CHAPTER 15

Marker-assisted selection in forestry species

Penny Butcher and Simon Southerton

SUMMARY

The primary goal of tree breeding is to increase the quantity and quality of wood products from plantations. Major gains have been achieved using recurrent selection in genetically diverse breeding populations to capture additive variation. However, the long generation times of trees, together with poor juvenile-mature trait correlations, have promoted interest in marker-assisted selection (MAS) to accelerate breeding through early selection. MAS relies on identifying DNA markers, which explain a high proportion of variation in phenotypic traits. Genetic linkage maps have been developed for most commercial tree species and these can be used to locate chromosomal regions where DNA markers co-segregate with quantitative traits (quantitative trait loci, QTL). MAS based on QTL is most likely to be used for within-family selection in a limited number of elite families that can be clonally propagated. Limitations of the approach include the low resolution of marker-trait associations, the small proportion of phenotypic variation explained by QTL and the low success rate in validating QTL in different genetic backgrounds and environments. This has led to a change in research focus towards association mapping to identify variation in the DNA sequence of genes directly controlling phenotypic variation (gene-assisted selection, GAS). The main advantages of GAS are the high resolution of marker-trait associations and the ability to transfer markers across families and even species. Association studies are being used to examine the adaptive significance of variation in genes controlling wood formation and quality, pathogen resistance, cold tolerance and drought tolerance. Single nucleotide polymorphisms (SNPs) in these gene sequences that are significantly associated with trait variation can then be used for early selection. Markers for SNPs can be transferred among individuals regardless of pedigree or family relationship, increasing opportunities for their application in tree breeding programmes in developing as well as developed countries. Significant reductions in genotyping costs and improved efficiencies in gene discovery will further enhance these opportunities.

INTRODUCTION

Tree breeding offers a unique set of challenges associated with long generation times, outcrossing breeding systems and a relatively short history of genetic improvement. Breeding populations are often only one or two generations from the wild state. This has the advantage over crop breeding of providing vast stores of genetic variation that can be utilized in tree improvement. Tree breeding programmes have generally relied on testing and selecting large numbers of genotypes derived from multiple genetic backgrounds, the maintenance of high genetic diversity in production forests, and sexual propagation and capture of additive genetic variation through recurrent selection (Strauss, Lande and Namkoong, 1992). Inbreeding depression and long generation intervals have precluded the use of inbred lines, although research into their development continues (Wu, Abarquez and Matheson, 2004). The greatest use of interspecific hybrids in operational tree breeding has been with introduced species; for example, *Pinus elliottii* x *P. caribaea* in Australia (Nikles, 1996), hybrid eucalypts in South Africa, Brazil and the Congo (Eldridge *et al.,* 1993), *Acacia mangium* x *A. auriculiformis* in Viet Nam (Kha, Hai and Vinh, 1998) and hybrid poplars in temperate regions. These programmes often rely on clonal propagation for deployment.

The goal of commercial tree breeding is to increase the quantity and quality of wood products from plantations. Production of industrial timber was estimated at 2.8 thousand million cubic metres in 2004 and has been increasing at an average annual rate of 2.4 percent since 1998 (FAOSTAT) with much of the recent increase being due to rapid economic growth in China. Consumption of fuelwood is increasing at a similar rate (Carson, Walter and Carson, 2004). Rising demand together with restrictions on the supply of timber from native forests mean that increases in forest productivity will be required. To date, increased production has been achieved by expanding the area of plantations, particularly in tropical regions where high growth rates can be achieved. Gains have also been made using conventional breeding, but further productivity increases are required to reduce pressure on native forests and limit the increases in land area required for plantations. MAS has the potential to enhance plantation productivity if the relationship between genetic variation in gene sequences and phenotypic variation in traits can be demonstrated.

The relatively long generation times and poor juvenile-mature trait correlations in forest trees have promoted interest in MAS to accelerate breeding through early selection. MAS relies on identifying DNA markers which explain a high proportion of additive variation in phenotypic traits. Initially, research focused on the use of DNA markers in genome-wide linkage analysis of progeny arrays (Lander and Botstein, 1989). By identifying patterns of co-segregation in complex traits and polymorphic markers (QTL), these studies aimed to reveal causative regions of the chromosome or gene that were inherited intact over a few generations. The QTL approach can be used for marker-aided breeding within families. The low success rate in validating QTL in different genetic backgrounds and environments (Neale, Sewell and Brown, 2002) led to a change in research focus towards population-level association mapping. This approach seeks to find alleles of genes that affect the phenotype directly (Neale and Savolainen, 2004), and relies on the retention of much

smaller regions of intact DNA over many generations. Candidate genes targeted in these studies can be identified by gene mapping, expressed sequence tag (EST) sequencing, gene expression profiling or functional studies (transgenics). If variation can be found in the sequence of these genes in different phenotypes, it allows MAS to be used for within- and between-family selection in forest trees.

The application of biotechnology in tree improvement research is currently taking different paths in developed and developing countries due to contrasting regulatory approval processes for genetically modified plants and differences in public acceptance of genetically modified organisms (GMOs). There is considerable resistance in developed countries towards transgenic trees, which has more to do with their possible effects on other plants and on the environment than with concerns about transgenic wood (Sedjo, 2004). Long-term field trials are needed to ensure the stability of any genetic modification and the absence of negative impacts on growth and resistance to environmental stresses before they can be incorporated into industrial plantations (Strauss *et al.,* 1998; Strauss *et al.,* 2004). Regulation costs, possible trade restrictions, lack of public acceptance of transgenics and lack of support by major forestry certification groups such as the Forest Stewardship Council (FSC) are currently barriers to the development of transgenics (Sedjo, 2004). Consequently, trials of transgenic trees in developed countries remain in the research phase, mostly conducted with young trees grown under glasshouse conditions (see Walter and Killerby, 2004 for review). Due to these problems, some research has shifted towards alternative methods of investigating gene function and incorporating desirable genes into breeding populations, mainly through association mapping.

Recent MAS research in forest trees has been greatly assisted by advances in our understanding of tree genomes. The complete sequencing of plant genomes such as *Arabidopsis* (*Arabidopsis* Genome Initiative, 2000) and rice (Yu *et al.,* 2002) is improving our understanding of the number of genes involved (25 000–55 000) in the development of different organs and the function of the genes.

The *Populus* genome was the first tree genome to be sequenced with 58 036 predicted genes (www.jgi.doe.gov/poplar) and efforts are under way to sequence the *Eucalyptus* genome (www.ieugc.up.ac.za), a genus of particular importance in countries with developing economies in Asia and South America. To date, partial coverage of the *E. camaldulensis* genome (600 Mb) has been completed by random shotgun sequencing, through collaboration between Oji Paper and the Kasuza DNA Research Institute in Chiba, Japan (S. Potter Ensis, personal communication). A draft sequence, based on four-fold coverage of the genome is expected to be available by mid-2007 (Poke *et al.,* 2005). The large genome size of conifers is currently a barrier to whole genome sequencing; however, comprehensive EST sequencing is likely to yield most genes expressed in target tissues.

Genomic resources and tools are now being established for important forest tree species. Rapidly growing numbers of ESTs are publicly available in a range of species including *Eucalyptus grandis, Pinus radiata*, *P. taeda*, *Picea abies*, *Populus trichocarpa*, *P. tremula x tremuloides* and *Cryptomeria japonica* (see Krutovskii and Neale, 2001 and Strabala, 2004 for reviews) and *Avicennia marina* (Mehta *et al.,* 2005).

Over 80 000 ESTs have been sequenced from pine (http://pinetree.ccgb.umn.edu/), over 130 000 from poplar (http://poppel.fysbot.umu.se/) and over 100 000 from spruce (www.arborea.ulaval.ca/en/; www.treenomix.ca/).

Comprehensive microarrays (Shena *et al.*, 1996) are now being used in many of these species, allowing transcription profiling of thousands of genes in contrasting phenotypes in a range of tissues under different environmental/stress/developmental regimes. Identification of candidate genes from expression profiling is based on the assumption that the genes showing genotype-specific differences in their level of expression are causing variation in that trait (Morgante and Salamini, 2003). Microarrays are being used to identify genes that are regulated differentially in individuals with contrasting wood traits in eucalypts (Moran *et al.*, 2002), symbiosis-regulated genes in *Eucalyptus globulus-Pisolithus tinctorius* ectomycorrhiza (Voiblet *et al.*, 2001) and genes involved in embryogenesis in pines (van Zyl *et al.*, 2003). Using microarrays, many genes involved in cell wall biosynthesis have been identified in loblolly pine (Whetten *et al.*, 2001; Pavy *et al.*, 2005), eucalypts (Paux *et al.*, 2004) and hybrid aspen (*Populus tremula x P. tremuloides*) (Hertzberg *et al.*, 2001). A combination of proteomics, which examines changes in protein expression in different tissue and developmental stages, and microarray technology is also being used to give a more complete picture of gene function, for example of drought tolerance in *Pinus pinaster* (Plomion *et al.*, 2004). This discovery work is uncovering large numbers of candidate genes that are excellent targets for both QTL mapping and association studies aimed at identifying markers for use in MAS.

FAMILY-BASED GENETIC LINKAGE MAPPING AND QTL ANALYSIS

Genetic linkage or recombination mapping relies on finding sufficient polymorphism using DNA markers in progeny arrays from full-sib pedigrees to identify associations between linked loci on a chromosome. Genetic linkage maps have been constructed for most of the commercially important forest tree genera (summarized in Table 1), and updated information on genetic linkage maps for forest trees is available at http://dendrome.ucdavis.edu/index.php. The number and location of chromosomal regions affecting a trait (QTL) and the magnitude of their effect can then be investigated by QTL mapping. QTL are identified by a statistical association between variation in a quantitative trait and segregation of alleles at a marker locus in a segregating population (mapping pedigree).

Most phenotypic traits of interest for tree breeding are characterized by continuous variation. Such traits are usually influenced by a number of genes with a small effect interacting with other genes and the environment. The main traits targeted for QTL mapping are wood properties and traits related to adaptation and growth (reviewed by Sewell and Neale, 2000). These include physical wood properties that affect the strength of sawn timber (e.g. wood density and microfibril angle), and properties that affect paper pulping, e.g. pulp yield, fibre length and the relative proportion of cellulose, hemicellulose and lignin, generally measured as percentage cellulose. In addition, QTL have been identified for disease resistance, growth, flowering, vegetative propagation, frost tolerance and leaf oil composition (Table 2).

The detection of QTL requires large sample sizes; the lower the heritability

TABLE 1
Genetic linkage maps constructed for forest trees, markers used and location of mapping pedigrees

Species	Markers[1]	Country	Reference
Acacia mangium	RFLP, SSR	Australia	Butcher and Moran, 2000
Cryptomeria japonica	RFLP, RAPD, isozyme	Japan	Mukai *et al.,* 1995
Eucalyptus camaldulensis	RAPD, RFLP, SSR	Egypt	Agrama, George and Salah, 2002
Eucalyptus globulus	RAPD, SSR	Australia	Bundock, Hayden and Vaillancourt, 2000
	Candidate genes, isozymes, ESTP, RFLP, SSR	Australia	Thamarus *et al.,* 2002
Eucalyptus grandis x *E. globulus*	AFLP	Uruguay	Myburg *et al.,* 2003
Eucalyptus grandis x *E. urophyllla*	RAPD	Brazil Congo	Grattapaglia and Sederoff, 1994; Verhaegen and Plomion, 1996
Eucalyptus nitens	Isozyme, RAPD, RFLP	Australia	Byrne *et al.,* 1995
Eucalyptus tereticornis x *E. globulus*	AFLP	Portugal	Marques *et al.,* 1998
Eucalyptus urophylla x *E. tereticornis*	RAPD	China	Gan *et al.,* 2003
Fagus sylvatica	AFLP, RAPD, SSR	Italy	Scalfi *et al.,* 2004
Hevea braziliensis x H. benthamiana (rubber tree)	AFLP, isozymes, RFLP, SSR	French Guyana	Lespinasse *et al.,* 1999
Larix decidua & L. kaempferi	AFLP, ISSR, RAPD	France	Arcade *et al.,* 2000
Picea abies	RAPD	Italy	Binelli and Bucci, 1994
	RAPD	Denmark	Skov and Wellendorf, 1998
	AFLP, SAMPL, SSR	Italy	Paglia, Olivieri and Morgante, 1998
Picea glauca	ESTP, RAPD, SCAR	Canada	Gosselin *et al.,* 2002
Pinus edulis	AFLP	USA	Travis *et al.,* 1998
Pinus elliottii var elliottii	RAPD	USA	Nelson, Nance and Doudrick, 1993
Pinus elliottii var. elliottii & P. caribaea var. hondurensis	AFLP, SSR	Australia	Shepherd *et al.,* 2003
Pinus palustris	RAPD	USA	Kubisiak *et al.,* 1996
Pinus pinaster	RAPD	France	Plomion, O'Malley and Durel, 1995
	AFLP, RAPD, protein	France	Costa *et al.,* 2000
	AFLP	France	Chagne *et al.,* 2002
Pinus radiata	RFLP, RAPD, SSR	Australia	Devey *et al.,* 1996
Pinus sylvestris	RAPD	Sweden	Yazdani, Yeh and Rimsha, 1995
Pinus taeda	Isozymes, RAPD, RFLP	USA	Devey *et al.,* 1994; Sewell, Sherman and Neale, 1999
	AFLP	USA	Remington *et al.,* 1999
Pinus thunbergii	AFLP, RAPD	Japan	Hayashi *et al.,* 2001
Populus	AFLP, candidate genes, isozymes, ISSR, RAPD, RFLP, STS, SSR	Belgium, France, USA	See review in Cervera *et al.,* 2004; Yin *et al.,* 2004
Pseudotsuga menziesii	RAPD, RFLP	USA	Jermstad *et al.,* 1998; Krutovskii *et al.,* 1998
Quercus robur	Isozyme, minisatellite, RAPD, SCAR, SSR. 5SrDNA	France	Barreneche *et al.,* 1998
Salix viminalis	AFLP, SSR	UK	Hanley *et al.,* 2002
Salix viminalis x S. schwerinii	AFLP, RFLP	Sweden	Tsarouhas, Gullberg and Lagercrantz, 2002

