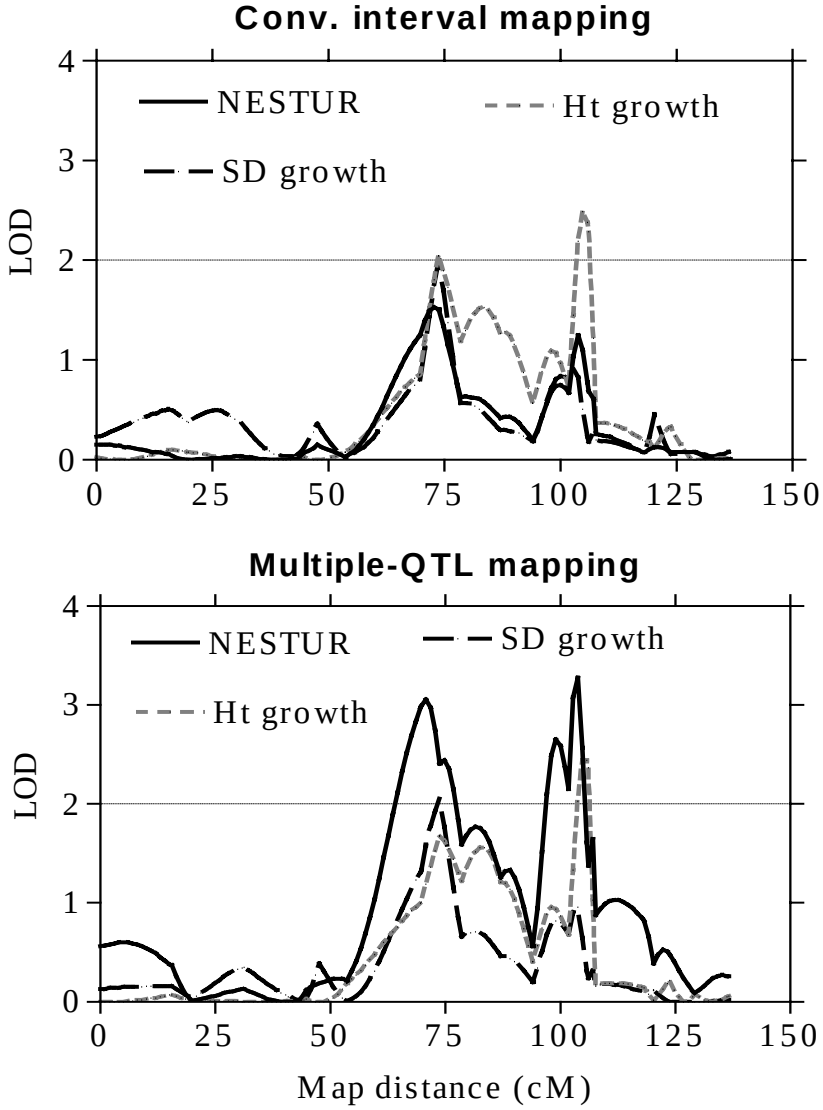


Figure 6.5 Logarithm of odds ratio (LOD) plot of significant genomic region in group 12, showing scan profiles for NESTUR, height growth (Ht growth), and diameter growth (SD growth). The LOD scores were calculated using MAPQTL programs for conventional and multiple-QTL interval mapping methods. Putative QTLs were inferred within an interval with a LOD score of ≥ 2.0.

QTL coincidence can be used to tentatively resolve some of the observed discrepancies between the two methods of interval mapping. For traits that are significantly correlated at the phenotypic level, coincidence of some of their QTL is likely to be present. The QTL for NESTUR located close to marker OPV-17c, had a significant effect on SD, SVol and Ht growths, explaining a similar

magnitude of the population variance for these traits (Table 6.3). The two putative QTL in linkage group 12 which were detected only by multiple-QTL mapping were coincidental with those for SD, SVol, and Ht growths (Figure 6.5). These may therefore not be artifactual. Moreover, markers closest to the two QTL were significantly ($P \leq 0.01$) associated with NESTUR by single-point analyses (Table 6.2).

6.3.5 Epistasis between QTL

Two-factor interaction analyses performed using markers closest to putative QTL positions showed no evidence of digenic interaction between the QTL and genomic regions associated with measured traits. The F-statistic showed a significance level of 0.07 to 0.98, suggesting that epistasis may not be a major factor contributing to observed phenotypic variation for these traits in this cross.

6.3.6 Trait correlations and pleiotropic effects

Marker loci in this study were also used to examine the relationships between pairs of measured traits. A correlation matrix of the various traits at the phenotypic level is presented in Table 6.4. All the observed correlations were positive and significant, except for the relationship of NESTUR with growth in needle volume (NVol growth), which was not significant. Magnitude of the correlation coefficients suggests that, in the current mapping population, high NESTUR was primarily associated with a high growth rate in stem diameter, stem volume, and height.

Table 6.4 Phenotypic correlation coefficients for NESTUR and associated seedling growth traits.

	NESTUR	SD growth	SVol growth	NVol growth
SD growth	0.76**	-		

SVol growth	0.70**	0.87**	-	
NVol growth	0.18	0.50**	0.70**	-
Ht growth	0.60**	0.53**	0.73**	0.55**

Table 6.5 LOD-score correlations for NESTUR and associated seedling growth traits. Correlations were calculated from 1022 LOD scores in 7 linkage groups.

	NESTUR	SD growth	SVol growth	NVol growth
SD growth	0.45**	-		
SVol growth	0.40**	0.76**	-	
NVol growth	-0.06	0.35**	0.54**	-
Ht growth	0.38**	0.47**	0.57**	0.35**

Correlation coefficients for pairs of measured traits were calculated using their LOD scores to assess pleiotropic effects of putative QTL-linked genomic regions. The LOD scores were obtained by scanning at 1-cM intervals the 7 linkage groups where putative QTLs were detected for all measured traits. As would be expected, the strongest genetic correlation was between SD growth and SVol growth (Table 6.5). The LOD scores for NESTUR and NVol growth were not significantly related, suggesting a lack of genetic relationship between the two traits. The LOD scores for NESTUR and stem growth in diameter, volume and height were significantly different from 0 and 1. Thus, some putative QTL positions for NESTUR were coincidental with those for its components, but there were some significant QTL not shared with the component traits.

The observed pattern of trait phenotypic and genetic correlations agreed well with that from QTL clustering and coincidence. The putative QTLs detected for NESTUR and stem growth variables occurred mainly in linkage groups 3, 4 and

12. In linkage group 3, putative QTL for SD growth, SVol growth and NVol growth were detected in similar or adjacent locations (Figure 6.6). There was no evidence of a putative QTL for NESTUR or Ht growth within the genomic region. The significant genomic region of linkage group 4 also exerted strong effects on SD growth, SVol growth and Ht growth, but not on NVol growth (Figure 6.4). In addition, in linkage group 12, QTL occurred in coincidental locations for NESTUR, SD growth, and Ht growth (Figure 6.5). The evidence was insufficient to indicate existence of a QTL for SVol or NVol growth within this region. Factors affecting SVol growth and NVol growth were detected in linkage groups 2, 5a and 6, in regions not shared with significant QTLs for any of the other traits (Figure 6.2).

6.4 Discussion

6.4.1 Linkage map calculations

In the present study, a linkage map generated from RAPD markers in a haploid population was used to establish co-segregation of markers and putative QTL affecting stem growth efficiency in radiata pine seedlings. The linkage map had previously been constructed with markers of unknown phase (see chapter 5).

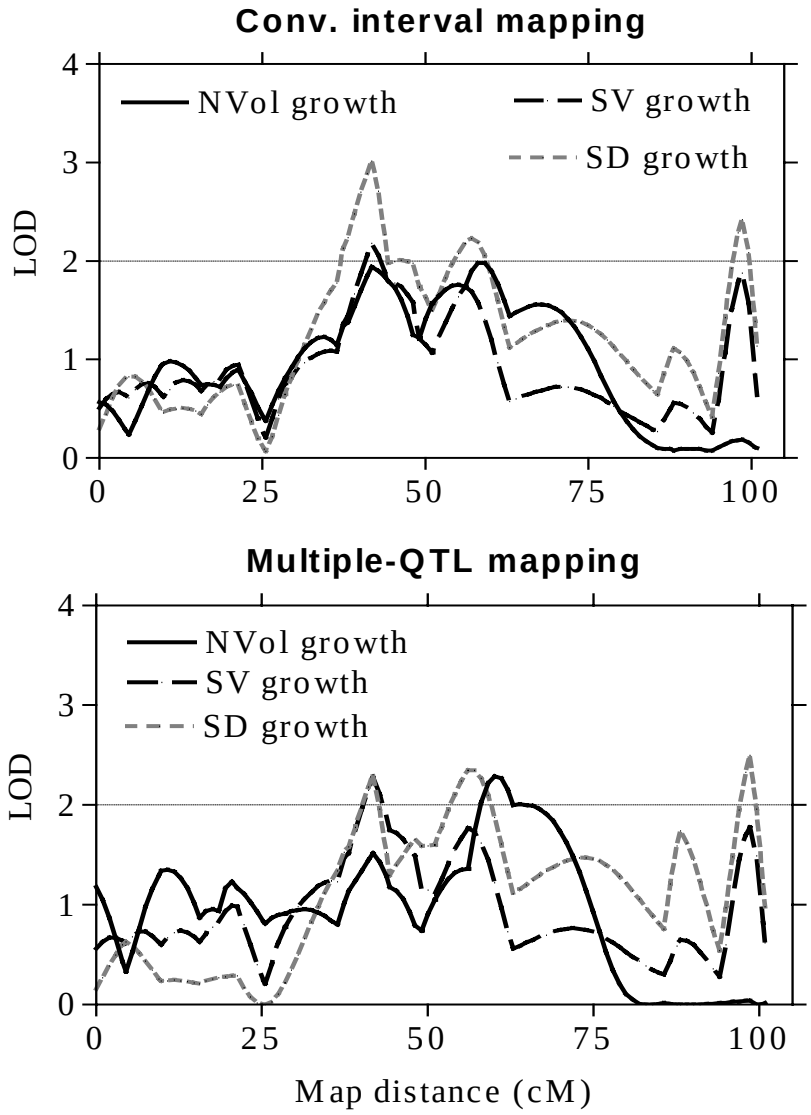


Figure 6.6 Logarithm of odds ratio (LOD) plot of significant genomic region in group 3, showing scan profiles for stem diameter growth (SD growth), volume growth (SVol growth) and needle volume growth (NVol growth). The LOD scores were calculated using MAPQTL programs for conventional [*dotted lines*] and multiple-QTL interval mapping methods. Putative QTLs were inferred within an interval with a LOD score of ≥ 2.0.

However, for QTL analyses using interval mapping methods, linkage phase of the markers must be known or inferred.

Linkage phase was inferred using MAPMAKER/EXP, and the map re-calculated. Marker groupings remained the same and total map length

increased by only 4% when the phase was inferred and map re-calculated. Counting each marker out of order in the previous map (with unknown phase) as an event (eg. a switch in the order of two adjacent markers counts as one event), only 4% of the markers changed order when linkage phase was inferred. Because the same locus order is expected, comparison of the two maps provided a measure of repeatability of the genomic linkage map construction. Two maps from the same individual genotype should closely resemble each other if the experimental methods used to produce the genetic markers and the statistical methods for constructing the genomic maps are sufficiently rigorous (Plomion *et al*. 1995b). Because of the large number of possible locus order (n! /2 for n loci), the comparison of genomic maps primarily depends on the accurate determination of locus order. A highly accurate order is important for gene isolation by map-based cloning. The current linkage map was constructed for the genetic analysis of quantitative traits, thus the accuracy of the locus order is adequate for this purpose.

6.4.2 Limitations of RAPDs for QTL dissection in out-breeding species

In this paper, we applied marker information derived from haploid populations to traits measured on diploid individuals. This was made possible by the use of RAPD markers, which are generally dominant, rather than co-dominant. Although dominant markers have commonly been used in linkage map construction of forest tree species, they are not very efficient for QTL analysis because only the additive mode of gene action can be inferred. Moreover, a dominant marker heterozygous in both parents will not be assayed since such markers will be monomorphic in the progeny. The occurrence of multiple alleles at a locus in offspring of out-breeding species (Byrne *et al*. 1997; Groover at al. 1994; van Eck *et al*. 1994) also limits the efficiency of QTL dissection with

dominant markers in an out-crossing species. Since only markers segregating from maternal parent are assayed, there is the possibility that paternally segregating alleles will confound maternal alleles through intra-locus interactions.

Thus, one of the potential problems identified for QTL mapping in haploid populations derived from megagametophytes is the appropriate significance threshold for declaring a significant marker-QTL linkage. Since only markers segregating from the maternal parent are assayed, theoretically only half of the total number of QTLs will be detectable when QTLs are segregating from both parents. With a population of less than 200 individuals, only QTLs with large additive effects (with explained variation > 10%) are likely to be detected at the commonly used LOD threshold of 3.0. QTLs of small effects, which may collectively explain a large proportion of genetic variation, will escape detection. Using a low LOD threshold, however, will increase the probability of Type 1 error (ie. accepting the hypothesis of marker-linked QTL when such an association does not exit).

Lander and Schork (1994) advocated the use of simulation-based tests to directly estimate the false positive rate appropriate for a specific experimental setting. Such tests have been applied to problems of genome-wide search in QTL mapping (Cheverud *et al*. 1996; Horvat and Medrano 1995), and described by Churchill and Doerge (1994). The method used in the present study was to randomly generate the original data set for each trait 100 times, and then apply interval mapping to each of the 100 replications. While the 99^{th} percentile value is quite stringent enough to control Type I error, reducing the probability of Type I error may increase the probability of Type II error (rejection of the hypothesis of no association between a marker and a QTL when in fact such an

association exists). The intent of this study is to form a basis for marker-assisted selection, and Type II errors are of greater importance than when the objective is to locate and clone a gene (such as in human genetics). The LOD score of 2.0 adopted for the present study, nevertheless, compares well with values suggested from several numerical calculations made on simulated data. Darvasi et al. (1993) estimated numerically that, at a marker spacing of < 10 cM and 11 markers per chromosome, a LOD score of 1.53-corresponding to a per-marker Type I error rate of 0.0084-ensures a 0.05 per chromosome Type I error. With an infinite number of markers per chromosome, a LOD score of 1.96 would be adequate to ensure an error rate of 5%. Recently, Doerge and Rebai (1996) came to a similar conclusion from simulated data for a backcross progeny, and determined that a threshold value of 1.71 was adequate to control Type I error at 5% probability.

6.4.3 Comparison of QTL mapping methods

Most QTL positions detected from interval mapping showed LOD peaks closest to markers identified from single-point analyses (Table 1 and 2). Similarity between single-point analyses and interval mapping in QTL detection has been observed previously by several researchers (Grattapaglia *et al*. 1995; Stuber *et al*. 1992).

On the other hand, results from the two methods of interval mapping showed some discrepancies in LOD scores and explained variation. This was most evident in linkage group 12 (Fig. 2). Two linked QTLs were detected for NESTUR by MQM mapping whereas LOD scores from conventional interval mapping were below the threshold value (Fig. 2). Both QTLs are linked in repulsion, a configuration that renders the QTLs difficult to detect by conventional interval mapping, because the effects cancel each other (Korol et

al. 1994; Ooijen 1992). The two putative QTLs detected in linkage group 4 were also linked in repulsion, but were detectable by conventional interval mapping because they were less tightly linked. Using simulated data, Utz and Melchinger (1994) showed that linked QTLs in repulsion phase can be detected and separated much more powerfully by MQM mapping with marker cofactors than by conventional interval mapping.

6.4.4 Perspective

In the present study, significant genomic regions were found to be associated with stem increment per unit needle volume growth (NESTUR). A similar finding was made from a bulked segregant analysis approach in an unrelated radiata pine family (Emebiri *et al*. in press). In that study, linkage to components of NESTUR (increments in stem diameter and stem volume) was demonstrated for one of the identified QTL, while the other was not linked to NESTUR components. Of the 5 QTLs detected for NESTUR in the present study, two were coincidental with those for SD growth and Ht growth, and one with that of SVol growth. It is possible that these multiple effects are due to multiple QTLs in the marker-linked genomic region, rather than being representative of pleiotropy. The present data are however, insufficient to distinguish between the two possibilities.

7 CHAPTER 7

Experiment 4: Marker prediction of NESTUR and seedlings with superior growth potential

7.1 Introduction

The identification of molecular markers linked to NESTUR was reported in chapter 6, as a first step in assessing their potential to enhance selection efficiency in radiata pine. Marker-linked QTLs, however, are statistically defined entities and have the potential to be artifacts. The detection of a QTL provides an opportunity to test several hypotheses regarding marker-assisted selection in specific crosses where QTLs have already been detected. Theoretically, selection can be made more efficient by using a set of molecular tags that can identify various regions of the genome influencing the character. These molecular markers serve as indirect selection criteria for the desired trait. Thus, markers associated with QTLs should be able to predict the phenotype. Secondly, because molecular markers are not affected by environmental factors, use of markers linked to early selection traits will provide better prediction of later-age performance.

Stem growth efficiency is one of the most important traits in forest tree breeding programs. Indirect early selection is usually applied to provide the maximum genetic gain per unit time in long-lived forest tree breeding programs. However, for most growth and volume traits in forest trees, the phenotypic value of an individual is a poor predictor of its genetic value. Differential responses to environmental factors reduce age-age genetic correlation between trait, and effective early selection.

The failure of measurements of single-process components of stem growth to predict field performance (Greenwood and Volkaert 1992) supports the view

that growth variables are under the control of large number of genes, each with a small additive effect (Neale and Williams 1991; Greenwood and Volkaert 1992). Derived traits, such as NESTUR, which is described by a function analogous to the same process underlying stem growth in mature trees (Matheson[1]), may provide more reliable indications of later-age growth performance.

A problem with selection based on NESTUR *per se* is that the trait requires a large amount of labour for its measurement. For practical breeding applications, thousands of progeny often need to be evaluated to identify superior individuals. Molecular-based selection procedures are particularly useful for traits that are difficult or tedious to measure. Detection of strong marker-trait associations afford the breeder an opportunity for the directed manipulation of specific genes through marker-assisted selection.

The aims of the present study were three-fold: (i) to assess predictive ability of six RAPD markers, which were detected in experiment 3 (chapter 6; Table 6.3 & 6.3) to be associated with NESTUR, (ii) compare accuracy of predicting seedlings with superior growth potential using NESTUR *per se* and linked markers, and (iii) compare growth of individuals with favourable and unfavourable NESTUR-linked markers under field conditions.

7.2 Materials and methods

7.2.1 Plant materials

The plant materials comprised the 284 progenies of the controlled cross (30040 x 80121) used in chapter 5. All progeny were measured for quantitative traits, but only ninety-three were genotyped for RAPD markers, as reported in section

[1] Matheson, A. C. CSIRO Forestry & Forest Products, Banks Street, Yarralumla, P. O. Box 4008, Queen

5.2.1. A linkage map constructed with RAPDs was used for QTL analysis as described in chapter 6. Details of six of the RAPD markers with significant linkage to NESTUR are shown in Table 6.2.

7.2.2 Field planting and trait measurements

Seedlings were assessed for NESTUR as described in chapter 4, section 4.2.1. The seedling trees were then planted out in the field, on a CSIRO Forestry and Forest Products research plot at Cotter Road, Weston, Canberra. Field preparation involved ploughing and ridging of the experimental plot, which was then fenced. The seedling trees were planted on the ridges at a spacing of 2.5 m, with a border row of radiata pine trees. Field maintenance comprised a combination of slash weeding and application of herbicide (glyphosate). Field irrigation was commenced intensively after planting, using overhead sprinklers, and maintained as necessary through the first year of growth.

Field growth performance was assessed after the first and second years of growth in the field. Total height was measured from ground level to the tip of the growing shoot. Basal diameter was measured as the average of two measurements taken at right angles from 20 mm above ground level. Stem volume was derived as total height x stem basal area. Data were also collected on the number of branch whorls, total number of branches and crown-height, measured as the distance from the first whorl to the growing tip.

7.2.3 Regression analyses

Since no single marker exerted sufficiently large effects to divide the progeny into high and low NESTUR classes, all the six RAPDs significantly ($\alpha \leq 0.01$) linked to NESTUR (see chapter 6 and Table 6.2) were used. The markers were

used in a multiple linear regression model to predict observed values of NESTUR. Each marker, scored as 1 for absence and 2 for presence, was treated as an independent variable in the regression model:

$Y = a + b_iM_i + ...+ b_nM_n + e$, which relates the variation in the dependent variable (NESTUR) to a linear function of the set of independent variables, $M_i...M_n$, representing the RAPD markers.

Prediction of growth in height, stem diameter and volume over a two year period was also assessed using either NESTUR *per se*, the six RAPD marker genotypes or a combination of markers + NESTUR as the independent variable (s). To control confounding effects of missing marker data, cases were excluded listwise, using only data points with valid values for all variables. The regression analysis was carried out with SPSS for Windows computer program (SPSS, Inc.) to derive values predicted from the regression model. The correlation between the predicted and observed value was calculated for each trait as an indication of prediction accuracy.

7.2.4 Comparative analyses

The comparative field growth of individual seedlings with (i) similar banding patterns for NESTUR-linked RAPD markers, and (ii) top 1% of the population (based on NESTUR values), was also assessed over a two-year period. Since no single marker was detected in experiment 3 that categorically divides the progeny into high- and low-NESTUR classes, individuals were compared which shared similar banding patterns associated with high or low NESTUR values. Presence of OPX-04a, OPV-17c and OPZ-09a was associated high NESTUR values; presence of OPE-09a, OPV-04c, and OPA-08e with low values. Thus, two sets of genotypes were generated from within the original experimental

population. Individuals in the high-NESTUR marker allele group had [+ + + - - -] banding pattern for the six markers, while the low-NESTUR group had [- - - + + +] pattern. Because of the loose linkage between the identified markers and QTLs, only individuals (sample size = 4) with all the increasing or decreasing marker alleles were compared.

In order to compare trait-based and marker-based selections, an additional set of genotypes was produced by selecting the top 1% (sample size = 4) of the original population, based on their phenotypic NESTUR values.

7.3 Results

7.3.1 Prediction of NESTUR from linked markers

Multiple linear regression using the six RAPD markers revealed that the mean observed NESTUR value of 7.8 was correctly predicted from the marker data. A plot of the observed and predicted NESTUR values from the multi-marker model analysis is presented in Figure 7.1. The correlation between observed and predicted values was 0.70.

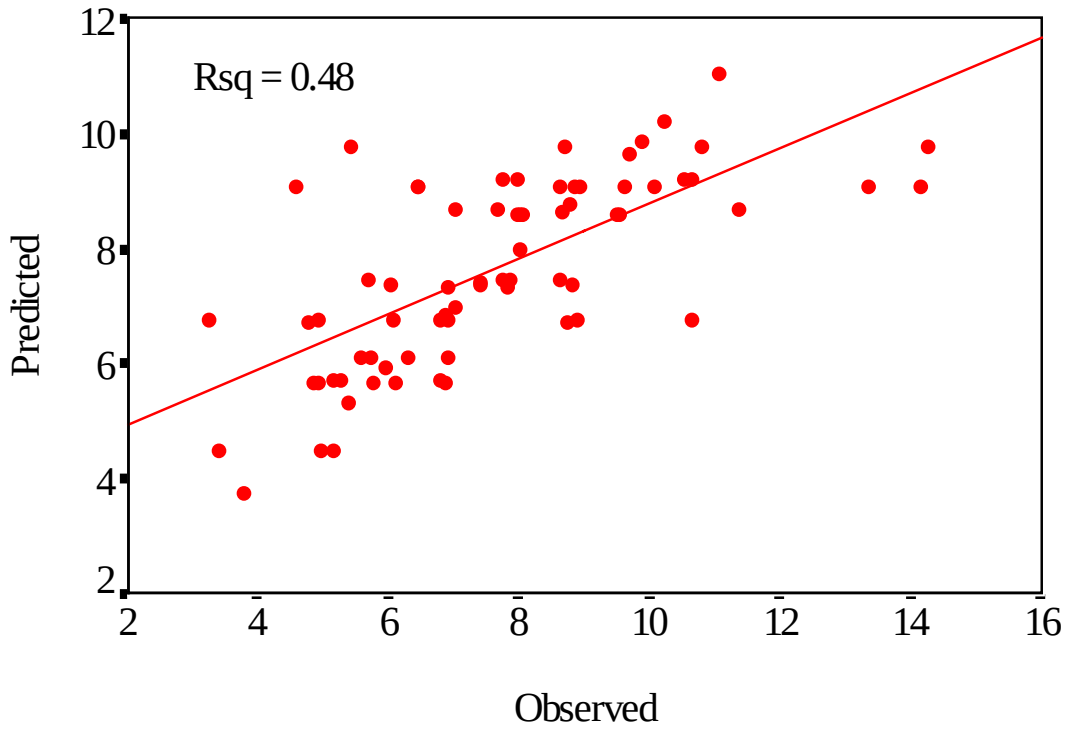


Figure 7.1 Plot of observed and predicted NESTUR values from linked RAPD markers using a multiple regression model.

A normal probability plot of the cumulative regression residual values (Figure 7.2) showed moderate departure from a line of unit slope. Overall, the values of the prediction errors were random with no systematic deviation from a mean of zero and variance of 1.0. The Durbin-Watson test of serially correlated residuals (Vinod and Ullah 1981) had a value of 1.94, indicating that residuals for consecutive observations were uncorrelated.

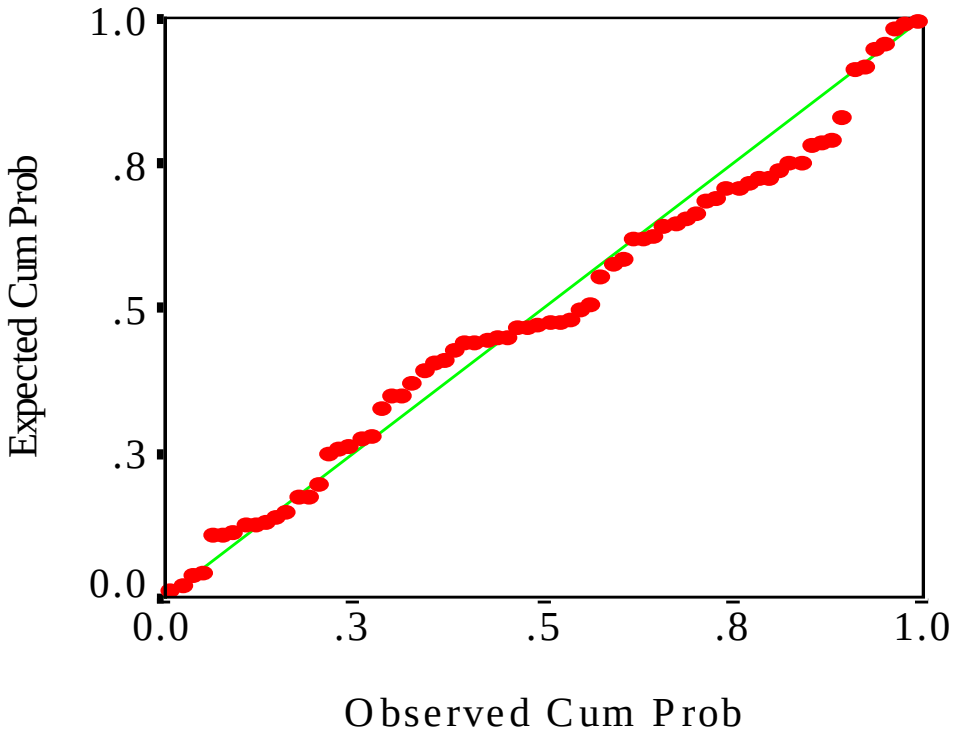


Figure 7.2 Normal probability plot of the cumulative distribution of residuals from predicting NESTUR using linked markers. The normal probability plot compared distribution of standardised residuals to a normal distribution.

7.3.2 Classification accuracy

Correct classification of progenies as either high or low NESTUR individuals may be of more use for selection purposes than specific values of correlation coefficients (Carter *et al*. 1990; Sulzer *et al*. 1993). Accuracy of classification was assessed by inspecting percentile plot of observed vs. predicted NESTUR values, shown in Figure 7.3. The plot showed a good agreement between observed ranking and that predicted using the six RAPD markers linked to NESTUR. The six RAPD markers more accurately predicted NESTUR values of individuals in the middle 50% of the percentiles than those in the upper and lower 25%. However, differences in observed vs. predicted values were not large and did not indicate serious misclassification, such as classification of an upper ranking individual into the lower group, or vice versa.

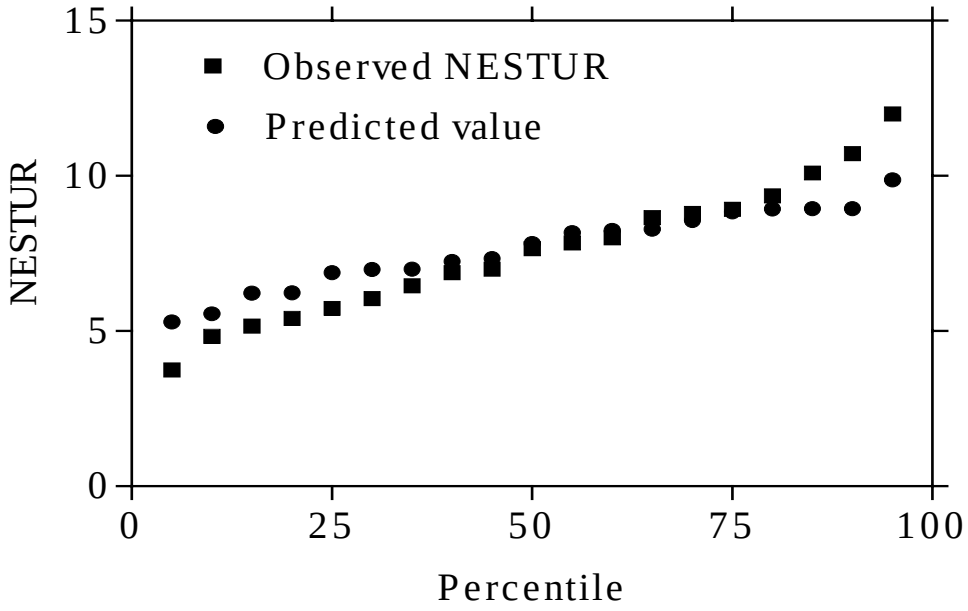


Figure 7.3 Percentile plot of observed vs. predicted NESTUR values. Observed NESTUR values were predicted using a multiple linear regression model that included six RAPD markers linked to NESTUR at $\alpha = 0.01$.

7.3.3 Comparison of NESTUR versus linked markers in predicting seedling growth

Ability of identified NESTUR-linked markers to predict stem growth potential was assessed using data measured over a two-year period, and compared with growth prediction using NESTUR values *per se*. Multiple linear regression was applied to height, stem diameter and stem volume, entered as the dependent variables, and NESTUR, markers, and NESTUR + markers, entered as the predictor variables. Correlation coefficients between observed and predicted values are presented in Figure 7.4.

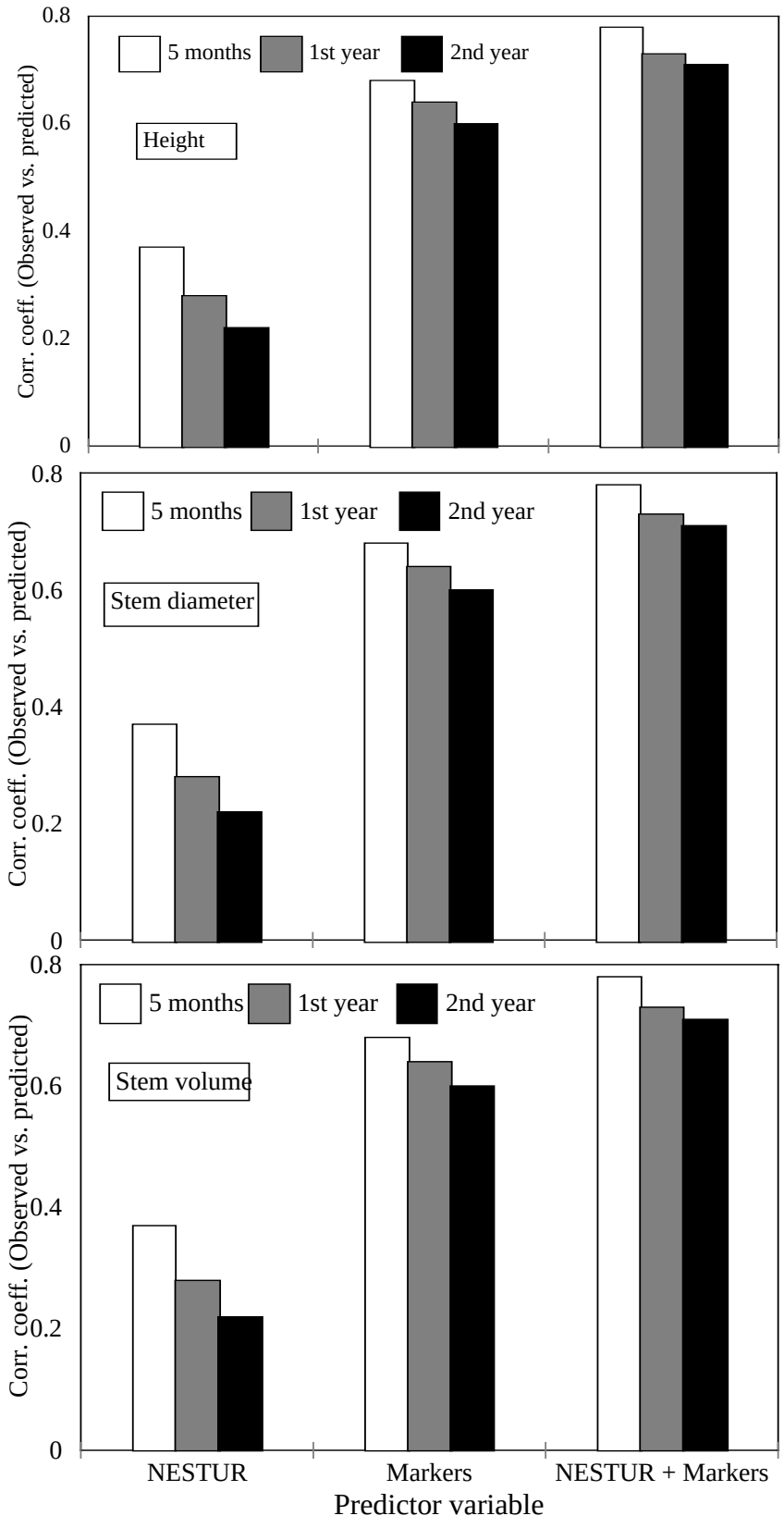


Figure 7.4 Correlation coefficient between observed stem growth over two years and predicted values using NESTUR, NESTUR-linked markers, and a combination of NESTUR + linked markers.