[1] AFLP = amplified fragment length polymorphism; ESTP = expressed sequence tagged polymorphism; ISSR = inter-simple sequence repeats; RAPD = random amplified polymorphic DNA; RFLP = restriction fragment length polymorphism; SAMPL = selective amplification of microsatellite polymorphic loci; SCAR = sequence characterized amplified regions; SSR = simple sequence repeat (microsatellite); STS = sequence-tagged sites

TABLE 2
Quantitative trait loci reported for forest tree species

Species	Markers[1]	QTL	Reference
Acacia mangium	RFLP, SSR	Disease resistance	Butcher, 2004
Cryptomeria japonica	Isozyme, RAPD, RFLP	Juvenile growth, flowering, vegetative propagation	Yoshimaru *et al.,* 1998
	RAPD	Wood quality	Kuramoto *et al.,* 2000
Eucalyptus globulus	Isozyme, RFLP, SSR	Wood density, pulp yield, microfibril angle	Thamarus *et al.,* 2004
Eucalyptus grandis	RAPD	Growth, wood density	Grattapaglia *et al.,* 1996;
	RAPD	Disease resistance	Junghaus *et al.,* 2003
E.grandis x *E. urophylla*	RAPD	Vegetative propagation	Grattapaglia, Bertolucci and Sederoff, 1995; Marques *et al.,* 1999
	RAPD	Growth, wood density	Verhaegen *et al.,* 1997
	RAPD	Leaf oil composition	Shepherd, Chaparro and Teasdale, 1999
Eucalyptus nitens	RFLP	Growth	Byrne *et al.,* 1997a
	RFLP	Frost tolerance	Byrne *et al.,* 1997b
Eucalyptus tereticornis x *E. globulus*	AFLP	Vegetative propagation	Marques *et al.,* 2002
Fagus sylvatica	AFLP, RAPD, SSR	Leaf traits, growth	Scalfi *et al.,* 2004
Hevea braziliensis x H. benthamiana	AFLP, isozyme, RFLP, SSR	Disease resistance	Lespinasse *et al.,* 2000
Pinus palustris x *P. elliottii*	RAPD	Juvenile growth	Weng *et al.,* 2002
Pinus pinaster	RAPD	Bud set, frost tolerance	Hurme *et al.,* 2000
	AFLP	Growth, water use efficiency	Brendel *et al.,* 2002
Pinus radiata	RAPD	Growth	Emebiri *et al.,* 1997,1998a,b
	AFLP, RAPD, SSR	Wood density	Kumar *et al.,* 2000
	RFLP, SSR	Growth, wood density, disease resistance	Devey *et al.,* 2004a,b
Pinus sylvestris	AFLP	Growth, cold acclimation	Lerceteau, Plomion and Andersson, 2000; Yazdani *et al.,* 2003
Pinus taeda	Isozymes, RAPD, RFLP	Growth	Kaya, Sewell and Neale, 1999
	RFLP	Physical wood properties	Groover *et al.,* 1994; Sewell *et al.,* 2000
	RFLP	Chemical wood properties	Sewell *et al.,* 2002
	ESTP, RFLP	Wood properties	Brown *et al.,* 2003
Populus	AFLP, ISSR, RAPD, RFLP, SCAR, SSR, STS	Growth, form, leaf architecture, leaf & bud phenology, disease resistance, wood quality	See review in Cervera *et al.,* 2004
Pseudotsuga menziesii	RAPD, RFLP	Vegetative bud flush	Jermstadt *et al.,* 2001a
	RAPD, RFLP	Cold hardiness	Jermstadt *et al.,* 2001b
	RAPD, RFLP	QTL x environment	Jermstadt *et al.,* 2003
Quercus robur	AFLP, RAPD, SCAR, SSR	Growth, bud burst	Saintagne *et al.,* 2004
Salix dasyclados x *S. viminalis*	AFLP	Growth, drought tolerance, bud flush	Tsarouhas, Gullberg and Lagercrantz, 2002, 2003; Rönnberg-Wästljung, Glynn and Weih, 2005

[1] AFLP = amplified fragment length polymorphism; ESTP = expressed sequence tagged polymorphism; ISSR = inter-simple sequence repeats; RAPD = random amplified polymorphic DNA; RFLP = restriction fragment length polymorphism; SCAR = sequence characterized amplified regions; SSR = simple sequence repeat (microsatellite); STS = sequence-tagged sites

of the trait and the larger the number of genes affecting the trait, the larger the sample size required (see Strauss, Lande and Namkoong, 1992). As shown by Brown *et al.* (2003), the use of small mapping populations of 100–200 segregating individuals, typical of most QTL studies in trees, is likely to cause an upward bias in the estimated phenotypic effect of QTL. Simulation and practical studies have shown that, in addition to sample size, QTL detection is affected by genetic background, environment and interactions among QTL. The location of QTL is imprecise as they can only be mapped to 5–10 cM. This may translate into a physical distance of several megabases, which may contain several hundred genes. The effect of a QTL is also likely to vary over time in perennial plants with changing biotic and abiotic factors (Brown *et al.,* 2003). This highlights the necessity of verifying QTL in different seasons, environments and genetic backgrounds (Sewell and Neale, 2000). The challenges of developing and genotyping the large progeny arrays required to locate QTL accurately in outbred pedigrees, and of verifying these QTL in different environments and ages, are such that MAS has not yet been applied in any commercial tree breeding programme.

In one of the most intensive studies on applying MAS to date, and based on data from over 1 300 individuals for wood density, 4 400 individuals for wood diameter from a single pedigree and using selective genotyping of the 50 highest and lowest scoring individuals for density and 100 of each for diameter, Devey *et al.* (2004a) were able to validate (in the same pedigree) two out of 13 QTL for diameter and eight out of 27 QTL for wood density in *Pinus radiata*. The effect of each QTL ranged from 0.8 to 3.6 percent of phenotypic variation, implying that these traits were controlled by a large number of genes, each of small effect. Using a different approach, Brown *et al.* (2003) used a verification population of 447 progeny (derived from re-mating the parents of the QTL pedigree) and an "unrelated population" of 445 progeny from the base pedigree to verify QTL in *Pinus taeda*. They found about half the QTL were detected in multiple seasons and fewer QTL were common to different populations.

An area where QTL mapping may assist breeders is in breaking linkages between negatively correlated traits. For example in *E. grandis* and *E. urophylla,* Verhaegen *et al.* (1997) reported co-location of QTL for the negatively correlated traits of wood density and growth. If these traits are controlled by tightly linked genes, markers could be used to select favourable recombinants.

Most markers used in QTL mapping have been anonymous markers that are unlikely to occur in a gene influencing a quantitative trait. In an attempt to increase the power of QTL mapping, candidate genes that may control the trait in question are being used as molecular markers. Candidate genes are typically sourced from the tissue of interest (e.g. xylem or leaves) and have either a known function intuitively related to the trait, or are of interest from studies of their expression using DNA microarrays. Comparative mapping and candidate gene approaches can utilize such information to search for homologous genes in different genomes. Candidate genes have been mapped to QTL for wood quality in *E. urophylla* and *E. grandis* (Gion *et al.,* 2000), *Pinus taeda* (Neale, Sewell and Brown, 2002), and *E. globulus* (Thamarus *et al.,* 2004). They have also been mapped to QTL for bud set and bud flush in *Populus deltoides* (Frewen *et al.,* 2000).

Kirst *et al.* (2004) measured transcript abundance in 2 608 genes in the differentiating xylem of 91 *E. grandis* backcross progeny. QTL analysis of lignin-related transcripts (expressed gene QTL [eQTL]) showed that their mRNA abundance is regulated by two genetic loci. Coordinated down-regulation of genes encoding lignin enzymes was observed in fast growing individuals, indicating that the same genomic regions are regulating growth and the lignin content and composition in the progeny. Comparative mapping has shown that gene content and gene order are conserved over long chromosomal regions among related species. Comparative maps are therefore likely to play an important role in enabling information on gene location and function to be transferred between species and genera. However, this will depend on orthologous genetic markers being mapped in each species (Krutovskii and Neale, 2001). To date, comparative maps have been published for *Populus* (Cervera *et al.,* 2001), *Pinus* (Devey *et al.,* 1999; Krutovsky *et al.,* 2004), *Quercus* and *Castanea* (Barreneche *et al.,* 2004).

MAS, based on QTL, is most likely to be used for within-family selection in a limited number of elite families that can be propagated clonally for deployment in large-scale industrial plantations. It is most suitable for traits that are expensive to measure or can only be detected after plants have been subjected to a particular stress or pathogen, and that have poor juvenile-mature correlations. Limitations of the approach include the low resolution of the marker-trait associations, the low proportion of phenotypic variation explained by QTL (generally less than 10 percent), and the low success rate in validating QTL in different genetic backgrounds and environments (Sewell and Neale, 2002). Recombination-based methodologies have been applied to inbred crop lines to positionally clone genes of interest in QTL regions (Salvi *et al.,* 2002); however, the use of this technique in forest trees is not practicable due to their outcrossing breeding system.

POPULATION-BASED ASSOCIATION STUDIES

Limitations of the QTL approach have led to a change in research focus towards identifying variations in the DNA sequence of genes directly controlling phenotypic variation, known by some as GAS. One of the main advantages of association genetics is the high resolution of marker-trait associations. As natural populations are used in association studies, recombinations that accumulate over many generations of the population break any long range associations between marker and trait leaving short stretches of the genome associated with the trait. If alleles or SNPs can be found that are strongly associated with superior phenotypes, they can be used for selection across breeding populations. This methodology is better suited to tree breeding programmes, which aim to maintain a broad genetic base, i.e. programmes with a large number of families. In contrast, the QTL approach is used for within-family selection. Spurious associations may be observed in association studies where there is undetected genetic structure in the breeding population that invalidates standard statistical tests. Strategies for dealing with population stratification have been developed to avoid these problems (Pritchard *et al.,* 2000; Wu and Zeng, 2001).

In the first association study published in forest trees, Thumma *et al.* (2005) identified 25 common SNP markers in the lignin biosynthesis gene *CCR* from *Eucalyptus nitens*. Using single-marker

and haplotype analyses in 290 trees from a natural population, they observed two haplotypes that were significantly associated with microfibril angle, a major determinant of timber strength. These results were confirmed in a full-sib family in *E. nitens* and in the related species *E. globulus.* In a powerful demonstration of the resolution of association genetics, Thumma *et al.* (2005) detected an alternatively-spliced variant of the *CCR* gene from the region of the significant haplotype, thereby revealing the probable molecular basis of the trait variation.

Association mapping is a particularly useful approach when genes are available that are likely to be functionally relevant to the trait of interest. Homologues of genes characterized in model species such as *Arabidopsis*, maize or rice, and poplar are excellent targets for association studies in forest species. In most cases, putative orthologues can be identified by comparing ESTs to gene sequences in public databases. In some cases, gene function may be determined by modulating the expression of selected genes using sense and antisense modification to up- and down-regulate gene expression, or intron RNA hairpin constructs to silence genes (Smith *et al.,* 2000). However, one of the advantages of association studies is the capacity to study a large number of genes simultaneously without the need for transformation (Peter and Neale, 2004).

There is considerable interest in understanding the genes controlling wood fibre cell wall development in forest trees as fibre microstructure is a major determinant of the commercial value of wood. For example, the angle of orientation of cellulose microfibrils (MFA) in fibre cell walls is known to affect timber strength and stiffness as well as fibre collapsibility, an important determinant of tensile strength in paper. Knowledge of cell wall biosynthesis would also assist in understanding and manipulating the development of abnormal wood, e.g. tension/compression wood (see Paux *et al.,* 2005; Pavy *et al.,* 2005), which is known to have an impact on wood stability, sawing patterns and pulpability. Wood is primarily composed of secondary xylem, and its properties are the product of sequential developmental processes from cambial cell division and expansion, to secondary wall formation and lignification. Genes expressed during xylogenesis determine the physical and chemical properties of wood. Important genes are now being identified that control the synthesis of the major constituents of the cell wall: cellulose, hemicellulose and lignin. Genes for cellulose synthesis (*CesA*) have been cloned in aspen (Joshi, Wu and Chiang, 1999; Wu, Joshi and Chiang, 2000), poplar (Djerbi *et al.,* 2005) and loblolly pine (Nairn and Haselkorn, 2005). Characterizing the *CesA* gene in aspen revealed strong similarity with a secondary cell wall protein in cotton, indicating that they serve similar functions in the two evolutionarily divergent genera. Transformation of cellulose synthase genes in aspen (Joshi, 2004) should further elucidate gene function. Each of the three loblolly *CesA* genes is orthologous to one of the three angiosperm secondary cell wall *CesA*s, suggesting functional conservation between angiosperms and gymnosperms. A search of the poplar genome revealed 18 distinct *CesA* gene sequences in *Populus trichocarpa* (Djerbi *et al.,* 2005). The *CesA* genes belong to a superfamily of *CesA*-like (*Csl*) genes, which includes a very large number of glycosyltransferases that are likely to be involved in the synthesis of the numerous non-cellulosic polysaccharides in plants (Liepman, Wilkerson and

Keegstra, 2005).