The ability of NESTUR *per se* to predict stem growth was strongly affected by seedling age, with low correlations between observed and predicted values. In contrast, markers linked to NESTUR showed a consistent ability to predict stem growth, irrespective of seedling age (Figure 7.4). The correlation between observed and predicted stem growth traits ranged from 0.60 to 0.66. The use of NESTUR + Markers in prediction did not significantly increase the correlations beyond that obtained by using markers alone.

When NESTUR was combined with linked markers to predict seedling growth, the accuracy of prediction increased slightly to a range of 0.70 to 0.74.

7.3.4 Growth performance of seedlings with high- and low-NESTUR marker alleles

A possible strategy for use of the markers to aid selection would be to identify individuals with a high frequency of NESTUR-increasing marker alleles. A comparison of stem growth was made between individual seedlings with (i) a high frequency of NESTUR-increasing and decreasing marker alleles, and (ii) the top 1% of the population (based on NESTUR values) and individuals with a high frequency of NESTUR- increasing marker alleles.

There was a significant (P = 0.05) difference in first year basal area and stem diameter with the frequency of similar NESTUR-linked marker alleles individual seedlings possessed. Genotypes with a 1.00 frequency of the high-NESTUR alleles showed a superiority of 9 to 59% in stemwood production during the first year of growth, compared with those with 0.00 frequency (low-NESTUR) (Table 7.1). Although growth performance was not statistically different amongst the two sets in the second year, the individuals with a 1.00 frequency of the high-NESTUR alleles showed a 14 to 51% better growth in height, basal area, stem diameter and stem volume.

Growth performance of individuals selected for favourable genotypic and phenotypic values is presented in Table 7.2. In the two years under consideration, there were significant differences (P = 0.01-0.1) in stem growth of the two sets of individuals (Table 7.2). Genotypes with a 1.00 frequency of NESTUR-increasing marker alleles showed a superiority of 19 to 103% in stemwood production, compared with the top 1% of the original population. The largest superior growth performance was observed for stem volume in both years (Table 7.2). Individuals selected for their genotypic values manifested a higher stem volume growth of 103% in the first year, and 76% in the second year, when compared with the top 1% of the original population.

Ontogenetic height growth of seedlings was monitored from 2 weeks after germination, and over a 2-year growth period. Comparison of the three classes of individuals (ie. those with a high frequency of high- and low-NESTUR marker alleles, and the top 1% of the progeny) was made by graphical analysis.

Results showed that height differences at two years of age still reflected differences established in the nursery stage (Figure 7.5). Individuals with a high frequency of NESTUR-increasing marker alleles maintained a superior height growth even after fielding planting.

7.4 Discussion

7.4.1 Marker prediction of phenotype

The six RAPD markers used to predict observed NESTUR explained 48% of the observed variation. Although the observed mean value of 7.8 was accurately predicted using these molecular markers, individual observed values was accurately predicted only 70% of the time. It is quite possible failure of the markers to better predict the phenotype may be due to inability to identify a

RAPD marker associated with a QTL of major effects. If this were the case, then the residuals would be expected to be larger for individuals, which were observed to be high in phenotypic value (Nienhuis *et al*. 1987). Inspection of residual and percentile plots (Figure 7.2 & 7.3), however, revealed that the errors associated with prediction were random.

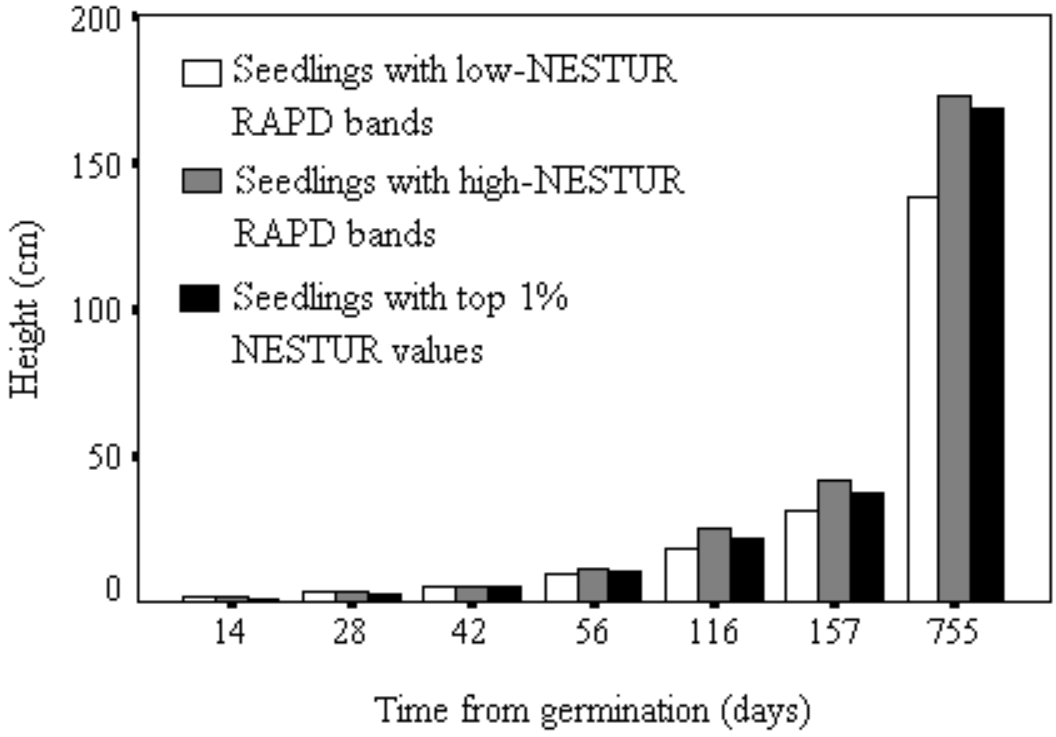


Figure 7.5 Mean height growth of seedlings trees with banding patterns for RAPD markers associated with low-and high-NESTUR, and the top 1% of the progenies (based on NESTUR values).

Failure of the RAPD markers to predict NESTUR better may be due to

(i) Inability to accurately assess NESTUR. Since many measurements were repeatedly made on the primary components, non-genetic variation and measurement errors would influence accuracy of prediction;

(ii) The effect of missing genotypic values. Due to incomplete amplification of DNA, ambiguity of genotyping RAPDs often occurred (reported in chapter 5).

Markers that could not be scored unambiguously were scored as missing data. In the multiple linear regression procedure, observations with missing genotypic data were discarded from analysis, resulting in lower prediction accuracy. Missing data is considered the cause of differences in explained variations derived from using interval and single-point statistical methods in QTL analyses (Plomion *et al*. 1996). Higher estimates of phenotypic variation explained by the joint action of QTLs are usually obtained with interval mapping than with single-point analytical methods because in interval mapping an inference is made based on data from flanking markers and the missing data point is replaced (Stuber *et al*. 1992); and

(iii) The confounding effects of the paternal contribution to the phenotype, since only the maternal parent contributed alleles for QTL detection;

7.4.2 Marker prediction of seedlings with superior growth potential

NESTUR as a phenotypic trait showed weak correlations with seedling growth characteristics during the first and second years of growth. The weak correlation during this period agrees with expectation (Matheson[1]) and with trends in age-age correlations established for other stem growth traits (Matheson *et al*. 1994). The weak correlation may be due to rank changes following field planting and micro-environment capture, which are most pronounced in the few years of growth. Lambeth (1980) noted that age-age correlations markedly improved after 3 years, with height at 1-3 years showing almost no correlation with height at later years. Delaying the age at trait measurements might improve early selection, but this raises practical problems, especially for stem growth efficiency traits (see Cannell *et al*. 1983; Velling and Tigerstedt 1984). An alternative is to find molecular markers associated with the

quantitative traits, and to use them as indirect selection criteria. In the present study, markers linked to NESTUR showed a consistent ability to predict stem growth, irrespective of seedling age (Figure 7.4). This is probably because, unlike quantitative traits, molecular markers are not affected by environmental factors. Genotypes with high-NESTUR marker allele frequency of 1.0 showed a significantly (P = 0.01-0.1) superior performance in stemwood production, compared with genotypes with the top 1% of NESTUR values. Thus, the use of RAPD markers as an indirect selection criterion proved to be more effective in identifying superior radiata pine seedlings than NESTUR *per se*. Comparison of ontogenetic growth of seedlings showed that height differences at two years of age still reflected differences established in the nursery stage (Figure 7.5). Individuals with a high frequency of NESTUR-increasing marker alleles maintained a superior height growth even after fielding planting.

[1] A. C. Matheson, CSIRO Forestry & Forest Products, Banks Street, Yarralumla, Canberra ACT, Australia. Postal address: PO Box 4008, Queen Victoria Terrace, Canberra ACT2600, Australia.

Table 7.1 Comparison of growth (over 2 years) of seedling trees which shared similar banding pattern (presence/absence) of 6 RAPD markers linked to NESTUR in the cross 30040 x 80121.

	First year growth performance of seedlings				Second year growth performance of seedlings			
Trait	Progenies with high-NESTUR marker alleles	Progenies with low-NESTUR marker alleles	F_p	% Diff	Progenies with high-NESTUR marker alleles	Progenies with low-NESTUR marker alleles	F_p	% Diff
Basal area (cm^2)	6.1	4.2	0.05	45.0	15.4	11.8	0.15	31.0
Stem diameter (mm)	27.7	22.9	0.05	20.8	44.0	38.5	0.17	14.2
Stem volume (cm^3)	656.2	411.6	0.06	59.4	3075.1	2043.4	0.13	50.5
Height (cm)	107.5	98.5	0.28	9.1	195.5	171.5	0.23	14.0
Number of branches	18.8	21.5	0.41	-14.7	36.8	34.0	0.64	8.1
Number of whorls	2.5	2.8	0.54	-10.0	5.0	5.0	1.00	0.0
Branches per whorl	7.1	8.0	0.89	-1.6	7.3	6.8	0.47	6.9
Crown-height (cm)	26.3	24.5	0.77	7.1	124.3	104.0	0.33	19.5
Inter-whorl length (cm)	10.2	9.0	0.32	13.5	24.6	20.8	0.19	18.4
NESTUR	9.2	6.1	0.01	50.7	-	-	-	-

Table 7. 2 Comparison of growth (over 2 years) of the top 1% of the progenies (based on NESTUR values) with seedling trees possessing high frequency of NESTUR-increasing RAPD marker alleles in cross 30040 x 80121.

	First year growth performance of seedlings				Second year growth performance of seedlings			
Trait	Progenies with high-NESTUR marker alleles	Top 1% of progenies (based on NESTUR *per se*)	F_p	% Diff	Progenies with high-NESTUR marker alleles	Top 1% of progenies (based on NESTUR *per se*)	F_p	% Diff
Basal area (cm^2)	6.1	3.6	0.01	70.2	15.4	10.5	0.07	46.6
Stem diameter (mm)	27.7	21.2	0.01	20.8	44.0	36.4	0.08	21.0
Stem volume (cm^3)	656.2	323.2	0.02	103.0	3075.1	1743.1	0.06	76.4
Height (cm)	107.5	90.3	0.07	19.1	195.5	165.8	0.09	18.0
Number of branches	18.8	17.3	0.31	8.7	36.8	34.3	0.66	7.3
Number of whorls	2.5	2.5	1.00	0.0	5.0	5.3	0.62	-4.8
Branches per whorl	7.7	7.1	0.55	7.6	7.3	6.6	0.47	10.3
Crown-height (cm)	26.3	21.5	0.52	22.1	124.3	104.0	0.27	19.5
Inter-whorl length (cm)	10.2	8.4	0.26	20.8	24.6	20.0	0.08	23.1
NESTUR	9.2	17.4	0.01	-88.7	-	-	-	-

7.4.3 Potential for marker-assisted selection

Several practical studies have demonstrated the effectiveness of marker-assisted selection for improving qualitative traits such as disease resistance. In apple, Gianfranceschi *et al*. (1996) reported the use of markers to assist in selection for resistance to the scab disease caused by *Venturia inaequalis*.

Experience to date with quantitative traits is inconclusive. In maize (*Zea mays* L.), Stuber *et al*. (1982) showed the effectiveness of allozyme selection for improving yield in a corn population. From the unselected Jarvis population, a selected sub-population was generated with gene frequencies at seven allozyme loci identical to those found in the 10th cycle of full-sib selection for increased yield. Field evaluation of this sub-population showed a yield response of about 6% over the unselected population. Other successful accounts of the use of molecular markers for yield enhancement in maize are detailed in Stuber (1994a 1994b; 1994c). In tomato (*Lycopersicon* spp.), Monteforte *et al*. (1996) have also successfully used molecular markers to develop salt-tolerant lines from an F_2 derived from a cross of *L*. *esculentum* and *L*. *pimpinellifolium*.

Although molecular marker locus-trait associations have been established for commercial forest tree species, very little information exists on use of markers to assist selection. Results of the present study will permit development of strategies that use the identified markers for early selection. However, making extrapolations from seedlings to mature trees may be misleading because stem volume is not a simple outcome of steady exponential growth. Progenies showing good growth in the initial stages may not necessarily be performing the same in subsequent years. It is suggested that identified marker-trait linkages be verified in an independent population, preferably a population of older trees,

in order to test predictive ability of detected markers. Because DNA markers do not vary among tissue types or developmental stages (Neale *et al*. 1992), the molecular markers should be detectable in older progenies of the same cross. However, genetic control of growth may vary with time, which will reduce accuracy of prediction. It will be necessary to monitor stability of stem growth QTLs over time in order to develop more accurate predictive variables for early selection of seedlings with superior growth potential.

8 CHAPTER 8

Experiment 7: Marker-based inferences on age-specific QTL expression in seedling growth of radiata pine

8.1 Introduction

Results of the experiment reported in chapter 7 show conformity with theoretical expectation, and increase enthusiasm about the potential of MAS as a breeding tool. However, QTL expression may not be stable across years of growth. Genetic control of growth may vary with time such that progenies showing good growth in the initial stages may not necessarily be performing the same in subsequent years.

Forest tree breeding programs differ from those of annual crops due to the longer life cycle. Age-related changes make it difficult to predict mature characteristics from juvenile traits. It has been proposed that age-related changes occur due to modification of gene expression (Kremer 1992). After each season, a portion of the set is replaced and several years later, the original set has been totally modified. Hodge and White (1992) gave a similar genetic interpretation in a study of many progeny tests of slash pine (*Pinus elliottii* Engelm. var. *elliottii*) measured at different ages. In contrast, Costa and Durel (1996) suggested growth could be determined by a more constant set of genes whose expression is only slightly modified during the course of tree development.

To determine whether QTLs at similar chromosomal locations account for observed phenotypic variation in the different years of growth, an analysis was carried using growth data measured for the same individual over a two-year period.

8.2 Materials and methods

8.2.1 Plant materials

The plant materials used for experiment 5 were the 93 mapping progeny used for linkage map construction in chapter 4 (section 4.2.1). The linkage map constructed in chapter 6 from RAPD markers was used for QTL analysis.

8.2.2 Trait measurements

Field planting and maintenance were as detailed in chapter 7, section 7.2.2. On each individual seedling tree, trait measurements were made at 5 months, 12 months and 24 months of age. The stem growth traits measured comprised total height, basal diameter, and total stem volume, calculated as height x basal area. Stem growth rates (increments) were assessed simply as the difference in periodic growth measurements. Three periodic increments were computed as $X_0 - X_1$, $X_1 - X_2$ and $X_0 - X_2$, where X_0 = trait measurement at 5 months of age, X_1 = 1st year growth measurement and X_2 = 2nd year trait measures.

8.2.3 QTL mapping by interval method

The chromosomal locations of putative QTL were determined by the method of conventional interval mapping (Lander and Botstein 1989), and are represented as QTL likelihood plots (Paterson *et al*. 1988). QTL analyses were performed using the computer software MAPQTL v. 3.0 (Ooijen and Maliepaard 1996), and likelihood maps drawn from scan profiles made every 1-cM. A LOD score of 2.0 was used to identify marker intervals containing putative QTLs. Correspondence in peak LOD scores was assessed within a 30-cM interval in significant genomic regions.

8.3 Results

8.3.1 QTL dissection of stem growth

At 5 months of age a total of seven QTLs were detected for stem volume, located on linkage groups 2, 4, 6 and 13 (Table 8.1). The QTLs explained 9 to 25% of the population variance. At 12 months of age, only two QTLs were found active, while at 24 months of age, only one was detected. These QTLs were located on linkage groups 6 and 13 (Figure 8.1).

Two putative QTLs were detected for stem diameter at 5 months of age, located on linkage groups 2 and 4. At 1 year of age, these QTLs were no longer active, but on linkage group 6, two QTLs were found, which collectively explained 27% of the population variance. At 2 years of age, only one QTL was found for stem diameter, located on linkage group 13 (Table 8.1).