Lignin biosynthesis is well understood at the molecular level in plants (reviewed by Boerjan, Ralph and Baucher, 2003 and Peter and Neale, 2004) and is of particular interest in forest trees as removal of lignin for paper-making has major economic and environmental costs. In some cases, genetic modification of lignin structure has been shown to improve delignification in pulp and paper-making (Jouanin and Goujon, 2004), and down regulation of lignin pathway enzymes has also been shown to increase cellulose content (Hu *et al.*, 1999). Gymnosperms and angiosperms share a common set of enzymes that are responsible for the formation of guaiacyl lignin, while angiosperms have evolved at least two enzymes that catalyse the production of syringyl lignin. Association studies are now being carried out in loblolly to examine the adaptive significance of sequence variation in monolignol biosynthetic genes (Peter and Neale, 2004) and other genes controlling wood properties (Brown *et al.*, 2004). Similar research (Table 3) aimed at identifying genes controlling wood formation is being undertaken in Douglas fir (Krutovsky *et al.*, 2005), maritime pine (Pot *et al.*, 2004), radiata pine (S.G. Southerton and G.F. Moran, personal communication), spruce (MacKay *et al.*, 2005) and eucalypts (Moran *et al.*, 2002).

The availability of genes linked to a range of other traits in model plants opens up new areas of investigation in association genetics. For example, association studies are in progress to identify genes controlling pathogen resistance (Ersoz *et al.*, 2004; MacKay *et al.*, 2005), drought tolerance (Ersoz *et al.*, 2004), cold tolerance (Krutovsky *et al.*, 2005) and bud set (Paoli and Morgante, 2005) (Table 3). Further opportunities exist for association studies aimed at identifying SNPs linked to important traits. Flowering is particularly well understood at the molecular level (Zik and Irish, 2003), and increasing numbers of genes controlling flowering have been cloned in angiosperm tree species including eucalypts (Southerton *et al.*, 1998; Watson and Brill, 2004), silver birch (Elo *et al.*, 2001), poplar (Rottmann *et al.*, 2000) and gymnosperm tree species including spruce (Tandre *et al.*, 1995; Rutledge *et al.*,1998) and pines (Mouradov *et al.*, 1998, 1999).

Another important technological advance that is making large-scale association studies possible is the recent development of rapid, high-throughput

TABLE 3
Association studies in progress for forest tree species

Species	Trait	Reference
Eucalyptus nitens	Wood properties	Moran *et al.*, 2002; Thumma *et al.*, 2005
Populus	Wood properties	MacKay *et al.*, 2005
	Disease resistance	MacKay *et al.*, 2005
Picea glauca	Wood properties	MacKay *et al.*, 2005
	Disease resistance	MacKay *et al.*, 2005
Picea abies	Bud set	Paoli and Morgante, 2005
Pseudotsuga menziesii	Wood properties	Krutovsky *et al.*, 2005;
	Cold hardiness	Krutovsky *et al.*, 2005;
Pinus radiata	Wood properties	Southerton and Moran unpub. data
Pinus taeda	Wood properties	Peter and Neale, 2004; Brown *et al.*, 2004
	Disease resistance	Ersoz *et al.*, 2004
	Drought tolerance	Ersoz *et al.*, 2004
Pinus pinaster	Wood properties	Pot *et al.*, 2004

genotyping techniques that have drastically reduced the cost of genotyping SNPs in association populations (www.illumina.com/products/prod_snp.ilmn). It is now feasible to genotype SNPs in hundreds of genes potentially associated with a trait.

QTL mapping remains largely a research tool to improve our understanding of the number, distribution and mode of action of genes controlling quantitative traits. QTL can also play a role in GAS as a vehicle for validating significant SNP correlations identified in association populations (Thumma *et al.,* 2005). In the near future, association studies promise to yield numerous SNP markers that could be used in breeding programmes for early selection of superior alleles associated with a wide range of traits. As the efficiency of techniques for microarray analysis, SNP discovery, genotyping and other molecular procedures improve further, the opportunities to incorporate molecular technologies into breeding programmes for forest trees will increase.

USE OF MAS TO ENHANCE BREEDING PROGRAMMES IN DEVELOPING COUNTRIES

The adoption of molecular techniques varies widely, not only between developed and developing countries but also among the less developed economies (Chaix and Monteuuis, 2004). Countries such as China, India, Indonesia, Malaysia, Thailand and Viet Nam have established molecular laboratories for genotyping. Molecular markers are being used routinely to assess the level of genetic diversity in breeding programmes and to monitor any changes following selection (Butcher, 2003; Marcucci Poltri *et al.,* 2003). They are also being used to estimate levels of contamination and inbreeding in open-pollinated seed orchards (Chaix *et al.,* 2003; Harwood *et al.,* 2004), to validate intra- and interspecies crosses and to determine error rates in clonal propagation or trial establishment (see, for example, Bell *et al.,* 2004). This has identified relatively high error rates in several breeding programmes, affecting calculations of heritability, breeding value and genetic gain. Genetic linkage maps have been published for eucalypts in China (Gan *et al.,* 2003) and Brazil is prominent in eucalypt mapping and genomics (Grattapaglia, Chapter 14). EST libraries have been developed for mangroves in India as a first step towards characterizing genes associated with salinity tolerance (Mehta *et al.,* 2005), while DNA markers have been used for backward selection to identify superior male parents in eucalypt seed orchards in Brazil (Grattapaglia *et al.,* 2004). This approach has some potential for accelerating the breeding cycle in open-pollinated breeding programmes, particularly with species that are difficult to hand pollinate (Butcher, Moran and DeCroocq, 1998). The application of QTL-MAS in developing countries remains limited, exceptions being selection of coconut parents for breeding (FAO, 2003) and identification of QTL for rubber tree improvement (Lespinasse *et al.,* 2000). More widespread application may depend on economic considerations. Reports on the financial viability of MAS differ, with Johnson, Wheeler and Strauss (2000) indicating that large areas would need to be planted with MAS-improved germplasm to justify initial investment, while Wilcox *et al.* (2001) suggest significant financial gains are possible even when selection is based on DNA markers linked to a few loci each of relatively small effect. Association mapping has the potential for more widespread application in developing countries due, in part, to the ability to transfer markers among individuals, regardless of pedigree

or family relationships. The possibility of transferring SNP markers among species has already been demonstrated in eucalypts (Thumma *et al.,* 2005).

Forest trees, including many species in developing countries, are near their wild state, and significant improvement can still be made quite rapidly based on selection among existing genotypes (FAO, 2003). Association studies are ideally suited to exploiting variation in natural populations and do not rely on the existence of extensive pedigrees from controlled crosses. Suitable populations should include a large number of unrelated individuals of the same age growing on the one site. It has been estimated that 500 individuals would be necessary to detect association between a quantitative nucleotide responsible for 5 percent or more of the phenotypic variance (Long and Langley, 1999). The development of such populations would provide a good foundation for future GAS research in developing countries. While the markers developed using this approach are likely to be more easily transferred between breeding programmes, the application of GAS would require the subsequent development of advanced breeding programmes where the selection of superior alleles could take place. However, publicly funded forestry research is suboptimal in many developing countries and development priorities do not necessarily include genetic improvement programmes (FAO, 2003).

The major costs of GAS are associated with identifying candidate genes potentially linked to the relevant traits, and discovering SNPs. In some cases, public databases may contain large numbers of genes from the target or closely related species but, if not, there would be additional costs associated with EST sequencing of genes from the relevant tissue (i.e. xylem genes for wood traits). These additional costs may be offset partially by EST sequencing clones from mixed cDNA libraries derived from a number of unrelated trees from the association population.

Previously, the cost of genotyping SNPs was prohibitive, but this has fallen dramatically in recent years as high-throughput technologies have been developed for the human HapMap project (International HapMap Consortium, 2003). The Illumina Beadstation technology (www.illumina.com) is particularly suited to smaller-scale genotyping projects such as those being undertaken in forest trees. Cost is certainly a limitation in many developing countries including much of Africa and some South American countries; however, in most Asian countries and countries such as Brazil where molecular genetic laboratories are already well established, costs would not be prohibitive. The full benefits of GAS would require development of efficient clonal propagation and deployment systems before it was applied routinely.

Less stringent regulatory approval processes and greater public acceptance of genetically modified plants have allowed Brazil and China to take a lead role in commercializing transgenic tree technology. China is the only country to announce the commercial release of transgenics (poplar) with 300–500 hectares being planted in 2002 (Wang, 2004). Regulatory approval for the release of *Bacillus thuriengensis* (Bt) insect resistant eucalypts in Brazil is pending (Sedjo, 2004). Given the difficulty of carrying out long-term transgenic field trials with long rotation conifers, transgenic approaches are likely to be restricted to modification of high-value traits such as wood fibre properties in short rotation species grown on a large scale. Conventional breeding, using either open or controlled

pollination in seed orchards, will continue to be the most important mechanism for developing new genotypes for increased genetic gain (Carson, Walter and Carson, 2004).

ACKNOWLEDGEMENTS

We thank Simon Potter and Reddy Thumma (ENSISjv) for constructive comments on the manuscript.

REFERENCES

Agrama, H.A., George, T.L. & Salah, S.F. 2002. Construction of genome map for *Eucalyptus camaldulensis* Dehn. *Silv. Genet*. 51: 201–206.

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

Arcade, A., Anselin, F., Faivre Rampant, P., Lesage, M.C., Paques, L.E. & Prat, D. 2000. Application of AFLP, RAPD and ISSR markers to genetic mapping of European and Japanese larch. *Theor. Appl. Genet*. 100: 299–307.

Barreneche, T., Bodénès, C., Lexer, C., Trontin, J.F., Fluch, S., Streiff, R., Plomion, C., Roussel, G., Steinkellner, H., Burg, K., Favre, J.M., Glössl, J. & Kremer, A. 1998. A genetic linkage map of *Quercus robur* L. (pedunculate oak) based on RAPD, SCAR, microsatellite, minisatellite, isozyme, and 5S rDNA markers. *Theor. Appl. Genet.* 97: 1090–1103.

Barreneche, T., Casosoli, M., Russell, K., Akkak, A., Meddour, H., Plomion, C., Villai, F. & Kremer, A. 2004. Comparative mapping between *Quercus* and *Castanea* using simple-sequence repeats (SSRs). *Theor. Appl. Genet.* 108: 558–566.

Bell, J.C., Powell, M., Devey, M.E. & Moran, G.F. 2004. DNA profiling, pedigree lineage analysis and monitoring in the Australian breeding program of radiata pine. *Silv. Genet*. 53: 130–134.

Binelli, G. & Buchi, G. 1994. A genetic linkage map of *Picea abies* based on RAPD markers, as a tool in population genetics. *Theor. Appl. Genet.* 88: 283–288.

Boerjan, W., Ralph, J. & Baucher, M. 2003. Lignin biosynthesis. *Ann. Rev. Plant Biol*. 54: 519–46.

Brendel, O., Pot, D., Plomion, C., Rozenberg, P. & Guehl, J.M. 2002. Genetic parameters and QTL analysis of d13C and ring width in maritime pine. *Plant Cell Environ.* 25: 945–953.

Brown, G.R., Bassoni, D.L., Gill, G.P., Fontana, J.R., Wheeler, N.C., McGraw, R.A., Davis, M.F., Sewell, M.M., Tuskan, G.A. & Neale, D.B. 2003. Identification of quantitative trait loci influencing wood property traits in loblolly pine (*Pinus taeda* L.) III. QTL verification and candidate gene mapping. *Genetics* 164: 1537–1546.

Brown, G.R., Gill, G.P., Kuntz, R.J., Beal, J.A., Nelson, D., Wheeler, N.C., Penttila, B., Roers, J. & Neale, D.B. 2004. Associations of candidate gene single nucleotide polymorphisms with wood property phenotypes in loblolly pine, Abstract W98. In *Plant & Animal Genomes XII Conference,* San Diego, CA, USA.

Bundock, P.C., Hayden, M. & Vaillancourt, R. 2000. Linkage maps of *Eucalyptus globulus* using RAPD and microsatellite markers. *Silv. Genet*. 49: 223–232.

Butcher, P.A. 2003. Molecular breeding of tropical trees. *In* A. Rimbawanto, & M. Susanto, eds. *Proc. Internat. Conf. on Advances in Genetic Improvement of Tropical Tree Species,* pp. 59–66. Centre for Forest Biotechnology and Tree Improvement, Yogyakarta, Indonesia.

Butcher, P.A. 2004. Genetic mapping in acacias. *In* S. Kumar & M. Fladung, eds. *Molecular genetics and breeding of forest trees*, pp. 411–427. New York, USA, Haworth Press.

Butcher, P.A. & Moran, G.F. 2000. Genetic linkage mapping in *Acacia mangium* 2. Development of an integrated map from two outbred pedigrees using RFLP and microsatellite loci. *Theor. Appl. Genet.* 101: 594–605.

Butcher, P.A., Moran, G.F. & DeCroocq, S. 1998. Use of molecular markers in domestication and breeding programs for acacias. *In* J.W.Turnbull, H.R. Crompton, & K. Pinyopusarek, eds. *Recent developments in acacia planting*, pp. 211–215. ACIAR Proc. No. 82, Canberra. ACIAR.

Byrne, M., Murrell, J.C., Allen, B. & Moran, G.F. 1995. An integrated genetic linkage map for eucalypts using RFLP, RAPD and isozyme markers. *Theor. Appl. Genet*. 91: 869–875.

Byrne, M., Murrell, J.C., Owen, J.V., Kriedemann, P., Williams, E.R. & Moran, G.F. 1997a. Identification and mode of action of quantitative trait loci affecting seedling height and leaf area in *Eucalyptus nitens*. *Theor. Appl. Genet*. 94: 674–681.

Byrne, M., Murrell, J.C., Owen, J.V., Williams, E.R. & Moran, G.F. 1997b Mapping of quantitative trait loci influencing frost tolerance in *Eucalyptus nitens*. *Theor. Appl. Genet*. 95: 975–979.

Carson, M., Walter, C. & Carson, S. 2004. The future of forest biotechnology. *In* R. Kellison, S. McCord, & K.M.A Gartland, eds. *Forest biotechnology in Latin America*, pp. 13–38. Proceedings, Workshop Biotecnolegia Forestal, Biotechnology Forum, Conceptión, Chile (available at www.forestbiotech.org/pdffiles/ChlePDFfinal31Jan2005.pdf).

Cervera, M.T., Sewell, M.M., Rampant, P.-F., Storme, V. & Boejan, W. 2004. Genome mapping in *Populus. In* S. Kumar, & M. Fladung, eds. *Molecular genetics and breeding of forest trees*, pp. 387–410. New York, USA, Haworth Press.

Cervera, M.T., Stormel, V., Ivens, B., Gusmão, J., Liu, B.H., Hostyn, V., Van Slyckenc, J., Van Montagu, M. & Boerjan, W. 2001. Dense genetic linkage maps of three *Populus* species (*Populus deltoides*, *P. nigra* and *P. trichocarpa* based on AFLP and microsatellite markers. *Genetics* 158: 787–809.

Chagne, D., Lalanne, C., Madur, D., Kumar, S., Frigerio, J.-M., Krier, C., Decroocq, S., Savoure, Al, Bou-Dagher-Karrat, M., Bertocchi, E., Brach, J. & Plomion, C. 2002. A high density linkage map for maritime pine based on AFLPs. *Ann. For. Sci*. 59: 627–636.

Chaix, G., Gerber, S., Razafimaharo, V., Vigneron, P., Verhaegen, D. & Hamon, S. 2003. Gene flow estimation in a Malagasy seed orchard of *Eucalyptus grandis*. *Theor. Appl. Genet.* 107: 705–712.

Chaix, G. & Monteuuis, O. 2004. Biotechnology in the forestry sector. In *Preliminary review of biotechnology in forestry including genetic modification*. FAO Forest Genetic Resources Working Paper FGR/59E. Rome (available at www.fao.org/docrep/008/ae574e/ae574e00.htm).

Costa, P., Pot, D., Dubos, C., Frigerio, J.M., Pionneau, C., Bodenes, C., Bertocchi, E., Cervera, M.-T., Remington, D.L. & Plomion, C. 2000. A genetic map of Maritime pine based on AFLP, RAPD and protein markers. *Theor. Appl. Genet*. 100: 39–48.

Devey, M.E., Fiddler, T.A, Liu, B.-H., Knapp, J. & Neale, D.B. 1994. An RFLP linkage map for loblolly pine based on a three-generation pedigree. *Theor. Appl. Genet*. 88: 273–278.

Devey, M.E., Bell, J.C., Smith, D.N., Neale, D.B. & Moran, G.F. 1996. A genetic linkage map for *Pinus radiata* based on RFLP, RAPD and microsatellite markers. *Theor. Appl. Genet*. 92: 673–679.

Devey, M.E., Sewell, M.M., Uren, T.L. & Neale, D. 1999. Comparative mapping in loblolly and radiata pine using RFLP and microsatellite markers. *Theor. Appl. Genet*. 99: 656–662.

Devey, M.E., Carson, S.D., Nolan, M.F., Matheson, A.C., Te Riini, C. & Hohepa, J. 2004a. QTL association for density and diameter in *Pinus radiata* and the potential for marker-aided selection. *Theor. Appl. Genet*. 108: 516–524.

Devey, M.E., Groom, K.A., Nolan, M.F., Bell, J.C., Dudzinski, M.J., Old, K.M., Matheson, A.C. & Moran, G.F. 2004b. Detection and verification of quantitative trait loci for resistance to *Dothistroma* needle blight in *Pinus radiata*. *Theor. Appl. Genet*. 108: 1056–1063.

Djerbi, S., Lindskog, M., Arvestad, L., Sterky, F. & Teeri, T.T. 2005. The genome sequence of black cottonwood (*Populus trichocarpa*) reveals 18 conserved cellulose synthase (*CesA*) genes. *Planta* 221: 739–746.

Eldridge, K.G., Davidson, J., Harwood, C.E. & van Wyk, G. 1993. *Eucalypt Domestication and Breeding*. Oxford, Clarendon Press.

Elo, A., Lemmetyinen, J., Turunen, M.-L., Tikka, T. & Sopanen, T. 2001. Three MADS-box genes similar to *APETALA* 1 and *FRUITFULL* from silver birch (*Betula pendula*). *Physiol. Plant* 112: 95–103.

Emebiri, L.C., Devey, M.E., Matheson, A.C. & Slee, M.U. 1997. Linkage of RAPD markers to NESTUR, a stem growth index in radiata pine seedlings. *Theor. Appl. Genet.* 95: 119–124.

Emebiri, L.C., Devey, M.E., Matheson, A.C. & Slee, M.U. 1998a. Age-related changes in the expression of QTLs for growth in radiata pine seedlings. *Theor. Appl. Genet.* 97: 1053–1061.

Emebiri, L.C., Devey, M.E., Matheson, A.C. & Slee, M.U. 1998b. Interval mapping of quantitative trait loci affecting NESTUR, a stem growth efficiency index of radiata pine seedlings. *Theor. Appl. Genet.* 97: 1062–1068.

Ersoz, E.S., Gonzalez-Martinez, S., Gill, G., Brown G, Morse, A., Davis, J., White, T. & Neale, D.B. 2004. SNP discovery in candidate genes for disease resistance (pitch canker and fusiform rust) and drought tolerance in loblolly pine, Abstract W95. In *Plant & Animal Genomes XII Conference*, San Diego, CA, USA.

FAO. FAOSTAT interactive database (available at http://faostat.fao.org/default.aspx).

FAO. 2003. *Molecular marker-assisted selection as a potential tool for genetic improvement of crops, forest trees, livestock and fish in developing countries*. Summary Document to Conference 10 (17 November to 14 December 2003) of the FAO Biotechnology Forum (available at www.fao.org/biotech/logs/C10/summary.htm).

Frewen, B.E., Chen, T.H.H., Howe, G.T., Davis, J., Rohde, A., Boerjan, W. & Bradshaw, H.D.Jr. 2000. Quantitative trait loci and candidate gene mapping of bud set and bud flush in *Populus*. *Genetics* 154: 837–845.

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

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

Gosselin, I., Zhou, Y., Bousquet, J. & Isabel, N. 2002. Megagametophyte-derived linkage maps of white spruce (*Picea glauca*) based on RAPD, SCAR and ESTP markers. *Theor. Appl. Genet.* 104: 987–997.

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

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

Grattapaglia, D., Bertolucci, L.G., Penchel, R. & Sederoff, R.R. 1996. Genetic mapping of quantitative trait loci controlling growth and wood quality traits in *Eucalyptus grandis* using a maternal half-sib family and RAPD markers. *Genetics* 144: 1205–1214.

Grattapaglia, D., Ribeiro, V.J. & Rezende, G.D.S.P. 2004. Retrospective selection of elite parent trees using paternity testing with microsatellite markers: an alternative short term breeding tactic for *Eucalyptus*. *Theor. Appl. Genet*. 109: 192–199.

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

Hanley, S., Barker, J.H.A., van Oijen, J.W., Aldam, C., Harris, S.L., Ahman, I., Larsson, S. & Karp, A. 2002. A genetic linkage map of willow (*Salix viminalis*) based on AFLP and microsatellite markers. *Theor. Appl. Genet*. 105: 1087–1096.

Harwood, C.E., Ha Huy Thinh, Tran Ho Quang, Butcher, P.A. & Williams, E.R. 2004. The effect of inbreeding on early growth of *Acacia mangium* in Vietnam. *Silv. Genet*. 53: 65–69

Hayashi, E., Kondo, T., Terada, K., Kuramoto, N. Goto, Y., Okamura, M. & Kawasaki, H. 2001. Linkage map of Japanese black pine based on AFLP and RAPD markers including markers linked to resistance against the pine needle gall midge. *Theor. Appl. Genet*. 102: 871–875.

Hertzberg, M., Aspeborg, H., Schrader, J., Andersson, A., Erlandsson, R., Blomqvist, K., Bhalerao, R., Uhlén, M., Teeri, T.T., Lundeberg, J., Sundberg, B., Nilsson, P. & Sandberg, G. 2001. A transcriptional roadmap to wood formation. *Proc. Nat Acad. Sci. USA* 98: 14732–14737.

Hu, W.-J., Harding, S.A., Lung, J., Popko, J.L., Ralph, J., Stokke, D.D., Tsai, C.-J. & Chiang, V.L. 1999. Repression of lignin biosynthesis promotes cellulose accumulation and growth. *Nature Biotech.* 17: 808–812.

Hurme, P., Sillanpaa, M.J., Arjas, E., Repo, T. & Savolainen, O. 2000. Genetic basis of climatic adaptation in Scots pine by Bayesian quantitative trait locus analysis. *Genetics* 156: 1309–1322.

International HapMap Consortium. 2003. The International HapMap Project. *Nature* 426: 789–796.

Jermstad, K.D., Bassoni, D.L., Jech, K.S., Ritchie, G.A., Wheeler, N.C. & Neale, D.B. 2003. Mapping of quantitative trait loci controlling adaptive traits in coastal Douglas fir. III. QTL by environment interactions. *Genetics* 165: 1489–1506.

Jermstad, K.D., Bassoni, D.L., Jech, K.S., Wheeler, N.C. & Neale, D.B. 2001a. Mapping of quantitative trait loci controlling adaptive traits in coastal Douglas fir. I. Timing of vegetative bud flush. *Theor. Appl. Genet*. 102: 1142–1151.

Jermstad, K.D., Bassoni, D.L., Wheeler, N.C., Anekonda, T.S., Aitken, S.N., Adams, W.T. & Neale, D.B. 2001b. Mapping of quantitative trait loci controlling adaptive traits in coastal Douglas fir. II. Spring and fall cold hardiness. *Theor. Appl. Genet*. 102: 1152–1158.

Jermstad, K.D., Bassoni, D.L., Wheeler, N.C. & Neale, D.B. 1998. A sex-averaged linkage map in coastal Douglas-fir (*Pseudotsuga menziesii* [Mirb.] Franco) based on RFLP and RAPD markers. *Theor. Appl. Genet*. 97: 762–770.

Johnson, G.R., Wheeler, N.C. & Strauss, S.H. 2000. Financial feasibility of marker-aided selection in Douglas fir. *Can. J. For. Res.* 30: 1942–1952.

Joshi, C.P. 2004. Molecular genetics of cellulose biosynthesis in trees. *In* S. Kumar & M. Fladung, eds. *Molecular genetics and breeding of forest trees,* pp. 141–165. New York, USA, Haworth Press.

Joshi, C.P., Wu, L. & Chiang, V.L. 1999. A novel xylem specific and tension stress responsive cellulose synthase gene from quaking aspen, Abstract p.19. *In* M. Campbell, ed. *Forest Biotechnology 99.* Oxford, UK, Oxford University Press.

Jouanin, L. & Goujon, T. 2004. Tuning lignin metabolism through genetic engineering in trees. *In* S. Kumar & M. Fladung, eds. *Molecular genetics and breeding of forest trees,* pp. 167–191. New York, USA, Haworth Press.

Junghaus, D.T., Alfenas, A.C., Brommonschenkel, S.H., Oda,S., Mello, E.J. & Grattapaglia, D. 2003. Resistance to rust (*Puccinia psidii* Winter) in *Eucalyptus*: mode of inheritance and mapping of a major gene with RAPD markers. *Theor. Appl. Genet*. 108: 175–180.

Kaya, Z., Sewell, M.M. & Neale, D.B. 1999. Identification of quantitative trait loci influencing annual height- and diameter-increment growth in loblolly pine. *Theor. Appl. Genet*. 98: 586–592.

Kha, L.D., Hai, N.D. & Vinh, H.Q. 1998. Clonal tests and propagation options for natural hybrids between *Acacia manguim* and *A. auriculiformis. In* J.W. Turnbull H.R. Crompton & K. Pinyopusarek, eds. *Recent developments in acacia planting,* pp. 203–210. ACIAR Proc. No. 82, Canberra, ACIAR.

Kirst, M., Myburg, A.A., De León, J.P.G., Kirst, M.E., Scott, J. & Sederoff, R. 2004. Coordinated genetic regulation of growth and lignin revealed by quantitative trait locus analysis of cDNA microarray data in an interspecific backcross of *Eucalyptus*. *Plant Physiol*. 135: 2368–2378.

Krutovskii, K.V. & Neale, D.B. 2001. Forest genomics for conserving adaptive genetic diversity. FAO Forest Genetic Resources Working Paper FGR/3 (E). Rome (available at www.fao.org/DOCREP/003/X6884E/X6884E00.HTM).

Krutovskii, K.V., Vollmer, S.S., Sorensen, F.C., Adams, W.T., Knapp, S.J. & Strauss, S.H. 1998. RAPD genome maps of Douglas fir. *J. Hered*. 89: 197–205.

Krutovsky, K.V., Troggio, M., Brown, G., Jermstad, K.D. & Neale, D.B. 2004. Comparative mapping in the Pinaceae. *Genetics* 168: 447–461.