For stem height, only one putative QTL was detected at 5 months of age, located on linkage group 4, and explaining 23% of the population variance. No QTLs were detected at 1 year of age. At 2 years of age, only one putative QTL was found, located on linkage group 13, and explaining 14% of the population variance.

The branch characteristics were measured only in the first and second years of growth. QTL were detected for only total number of branches, number of whorls and branches per whorl. No QTLs were found for branches per whorl at 1 year of age, but at 2 years, a single QTL was detected on linkage group 14, explaining 11% of the population variance.

A single QTL detected for total number of branches at 1 year of age was inactive at 2 years of age. Similarly, the single QTL detected for number of whorls at 1 year of age could not be found at 2 years of age (Table 8.1).

Of the 14 putative QTL positions detected for stem growth traits (height, diameter and volume) at LOD scores ≥ 2.0, none was strongly expressed at all

the three stages of development (5 months, 1 year and 2 years after germination). Three (or 21%) of the positions were detected at 5 months of age and at 1 year of age. However, no position detected at 1 year of age was still strongly influential at 2 years of age (Table 8.1). Age-specificity of detected QTL can be summarised as follows (Figure 8.1):

(i) QTLs active at 5 months of age gradually become down regulated as the seedlings aged, and at 2 years of age, had lost their influence on the character ('A' in Figure 8.1). Sixty-four percent of the putative QTLs for stem growth positions fit this pattern. The LOD scores at 5 months of age ranged from 2.2 to 3.1. At 1 year of age, peak LOD scores at these locations ranged from 0.89 to 1.76, and at 2 years of age, the LOD scores had dropped to a mean value of 0.5.

Table 8.1 Biometrical parameters of detected QTLs affecting seedling growth at 5 months, 1 yr and 2 yrs of age in the radiata pine family 30040 x 80121.

Trait	Linkage group	Marker interval	Peak LOD			Var. Explained (%)		
			5 months	1 year	2 year	5 months	1 year	2 year
Stem volume	2	OPAB-12b--OPB-03a	2.58	1.62	0.59	16.0	10.9	4.9
	4	OPAM-12b--OPC-10a	2.20	1.13	0.45	10.4	5.7	3.9
	4	OPB-11d--OPV-17c	2.93	0.89	0.37	25.1	7.3	1.9
	6	UBC-119b--OPK-03d	2.99	2.20	0.32	18.9	14.9	2.5
	6	OPAJ-12a--OPB-11a	3.08	1.76	0.32	19.6	11.7	3.1
	6	OPA-11b--UBC-295a	3.17	2.37	1.12	14.5	11.5	5.7
	13	OPAB-16b--OPAI-20e	0.52	1.09	2.45	2.8	7.0	14.2
Stem diameter	2	UBC-533c--OPB-03a	2.76	1.38	0.53	18.0	9.0	3.8
	4	OPB-11d--OPV-17c	2.28	1.31	0.63	19.3	9.8	3.3
	6	UBC-119b--UBC-268a	1.76	2.38	0.49	10.2	16.2	4.3
	6	OPB-11a--UBC-295a	1.51	2.29	0.92	8.2	11.1	4.7
	13	OPAB-16b--OPAI-20e	0.43	0.67	2.07	2.6	4.4	13.1
Height	4	OPB-11d--OPV-17c	2.70	1.31	0.43	23.2	9.8	3.6
	13	OPAB-16b--OPAI-20e	1.12	1.09	2.45	8.8	7.0	14.2
No. Branches	8	OPX-04c--OPAB-18a	-	2.62	1.64	-	14.0	8.7
No. Whorls	12	UBC-366e--OPV-04b	-	2.21	1.15	-	12.9	6.9
Branches/whorl	13	UBC-519a--UBC-479a	-	0.65	2.19	-	3.4	10.9

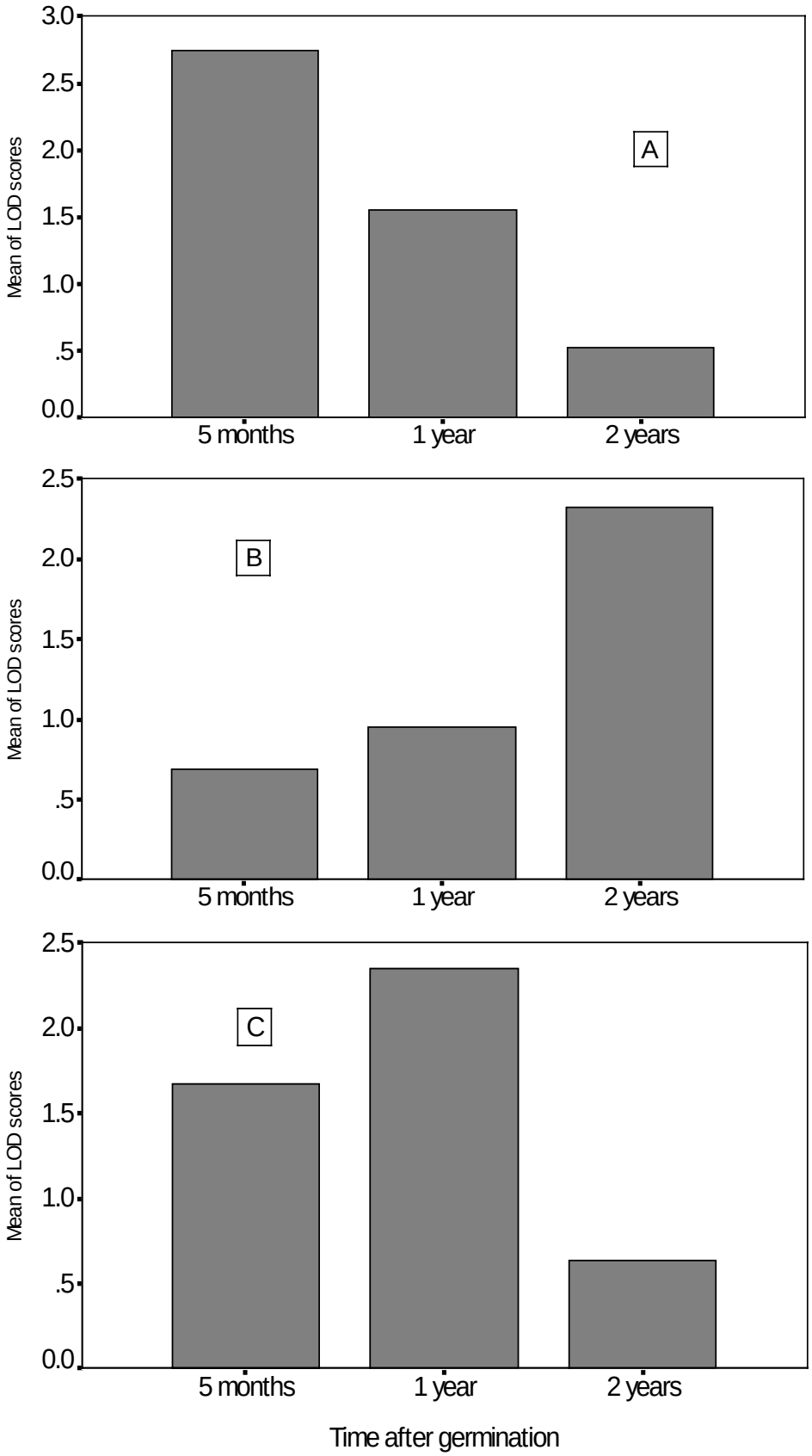


Figure 8. 1 Mean of LOD scores for stem growth traits (height, diameter and volume) measured at various growth stages (time from germination) in radiata pine seedlings grown in glasshouse and under field conditions.

In some instances, the QTLs expressed at 5 months of age were still strongly expressed after 1 year of growth (Table 8.1). Three of the detected putative QTLs for stem volume in linkage group 6 showed a LOD score of 3.0 to 3.2 at 5 months of age. At 1 year of age, these positions still showed strong association with the trait (LOD scores ranged from 2.0 to 2.4). The effects, however, were no longer detectable after 2 years of growth.

(ii) QTLs at some locations showed gradual linear increase in influence from 5 months of age but became detectable at the second year of age ('B' in Figure 8.1). For stem volume, a single QTL detected on linkage group 13 between markers OPAB-16b and OPAI-20e, and which was coincidental with those for stem diameter and height, showed this pattern of influence. The QTL exerted a similar pattern of expression on these characters (Table 8.1). The peak LOD score at 5 months of age was an average of 0.6. This increased linearly to 2.3 at 2 years of age.

(iii) Some QTLs showed gradual curvilinear increase in effect from 5 months of age, reaching their peak expression at 12 months, but were not detected at 2 years of age ('C' in Figure 8.1). Two of the putative QTLs detected for stem diameter on linkage group 7 (Table 8.1) had LOD scores of 1.5 to 1.8 at 5 months of age, which increased to 2.3 to 2.4 after the first year, and declined subsequently to 0.5 to 0.9.

8.3.2 Coincidence of stem growth QTLs

Comparison of the presence or absence of QTLs, and scan profile of LOD scores, showed that QTLs detected for the final phenotype (ie the phenotype at time of trait measurement) may be the same as that which controlled its increment. In linkage group 6, for instance, the two linked QTLs for 1st year stem diameter and stem volume showed similar location and LOD scan profiles

with those observed for their growth rate or increments, ie. increments from 0-1 year of growth (Figure 8.2).

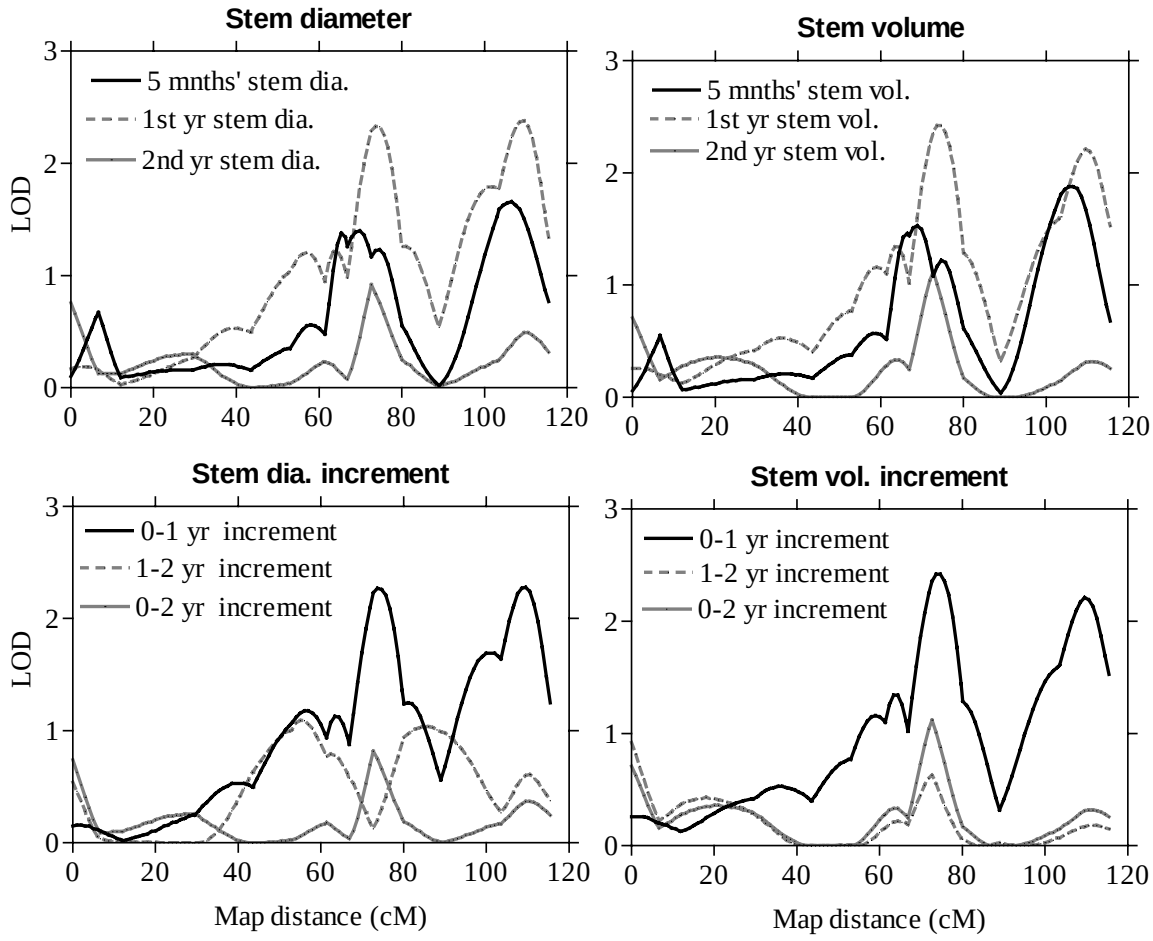


Figure 8.2 Logarithm of odds ratio (LOD) plot of significant genomic region in group 6, showing scan profiles for stem diameter, stem volume and their increments over three successive periods. Putative QTLs were inferred within an interval with a LOD score of ≥ 2.0.

In linkage group 13, a single putative QTL detected for 2nd year height showed similar location and scan profile for a QTL detected for 2nd year height increment (ie the difference between 1st and 2nd year heights) (Figure 8.3).

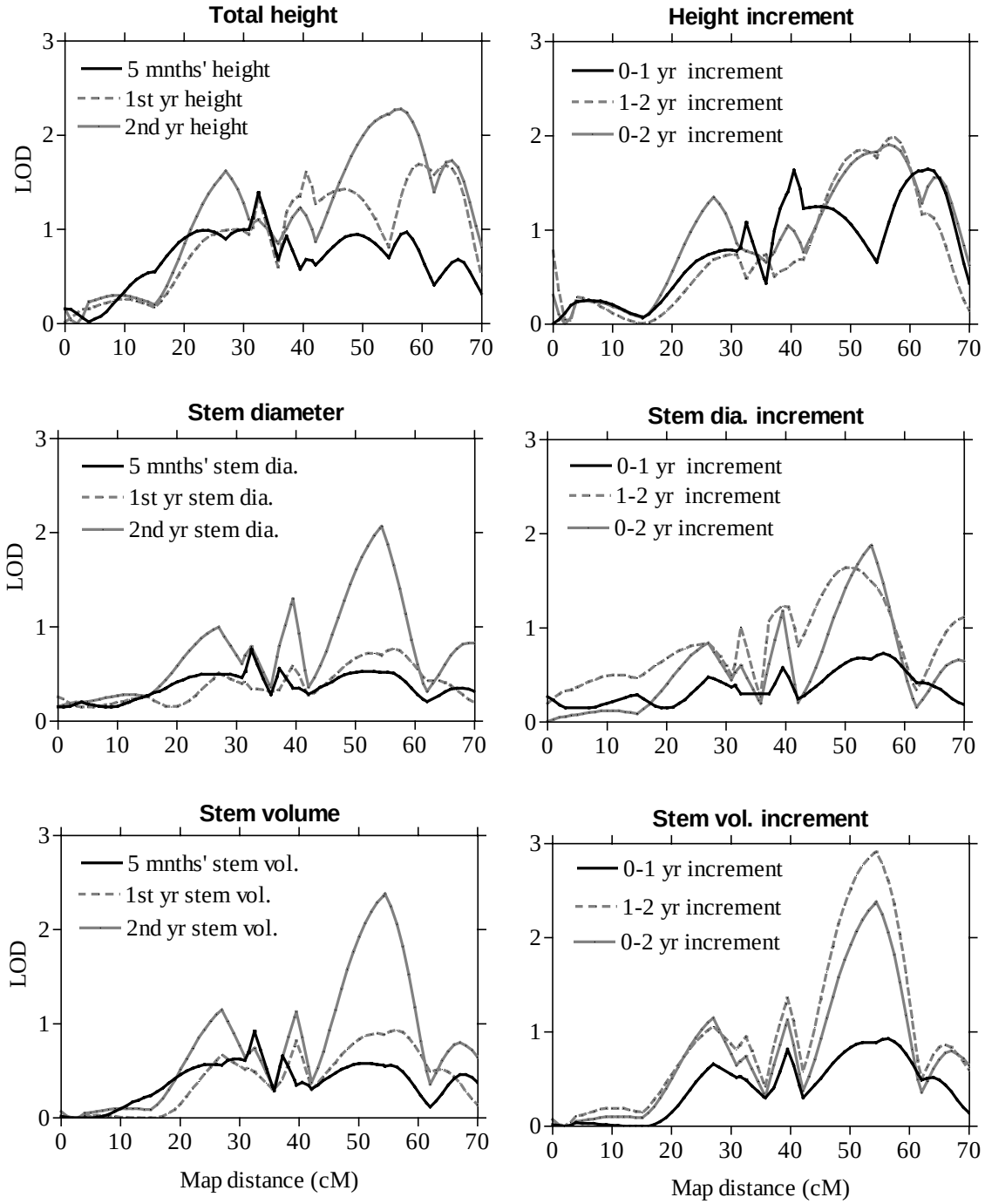


Figure 8.3 Logarithm of odds ratio (LOD) plot of significant genomic region in group 13, showing scan profiles for height, diameter, stem volume and their increments over three successive periods. Putative QTLs were inferred within an interval with a LOD score of ≥ 2.0.