Krutovsky, K.V., Pande, P., Jermstad, K.D., Howe, G.T., St. Clair, J.B., Wheeler, N.C. & Neale, D.B. 2005. Nucleotide diversity and linkage disequilibrium in adaptive trait related candidate genes in Douglas fir, Abstract W110. In *Plant & Animal Genomes XII Conference,* San Diego, CA, USA.

Kubisiak T.L., Nelson, C.D., Nance, W.L. & Stine, M. 1996. Comparison of RAPD linkage maps constructed for a single longleaf pine from both haploid and diploid mapping populations. *For. Genet*. 3: 203–211.

Kumar, S., Spelman, R.J., Garrick, D.J., Richardson, T.E., Lausberg, M. & Wilcox, P.L. 2000. Multiple marker mapping of wood density loci in an outbred pedigree of radiata pine. *Theor. Appl. Genet*. 100: 926–933.

Kuramoto, N., Kondo, T., Fujisawa, Y., Nakata, R., Hayashi, E. & Goto, Y. 2000. Detection of quantitative trait loci for wood strength in *Cryptomeria japonica*. *Can. J. For. Res.* 30: 1525–1533.

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

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

Lespinasse, D., Grivet, L., Troispoux, V., Rodier-Goud, M., Pinard, F. & Seguin, M. 2000. Identification of QTLs involved in resistance to South American leaf blight (*Microcyclus ulei*) in the rubber tree. *Theor. Appl. Genet*. 100: 975–984.

Lespinasse, D., Rodier-Goud, M., Grivet, L., Leconte, A., Legnate, H. & Seguin, M. 1999. A saturated linkage map of rubber tree (*Hevea* spp.) based on RFLP, AFLP, microsatellite and isozyme markers. *Theor. Appl. Genet*. 100: 127–138.

Liepman, A.H., Wilkerson, C.G. & Keegstra, K. 2005. Expression of cellulose synthase-like (Csl) genes in insect cells reveals that CslA family members encode mannan synthases. *Proc. Nat. Acad. Sci. USA* 102: 2221–2226.

Long, A.D. & Langley, C.H. 1999. The power of association studies to detect the contribution of candidate genetic loci to variation in complex traits. *Gen. Res.* 9: 720–731.

MacKay, J., Bousquet, J., Cooke, J., Isabel, N., Pavy, N., Regan, S., Retzel, E. & Séguin, A. 2005. Functional genomic investigations of wood formation and defense response in trees, and beyond, Abstract W101. In *Plant & Animal Genomes XII Conference*, San Diego, CA, USA.

Marcucci Poltri, S.N., Zelena, N., Rodriguez Traverso, J., Gelid, P. & Hopp, H.E. 2003. Selection of a seed orchard of *Eucalyptus dunnii* based on genetic diversity criteria calculated using molecular markers. *Tree Physiol.* 23: 625–632.

Marques, C.M., Araújo, J.A., Ferreira, J.G., Whetten, R., O'Malley, D.M., Liu, B.-H. & Sederoff, R. 1998. AFLP genetic maps of *Eucalyptus globulus* and *E. tereticornis*. *Theor. Appl. Genet*. 96: 727–737.

Marques, C.M., Vasquez-Kool, J., Carocha, V.J., Ferreira, J.G., O'Malley, D.M., Liu, B.-H. & Sederoff, R. 1999. Genetic dissection of vegetative propagation traits in *Eucalyptus tereticornis* and *E. globulus*. *Theor. Appl. Genet*. 99: 936–946.

Marques, C.M., Brondani, R.P.V., Grattapaglia, D. & Sederoff, R. 2002. Conservation and synteny of SSR loci and QTLs for vegetative propagation in four *Eucalyptus* species. *Theor. Appl. Genet*. 105: 474–478.

Mehta, P., Sivaprakash, K., Parani, M., Venkataraman, G. & Parida, A.K. 2005. Generation and analysis of expressed sequence tags from the salt tolerant mangrove species *Avicennia marina* (Forsk.) Vierh. *Theor. Appl. Genet*. 110: 416–424.

Moran, G.F., Thamarus, K.A., Raymond, C.A., Qiu, D., Uren, T. & Southerton, S. 2002. Genomics of *Eucalyptus* wood traits. *Ann. For. Sci.* 59: 645–650.

Morgante, M. & Salamini, F. 2003. From plant genomics to breeding practice. *Curr. Opin. Biotech*. 14: 214–219.

Mouradov, A., Glassick, T., Hamdorf, B., Murphy, L., Fowler, B., Marla, S. & Teasdale, R.D. 1998. *NEEDLY*, a *Pinus radiata* ortholog of *FLORICAULA/LEAFY* genes, expressed in both reproductive and vegetative meristems. *Proc. Nat. Acad. Sci. USA* 95: 6537–6542.

Mouradov, A., Hamdorf, B., Teasdale, R.D., Kim, J.T. Winter, K.U. & Theißen, G.A. 1999. *DEF/GLO*-like MADS-Box gene from a gymnosperm: *Pinus radiata* contains an ortholog of angiosperm B class floral homeotic genes. *Dev. Genet.* 25: 245–252.

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

Myburg, A.A., Griffin, A.R., Sederoff, R.R. & Whetten, R.W. 2003. Comparative genetic linkage maps of *Eucalyptus grandis*, *Eucalyptus globulus* and their F1 hybrid based on a double pseudo-backcross mapping approach. *Theor. Appl. Genet*. 107: 1028–1042.

Nairn, C.J. & Haselkorn, T. 2005. Three loblolly pine *CesA* genes expressed in developing xylem are orthologous to secondary cell wall *CesA* genes of angiosperms. *New Phytologist* 166: 907–915.

Neale, D.B. & Savolainen, O. 2004. Association genetics of complex traits in conifers. *Trends Plant Sci.* 9: 325–330.

Neale, D.B., Sewell, M.M. & Brown, G.R. 2002. Molecular dissection of the quantitative inheritance of wood property traits in loblolly pine. *Ann. For. Sci.* 59: 595–605.

Nelson, C.D., Nance, W.L. & Doudrick, R.L. 1993. A partial genetic linkage map of Slash pine (*Pinus elliottii* Englem var. *elliottii*) based on random amplified polymorphic DNAs. *Theor. Appl. Genet.* 87: 145–151.

Nikles, D.G. 1996. The first 50 years of the evolution of forest tree improvement in Queensland. *In* M.J. Dieters, A.C. Matheson, D.G. Nikles, C.E. Harwood, & S.M. Walker, eds. *Tree improvement for sustainable tropical forestry*, pp. 51–64. Proc. QFRI-IUFRO Conference, Caloundra, Queensland, Australia.

Paglia, G.P., Olivieri, A.M. & Morgante, M. 1998. Towards second-generation STS (sequence-tagged sites) linkage maps in conifers: a genetic map of Norway spruce (*Picea abies* K.). *Mol. Gen. Genet*. 258: 466–478.

Paoli, E.D. & Morgante, M. 2005. Nucleotide diversity and linkage disequilibrium in a set of candidate genes for time of bud set in *Picea abies,* Abstract W106. In *Plant & Animal Genomes XII Conference*, San Diego, CA, USA.

Paux, E., Tamasloukht, M., Ladouce, N., Sivadon, P. & Grima-Pettenati, J. 2004. Identification of genes preferentially expressed during wood formation in *Eucalyptus. Plant Mol. Biol.* 55: 263–280.

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

Pavy, N., Laroche, J., Bousquet, J. & MacKay, J. 2005. Large-scale statistical analysis of secondary xylem ESTs in pine. *Plant Mol. Biol.* 57: 203–224.

Peter, G. & Neale, D. 2004. Molecular basis for the evolution of xylem lignification. *Curr. Opin. Plant Biol.* 7: 737–742.

Plomion, C., O'Malley, D.M. & Durel, C.E. 1995. Genomic analysis in maritime pine (*Pinus pinaster*). Comparison of two RAPD maps using selfed and open-pollinated seeds of the same individual. *Theor. Appl. Genet*. 90: 1028–1034.

Plomion, C., Bahrman, N., Costa, P., Dubos, C., Frigerio, J.-M., Lalanne, C., Madur, D., Pionneau, C. & Gerber, S. 2004. Proteomics for genetic and physiological studies in forest trees: application in maritime pine. *In* S. Kumar & M. Fladung, eds. *Molecular genetics and breeding of forest trees*, pp. 53–79. New York, USA, Haworth Press.

Poke, F.S., Vaillancourt, R.E., Potts, B.M. & Reid, J.B. 2005. Genomic research in *Eucalyptus*. *Genetica* 125: 79–101.

Pot, D., Eveno, E., Garnier-Géré, P. & Plomion, C. 2004. Association studies for wood properties in maritime pine, Abstract W88. In *Plant & Animal Genomes XII Conference*, San Diego, CA, USA.

Pritchard, J.K., Stephens, M., Rosenberg, N.A. & Donnelly, P. 2000. Association mapping in structured populations. *Am. J. Hum. Genet.* 67: 170–181.

Remington, D.L., Whetten, R.W., Liu, B.-H. & O'Malley, D.M. 1999. Construction of an AFLP genetic map with nearly complete coverage in *Pinus taeda*. *Theor. Appl. Genet*. 98: 1279–1292.

Rönnberg-Wästljung, A.C., Glynn, C. & Weih, M. 2005. QTL analysis of drought tolerance and growth for a *Salix dasyclados* x *Salix viminalis* hybrid in contrasting water regimes. *Theor. Appl. Genet*. 110: 537–549.

Rottmann, W.H., Meilan, R., Sheppard, L.A. Brunner, A.M., Skinner, J.S., Ma, C., Cheng, S., Jouanin, L., Pilate, G. & Strauss, S.H. 2000. Diverse effects of overexpression of *LEAFY* and *PTFL*, a poplar (*Populus*) homolog of *LEAFY/FLORICAULA*, in transgenic poplar and *Arabidopsis. Plant J.* 22: 235–245.

Rutledge, R., Regan, S., Nicolas, O., Fobert, P., Côté, C., Bosnich, W., Kauffeldt, C., Sunohara, G., Séguin, A. & Stewart, D. 1998. Characterization of an *AGAMOUS* homologue from the conifer black spruce (*Picea mariana*) that produces floral homeotic conversions when expressed in *Arabidopsis. Plant J.* 15: 625–634.

Saintagne, C.-S., Bodenes, C., Barreneche, T., Bertocchi, E., Plomion, C. & Kremer, A. 2004. Detection of quantitative trait loci controlling bud burst and height growth in *Quercus robur* L. *Theor. Appl. Genet.* 109: 1648–1659.

Salvi, S., Tuberosa, R., Chiapparrino, E., Maccaferri, M., Veillet, S., van Beuningen, L. Isaac, P., Edwards, K. & Phillips, R.L. 2002. Towards positional cloning of Vgt1, a QTL controlling the transition from the vegetative to the reproductive phase in maize. *Plant Mol. Biol*. 48: 601–613.

Scalfi, M., Troggio, M., Piovani, P., Leonardi, S., Magnaschi, G., Vendramin, G.G. & Menozzi, P. 2004. A RAPD, AFLP and SSR linkage map, and QTL analysis in European beech (*Fagus sylvatica* L.). *Theor. Appl. Genet.* 108: 433–441.

Sedjo, R.A. 2004. Transgenic trees and trade: problems on the horizon? *Resources* 155: 8–13.

Sewell, M.M. & Neale, D.B. 2000. Mapping quantitative traits in forest trees. *In* S.M. Jain & S.C. Minocha, eds. *Molecular biology of woody plants*, Vol. 1, pp. 407–423. Dordrecht, Netherlands, Kluwer.

Sewell, M.M., Sherman, B.K. & Neale, D.B. 1999. A consensus map for loblolly pine (*Pinus taeda* L.). I. Construction and integration of individual linkage maps from two outbred three-generation pedigrees. *Genetics* 151: 321–330.

Sewell, M.M., Bassoni, D.L., Megraw, R.A., Wheeler, N.C. & Neale, D.B. 2000. Identification of QTLs influencing wood property traits in loblolly pine (*Pinus taeda* L.) I. Physical wood properties. *Theor. Appl. Genet.* 101: 1273–1281.

Sewell, M.M., Davis, M.S., Tuskan, G.A., Wheeler, N.C., Elam, C.C., Bassoni, D.L. & Neale, D.B. 2002. Identification of QTLs influencing wood property traits in loblolly pine (*Pinus taeda* L.) II. Chemical wood properties. *Theor. Appl. Genet.* 104: 214–222.

Shena, M., Shalon, D., Davis, R.W. & Brown, P.O. 1996. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270: 467–470.

Shepherd, M., Chaparro, J.X. & Teasdale, R. 1999. Genetic mapping of monoterpene composition in an interspecific eucalypt hybrid. *Theor. Appl. Genet*. 99: 1207–1215.

Shepherd, M., Cross, M., Dieters, M.J. & Henry, R. 2003. Genetic maps for *Pinus elliottii* var. *elliottii* and *P. caribaea* var. *hondurensis* using AFLP and microsatellite markers. *Theor. Appl. Genet*. 106: 1409–1419.

Skov, E. & Wellendorf, H. 1998. A partial linkage map of *Picea abies* clone V6470 based on recombination of RAPD-markers in haploid megagametophytes. *Silv. Genet*. 47: 273–282.

Smith, N.A., Singh, S.P., Wang, M.B., Stoutjesdijk, P.A., Green, A.G. & Waterhouse, P.M. 2000. Total silencing by intron-spliced hairpin RNAs. *Nature* 407: 319–320.

Southerton, S.G., Strauss, S.H., Olive, M.R., Harcourt, R.L., Decroocq, V., Zhu, X., Llewellyn, D.J., Peacock, W.J. & Dennis, E.S. 1998. *Eucalyptus* has a functional equivalent of the *Arabidopsis* floral meristem identity gene *LEAFY. Plant Mol. Biol.* 37: 897–910.

Strabala, T.J. 2004. Expressed sequence tag databases from forestry tree species. *In* S. Kumar & M. Fladung, eds. *Molecular genetics and breeding of forest trees,* pp. 19–51. New York, USA, Haworth Press.

Strauss, S.H., Lande, R. & Namkoong, G. 1992. Limitations of molecular-marker-aided selection in forest tree breeding. *Can. J. For. Res.* 22: 1050–1061.

Strauss, S.H., Brunner, A.M., DiFazio, S.P. & James, R.R. 1998. Environmental effects of genetically engineered woody biomass crops. *Biomass & Bioenergy* 14: 403–414.

Strauss, S.H., Brunner, A.M., Busov, V.B., Caiping, M.A. & Meilon R. 2004. Ten lessons from 15 years of transgenic *Populus* research. *Forestry* 77: 455–465.

Tandre, K., Albert, V.A., Sundas, A. & Engstrom, P. 1995. Conifer homologues to genes that control floral development in angiosperms. *Plant Mol. Biol.* 27: 69–78.

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

Thamarus, K.A., Groom, K., Bradley, A., Raymond, C.A., Shimleck, L.R., Williams, E.R. & Moran G.F. 2004. Identification of quantitative trait loci for wood and fibre properties in two full-sib pedigrees of *Eucalyptus globulus*. *Theor. Appl. Genet*. 109: 856–864.

Thumma, B.R., Nolan, M.F., Evans R. & Moran G.F. 2005. Polymorphisms in *cinnamoyl CoA reductase* (*CCR*) are associated with variation in microfibril angle in *Eucalyptus* spp. *Genetics* 171: 1257–1265.

Travis, S.E., Ritland, K., Whitham, T.G. & Keim, P. 1998. A genetic linkage map of pinyon pine (*Pinus edulis*) based on amplified fragment length polymorphisms. *Theor. Appl. Genet*. 97: 871–880.

Tsarouhas, V., Gullberg, U. & Lagercrantz, U. 2002. An AFLP and RFLP linkage map and quantitative trait locus (QTL) analysis of growth traits in *Salix*. *Theor. Appl. Genet.* 105: 277–288.

Tsarouhas, V., Gullberg, U. & Lagercrantz, U. 2003. Mapping of quantitative trait loci controlling timing of bud flush in *Salix*. *Hereditas* 138: 172–178.

van Zyl, L., Bozhkov, P.V., Clapham, D.H., Sederoff, R.R. & von Arnold, S. 2003. Up, down and up again is a signal global gene expression pattern at the beginning of gymnosperm embryogenesis. *Gene Expr. Patterns* 3: 83–91.

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

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

Voiblet, C., Duplessis, S., Encelot, N. & Martin, F. 2001. Identification of symbiosis-regulated genes in *Eucalyptus globulus – Pisolithus tinctorius* ectormycorrhiza by differential hybridization of arrayed cDNAs. *Plant J.* 25: 181–191.

Walter, C. & Killerby, S. 2004. A global study on the state of forest tree genetic modification. In *Preliminary review of biotechnology in forestry including genetic modification*. FAO Forest Genetic Resources Working Paper FGR/59E. Rome (available at www.fao.org/docrep/008/ae574e/ae574e00.htm).

Wang, H. 2004. The state of genetically modified trees in China. In *Preliminary review of biotechnology in forestry, including genetic modification*, pp. 96–105. FAO Forest Genetic Resources Working Paper FGR/59E. Rome (available at www.fao.org/docrep/008/ae574e/ae574e00.htm).

Watson, J.M. & Brill, E.M. 2004. *Eucalyptus grandis* has at least two functional *SOC 1*-like floral activator genes. *Funct. Plant Biol.*, 31: 225–234.

Weng, C., Kubisiak, T.L., Nelson, C.D. & Stine, M. 2002. Mapping quantitative trait loci controlling early growth in a (longleaf pine x slash pine) x slash pine BC1 family. *Theor. Appl. Genet.* 104: 852–859.

Whetten, R.W., Sun, Y.H., Zhang, Y. & Sederoff, R. 2001. Functional genomics and cell wall biosynthesis in loblolly pine. *Plant Mol. Biol.* 47: 275–291.

Wilcox, P.L., Carson, S.D., Richardson, T.E., Ball, R.D., Horgan, G.P. & Carter, P. 2001. Benefit-cost analysis of DNA marker-based selection in progenies of *Pinus radiata* seed orchard parents. *Can. J. For. Res*. 31: 2213–2224.

Wu, R. & Zeng, Z-B. 2001. Joint linkage and linkage disequilibrium mapping in natural populations. *Genetics* **157:** 899–909.

Wu, H.X., Abarquez, A. & Matheson, A.C. 2004. Inbreeding in *Pinus radiata* V. The effects of inbreeding on fecundity. *Silv. Genet*. 53: 80–87.

Wu, L., Joshi, C.P. & Chiang, V.L. 2000. A xylem specific cellulose synthase gene from aspen (*Populus tremuloides*) is responsive to mechanical stress. *Plant. J.* 22: 495–502.

Yazdani, R., Nilsson, J.E., Plomion, C. & Mathur, G. 2003. Marker trait association for cold acclimation and growth rhythm in *Pinus sylvestris*. *Scand. J. For. Res*. 18: 29–38.

Yazdani, R., Yeh, F.C. & Rimsha, J. 1995. Genomic mapping of *Pinus sylvestris* (L.) using random amplified polymorphic markers. *For. Genet*. 2: 109–116.

Yin, T.M., DiFazio, S.P., Gunter, L.E., Riemenshneider, D. & Tuskan, G.A. 2004. Large scale heterospecific segregation distortion in *Populus* revealed by a dense genetic linkage map. *Theor. Appl. Genet*. 109: 451–463.

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

Yu, J. Hu, S., Wang, J., Wong, G.K-S., Li, S., Liu, B., Deng, Y., Dai, L., Zhou, Y., Zhang, X., Cao, M., Liu, J., Sun, J., Tang, J., Chen, Y., Huang, X., Lin, W., Ye, C., Tong, C., Cong, L., Geng, J., Han, Y., Li, L., Li, W., Hu, G., Huang, X, Li, W., Li, J., Liu, Z., Li, L., Liu, J., Qi, O., Liu, J., Li, L., Li, T., Wang, X., Lu, H., Wu, T., Zhu, M., Ni, P., Han, H., Dong, W., Ren, X., Feng, X., Cui, P., Li, X., Wang, H., Xu, X., Zhai, W., Xu, Z., Zhang, J., He, S., Zhang, J., Xu, J., Zhang, K., Zheng, X., Dong, J., Zeng, W., Tao, L., Ye, J., Tan, J., Ren, X., Chen, X., He, J., Liu, D., Tian, W., Tian, C., Xia, H., Bao, O., Li, G., Gao, H., Cao, T., Wang, J., Zhao, W., Li, P., Chen, W., Wang, X., Zhang, Y., Hu, J., Wang, J., Liu, S., Yang, J., Zhang, G., Xiong, Y., Li, Z., Mao, L., Zhou,C., Zhu, Z., Chen, R., Hao, B., Zheng, W., Chen, S., Guo, W., Li, G., Liu, S., Tao, M., Wang, J., Zhu, L., Yuan, L. & Yang, H. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* 296: 79–92.

Zik, M. & Irish, V.F. 2003. Flower development: initiation, differentiation, and diversification. *Annu. Rev. Cell Dev. Biol.* 19: 119–140.

SECTION V

Marker-assisted selection in fish – case studies

CHAPTER 16

Possibilities for marker-assisted selection in aquaculture breeding schemes

Anna K. Sonesson

SUMMARY

FAO estimates that there are around 200 species in aquaculture. However, only a few species have ongoing selective breeding programmes. Marker-assisted selection (MAS) is not used in any aquaculture breeding scheme today. The aim of this chapter, therefore, is to review briefly the current status of aquaculture breeding schemes and to evaluate the possibilities for MAS of aquaculture species. Genetic marker maps have been published for some species in culture. The marker density of these maps is, in general, rather low and the maps are composed of many amplified fragment length polymorphism (AFLP) markers anchored to few microsatellites. Some quantitative trait loci (QTL) have been identified for economically important traits, but they are not yet mapped at a high density. Computer simulations of within-family MAS schemes show a very high increase in genetic gain compared with conventional family-based breeding schemes, mainly due to the large family sizes that are typical for aquaculture breeding schemes. The use of genetic markers to identify individuals and their implications for breeding schemes with control of inbreeding are discussed.

INTRODUCTION

Aquaculture is an expanding industry, with a total global value of US$61 billion (FAO, 2003). FAO estimates that there are around 200 species in culture, of which carps and oysters have the largest worldwide production. However, only a few species have ongoing selective breeding programmes.

MAS is not used in any aquaculture breeding scheme today. The aim of this chapter, therefore, is to review briefly the current status of aquaculture breeding schemes and to evaluate the possibilities for MAS of aquaculture species.

TRAITS OF BREEDING INTEREST

Growth rate is the most important trait for most aquaculture species under selection. It is recorded on the selection candidates, and can easily be improved using mass selection. Sexual maturation is a trait that leads to reduced growth, reduced feed conversion efficiency and reduced fillet quality in several aquaculture species. Therefore, selection against early maturation is often performed, i.e. the individuals that become sexually mature before market size are discarded as selection candidates. In tilapia, late maturity is desirable because of excessive spawning that results in overcrowding of ponds and reduced size of the fish.

For many other traits, however, accurate measurement techniques for live individuals are inadequate. Hence, selection must be based on information from other family members, e.g. on siblings. MAS would be especially valuable for traits that are difficult and/or costly to measure on the selection candidate or for traits that are measured late in life or at slaughter (Lande and Thompson, 1990; Poompuang and Hallerman, 1997). Examples of these important traits are:

- *Disease resistance*. Challenge tests exist for viral (e.g. white spot syndrome in shrimps and infectious pancreatic necrosis in most marine fishes) and bacterial (e.g. furunculosis and vibriosis) diseases, as well as for parasites (e.g. sea lice). When challenge tests are used in breeding programmes, however, surviving individuals cannot enter the breeding nucleus because of the risk that they will introduce the disease to the nucleus. Therefore, these individuals cannot be selection candidates and only their sibs, who have no records for these traits, are candidates.
- *Fillet quality traits*. To this group of traits belong colour, texture, gaping, different fat-related traits (e.g. fat percentage and distribution) and dressing percentage. Accurate measurements of these traits are available only for slaughtered individuals. Techniques for measuring fillet colour on live fish are under development.
- *Feed conversion efficiency* is a trait that can be measured practically only at the family level at a young age in the breeding nucleus, but not at the individual level or in grow-out operations. The value of such early records is rather limited because of the unknown correlation with feed conversion efficiency at a later age. Feed intake is, in general, a difficult trait to measure in aquaculture species due to unequal feed intake over days. No active selection programme for aquaculture species selects directly for feed conversion efficiency; rather, indirect selection is practised by selecting for growth.
- *Salinity and low temperature tolerance* are two traits of interest for tilapia breeding programmes. Today, tilapias are produced in freshwater in tropical and subtropical areas. The purpose of selecting for salinity and temperature tolerance

is to develop fish that could reproduce and grow in areas of higher salinity and lower temperature, i.e. to increase the production area for tilapias. Although these traits could be measured early in the life of the fish and therefore could be selected, sexually mature fish may respond differently to the temperature and salinity conditions.

Gjedrem and Olesen (2005) provide a more complete list of aquaculturally relevant traits and their heritabilities and correlations.

STRUCTURE OF BREEDING SCHEMES

Most aquaculture species are currently bred in mono- or polyculture systems (i.e. with one or several species reared together) using mass selection for growth rate. Only about 30 family-based breeding programmes worldwide utilize sib information in the estimation of breeding values (B. Gjerde, personal communication). The main part of a family-based breeding programme is a closed breeding nucleus, with trait information from sibs coming from test stations. Breeding programmes for species with limited reproductive ability (e.g. salmonids as opposed to several highly fecund marine species such as Atlantic cod) have a multiplier unit, where genetic material from the nucleus is used to produce eggs or fry for the grow-out producers. The limiting factor for the breeding nucleus is often the number of tanks, where the fry of a certain full-sib family are kept until individuals are large enough to be physically tagged. The number of offspring per full-sib family is large, such that a very high selection intensity can be achieved. Generally, each male is mated to two females in order that the tank effect can be estimated separately from the additive genetic effects.

The high intensity of selection within the nucleus can easily result in high rates of inbreeding. Introduction of unrelated wild stock is often practised to reduce the rates of inbreeding. However, introduction of wild stock also leads to reduced genetic gain, and should generally be avoided in ongoing breeding schemes. Optimum contribution is an approach that maximizes genetic gain while restricting the rates of inbreeding for schemes with discrete (Meuwissen, 1997; Grundy, Villanueva and Woolliams, 1998) or overlapping (Meuwissen and Sonesson, 1998; Grundy, Villanueva and Woolliams, 2000) generation structures or for traits with a polygenic effect and the effect of a known single gene (Meuwissen and Sonesson, 2004). The key determination is the number of matings (full-sib families) per selected individual. One practical constraint in some marine species is that matings are volitional (natural mating in, for example, a tank) and thus depend on the availability of individuals ready to spawn at a certain moment. Hence, for these species, the number of matings per male or female is restricted for each spawning. The use of frozen milt makes the use of males more flexible. Milt from many species including salmonids, carp and shrimps can be frozen (Stoss, 1983), but the practical use of cryopreserved sperm in aquaculture breeding programmes has not been fully utilized.