In addition, QTLs for stem diameter and height were often coincident with those detected for stem volume. As shown in the plot of LOD scores for significant genomic region in linkage group 6, the two linked QTL for 1st year stem

diameter and stem volume showed similar location and LOD scan profiles (Figure 8.2). In linkage group 13, the single QTL position detected for 2nd year height also significantly affected 2nd year stem volume (Figure 8.3). Co-location of the QTLs and the similar magnitude of effects (Table 8.1) suggest that these occurrences represent QTLs with pleiotropic effects on the size and stem growth traits.

8.4 Discussion

8.4.1 Stability of stem growth QTLs with development stage

Although there is no period of absolute dormancy in growth of radiata pine trees, a seasonal pattern of variable growth has been shown by Pawsey (1964). The genetic control of stem growth may change with stand development due to the varying effects of the environment (eg. competition for resources) and physiological demands of maturation (Matheson *et al*. 1994; Costa and Durel 1996; Williams 1987). There was only weak evidence of QTL stability during the two years of growth under this study. As shown in Table 8.1 and in Figures 8.1, 8.2, and 8.3, individual QTL became active at various stages and were gradually down regulated with development. Some consistency was observed for QTLs involved in stem volume growth, which were active at 5 months of age and still detectable at 12 months of age.

Results of this study suggest that although several genes might control stem growth, they do not necessarily act at the same time. The proposition of Kremer (1992) that a given set of genes controlling annual growth effects is progressively modified might involve gradual down-regulation of active QTLs. Overall, the observed pattern of QTL expressions for height suggests that significant genomic regions showed a progressive shift in location of expressed QTLs during the two years of seedling growth. In linkage group 13, for instance

where a putative QTL was detected for 2nd year height, adjacent LOD scores suggest QTL for height at 5 months and 1 year that might have once been active, but now redundant (Figure 8.3).

Compared with the pattern observed for height, LOD scan profiles for diameter and volume showed coincidence of smaller peaks (LOD < 2.0) at the same location with the maximum LOD score, without suggestion of a progressive shift (Figures 8.2 and 8.3). These results suggest that the genetic control of height growth is probably more unstable than that of diameter. This agrees with the suggestion of Costa and Durel (1996) that diameter growth could be determined by a more constant set of genes whose expression is only slightly modified during the course of tree development. St. Clair (1994) has also shown that heritability for diameter generally increases at a higher rate with age than heritability for height.

8.4.2 Similar QTLs control stemwood variables and increments

Results obtained for other perennial plant species agree well with those from this study. In an intervarietal cross of Almond, Asins *et al*. (1994) found only 3 out of 17 detected QTLs affecting morphological traits were homogenous across three years of growth. Sondur *et al*. (1995) found QTLs affecting growth in Papaya (*Carica papaya* L.) to be environmental or developmentally sensitive, influencing growth during different seasons.

However, observations of the present study differ from those of Sondur *et al*. (1995) who found that QTLs affecting growth rate have no detectable effect on the final phenotype. As shown in Figures 8.2 and 8.3, QTLs controlling the phenotype at the time of measurement (ie the final phenotype) showed similar magnitude of effects on the its increments (or growth rate). This agrees with results in quantitative genetics where size and increment measurements have

been shown to be highly interrelated, both phenotypically and genetically (St. Clair 1994).

8.4.3 Perspective on marker-assisted selection

The present study considered the early establishment stage of seedling trees, corresponding to the stage of microenvironment capture after planting. It is difficult at this stage to distinguish whether the differential expression of stem growth QTLs is related to maturational effects or differential ability to adapt to a new environment. However, results of the present study are similar to those obtained with maritime pine seedlings grown under accelerated growth conditions (Plomion *et al*. 1996), and suggest that maturation effects were probably predominant. For marker-assisted selection to become feasible, an understanding of the pattern and stability of QTL expression in the course of tree growth is fundamental. Since the progeny used in the present study were planted out, it is suggested that trait measurements be continued over the years to provide additional insights on pattern and stability of stem growth QTLs.

9 CHAPTER 9: General discussion

9.1 Stem growth efficiency and early selection

Tree breeders face two challenges, (i) maximising unit-area wood production through efficient conversion of the sun's energy into wood, and (ii) shortening the generation interval associated with selection in trees. Growth assessment and selection at a stage well before rotation would markedly improve the efficiency of tree improvement. However, efforts at developing early testing criteria have met with limited success because of low age-age correlations due to maturational changes in growth behaviour.

Models of tree growth and dry matter production, however, suggest an "inherent ability to grow rapidly" can be detected at early ages. If stem volume at t years (V_t), for instance, was a simple outcome of steady exponential growth over the period, then $V_t = V_o e^{rt}$, where V_o = stem volume at planting and r = relative rate of stem volume increase (years^{-1}) (Matheson *et al.* 1995.; Cannell 1978). Variation in r and/or V_o due to either genetic or environmental effects would thus influence V_t. Either a high r or an initially large V_o would result in a substantially larger V_t after a few years. Selection for seedling size, however, has not proved a useful indicator of later-age stem volume.

Derived traits, described by a function involving two or more primary traits, may provide opportunities for progress towards improved early selection. What may be required is a stem growth index in seedling trees analogous, in functional terms, to the same process underlying stem growth in mature trees. Using a measure of stem:needle mass in seedlings, Matheson *et al*. (1995) obtained genetic correlations of as high as 0.98 with the relative growth rate (RGR) of

stem diameter in 8-11 year-old radiata pine trees growing in several plantations in Australia. However, the methods of trait measurement were destructive, precluding further use of valuable germplasm. An index of stem growth efficiency was developed by staff at the CSIRO Forestry & Forest Products to measure needle-to-stem unit rate (NESTUR) in conifer seedlings. The index is similar to relative growth rate (RGR), but takes into account the influences of both photosynthetic capacity and carbon allocation on stemwood production, and is a potentially useful early selection trait.

Measurements of NESTUR are non-destructive. This permits (i) a time-course tracking of tree growth dynamics and (ii) the conservation of valuable genetic materials. NESTUR also appears to allow early selection of genotypes capable of efficient conversion of the sun's energy into stemwood production. However, NESTUR is time-consuming to measure, and like most derived traits, measurement of many primary components may be subject to errors. Marker-based selection procedures are particularly useful for such traits.

9.2 Using marker-assisted selection to improve stem growth efficiency

Genetic markers range from morphological mutants to DNA sequence variation. In plant breeding applications morphological markers can have undesirable effects on plant phenotype and their use in tree improvement has been limited. Also, morphological markers identified by macromutant alleles are rare in natural populations. Without allelic variation there is no segregation, and without segregation no linkage tests can be performed to detect QTLs (Tanksley 1993).

DNA-based markers, on the other hand, are phenotypically neutral and abundant in existing natural populations. Phenotypically neutral markers not

only make it easier to detect linkage between the segregating marker and the QTL, but also provide an unbiased way to estimate the phenotypic effect of each QTL without interference by the marker locus (Tanksley 1993). Also, there is a high level of existing natural polymorphisms at the DNA level, and several laboratory techniques are now available for detecting such variations.

The most commonly used technique for assaying polymorphisms at the DNA level is the random amplified polymorphic DNA (RAPD) marker system (Williams *et al*. 1990; Welsh and McClelland 1990). RAPD markers are dominant genetic markers because the homozygote from the amplification product cannot be distinguished from the heterozygote. In conifers, however, the problem of dominance can be overcome if RAPD markers are assayed from haploid megagametophyte tissue (Neale *et al*. 1992). The advantages of RAPDs over other types of markers such as RFLP include technical simplicity of the procedure, which requires no prior sequence information, only a small amount of DNA, and a detection system based on gel electrophoresis. In addition, RAPDs provide markers in regions of the genome inaccessible to RFLP analysis due to presence of repetitive DNA sequence (Williams *et al*. 1990).

9.2.1 Detection of QTLs influencing NESTUR by bulked segregant analysis

To determine whether RAPD analysis could be successfully extended to the development of molecular markers for NESTUR in radiata pine, bulked segregant analysis was applied to haploid DNA from the megagametophytes of a full-sib radiata pine cross (12038 x 10946). Using the approach, 23 of 933 primers displayed putative linkage to factors controlling NESTUR. Based on the genotypic analysis of 174 individuals, two quantitative trait loci (QTLs)

controlling NESTUR were identified at ANOVA P-levels of 0.01-0.001. The QTLs were identified by RAPD markers OPE-06_{450} and OPA-10_{1200}, which were linked to each other (r = 7%) and UBC-333_{550} which was not linked to the other two. Linkage to components of NESTUR (increments in stem diameter and stem volume) was demonstrated for UBC-333_{550}, while the others were not linked to NESTUR components.

The methods used served to establish marker-QTL linkage for NESTUR. However, it is difficult to determine the extent of genome coverage achieved with the use of bulked segregant analysis. Some QTL-linked markers, which would have been detected had all 174 progeny been genotyped with the 933 primers, were probably missed by the use of bulked segregant analysis. There is no chance of detecting QTLs in chromosomal regions devoid of segregating markers.

Also, explained variation associated with detected RAPD markers was not large enough to permit a realistic judgement about the prospects of marker-assisted selection. The amount of variation explained by a marker-linked QTL depends on the distance between the QTL and the marker. If enough segregating markers are scattered throughout the genome, it is theoretically possible to detect a marker(s) with sufficiently large effects on a quantitative trait to permit use of marker-assisted selection.

9.2.2 Genomic mapping of NESTUR-linked QTLs

A linkage map was constructed from RAPD markers segregating in a different haploid family of radiata pine (30040 x 80121). Initially, six megagametophyte DNA samples of the cross were screened with 432 primers. Primers found to be informative were further characterised within a sample of 93

megagametophytes, including the six already tested and used this time as a repeatability assay. Two hundred and sixty-two (262) fragments segregating 1:1, present-to-absent, ratio were classified as Mendelian markers and mapped using multipoint linkage analysis. The analysis yielded 14 linkage groups of at least 7 markers, ranging in size from 39 to 183 cM. The 14 linkage groups covered approximately 1511 cM of genetic map distance. Using a 30-cM map scale and including the ends of the 14 linkage groups, the set of RAPD markers covered 1571 cM, or 78% of the radiata pine genome.

The linkage map was used to search for QTLs controlling NESTUR, as well as increments in seedling stem diameter, volume, and height and needle volume. The association of markers with trait-causing genetic factors was assessed using both single-point analysis of variance (ANOVA) and interval mapping. Interval mapping technique, which localises the effect of a QTL to a genomic segment located between pairs of markers rather than associating the effect with an individual locus, provides a more powerful method of QTL analysis for a complex trait (Lander and Botstein 1989). However, because of the density of markers used in the present study, both the ANOVA and interval mapping methods were expected to give similar results, and therefore provide some sort of checks on QTL detection. Interval mapping was carried out using both the conventional and multiple-QTL mapping (MQM) methods.

Altogether, five putative QTLs were detected for NESTUR, with explained variation ranging from 9 to 22%. Of the five QTLs detected, 3 were coincidental with those for stem growth in height, diameter and volume. QTLs for component traits flanked the two QTL positions that were unique to NESTUR. In linkage group 4, a unique QTL position was detected next to a QTL for Height growth. Such coincidence and clustering provides biological validation

for the detected QTL, and may be caused by pleiotropic action of the same gene or segregation of other genes.

Results from single-point analyses of variance agreed well with those from interval mapping. However, the two interval mapping methods showed slight differences in results. The main difference was in the magnitude of LOD scores and explained variation at putative QTL positions. The MQM mapping method gave slightly higher LOD scores, and lower values of explained variation than conventional interval mapping. A discrepancy was observed in a region of linkage group 12 where the LOD scores for two putative QTLs that were linked in repulsion was below the LOD threshold value from conventional interval mapping. Both QTL positions were detected from single-point analysis and MQM mapping. Failure of conventional interval mapping to detect both putative QTL positions highlights the limitations of conventional interval mapping, and the need for complementary methods of QTL detection.

9.2.3 Marker prediction of stem growth potential

As quantitative trait loci are located and mapped, questions about their nature will surely receive some attention. Do they represent structural loci, or loci that encode regulatory proteins? The present data are not adequate to definitively test the possibility that QTL detected for NESTUR represent regulatory loci encoding proteins for stem volume growth. However, the detection of putative QTLs influencing stem growth increment per needle volume growth is consistent with the carbon allocation pattern of trees determined by physiologists and quantitative geneticists. Past studies have shown that stemwood formation results from cambial activities that are mediated by leaf development (Kozlowski 1971). The greater the number of leaves, the greater

the wood production. However, stem increment per unit amount of foliage decreases with an increasing amount of foliage per tree. This is largely because the amount of branchwood produced by a unit amount of foliage increases with an increasing amount of foliage per tree. If respiratory losses by nongreen tissues were held constant, differences in foliage efficiency of trees in producing wood would largely be caused by differences in distribution of dry matter in various parts of the trees. While the quantitative genetic basis of such differences has long been established for the conifers (St. Clair, 1994; Pulkkinen and Poykko 1990; Pulkkinen *et al*. 1989; Sheppard and Ford 1986; Velling and Tigerstedt 1984; Cannell *et al*. 1983; Drew and Bazzaz 1978), this is the first report of efforts to map the QTLs using molecular markers.

Marker-linked QTLs, however, are statistically defined entities and have the potential to be artifacts. The detection of a QTL provides an opportunity to test hypotheses regarding marker-assisted selection in specific crosses where QTLs have already been detected. One of such hypotheses (Grattapaglia *et al*. 1993) is that markers associated with QTLs should be able to predict the phenotype. Using markers closest to putative QTL in a predictive model, the rank correlation between observed and predicted NESTUR values in the present study was 0.70, suggesting that the ranking of 70% of the individuals were correctly predicted from their marker information, and that identified markers can successfully be used as indirect selection criteria for NESTUR.

The field growth of individual seedlings with (i) similar banding patterns for NESTUR-linked RAPD markers, and (ii) top 1% of the population (based on NESTUR values), was also assessed over a two-year period. Compared with selection based on NESTUR *per se*, NESTUR-linked markers provided a better

prediction of individuals with superior growth potential. The weak correlation between NESTUR and stem growth traits during this period may be due to rank changes following field planting and micro-environment capture, which are most pronounced in the few years of growth. Because molecular markers are not affected by environmental factors, use of markers linked to early selection traits of interest provides better prediction of later-age performance than phenotype-based selection.

Molecular markers associated with mature traits may allow superior individuals to be identified as early as the embryo stage (Chaparro *et al*. 1995), since DNA markers do not vary among tissue types or developmental stages (Neale *et al*. 1992). Comparison of three classes of individuals (ie. those with a high frequency of NESTUR-increasing and -decreasing alleles, and the top 1% of the progeny) for ontogenetic height growth showed differences at two years of age still reflected differences established in the nursery stage (Figure 7.5). Individuals with a high frequency of NESTUR-increasing marker alleles maintained a superior height growth even after fielding planting.

9.2.4 Stability of QTL expression with development stage

These results show conformity with theoretical expectation, and increase enthusiasm about the potential of MAS as a breeding tool. However, QTL expression may not be stable across years of growth. Genetic control of growth may vary with time such that progenies showing good growth in the initial stages may not necessarily be performing the same in subsequent years. To determine whether QTL at similar chromosomal locations account for observed phenotypic variation in the different years of growth, an analysis was carried using growth data measured for the same individual over a two-year period.

The results suggest that different QTLs are involved in the cyclic growth pattern, but such conclusions are premature because marker-locus x age interaction could result from residual variations caused by non-genetic factors, which differ largely over years of growth. The magnitude of the non-genetic variation directly influences the explained variance of a QTL, and thus its LOD score (Ooijen 1992). Measurement errors, as well as the number of replications (in this case, size of the mapping population) influence residual non-genetic factors. Given the importance of QTL stability for marker-aided selection, additional studies are necessary, in which adequate replication permits estimation of the non-genetic residual variance across years of growth.

9.2.5 Potential applications of detected markers in radiata pine improvement

The detection of distinct genetic systems controlling NESTUR will permit development of strategies that use the identified markers for selection. Marker-assisted selection (MAS) for NESTUR would be expected to increase efficiency of early selection, because of the higher heritability of markers, and the higher selection intensities implied. Selecting for NESTUR-linked markers may permit the pyramiding of favourable QTL for stemwood production in a genotype because of the positive pleiotropic effects of putative QTL on height, diameter and volume growths. An improvement scenario utilising MAS would be to increase the efficiency of within-family selection. With the availability of clonal propagation techniques for radiata pine, genotypes with favourable marker alleles could be amplified for commercial use. Large gains could also be made for additive genetic variance if single genotypes with favourable NESTUR-linked markers were control-pollinated with one another, or placed individually in seed

orchards with marker-selected representatives of other families (Strauss *et al.* 1992).

10 REFERENCES

Adams, W. T. and Joly, R. J. (1980) Linkage relationships among twelve allozyme loci in loblolly pine. *Journal of Heredity*. **71**: 199-202.

Allard, R. W. (1960) 'Principles of Plant Breeding'. John Wiley and Sons, New York.

Asins, M. J., Mestre, P., Garcia, J. E., Dicenta, F. and Carbonell, E. A. (1994) Genotype x environment interaction in QTL analysis of an intervarietal almond cross by means of genetic markers. *Theoretical & Applied Genetics*. **89**: 258-264.

Ayliffe, M. A., Lawrence, G. J., Ellis, J. G. and Pryor, A. J. (1994) Heteroduplex molecules formed between allelic sequences cause nonparental RAPD bands. *Nucleic Acids Research*. **22**:1632-1636.

Bahrman, N. and Damerval, C. (1989) Linkage relationships of loci controlling protein amounts in maritime pine (*Pinus pinaster* Ait.). *Heredity*. **63**: 267-274.

Bartels, D., Gebhardt, C., Knapp, S., Rohde, W., Thompson, R., Uhrig, H. and Salamini, F. (1989) Combining conventional plant breeding procedures with molecular based approaches. *Genome*. **31**: 1014-1026.

Bassam, B. J., Caetano-Anolles, G. and Gresshoff, P. M. (1991) Fast and sensitive silver staining of DNA in polyacrylamide gels. *Analytical Biochemistry*. **196**: 80-83.

Bateson, W. (1907) The progress of genetics since the rediscovery of Mendel's paper. *Prog. Rei. Bot*. **1**: 368 (In, Tanksley 1993).

Beavis, W. D., Grant, D., Albertsen, M. and Fincher, R. (1991) Quantitative trait loci for plant height in four maize populations and their associations with qualitative genetic loci. *Theoretical & Applied Genetics*. **83**: 141-145.

Benet, H., Guries, R., Boury, S. and Smalley, E. B. (1995) Identification of RAPD markers linked to a black leaf spot resistance gene in Chinese elm. *Theoretical & Applied Genetics*. **90**: 1068-1073.

Binelli, G. and 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.

Boscherini, G., Morgante, M., Rossi, P., Vendramin, G. G. and Vicario, F. (1994) Detection of DNA polymorphism in *Pinus leucodermis* Ant. using random amplification. *Forest Genetics*. **1**: 131-137.

Botstein, D., White, R. L., Skolnik, M. and Davis, R. W. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphism. *American Journal of Human Genetics*. **32**: 314-331.

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

Bradshaw, H. D. and Stettler, R. F. (1994) Molecular genetics of growth and development in *Populus*. II. segregation distortion due to genetic load. *Theoretical & Applied Genetics*. **89**: 551-556.

Bradshaw, H. D. and Stettler, R. F. (1995) Molecular genetics of growth and development in *Populus*. IV. Mapping QTL with large effects on growth, form, and phenology in a forest tree. *Genetics*. **139**: 963-973.

Bradshaw, H. D. Jr and Foster, G. S. (1992) Marker-aided selection and propagation systems in trees: advantages of cloning for studying quantitative inheritance. *Canadian Journal of Forest Research*. **22**: 1044-1049.

Bradshaw, H. D., Villar, M., Watson, B. D., Otto, K. G., Stewart, S. and Stettler, R. F. (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.

Bucci, G. and Menozzi, P. (1993) Segregation analysis of random amplified polymorphic DNA in *Picea abies* Karst. *Molecular Ecology*. 22: 227-232.

Burke, T., Hanotte, O, Bruford, M. W. and Cairns, E. (1992) Multilocus and single locus minisatellite analysis in population biological studies. In: DNA fingerprint: Approaches and Applications (Eds. Burke, T, Dolf, G., Jeffreys, A. L. and Wollf, R), Birkhauser Verlag, Basel, pp 154-168.

Burr, B. and Burr, F. A. (1991) Recombinant inbreds for molecular mapping in maize. *Trends in Genetics*. **7**: 55-60.

Byrne, M., Marquez-Garcia, M. I., Uren, T., Smith, D and Moran, G. F. (1996) Conservation and genetic diversity of microsatellites loci in the genus *Eucalyptus*. *Australian Journal of B*otany. **44**: 331-341.

Byrne, M., Murell, J. C., Allen, B. and Moran, G. F. (1995) An integrated genetic linkage map of eucalypts using RFLP, RAPD and isozyme markers. *Theoretical & Applied Genetics*. **91**: 869-875.

Byrne, M., Murrell, J. C., Owens, J. V., Kriedemann, P., Williams, E. R. and Moran, G. F. (1997) Identification and mode of action of quantitative trait loci affecting seedling height and leaf area in *Eucalyptus nitens*. *Theoretical & Applied Genetics*. **94**: 674-681.

Cai, Q., Guy, C. L. and Moore, G. A. (1994) Extension of the linkage map in Citrus using random amplified polymorphic DNA (RAPD) markers and RFLP mapping of cold-acclimation-responsive loci. *Theoretical & Applied Genetics*. 89: 606-614.

Cannell, M. G. R. (1978) Improving per hectare forest productivity *In*: 'Proceedings of the 5th North American Forest Biology Workshop' (Ed. C. A. Hollis & A. E. Squillace). pp. 120-148.

Cannell, M. G. R. (1989) Physiological basis of wood production: a review. *Scandinavian Journal of Forest Research*. **4**: 459-490.

Cannell, M. G. R., Sheppard, L. J. and Ford, E. D. (1983) Clonal differences in dry matter distribution, wood specific gravity and foliage efficiency in *Picea sitchensis* and *Pinus contorta*. *Silvae Genetica*. **32**: 5-6.

Carlson, J. E., Tulsieram, L. K., Galubitz, J. C., Luk, V. W. K., Kauffeldt, C. and Rutledge, R. (1991) Segregation of random amplified DNA markers in F_1 progeny of conifers. *Theoretical & Applied Genetics*. **83**: 194-200.

Carson, S. D., Richardson, T. E., Corbett, G. E., Lee, J. R. and Wilcox, P. L. (1995) Verification of statistically significant linkages for wood density in *Pinus radiata*. Abstract of paper presented at the 6th Australasian Gene Mapping Workshop and New Zealand Genetical Society Conference, Nov. 27 - 1 Dec., University of Otago, Dunedin, New Zealand.

Carter, K. K., Adams, G. W., Greenwood, M. S. and Nitschke, P. (1990) Early family selection in jack pine. *Canadian Journal of Forest Research.* **20**: 285-291.

Chakravarti, A., Lasher, L. A. and Reefer, J. E. (1991) A maximum likelihood method for estimating genome length using genetic linkage data. *Genetics*. **128**: 175-182.

Chaparro, J. X., Werner, D. J., O'Malley, D. and Sederoff, R. R. (1994) Targeted mapping and linkage analysis of morphological, isozyme, and RAPD markers in peach. *Theoretical & Applied Genetics*. 87: 805-815.

Chaparro, J., Shepherd, M., Vogel, H. M., Gibbings, M. M. and Teasdale, R. D. (1995) Detection of vigour loci in Eucalypt interspecific hybrids. Abstract of paper presented at the 7th international Conference of Conifer Biotechnology Working Group, Surfers Paradise, Queensland, Australia, 26-30 June.

Charcosset, A., Causse, M., Moreau, L. and Gallais, A. (1994) Investigation into the effect of genetic background on QTL expression using three connected maize inbred lines (RIL) populations *In*: 'Biometrics in Plant Breeding: Applications of Molecular Markers', (Ed. Ooijen, J. W. and J. Jansen), Wageningen. pp. 75-84.

Cheliak, W. M. and Rogers, D. L. (1990) Integrating biotechnology into tree improvement programs. *Canadian Journal of Forest Research*. **20**: 452-463.

Cheng, F. S. H. and Roose, M. L. (1995) Origin and inheritance of dwarfing by the citrus rootstock *Poncirus trifoliata* flying dragon. *Journal of American Society of Horticulture*. 120: 286-291.

Cheng, F. S., Weeden, N. F. and Brown, S. K. (1996) Identification of co-dominant RAPD markers tightly linked to fruit skin colour in apple. *Theoretical & Applied Genetics.* **93**: 222-227.

Chevereud, J. M., Routman, F. A., Duarte, F. A. M., Swinderen, K., Cothran, K., and Perel, C. (1996) Quantitative trait loci for Murine growth. *Genetics*. **142**: 1305-1319.

Chou, Q., Russell, M. Birch, D. E., Raymond, J. and Bloch, W. (1992) Prevention of pre-PCR mis-priming and primer dimerization improves low copy-number amplifications. *Nucleic Acids Research*. **20**: 1717-1723.

Churchill, G. A. and Doerge, R. W. (1994) Empirical threshold values for quantitative trait mapping. *Genetics*. 138: 963-971.

Clark, A. and Lanigan, C. M. S. (1993) Prospects for estimating nucleotide divergence with RAPDs. *Molecular Biology & Evolution*. **10**: 1096-1111.

Conkle, M. T. (1981) Isozyme variation and linkage in six conifer species *In*, 'Proceedings of the symposium on isozymes in North American Trees and forest insects' (Pacific south-west experiment station, USDA Genetics Technical Report). PSW-**48**: 11-17.

Costa, P. and Durel, C. E. (1996) Time trends in genetic control over height and diameter in Maritime pine. *Canadian Journal of Forest Research*. **26**: 1209-1217.

Crouzillat, D., Lerceteau, E., Petiard, V., Morera, J., Rodriguez, H., Walker, D., Phillips, W., Ronning, C., Schnell, R., Osei, J. and Fritz, P. (1996) *Theobroma cacao* L.: a genetic linkage map and quantitative trait loci analysis. *Theoretical & Applied Genetics. 93*: 205-214.

Darvasi, A. and Soller, M. (1994) Selective DNA pooling for determination of linkage between a molecular marker and a quantitative trait locus. *Genetics*. **138**: 1365-1373.

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

Davinregli, A., Abed, Y., Charrel, R. N., Bollet, C. and Demicco, P. (1995) Variations in DNA concentrations significantly affect the reproducibility of RAPD fingerprint patterns. *Research in Microbiology*. **146**: 561-568.

Davis, T. M., Yu, H., Haigis, K. M. and McGowan, P. J. (1995) Template mixing: a method of enhancing detection and interpretation of co-dominant RAPD markers. *Theoretical & Applied Genetics*. **91**: 582-588.

Dekoning, G. J. and Weller, J. I. (1994) Efficiency of direct selection on quantitative trait loci for a two-trait breeding objective. *Theoretical & Applied Genetics*. 88: 669-677.

Dellaporta, S. L., Wood, J. and Hicks, J. B. (1983) A plant DNA minipreparation: Version II. *Plant Molecular Biology Reporter*. **1**: 19-22.

Devey, M. E., Bell, J. C., Smith, D. N., Neale, D. B. and Moran, G. F. (1996) A genetic linkage map of radiata pine based on RFLP, RAPD and microsatellite markers. *Theoretical & Applied Genetics*. **92**: 673-679.

Devey, M. E.; Delfino-Mix, A.; Kinloch, B. B. and Neale, D. B. (1995) Random amplified polymorphic DNA markers tightly linked to a gene for resistance to

white pine blister rust in sugar pine. *Proceedings of the National Academy of Science, USA*. **92**: 2066-2070.

Doerge, R. W. and Rebai, A. (1995) Significance threshold for QTL interval mapping tests. *Heredity*. 76: 459-464.

Drew, A. P. and Bazzaz, F. A. (1978) Variation in distribution of assimilates among plant parts in three populations of *Populus deltoides*. *Silvae Genetica*. **27**: 189-193.

Dudley, J. W. (1993) Molecular markers in plant improvement: manipulation of genes affecting quantitative traits. *Crop Science*. **33**: 660-668.

Dweikat, I., Mackenzie, S., Levy, M. and Ohm, H. (1992) Pedigree assessment using RAPD-DGGE in cereal crops. *Theoretical & Applied Genetics*. **85**: 497-505.

Eck HJ van, Jacobs JME, Stam P, Ton J, Stiekema WJ, Jacobsen E (1994) Multiple alleles for tuber shape in diploid potato detected by quantitative genetic analysis using RFLPs. *Genetics***. 137**: 303-309

Edwards, M. D., Stuber, C. W. and Wendel, J. F. (1987) Molecular-marker facilitated investigations of quantitative trait loci in maize. I. Numbers, genomic distribution and types of gene action. *Genetics*. **116**: 113-125.

Ellsworth, D. L., Rittenhouse, D. and Honeycutt, L. R. (1993) Artifactual variation in randomly amplified polymorphic DNA banding patterns. *BioTechniques*. **14**: 214-217.

Emebiri. LC, Devey, MI, Matheson, AC and Slee, M. (1997) Linkage of RAPD markers to NESTUR, a stem growth index of radiata pine seedlings. *Theoretical & Applied Genetics* (in press).

Erlich, H. A. and Arnheim, N. (1992) Genetic analysis using polymerase chain reaction. *Annual Review of Genetics*. **26**: 479-506.

Erlich, H. A., Gelfand, D. and Sninsky, J. J. (1991) Recent advances in the polymerase chain reaction. *Science*. **252**: 1643-1651.

Falconer, D. S. (1989) Introduction to quantitative genetics. Longman Scientific & Technical, Essex, England.

Fjellstrom, R. G. and Parfitt, D. E. (1994) RFLP inheritance and linkage in walnut. *Theoretical & Applied Genetics*. **89**: 665-670.

Forrest, G. I. (1994) Biochemical markers in tree improvement programmes. *Forestry Abstracts*. **55**: 123-153.

Gerber, S., Rodolphe, F., Bahrman, N. and Baradat, P. (1993) Seed-protein variation in maritime pine (*Pinus pinaste*r Ait.) revealed by two-dimensional electrophoresis: genetic determinism and construction of a linkage map. *Theoretical & Applied Genetics*. **85**: 521-528.

Gianfranceschi, L. Koller, B., Seglias, N., Kellerhals, M. and Gessler, C. (1996) Molecular selection in apple for resistance to scab caused by *Venturia inaequalis*. *Theoretical & Applied Genetics.* **93**: 199-204.

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

Grattapaglia, D., Bertolucci, F. L. and Sederoff, R. R. (1995) Genetic mapping of QTL 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, F. L., Penchel, R. and Sederoff, R. R. (1995) Advances in genetic mapping of *Eucalyptus grandis*. In: 'Eucalypt plantations: improving fiber yield and quality', proceedings of CRCTHF-IUFRO Conference, Hobart, Australia, 19-24 February, pp. 392-397.

Grattapaglia, D., Chaparro, J., Wilcox, P, McCord, S., Crane, B., Amerson, H., Werner, D., Liu, B. H., O'Malley, D., Whetten, R., McKeand, S., Goldforb, B., Greenwood, M., Kuhlman, G., Bridgewater, F. and Sederoff, R. (1993) Application of genetic markers to tree breeding. *In*: 'Proceedings of the 22nd Southern Forest Tree Improvement Conference', Atlanta, Georgia. **44**: 452-463.

Greenwood, M. S. and Volkeart, H. A. (1992) Morphological traits as markers for the early selection of conifer genetic families. *Canadian Journal of Forest Research*. **22**: 1001-1008.

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

Guilford, P., Prakash, S., Zhu, J. M., Rikkerink, E., Gardiner, S., Basett, H. and Forster, R. (1997) Microsatellites in Malus x domestica (apple): abundance, polymorphism and cultivar identification. *Theoretical & Applied Genetics.* **94**: 249-254.

Haldane, J. B. S. (1919) The combination of linkage values and the calculation of distances between the loci of linked factors. *Journal of Genetics.* **8**: 299-309.

Haley, C. S. and Knott, S. A. (1992) A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity*. **69**: 315-324.

Haley, S. D., Afanador, L. and Kelly, J. D. (1994) Selection for monogenic pest resistance with coupling-and repulsion-phase RAPD markers. *Crop Science*. **34**: 1061-1066.

Hallden, C., Hansen, M., Nilsson, N. -O., Hjerdin, A. and Sall, T. (1996) Competition as a source of errors in RAPD analysis. *Theoretical & Applied Genetics*. **93:** 1185-1192**.**

Hallden, C., Hjerdin, A., Rading, I. M., Säll, T., Fridlundh, B., Johannisdottir, G., Tuvesson, S., Akesson, C. and Nilsson, N. -O. (1996) A high density RFLP linkage map of sugar beet. *Genome*. **39**: 634-645.

Han, K-H. Bradshaw, H. D. and Gordon, M. P. (1994) Adventitious root and shoot regeneration *in vitro* is under major gene control in an F_2 family of hybrid Poplar (*Populus trichocarpa* x *P. deltoides*). *Forest Genetics*. **1**: 139-146.

He, S., Ohim, H. and Mackenzie, S. (1992) Detection of DNA sequence polymorphisms among wheat varieties. *Theoretical & Applied Genetics*. 84: 573-578.

Heinze, B. and Geburek, Th. (1995) Searching for DNA markers linked to leaf colour in copper beech, *Fagus sylvatica* L. var. *atropunicea*. *Silvae Genetica*. **44**: 339-343.

Helsel, D. B. and Frey, K. J. (1983) Genetic control of vegetative growth index in oats (*Avena* spp.) *Theoretical & Applied Genetics*. **65**: 219-223.

Hemmat, M., Weeden, N. F., Manganaris, A. G. and Lawson, D. M. (1994) Molecular marker linkage map for apple. *Journal of Heredity*. **85**: 4-11.

Heun, M. and Helentjaris, T. (1993) Inheritance of RAPDs in F_1 hybrids of corn. *Theoretical & Applied Genetics*. **85**: 961-968.

Heun, M., Kennedy, A. E., Anderson, J. A., Lapitan, N. L. V., Sorrells, M. E. and Tanksley, S. D. (1991) Construction of a restriction fragment length polymorphism map for barley (*Hordeum vulgare*). *Genome*. **34**: 437-447.

Hormaza, J. I., Dollo, L. and Polito, V. S. (1994) Identification of RAPD markers linked to sex determination in *Pistacia vera* using bulked segregant analysis. *Theoretical & Applied Genetics*. **89**: 9-13.