GENETIC MARKER MAPS

A genetic marker map is an ordered listing of the genes or molecular markers occurring along the length of the chromosomes in the genome. Distances between genes or markers are estimated in terms of how frequently recombination occurs between them. Genetic marker maps are available for some aquaculture species (Table 1). Most of these genetic maps are constructed using amplified fragment length polymorphism

(AFLP) markers (Vos *et al.*, 1995), which are generally anchored to a smaller collection of microsatellites. The marker density of the maps is currently rather low, and the markers are spread unevenly over the genome, which may explain why the number of linkage groups found in, for example, channel catfish (Waldbieser *et al.*, 2001) or white shrimp (Pérez *et al.*, 2004) does not correspond to the number of chromosomes. In rainbow trout, tetraploidy has been found for 20 chromosome arms (Sakamoto *et al.*, 2000). Recombination rates can differ between males and females, and hence marker map lengths can differ considerably between males and females. In salmonids, females have the higher recombination rate, which implies that most information comes from the females when constructing the marker map. The ratio between female and male recombination rates was 3.25:1.00 for rainbow trout (Sakamoto *et al.*, 2000), 1.69:1.00 for Arctic char (Woram *et al.*, 2004) and 8.25:1.00 for Atlantic salmon (Moen *et al.*, 2004a). However, in other aquaculture species, males have the higher recombination rate. For example, the ratio between male and female recombination rates was 7.4:1.00 in Japanese flounder (Coimbra *et al.*, 2003). There are also differences in recombination rate over the length of the chromosomes in males, i.e. recombination rate was higher in telomeric regions than in proximal regions in rainbow trout (Sakamoto *et al.*, 2000).

TABLE 1
Aquaculture species for which there are genetic marker maps

Species	Reference
Scallop	Li *et al.* (2005)
	Wang *et al.* (2004)
Pacific oyster	Hubert and Hedgecock (2004)
Eastern oyster	Yu and Guo (2003)
White shrimp	Pérez *et al.* (2004)
Kuruma prawn	Li *et al.* (2003)
Black tiger shrimp	Wilson *et al.* (2002)
Kuruma prawn	Moore *et al.* (1999)
Atlantic salmon	Moen *et al.* (2004a)
	Gilbey *et al.* (2004)
Arctic char	Woram *et al.* (2004)
Rainbow trout	Nichols *et al.* (2003)
	Sakamoto *et al.* (2000)
	Young *et al.* (1998)
Salmonids	May and Johnson (1990)
Common carp	Sun and Liang (2004)
European sea bass	Chistiakov *et al.* (2005)
Channel catfish	Waldbieser *et al.* (2001)
	Liu *et al.* (2003)
Tilapia	Lee *et al.* (2005)
	McConnell *et al.* (2000)
	Agresti *et al.* (2000)
	Kocher *et al.* (1998)
Japanese flounder	Coimbra *et al.* (2003)

MAPPING OF QTL

QTL are loci whose variability underlies variation in expression of a quantitative character (Geldermann, 1975). Detection of QTL would help in understanding the genetic architecture of the trait, i.e. the numbers and relative effects of genes that determine expression of a trait. A small, but growing, number of QTL for important traits have been identified in farmed aquatic species (Table 2). In tilapias, QTL have been identified for cold tolerance (Cnaani *et al.*, 2003; Moen *et al.*, 2004b). QTL for upper temperature tolerance (Jackson *et al.*, 1998; Danzmann, Jackson and Ferguson, 1999; Perry, Ferguson and Danzmann, 2003; Somorjai, Danzmann and Ferguson 2003), and for resistance to different disease traits have been found in salmonids, e.g. for infectious hematopoietic necrosis virus (Rodriguez *et al.*, 2005), infectious pancreatic necrosis virus (Ozaki *et al.*, 2001), *Ceratomyxa shasta* (Nichols, Bartholomew and Thorgaard, 2003) and infectious salmon anemia (Moen *et al.*, 2004c, 2006). QTL for general disease resistance and immune response have been found in tilapias (Cnaani *et al.*, 2004) and

TABLE 2
Known marker-QTL linkages in aquaculture species

Trait	Reference
Salmonids	
Spawning time	Leder, Danzmann and Ferguson (2006)
Early development	Martinez *et al.* (2005)
Pyloric caeca number	Zimmerman *et al.* (2005)
Natural killer cell-like activity	Zimmerman *et al.* (2004)
Hematopoetic necrosis resistance	Rodriguez *et al.* (2005)
Development rate	Sundin *et al.* (2005)
Infectious salmon anemia resistance	Moen *et al.* (2004c, 2006)
Ceratomyxa shasta resistance	Nichols, Bartholomew and Thorgaard (2003)
Infectious pancreatic necrosis resistance	Ozaki *et al.* (2001)
Infectious hematopoietic necrosis resistance	Khoo *et al.* (2004)
Body weight and condition factor	Reid *et al.* (2005)
Spawning date and body weight	O'Malley *et al.* (2003)
Growth and maturation	Martyniuk *et al.* (2003)
Temperature tolerance	Somorjai, Danzmann and Ferguson (2003)
Meristic traits	Nichols, Wheeler and Thorgaard (2004)
Embryonic development	Robison *et al.* (2001)
Albinism	Nakamura *et al.* (2001)
Development rate	Nichols *et al.* (2000)
Spawning time	Sakamoto *et al.* (1999)
Upper temperature tolerance, size	Perry *et al.* (2001, 2003)
Upper temperature tolerance	Danzmann, Jackson and Ferguson (1999)
Upper temperature tolerance	Jackson *et al.* (1998)
Tilapia	
Cold tolerance	Moen *et al.* (2004b)
Cold tolerance and fish size	Cnaani *et al.* (2003)
Stress and immune response	Cnaani *et al.* (2004)
Colour pattern	Streelman, Albertson and Kocher (2003)
Early survival	Palti *et al.* (2002)
Sex determination	Lee, Penman and Kocher (2003); Lee, Hulata and Kocher (2004)

salmonids (Zimmerman *et al.*, 2004). The data used for quantifying disease resistance and temperature tolerance traits are often based on challenge tests, for which models accounting for non-normality of data and special algorithms that take account of censored data (survival models) are used in combination with the QTL mapping methods (e.g. Moen *et al.*, 2006). In salmonids, QTL have been found related to body weight and size (Martyniuk *et al.*, 2003; O'Malley *et al.*, 2003; Reid *et al.*, 2005), for colouration pattern (Streelman, Albertson and Kocher, 2003) and for one form of albinism (Nakamura *et al.*, 2001). Zimmerman *et al.* (2005) found QTL for pyloric caeca number, a trait related to feed conversion efficiency. Epistasis has been found for upper temperature tolerance and body length in rainbow trout (Danzmann, Jackson and Ferguson, 1999; Perry, Ferguson and Danzmann, 2003); the epistasis depended on the genetic background, which would result in reduced effectiveness of MAS (Danzmann, Jackson and Ferguson, 1999).

The genetic diversity of wild and cultured populations, high fecundity, and the possibility of interspecific hybridization and reproductive manipulation of aquaculture species can be exploited in QTL mapping studies. These features have resulted in a

wide diversity of experimental populations being used in such studies. Double haploids (see below) have been used for QTL mapping in salmonids (by androgenesis; Robison *et al.,* 2001; Zimmerman *et al.,* 2005) and tilapias (by gynogenesis; Palti *et al*., 2002). Backcross populations have been set up where strains with large phenotypic differences in the trait of interest are crossed, e.g. for temperature tolerance in rainbow trout (Danzmann, Jackson and Ferguson, 1999). The strains can come from one species, but crosses between species also have been used (Streelman, Albertson and Kocher, 2003 for tilapia; Rodriguez *et al.,* 2005 for salmonids). Finally, families derived from a breeding nucleus have been used for QTL mapping. (e.g. Moen *et al*., 2006 for Atlantic salmon).

Most QTL mapping is based on marker association studies (e.g. Sakamoto *et al*., 1999) or on a marker association study followed by interval mapping (Moen *et al.,* 2006). Combined linkage/linkage disequilibrium methods (e.g. Meuwissen *et al*., 2002) have high precision for mapping QTL in outbred populations, but have not been used for QTL mapping in aquaculture species. This lack could be explained by the sparcity of genetic marker maps for the species under study and by many of the studies having used special crosses as mentioned above, where linkage is the main source of information for mapping the QTL.

Various reproductive manipulations may be applied to aquaculture species. One interesting reproductive manipulation technique for outbred populations for marker and QTL mapping is gyno- and androgenetic double haploids (Chourrout, 1984). In gynogenesis, the sperm's chromosomes are inactivated by irradiation and following fertilization, diploidy is restored by applying a temperature or hydrostatic pressure shock. The result is an individual that is double haploid with only the female's chromosomes. Depending on when diploidy is restored, two types of double haploid individuals can be produced. If an early shock is applied, extrusion of the second polar body is suppressed and, when the two maternal chromosome sets fuse, some heterozygosity is retained. If a late shock is applied, the first mitotic cleavage of the zygote is suppressed and, when the two maternal chromosome sets fuse, the resulting individual is virtually homozygous. In androgenesis, the egg is irradiated and, after "fertilization", the egg is shocked to suppress the first mitosis (Parsons and Thorgaard, 1984). The result is an individual that is a double haploid with only the male's chromosomes and that is virtually homozygous.

The power of an experiment to detect QTL depends on the effect of QTL alleles, the recombination rate among the marker and QTL loci, and the sample size of the mapping population. The effect of QTL genotypes is higher for double haploid than for full- or half-sib family designs in an outbred population because the QTL genotypes occur only in a homozygous form in double haploids (i.e. in their most extreme form). The relative advantage of a population of mitotic double haploids, where the two chromosome sets are copies of each other, is largest when the QTL has a small effect (Martinez, Hill and Knott, 2002). In meiotic haploid individuals, the two chromosome sets in an individual have been recombined. The power of QTL detection in these meiotic double haploid populations is therefore expected to resemble that of selfed populations (Martinez, Hill and Knott, 2002). Double haploids have been used for genetic marker and QTL mapping, as noted above.

On the other hand, the extremely large full-sib family size that is possible for aquaculture species may make use of special reproductive techniques for marker and QTL mapping studies unnecessary. For example, in Atlantic cod, both males and females produce millions of gametes at spawning. Also, Atlantic cod and halibut are examples of repeat spawners, in contrast to salmonids that normally die after a single spawning. Repeated spawning makes the mating structure more flexible, e.g. a certain pairing can be repeated or an individual may be used in multiple pairings. Another disadvantage of using double haploids is that they are fully inbred individuals and may therefore express the trait of interest differently than non-inbred animals. An example of this is the environmental variance for wing length in *Drosophila melanogaster*, which has been shown to be larger for inbred individuals than non-inbred individuals (Falconer and Mackay, 1996). Traits with dominant expression may be expressed differently because of inbreeding depression.

MAS SCHEMES

After QTL detection experiments, breeders will have knowledge of marker-QTL linkages and an estimate of the respective effects of QTL alleles on the trait in the population. This knowledge may be applied using MAS of spawners for producing the next generation. Generally, QTL detection would be carried out in one experiment and MAS in another (Poompuang and Hallerman, 1997). For within-family MAS schemes, the phase of marker and QTL alleles needs to be established for all families on which selection will be practised. Two within-family MAS schemes have been well studied for livestock populations. The first is a three-generation scheme, which is suitable for breeding schemes with progeny testing (Kashi, Hallerman and Solomon, 1990). Progeny tests are not, however, currently performed upon aquaculture species. The second is a two-generation scheme (Mackinnon and Georges, 1998), which may be modified for application to fish populations.

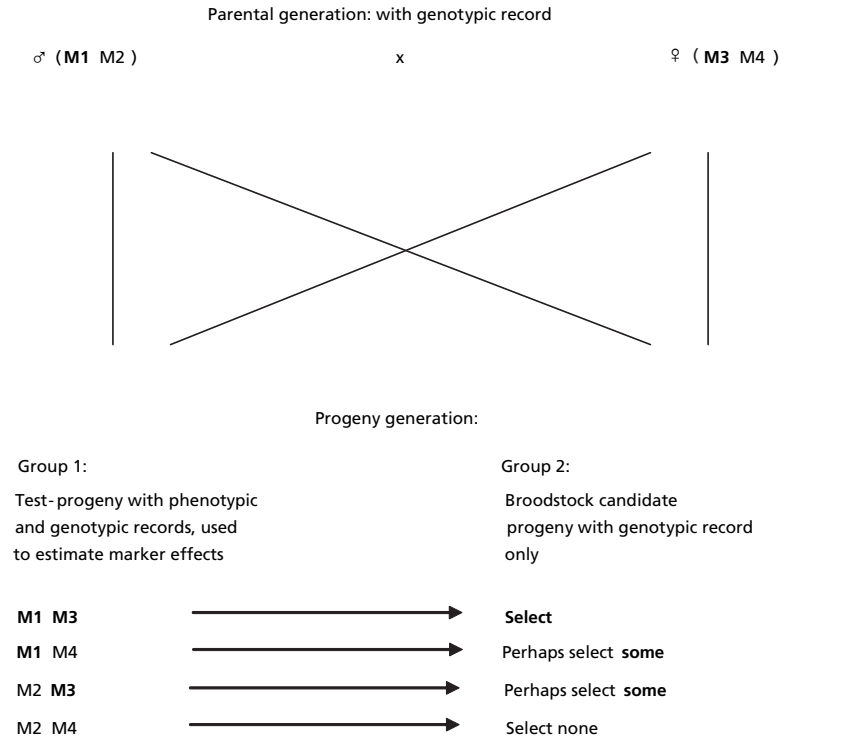
In the two-generation scheme (Figure 1), it is assumed that both sires and dams have genotypic records for markers linked to the QTL and that there are two groups of progeny from these parents. The group of test progeny has both phenotypic performance and genotypic records, while the group of selection candidates only has genotypic records. Phenotypic evaluation often implies that these individuals cannot be used later for breeding, perhaps because they were used in a challenge test or were slaughtered to obtain sib information for carcass traits. The genetic markers of the sire are denoted M1 and M2, and those of the dam M3 and M4, with M1 and M3 linked with the performance-increasing QTL alleles. With these linkage relationships, M1M3-bearing progeny would be selected while some M1M4- and M2M3-, and no M2M4-bearing progeny would be selected. The proportions of each genotype selected would vary depending upon selection intensity.