Horvat, S. and Medrano, J. F. (1995) Interval mapping of *high growth* (*hg*), a major locus that increases weight gain in Mice. *Genetics*. **139**: 1737-1748.

Hu, J., Quiros, C., Arus, P., Sruss, D. and Robbelen, G. (1995) Generation of DNA-based markers in specific genome regions by two-primer RAPD reactions. *PCR Methods & Applications*. **4**: 346-351.

Hulbert, S. H., Ilott, T. W., Legg, E. J., Lincoln, S. E., Lander, E. S. and Michelmore, R. W. (1987) Genetic analysis of the fungus, *Bremia lactucae*, using restriction fragment length polymorphism. *Genetics*. **120**: 947-958.

Hunt, G. J. and Page, R. E. (1992) Patterns and inheritance with RAPD molecular markers reveal novel types of polymorphism in the honey bee. *Theoretical & Applied Genetics*. **85**: 15-20.

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

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

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

Jermstad, K. D., Reem, A. M., Henifin, J. R., Wheeler, N. C. and Neale, D. B. (1994) Inheritance of restriction fragment length polymorphisms and random amplified polymorphic DNAs in coastal Douglas-fir. *Theoretical & Applied Genetics*. **89**: 758-766.

Jett, J. B. (1988) Thirty-five years later: an overview of tree improvement in the southeastern United States *In*: 'Proceedings of the Southern Forest Nursery Association', 25-28 July 1988, Charleston, S. C. Southern Forest Nursery Association. pp. 98-106.

Johns, M. A. (1992) Comparison of RAPD maps for two sets of maize recombinant inbred lines *In*: 'Abstract Poster sessions, Plant Genome I meeting', San Diego, CA, p 32.

Kaback, D. B., Guacci, V., Barber, D. and Mahon, J. W. (1992) Chromosome size-dependent control of meiotic recombination. *Science*. 256: 228-232.

Kata, S. R., Taylor, B. H., Bockholt, A. J. and Smith, J. D. (1994) Identification of opaque-2 genotypes in segregating populations of quality protein maize by analysis of restriction fragment length polymorphisms. *Theoretical & Applied Genetics*. **89**: 407-412.

Kesseli, R. V., Paran, I. and Michelmore, R. W. (1994) Analysis of detailed genetic linkage map of *Latuca sativa* (Lettuce) constructed from RFLP and RAPD markers. *Genetics*. **136**: 1435-1446.

Korol, A. B., Preygel, I. A. and Preygel, S. I. (1994) 'Recombination variability and evolution: algorithms of estimation and population-genetic models', Chapman & Hall, Melbourne. pp 354.

Kosambi, D. D. (1944) The estimation of recombination values. *Annals of Eugenics*. **12**: 172-175.

Kozlowski, T. T. (1971) 'Growth And Development of Trees'. Academic press, New York.

Kremer, A. (1992) Prediction of age-age correlations of total height based on serial correlations between height increments in Maritime pine (*Pinus pinaster* Ait.). *Theoretical & Applied Genetics*. **85**: 152-158.

Krugman, S. L. (1985) Traditional forest genetics vs. biotechnological and physiological approaches *In*: 'New ways in forest genetics', proceedings of the 20th meeting of Canadian Tree Improvement Association (Eds. Caron, F., Corriveau, A. G. and Boyle, T. J. B.), pp. 62-67.

Kubisiak, T. L., Nelson, C. D., Nance, W. L. and Stine, M. (1995) RAPD linkage mapping in a longleaf x slash pine F1 family. *Theoretical & Applied Genetics*. **90**: 1119-1127.

Lanaud, C., Risterucci, A. M., N'Goran, A. K. J., Clement, D., Flament, M. H., Laurent, V. and Falque, M. (1995) A genetic linkage map of *Theobroma cacao* L. *Theoretical & Applied Genetics*. **91**: 987-993.

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

Lander, E. S. and Botstein, D. (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*. **121**: 185-199.

Lander, E. S. and Schork, N. J. (1994) Genetic dissection of complex traits. *Science*. **265**: 2037-2048.

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

Lebowitz, R. J., Soller, M. and Beckman, J. S. (1987) Trait-based analysis for the detection of linkage between marker loci and quantitative trait loci crosses between inbred lines. *Theoretical & Applied Genetics*. **73**: 556-562.

Lebreton, C., Lazic-Jancic, V., Steed, A., Pekic, S. and Quarrie, S. A. (1995) Identification of QTL for drought responses in maize and their use in testing causal relationships between traits. *Journal of Experimental Botany*. 46: 853-865.

Lehner, A., Campbell, M. A., Wheeler, N. C., Poykko, T., Glossl, J., Kreike, J. and Neale, D. B. (1995) Identification of a RAPD marker linked to the pendula gene in Norway spruce (*Picea abies* (L.) Karst. f. pendula). *Theoretical & Applied Genetics*. **91**: 1092-1094.

Lin, H. -X., Qian, H. -R., Zhuang, J. -Y., Lu, J., Min, S. -K., Xiong, Z. -M., Huang, N. and Zheng, K. -L. (1996) RFLP mapping of QTL for yield and related characters in rice (*Oryza sativa* L.). *Theoretical & Applied Genetics*. **92**: 920-927.

Lin, J. -Z. and Ritland, K. (1996) Construction of a genetic linkage map in the wild plant *Mimulus* using RAPD and isozyme markers. *Genome*. **39**: 63-70.

Lu, M., Szmidt, A. E. and Xiao-Ru, W. (1995) Inheritance of RAPD fragments in haploid and diploid tissue of *Pinus sylvestris* (L.). *Heredity*. **74**: 582-589.

Lynch, P. J., Rattunde, K. J. and Frey, K. J. (1995) Inheritance of vegetative growth index and related traits in Pearl millet. *Crop Science*. **35**: 394-396.

Macpherson, J. M., Eckstein, P. E., Scoles, G. J. and Gajadhar, A. A. (1993) Variability of the random amplified polymorphic DNA assay among thermal cyclers, and effects of primer and DNA concentration. *Molecular & Cellular Probes*. **7**: 293-299.

Marino, M. A., Turni, L. A., Delrio, S. A. and Williams, P. E. (1994) Molecular size determinations of DNA restriction fragments and polymerase chain

reaction products using capillary gel electrophoresis. *Journal of Chromatography*. **676**: 185-189.

Mather, K. and Jinks, J. L. (1982) 'Biometrical Genetics: the study of continuos variation'. Chapman & Hall, New York.

Matheson, A. C., Spencer, D. J. and Kriedmann, P. E. (1995) Age-Age correlation and early selection in radiata pine. 1. Family x environment interactions in plantation and greenhouse. *Australian Forestry*. **58**: 35-43.

Matheson, A. C., Spencer, D. J. and Magnussen, D. (1994) Optimum age for selection in *Pinus radiata* using basal area under bark for age: age correlations. *Silvae Genetica*. **43**: 352-357.

Matise, T. C., Perlin, M. and Chakravarti, A. (1994) Automated construction of genetic linkage maps using expert system (Multimap): A human genome linkage map. *Nature Genetics*. **6**: 384-390.

Meagher, R. B., McLean, M. and Arnold, J. (1988) Recombination within a subclass of restriction fragment length polymorphisms may help link classical and molecular genetics. *Genetics*. **120**: 809-818.

Mehra, P. N. and Khoshoo, T. N. (1956) Cytology of conifers. I. *Journal of Genetics*. 54: 165-180.

Meunier, J. R. and Grimont, P. A. D. (1993) Factors affecting reproducibility of random amplified polymorphic DNA fingerprinting. *Research in Microbiology*. **144**: 373-379.

Mhameed, S., Hillel, J., Lahav, E., Sharon, D. and Lavi, U. (1995) Genetic association between DNA fingerprint fragments and loci controlling agriculturally important traits in Avocado (*Persea americana*). *Euphytica*. **84**: 81-87.

Michelmore, R. W., Paran, I. and Kesseli, R. V. (1991) Identification of markers linked to disease resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions using segregating populations. *Proceedings of the National Academy of Science, USA*. **88**: 9828-9832.

Monforte, A. J., Asins, M. J. and Carbonell, E. A. (1996) Salt tolerance in Lycopersicon species. IV. Efficiency of marker-assisted selection for salt tolerance improvement. *Theoretical & Applied Genetics*. **93**: 765-772.

Moran, G. F. and Bell, J. C. (1987) The origin and genetic diversity of *Pinus radiata* in Australia. *Theoretical & Applied Genetics*. **73**: 616-622.

Muralidharan, K. and Wakeland E. K. (1993) Concentration of primer and template qualitatively affects products in random-amplified polymorphic DNA PCR. *BioTechniques*. **14**: 362-364.

Namkoong, G. and Conkle, M. T. (1976) Time trends in genetic control of height growth in Ponderosa pine. *Forest Science*. **22**: 2-12.

Nance, W. L. and Nelson, C. D. (1989) Restriction fragment length polymorphisms and their use in marker assisted selection in southern pine improvement *In*: 'Proceedings of the 20th Southern Forest Tree Improvement Conference', Charleston, South Carolina. The National Technical Information Service, Springfield, Virginia. pp. 50-59.

Nance, W. L., Tuskan, G. A., Nelson, C. D. and Doudrick, R. L. (1992) Potential applications of molecular markers for genetic analysis of host-pathogen systems in forest trees. *Canadian Journal of Forest Research.* **22**: 1036-1043.

Neale, D. B. and Sederoff, R. R. (1996) Genome mapping in gymnosperms: a case study in loblolly pine (*Pinus taeda* L.). *In*: 'Genome mapping in plants', (Ed. Andrew H. Paterson), R. G. Landes Company. pp. 309-319.

Neale, D. B. and Williams, C. G. (1991) Restriction fragment length polymorphism mapping in conifers and applications to forest genetics and tree improvement. *Canadian Journal of Forest Research*. **21**: 545-554.

Neale, D. B., Devey, M. E., Jermstad, K. D., Ahuja, M. R., Alosi, M. C. and Marshall, K. A. (1992) Use of DNA markers in forest tree improvement research. *New Forests*. **6**: 391-407.

Nelson, C. D., Kubisiak, T. L., Stine, M., and Nance, W. L. (1994) A genetic linkage map of Longleaf pine (*Pinus palustris* Mill) based on random amplified polymorphic DNAs. *Journal of Heredity*. **85**: 433-439.

Nelson, C. D., Nance, W. L. and Doudrick, R. L. (1993) A partial linkage map of Slash pine (*Pinus elliottii* Engelm var. *Elliottii*) based on random amplified polymorphic DNAs. *Theoretical & Applied Genetics*. **87**: 145-151.

Nienhuis, J., Helentjaris, T., Slocum, M., Ruggero, B. and Schaefer, A. (1987) Restriction fragment length polymorphism analysis of loci associated with insect resistance in tomato. *Crop Science*. **27**: 797-803.

Novy, R. G. and Vorsa, N. (1996) Evidence for RAPD heteroduplex formation in cranberry: implications for pedigree and genetic-relatedness studies and a source of co-dominant RAPD markers. *Theoretical & Applied Genetics*. **92**: 840-849.

O'Malley, D. M. and McKeand, S. E. (1994) Marker assisted selection for breeding value in forest trees. *Forest Genetics*. 14: 207-218.

Ohri, D. and Khoshoo, T. N. (1986) Genome size in gymnosperms. *Plant systematics and Evolution*. **153**: 119-131.

O'Malley, D. M. and McKeand, S. E. (1994) Marker assisted selection for breeding value in forest trees. *Forest Genetics*. **1**: 207-218.

Ooijen, J. W. (1992) Accuracy of mapping quantitative trait loci autogamous species. *Theoretical & Applied Genetics*. **84**: 803-811.

Ooijen, J. W. (1994a) DRAWMAP: a computer program for drawing genetic linkage maps. *Journal of Heredity*. **85**: 66.

Ooijen, J. W. (1994b) Comparison of a single-QTL model with an approximate multiple-QTL model for QTL mapping *In*: 'Biometrics in Plant Breeding: Applications of Molecular Markers', (Ed. J. W. van Ooijen and J. Jansen), Proceedings of the 9th meeting of EUCARPIA section Biometrics in plant breeding, Wageningen, the Netherlands. pp 205-212.

Ooijen, J. W. and Maliepaard, C. (1996) MAPQTL (tm) version 3.0: software for calculation of QTL positions on genetic maps. DLO-CPRO, Wageningen.

Pammi, S., Schertz, K., Xu, G., Hart, G. and Muller, J. E. (1994) Random-amplified polymorphic DNA markers in sorghum. *Theoretical & Applied Genetics*. **89**: 80-88.

Paran, I. and Michelmore, R. W. (1993) Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. *Theoretical & Applied Genetics*. **85** 985-993.

Park, Y and Kohel, R. J. (1994) Effect of concentration of $MgCl_2$ on random-amplified DNA polymorphism. *BioTechniques*. 16: 652-655.

Paterson, A. H., Lander, E. S., Hewitt, J. D., Peterson, S., Lincoln, S. E. and Tanskley, S. D. (1988) Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. *Nature*. **335**: 721-726.

Paterson, A., Damon, S., Hewitt, J. D., Zamir, D., Rabinowitch, H. D., Lincoln, S. E., Lander, S. E. and Tanksley, S. D. (1991) Mendelian factors underlying quantitative traits in tomato: comparison across species, generations, and environments. *Genetics*. **127**: 181-197.

Pawsey, C. K. (1964) Height and diameter growth cycles in radiata pine. *Australian Forest Research*. **1**: 3-8.

Payne, F. (1918) An experiment to test the nature of the variation on which selection acts. *Indiana University Studies*. **5**: 1-45.

Payne, P. I. (1987) Genetics of wheat storage proteins and the effect of allelic variation on bread-making quality. *Annual Review of Plant Physiology*. **38**: 141-153.

Penner, G. A., Bush, A., Wise, R., Kim, W., Domier, L., Kasha, K., Laroche, A., Scoles, G., Molnar, S. L. and Fedak, G. (1993) Reproducibility of random amplified polymorphic DNA, RAPD, analysis among laboratories. *PCR Methods and Applications*. **2**: 341-345.

Plomion, C., Baahrman, N., Durel, C. E. and O'Malley D. M. (1995a) Genomic mapping in maritime pine (*Pinus pinaster*) using RAPD and protein markers. *Heredity*. **74**: 661-668.

Plomion, C., Durel, C. E. and O'Malley, D. M. (1996) Genetic dissection of plant height in maritime pine seedlings raised under accelerated growth conditions. *Theoretical & Applied Genetics*. **93**: 849-858.

Plomion, C., O'Malley, D. M. and Durel, C. E. (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.

Pooler, M. R. and Scorza, R. (1995) Aberrant transmission of RAPD markers in haploids, doubled haploids and F1 hybrids of peach-observations and speculation on causes. *Scientia Horticulturae*. **64**: 233-241.

Powell, W., Machray, G. C.and Provan, J. (1996) Polymorphism revealed by simple sequence repeats . *Trends in Plant Science.* **1**: 215-222.

Pulkkinen, P. and Poykko, T. (1990) Inherited narrow crown form, harvest index and stem biomass production in Norway spruce, *Picea abies*. *Tree Physiology*. **6**: 381-391.

Pulkkinen, P., Poykko, T., Tigerstedt, P. M. A. and Velling. (1989) Harvest index in northern temperate cultivated conifers. *Tree Physiology*. **5**: 873-898.

Qi, X., Stam, P. and Lindhout, P. (1996) Comparison and integration of four barley genetic maps. *Genome*. **39**: 379-394.

Rafalski, J. A. and Tingey, S. V. (1993) Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. *Trends in Genetics*. **9**: 275-279.

Ragot, M. and Hoisinton, D. A. (1993) Molecular markers for plant breeding: comparisons of RFLP and RAPD genotyping costs. *Theoretical & Applied Genetics*. **86**: 974-985.

Ragot, M., Sisco, P. H., Hoinsinton, D. A. and Stuber, C. W. (1995) Molecular-marker-mediated characterisation of favourable exotic alleles at quantitative trait loci in maize. *Crop Science*. **35**: 1306-1315.

Rajapakse, S., Belthooff, L. E., He, G., Estager, A. E., Scorza, R., Verde, I., Ballard, R. E., Baird, W. V., Callahan, A., Monet, R. and Abbott, A. G. (1995) Genetic linkage mapping in peach using morphological, RFLP and RAPD markers. *Theoretical & Applied Genetics*. **90**: 503-510.

Reed, D. D., Pregitzer, K. S., Liechty, H. O., Burton, A. J. and Mroz, G. D. (1994) Productivity and growth efficiency in sugar maple. *Forest Ecology & Management*. **70**: 319-327.

Rees, A. V. and Durrant, A. (1986) Recombination and genome size. *Theoretical & Applied Genetics*. 73: 72-76.

Reiter, R. S., Williams, J., Feldman, K., Rafalski, J. A., Tingey, S. V. and Scolnik, P. A. (1992) Global and local genome mapping in *Arabidopsis thaliana* recombinant inbred lines and random amplified polymorphic DNAs. *Proceedings of the National Academy of Science, USA*. **89**: 1477-1481.

Riedel, K.-H. J., Winfield, B. D. and Britz, T. J. (1992) Combined influence of magnesium concentration and polymerase chain reaction specificity enhancers. *FEMS Microbiology Letters.* **92**: 69-72.

Riemenschneider, D. E., Heissig, B. E. and Bingham, E. T. (1987) Integrating biotechnology into woody plant breeding programs. *In*: 'Genetic manipulation of woody plants' (Ed. J. W. Hanover and D. E. Keathley), Pleum press, New York. pp. 433-450.

Ronning, C. M., Schnell, R. J. and Kuhn, D. N. (1995) Inheritance of random amplified polymorphic DNA (RAPD) markers in *Theobroma cacao* L. *Journal of the American Society for Horticultural Science*. **120**: 681-686.

Rowland, L. J. and Levi, A. (1994) RAPD-based genetic linkage map of blueberry derived from a cross between diploid species (*Vaccinium darrowi* and *V. elliottii*). *Theoretical & Applied Genetics*. **87**: 863-868.

Roy, A., Frascaria, N., Mackay, J. and Bonsquet, J. (1992) Segregating random amplified polymorphic DNAs (RAPDs) in *Betulla alleghanionsis*. *Theoretical & Applied Genetics*. 85: 173-180.

Ruano, G., Pagliaro, E. M., Schwartz, T. R., Lamy, K., Messina, D. (1992) Heat-soaked PCR: an efficient method for DNA amplification with applications to forensic analysis. *BioTechniques*. **13**: 266-274.

Saiki, R. K., Gelefond, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, B. T., Mullis, K. B. and Erlich, H. A. (1988) Primer directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*. **239**: 487-491.

Säll, T. and Nilsson, N.-O. (1994) The robustness of recombination frequency estimates in intercrosses with dominant markers. *Genetics*. 137: 589-596.

Sambrook, J., Fritsch, F. F. and Maniatis, T. (1989) 'Molecular cloning: a laboratory manual'. Cold Spring Harbor press, Cold Spring Harbor, New York.

Saylor, L. C. and Smith, B. W. (1966) Meiotic irregularity in specise and intraspecific hybrids of *Pinus*. *American Journal of Botany*. **53**: 453-468.

Schierwater, B. and Ender, A. (1993) Different thermostable DNA polymerases may amplify different DNA products. *Nucleic Acids Research*. **21**: 4647-4648.

Schon, C. C., Melchinger, A. E., Boppenmaier, J., Brunklaus-Jung, E., Hermann, R. G. and Seitzer, J. F. (1994) RFLP mapping in maize: quantitative trait loci affecting testcross performance of elite European flint lines. *Crop Science*. **34**: 378-389.

Schweder, M. E., Shatters, G. R., West, S. H. and Smith, R. L. (1995) Effect of transition interval between melting and annealing temperatures on RAPD analyses. *BioTechniques*. **19**: 38-42.

Serebrovsky, A. S. (1936) Experience of applying a new methods of genetic analysis to quantitative traits. *Dokl. Akad. Nauk. SSSR.*, **2**: 45-47 (In Korol *et al*. 1994).

Serebrovsky, A. S. (1970) 'Genetic analysis,. Nauka, Moscow.

Sheppard, L. J. and Ford, E. D. (1986) Genetic and environmental control of crown development in *Picea sitchensis* and its relation to stem wood production. *Tree Physiology*. **1**: 341-352.

Siddique, K. H. M and Whan, B. R. (1994) Ear-stem ratios in breeding populations of wheat: significance for yield improvement. *Euphytica*. **73**: 241-254.

Skroch, P. W. and Nienhius, J. (1995) Qualitative and quantitative characterisation of RAPD variation among snap bean (*Phaseolus vulgaris*) genotypes. *Theoretical & Applied Genetics*. **91**: 1078-1085.

Smith, D and Devey, M. E. (1994) Occurrence and inheritance of microsatellites in *Pinus radiata*. *Genome*. 37: 977-983.

Smith, J. S. C. and Williams, J. G. K. (1994) Arbitrary primer mediated fingerprinting in plants: Case studies in plant breeding, taxonomy and phylogeny. *In*: 'Molecular Ecology and Evolution: Approaches and Applications', (Ed. B. Schiewater, B. Streit, G. P. Wagner & R. DeSalle). Birkhauser Verlag Basel/Switzerland. pp. 5-15.

Snedecor, G. W. and Cochran, W. G. . (1980) 'Statistical Methods', Iowa State University press, Ames.

Snyder, F. W. and Carlson, G. E. (1984) Selecting for partitioning of photosynthetic products in crops. *Advances in Agronomy*. **37**: 47-72.

Soller, M. and Beckmann, J. S. (1983) Genetic polymorphism in varietal identification and genetic improvement. *Theoretical & Applied Genetics*. **67**: 837-843.

Sondur, S. N., Manshardt, R. M. and Stiles, J. I. (1995) Genetics of growth rate and flowering time in Papaya (*Carica papaya* L.). *Journal of Quantitative Trait Loci.* URL: http://probe.nalusda.gov:8000/.

Southern, E. M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology*. **98**: 503-517.

Spickett, S. G. and Thoday, J. M. (1966) Regular responses to selection. 3. Interaction between located polygenes. *Genetical Research*. **7**: 96-121.

St. Clair, J. B. (1994) Genetic variation in tree structure and its relation to size in Douglas-Fir. 1. Biomass partitioning, foliage efficiency, stem form and wood density. *Canadian Journal of Forest Research*. **24**: 1226-1235.

Stam, P. (1993) Construction of integrated genetic linkage maps by means of a new computer package: JOINMAP. *Plant Journal*. **3**: 739-744.

Stam, S. and Ooijen, J. W. (1995) 'JOINMAP (tm) version 2.0. Software for the calculation of genetic linkage maps', CPRO-DLO, Wageningen.

Staub, J. E. and Serquen, F. C. (1996) Genetic markers, map construction, and their application in plant breeding. *HortScience*. **31**: 729-740.

Stoneman, G. L. and Whitford, K. (1995) Analysis of the concept of growth efficiency in *Eucalyptus marginata* (Jarrah) in relation to thinning, fertilising and tree characteristics. *Forest Ecology & Management.* **76**: 47-53.

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

Stromberg, L. D., Dudley, J. W. and Rufener, G. K. (1994) Comparing conventional early generation selection with molecular marker assisted selection in maize. *Crop Science*. **34**: 1221-1225.

Stuber, C. F., Edwards, M. D. and Wendel, J. F. (1987) Molecular marker facilitated investigations of quantitative trait loci in maize: II. Factors influencing yield and its component traits. *Crop Science*. **27**: 639-648.

Stuber, C. W. (1994a). Breeding multigenic traits. *In*: 'DNA-based Markers in Plants', (Ed. R. L. Phillips and I. K. Vasil), Kluwer Academic publishers. pp 97-115.

Stuber, C. W. (1994b) Enhancement of grain yield in maize hybrids using marker-facilitated introgression of QTL. *In*: 'Proceedings of the American society for horticultural science/ Crop science society of America joint plant breeding symposium', 5-6 August, Oregon, pp. 44-46.

Stuber, C. W. (1994c) Successes in the use of molecular markers for yield enhancementin corn. *In*: 'Proceedings of the 49th corn & sorghum industry research conference', American Seed Trade Association, Chicago, Illinois, Dec. 8, (in press).

Stuber, C. W. (1995) Mapping and manipulating quantitative traits in maize. *Trends In Genetics*. **11**: 477-481.

Stuber, C. W., Goodman, M. M. and Moll, R. H. (1982) Improvement of yield and ear number resulting from selection at allozyme loci in a maize population. *Crop Science*. **22**: 737-740.

Stuber, C. W., Lincoln, S. E., Wolff, D. W., Helentjaris, T. And Lander, E. S. (1992) Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. *Genetics*. **132**: 823-839.

Stuber, C. W., Moll, R. H., Goodman, M. M., Schaffer, H. and Weir, B. S. (1980) Allozyme frequency changes associated with selection for increased grain yield in maize. *Genetics*. **95**: 225-236.

Sulzer, A. M., Greenwood, M. S. and Livinston, W. H. (1993) Early selection of black spruce using physiological and morphological criteria. *Canadian Journal of Forest Research.* **23**: 657-664.

Tanhuanpaa, P. K., Vilkki, J. P. and Vilkki, H. J. (1994) Segregation and linkage analysis of DNA markers in microspore derived and F_2 populations of oilseed rape (*Brassica napus* L.). *Euphytica*. **74**: 59-65.

Tanksley, S. D. (1993) Mapping polygenes. *Annual Review of Genetics*. **27**: 205-233.

Tanksley, S. D., Ganal, M. W., Prince, J. P., de Vincente, M. C., Bonierbale, M. W., Broun, P., Fulton, T. M., Giovanni, J. J., Grandillo, S., Martin, G. B., Messeguer, R., Miller, J. C., Miller, L., Paterson, A. H., Pineda, O., Roder, M., Wing, R. A., Wu, W. and Young, N. D. (1992) High density molecular linkage maps of the tomato and potato genomes: biological inferences and practical applications. *Genetics*. **132**: 1141-1160.

Tanksley, S. D., Young, N. D., Paterson, A. H. and Bonierbale, M. W. (1989) RFLP. mapping in plant breeding: new tools for an old science. *Bio/Technology*. **7**: 257-264.

Tauer, C. G. and Hallgren, S. W. (1992) Using marker-aided selection to improve tree growth response to abiotic stress. *Canadian Journal of Forest Research.* 22: 1018-1030.

Thoday, J. M. (1961) Location of polygenes. *Nature*. **191**: 368-370.

Thormann, C. E. and Osborn, T. L. (1992) Use of RAPD and RFLP markers for germplasm evaluation. In: 'Applications of RAPD Technology to plant breeding', Joint Plant Breeding Symposia Series, CSSA/ASHC/AGA. pp. 9-11.

Tinker, N. A. and Mather, D. E. (1995) MQTL: software for simplified composite interval mapping of QTL in multiple environments. *Journal of Quantitative Trait Loci,* URL: http://probe.nalusda.gov:8000/.

Tinker, N. A., Mather, D. E. and Fortin, M. G. (1994) Pooled DNA for linkage analysis: practical and statistical considerations. *Genome*. **37**: 999-1004.

Tulsieram, L. K., Glaubitz, J. C., Kiss, G. and Carlson, J. E. (1992) Single tree genetic linkage mapping in conifers using haploid DNA from megagametophytes. *Bio/Technology*. **10**: 686-690.

Utz, H. F. and A. E. Melchinger. (1994) Comparison of different approaches to interval mapping of quantitative trait loci. In: 'Biometrics in Plant Breeding: Applications of Molecular Markers', (Ed. Ooijen, J. W. and J. Jansen), Wageningen. pp. 195-204.

Vaillancourt, R. E., Potts, B. M., Manson, A., Eldridge, T. and Reid, J. B. (1995) Using RAPDs to detect QTL in an interspecific F_2 hybrid of Eucalyptus. In: 'Eucalypt Plantations: Improving Fiber Yield And Quality', CRCTHF-IUFRO conference, Hobart, Australia. pp. 430-433.

Vandezande, L. and Bijlsma, R. (1995) Limitations of RAPD technique in phylogeny reconstruction in Drosophila. *Journal of Evolutionary Biology*. **8**: 645-656.

Velling, P. and Tigerstedt, P. M. A. (1984) Harvest index in a progeny test of Scots pine with reference to the model of selection. *Silva Fennica*. **18**: 21-32.

Vinod, H. D. and Ullah, A. (1981) 'Recent advances in regression methods'. Marcel Dekker, Inc., New York, pp 361.

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

Wagner, D. B. (1992) Nuclear, chloroplast and mitochondrial DNA polymorphisms as biochemical markers in population genetic analyses of forest trees. *New Forests*. **6**: 373-390.

Wakamiya, I., Newton, R. J., Johnston, R. S. and Price, H. J. (1993) Genome size and environmental factors in the genus *Pinus*. *American Journal of Botany*. **80**: 1235-1241.

Wang, G. L. And Paterson, A. H. (1994) Assessment of DNA pooling strategies for mapping QTL. *Theoretical & Applied Genetics*. 88: 355-361.

Warburton, M. L., Becerra-Vasques, V. L., Goffreda, J. C. and Bliss, F. A. (1996) Utility of RAPD markers in identifying genetic linkages to genes of economic interest in peach. *Theoretical & Applied Genetics*. **93**: 920-925.

Weber, D. and Helentjaris, T. (1989) Mapping RFLP loci in maize using B-A translocations. *Genetics*. **121**: 583-590.

Weeden, N. F., Hemmat, M., Lawson, D. M., Lodhi, M., Bell, R. L., Manganaris, A. G., Reisch, B. I., Brown, S. K. and Ye, G. N. (1994) Development and application of molecular marker linkage maps in woody fruit crops. *Euphytica*. **77**: 71-75.

Welsh, J. and McClelland, M. (1990) Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research*. 18: 7213-7218.

Wettstein, D. von, Jende-Strid, B., Ahrenst-Larsen, B. and Sorensen, J. A. (1977) Biochemical mutant in barley renders chemical stabilization of beet superfluous. *Carlsberg Research Communications*. **42**: 341-351.

Wettstein, D. von, Nilan, R. A., Ahrenst-Larsen, B. Erdal, K., Ingversen, J., Jende-Strid, B., Nyegaard, K. K., Larsen, J., Outtrup, H. and Ullrich, S. E. (1985) Proanthocyanidin-free barley for brewing: progress in breeding for high yield and research tool in polyphenol chemistry. *MBAA Technical Quarterly*. **22**; 41-52.

Wilcox, P. L., Amerson, H. V., Kuhlman, E. G., Liu, B. H. and O'Malley, D. M. (1996) Detection of a major gene for resistance to fusiform rust disease in loblolly pine by genomic mapping. *Proceedings of the National Academy of Science, USA*.**93**: 3859-3864.

Wilkie, S. E., Isaac, P. G. and Slater, R. J. (1993) Random amplified polymorphic DNA (RAPD) markers for genetic analysis in *Allium*. *Theoretical & Applied Genetics*. **86**: 497-504.

Williams, C. G. (1987) The influence of shoot ontogeny on juvenile-mature correlations in loblolly pine. *Forest Science*. **33**: 411-422.

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

Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, J. A. and Tingey, S. V. (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research*. 18: 6531-6535.

Williams, J. G. K., Reiter, R. S., Young, R. M. and Scolnik, P. A. (1993) Genetic mapping of mutations using phenotypic pools and mapped RAPD markers. *Nucleic Acids Research*. **21**: 2697-2702.

Woeste, K., McGranaham, G. H. and Bernatzky, R. (1996) Randomly amplified polymorphic DNA loci from a walnut backcross (*Juglans hindsii* x *J. regia*) x *J. regia*. *Journal of American Society of Horticultural Science*. **121**: 358-361.

Wolff, K., Schoen, E. D. and Peters-Van Rijn, J. (1993) Optimising the generation of random amplified polymorphic DNAs in chrysanthemum. *Theoretical & Applied Genetics*. **86**: 1033-1037.

Yang, H. And Kruger, J. (1994) Identification of a RAPD marker linked to the Vf gene for scab resistance in apples. *Euphytica*. **77**: 83-87.

Yazdani, R., Yeh, F. C. and Rimsha, J. (1995) Genomic mapping of *Pinus sylvestris* (L.) using random amplified polymorphic DNA markers. *Forest Genetics*. **2**: 109-116.

Young, N. D. (1994) Constructing a plant genetic linkage map with DNA markers. *In*: 'DNA-based Markers in Plants', (Eds. Phillips, R. L. and Vasil, I. K.), Kluwer Academic Pub., pp. 39-57.

Yu, K. F., Deynze, A. V. and Pauls, K. P. (1993) Random amplified polymorphic DNA (RAPD) analysis. *In*: 'Methods in Plant Molecular Biology and Biotechnology', (Ed. Glick, B. R. and Thompson, J. E.), CRC press Inc. pp. 287-301.

Zeng, Z-B. (1993) Theoretical basis of separation of multiple linked gene effects on mapping quantitative trait loci. *Proceedings of the National Academy of Science, USA*. **90**: 10972-10976.

Zeng, Z-B. (1994) Precision mapping of quantitative trait loci. *Genetics*. **136**: 1457-1468.

Zeven, A. C., Knott, D. R. and Johnson, R. (1983) Investigation of linkage drag in near isogenic lines of wheat by testing for seedling reaction to races of stem rust, leaf rust and yellow rust. *Euphytica*. **32**: 319-327.

Zobel, B. J. and Talbert, J. T. (1984) 'Applied forest tree improvement', John Wiley & Sons, Inc., New York.

11 Appendix 4.1

Nutrient composition

Major Elements

Hydrated Calcium Nitrate [$Ca(NO_3)_2.4H_2O$]	950g/1000litres
Potassium Nitrate [KNO_3]	610g/1000litres
Hydrated Magnesium Sulphate [$MgSO_4.7H_2O$]	490g/1000litres
Ammonium Biphosphate [$NH_4.H_2PO_4$]	120g/1000litres

Trace Elements

Boric Acid [H_3B0_3]	0.6g/1000litres
Hydrated Manganase Chloride [$MnCl_2.4H_2O$]	0.4g/1000litres
Hydrated Zinc Sulphate [$ZnSO_4.7H_2O$]	0.09g/1000litres
Hydrated Copper Sulphate [$CuSO_4.5H_2O$]	0.05g/1000litres
Hydrated Molybdic Acid [$H_2MoO_4.4H_2O$]	0.02g/1000litres
Hydrated Cobalt Nitrate [$Co(NO_3)_2.6H_2O$]	0.025g/1000litres
EDTA	33.2g/1000litres
Hydrated Ferrous Sulphate [$FeSO_4.7H_2O$]	24.9g/1000litres