When QTL are mapped densely (up to 5 cM between markers), both linkage disequilibrium within families and population-wide linkage disequilibrium can be used in the MAS scheme (Smith and Smith, 1993; Dekkers and Hospital, 2002).

Simulation of two-generation within-family MAS schemes

The attractive feature of MAS is the potential for increasing the genetic gain in a selective breeding programme. Lande and

FIGURE 1
Two-generation marker evaluation and selection scheme for aquaculture species



Individuals with marker genotypes shown in **bold** are selected. The proportion of each genotype selected will vary with selection intensity.

Thompson (1990) showed that the efficiency of MAS relative to conventional selection alone depends upon the heritability of the trait under selection, the proportion of genetic variance associated with marker loci and the particular selection scheme at issue.

Stochastic simulations of two-generation within-family MAS schemes in a typical aquaculture breeding programme were carried out by Sonesson (2006) for selection for one trait. Selection was truncated based upon best linear unbiased prediction (BLUP) breeding values including information of the genetic marker (Fernando and Grossman, 1989). The heritability, h^2, of the trait under selection was 0.06 or 0.12, and 20 percent of the genetic variance was accounted for by the QTL.

Genetic gain was 0.202 for MAS and 0.176 for conventional breeding, after selection in generation 1 (Table 3), i.e. MAS had 15 percent higher genetic gain than conventional breeding. After selection in generation 2, genetic gain was 68 percent higher for MAS than conventional breeding. The performance-increasing QTL allele was then almost fixed (i.e. its frequency approached 1). The increase in genetic gain is mainly due to the increased frequency of the positive QTL allele for the MAS scheme, where a higher QTL frequency implies more

TABLE 3
Genetic gain in schemes with heritability (h^2) of 0.06 or 0.12

	Generation number			
	1	2	3	4
		$h^2 = 0.06$		
Conventional	0.176	0.121	0.134	0.117
MAS	0.202	0.203	0.135	0.114
		$h^2 = 0.12$		
Conventional	0.337	0.206	0.196	0.191
MAS	0.361	0.318	0.206	0.177

genetic variance (as long as the frequency of the positive allele is less than 0.5). For situations with a higher h^2 of 0.12, genetic gain was 7 percent higher after selection in generation 1 and 54 percent higher after selection in generation 2, i.e. the superiority of MAS was somewhat lower for schemes with the higher heritability.

IDENTIFICATION OF INDIVIDUALS USING GENETIC MARKERS

In traditional family-based breeding programmes, individuals from the same full-sib families are reared separately (e.g. in tanks) until they are large enough to be tagged physically. This mode of rearing is very costly, and the number of full-sib families therefore limits the size of the breeding nucleus. In addition, separate rearing of full-sib families results in common environmental (tank) effects that need to be estimated, which in turn affects other parts of the design and analysis of the breeding programme. That is, in the mating design, a sire has to be mated to several dams (or *vice versa*) in order to be able to separate analytically these common environmental effects from additive genetic effects. The tank effect is generally higher for newly domesticated species, where feed and other environmental effects are not yet standardized. For example, the tank effect was 3–12 percent for juvenile Atlantic cod (Gjerde *et al.*, 2004), and the nursery pond effect was 32 percent for rohu carp (Gjerde *et al.*, 2003) compared with, for example, a tank effect of 2–6 percent for Atlantic salmon and rainbow trout (Rye and Mao, 1998; Pante *et al.*, 2002). Were it possible to pool progeny groups into one tank, tank effects would be eliminated, a smaller number of tanks would be needed per spawner and more pairs could be spawned.

Parentage assignment using molecular markers is useful for tracing pedigrees in breeding programmes, and can be used to identify parents in breeding schemes where progeny groups are pooled. Parentage assignment implies that parents and offspring are all genotyped for a number of genetic markers that are well spread over the genome and that the information on genotypes is used to assign individual progeny to the correct parental pair. The parent-offspring relationship is such that each offspring inherits one allele from each parent, making it possible to exclude possible parents when this condition is not fulfilled.

Exclusion of parents is the basic method of assigning parents. The exclusion probability per locus (E_l) can be calculated according to the formulae of Hanset (1975) and Dodds *et al.* (1996). The global exclusion probability over loci (E_g) is:

$$E_g = 1 - \prod_{l=1}^{l=L} (1 - E_l)$$

where L is the total number of marker loci screened. Genotyping errors can result in the true parents being excluded, because uncertainty is not accounted for with this method. Even a low error rate reduces the correct assignment rate, which increases with the number of loci and number of alleles. SanCristobal and Chevalet (1997) derived a likelihood and a Bayesian-based method that could take account of genotyping errors. They used the likelihood method on data for 50 parental pairs and their offspring and a genotyping error rate of 2 percent, but with no error rate accounted for in the likelihood calculation. The correct parentage assignment rate was only 88 percent using a system of five loci with five alleles, and 83 percent using a system of eight loci with five alleles. After including a genotyping error rate of 10^{-3} in the same calculations, the correct assignment rate increased to nearly 100 percent.

The number of loci and alleles per loci needed for correct parent assignment was estimated deterministically and validated stochastically by Villanueva, Verspoor and Visscher (2002). They showed that nine loci with five alleles per locus or six loci with ten alleles would assign 99 percent of offspring to the correct parents from 100 or 400 crosses. Similar results were found by Bernatchez and Duchesne (2000). These results agree well with the results from empirical studies of aquaculture populations. For example, Herbinger *et al.* (1995) assigned 66 percent of the offspring to correct parental pairs in a complete factorial cross between ten males x ten females (i.e. 100 parental pairs) of rainbow trout using four loci with four to ten alleles per locus. Perez-Enriquez, Takagi and Taniguchi (1999) reported 73 percent correct assignment of parental pairs using five microsatellites when 7 800 parental pairs were possible for a population of red sea bream. Norris, Bradley and Cunningham (2000) assigned over 95 percent of the parental pairs correctly using eight polymorphic loci with 10–29 alleles per locus in Atlantic salmon when the number of possible parental pairs was over 12 000. Jackson, Martin-Robichaud and Reith (2003) assigned 96–100 percent of the progeny to the correct parental pair in F_1 Atlantic halibut populations using five microsatellite loci. Castro *et al.* (2004) assigned over 99 percent of all parental pairs correctly with six microsatellites for 176 full-sib families of turbot. Vandeputte *et al.* (2004) assigned 95.3 percent of all parental pairs in a complete factorial cross of 10 dams x 24 sires of common carp using eight microsatellites. In addition to assessing parentage, full- and half-sib relationships have also been estimated using genetic markers in aquaculture stocks, e.g. Atlantic salmon (Norris, Bradley and Cunningham, 2000) and rainbow trout (McDonald, Danzmann and Ferguson, 2004).

There are underlying assumptions that affect experimental power for assigning parents. Some of these assumptions are:

- Hardy-Weinberg equilibrium (HWE). Small effective population sizes, non-random mating and unequal family size will lead to deviations from HWE. HWE is not an issue in assigning parental pairs to offspring, but affects the assignment of offspring genotypes to the parents (Estoup *et al.*, 1998), i.e. some genotypes of parents might be more difficult than others from which to assign offspring. If these genotypes are present in large families, parental assignment rate will be reduced relative to what would occur if they are present in small families.
- Zero mutation rate and no scoring errors. Castro *et al.* (2004) reported a mutation

rate of 6.7×10^{-4} when 13 464 gametes and marker information from 12 loci were screened in a turbot population. Mutations and scoring error have the same effect on excluding potential parents, giving rise to incorrect assignments.

- Unlinked loci and linkage equilibrium. Linkage and linkage disequilibrium between the loci will reduce the effective number of loci used for the parental assignment. Note that the power to assign parental pairs correctly can thereby differ between different sets of microsatellites. Estoup *et al.* (1998) quantified the difference in the power of microsatellite marker sets used to assign parents correctly in turbot (eight loci, eight alleles per loci) and rainbow trout (eight loci, four alleles per loci) populations by calculating the frequency of good and unique parent assignments (f_{gu}). The more variable set of microsatellites resulted in higher f_{gu} for larger maximal mating schemes for turbot than the less variable set of microsatellites for rainbow trout. In general, a set of loci with an equal number of alleles has the highest exclusion probability (Weir, 1996; Jamieson and Taylor, 1997).

WALK-BACK SELECTION SCHEMES

Practical breeding schemes using molecular marker-based parental assignment have been reported. Doyle and Herbinger (1994) proposed carrying out parentage assignment for individuals using genetic markers, such that full-sib families would not have to be kept separately until tagging but, rather, would be held in one large tank. Note that individuals that are genotyped also need to be physically tagged, so that genotyping results can be traced back to particular individuals. At the time for selection, fish in the tank were first ranked on their phenotypic value, assuming that selection was for only one trait that could be measured on the selection candidates (e.g. weight). Then the individual with the highest phenotypic value was selected and genotyped for family identification. Thereafter, the individual with the second highest phenotypic value was genotyped and selected if it was not a full- or half-sib of other, already-selected individuals, such that within-family selection was performed. This procedure continued until the desired number of brood stock was selected. This approach of progressing through the performance ranking, genotyping and selecting the best-performing individuals within families was termed "walk-back" selection. Matings subsequently would be made between families, a strategy preferred because it would keep the rate of inbreeding low (Falconer and Mackay, 1996). Herbinger *et al.* (1995) reported setting up a rainbow trout breeding programme using the walk-back selection programme of Doyle and Herbinger (1994) and genetic markers to estimate full/half relationships among the candidates using a likelihood ratio method and thereafter performing within-family selection.

Using stochastic simulations, Sonesson (2005) studied a combination of optimum contribution selection and walk-back selection. Optimum contribution is a selection method that maximizes genetic gain with a restriction on the rate of inbreeding (see earlier in this chapter). Hence, the combination of optimum contribution and walk-back selection ensures that the rate of inbreeding is under control, while the genetic gain is higher than in the within-family selection schemes of Doyle and Herbinger (1994) because selection is both within and between families. In the study by Sonesson (2005), batches of candidates were pre-selected from a single tank on

their phenotypic values and the batch size varied from 50 to all (1 000, 5 000 or 10 000) candidates. Relatively small batches of fish were genotyped at any one time to minimize genotyping costs. Thereafter, BLUP breeding values (Henderson, 1984) were estimated and the optimum contributions of the candidates calculated using the method of Meuwissen (1997). If the constraint on the rate of inbreeding could not be achieved, another batch of fish was genotyped and included in the total number of candidates. Results showed that with a batch size of 100, 76–92 percent of the genetic gain was achieved compared with schemes where all 1 000, 5 000 or 10 000 fish were genotyped to provide candidates for the optimum contribution selection algorithm. Hence, high genetic gain was achieved at a fixed rate of inbreeding with low genotyping costs.

The main practical advantage of these marker-assisted breeding schemes is that expenses associated with separate rearing of full-sib families are not incurred, which decreases start-up and operational costs for the breeding scheme. The most important trait at the start of a breeding programme is probably growth, which can easily be measured on the candidate. Use of the optimum contribution selection algorithm keeps the rate of inbreeding under control, which is especially important in breeding programmes for aquaculture species where selection intensity can be very high due to the large family sizes. In combination with BLUP estimated breeding values, which have a high within-family correlation such that individuals with the highest breeding values will tend to come from only a few families, high rates of inbreeding can result (Sonesson, Gjerde and Meuwissen, 2005). However, there remain unsolved issues with the combined optimum contribution and walk-back selection method:

- Biased BLUP breeding values lead to a reduction in accuracy of selection, because not all selection candidates are included in the estimation of breeding values.
- Low and unequal survival of families may lead to reduced genetic variation and thus increased rate of inbreeding. However, the optimum contribution selection will correct for some of this loss by selecting from more families to keep the genetic base broader than when selecting only for the BLUP estimated breeding values. One option for reducing this problem of unequal and low survival is to pool an equal number of individuals from each family after the main period of early mortality is over. In general, it is possible to use more parents in these programmes compared with conventional family-based selection programmes, which could compensate for some of the loss of families contributing to the next generation due to low and unequal survival.
- Multitrait selection is probably the largest practical problem to solve. One alternative could be to base the pre-selection on only one or two traits that are inexpensive to measure on the candidate. Techniques for measuring more traits on live selection candidates are steadily evolving (e.g. fat content in Atlantic salmon, Solberg *et al*., 2003), such that the sib-testing system might be unnecessary for these traits in the future.

INTROGRESSION SCHEMES

Many fish breeding schemes have been started with a relatively narrow genetic base, selecting for only one or two traits in relatively few animals. However, because all farmed aquatic species still have wild ancestors, introgression of genes (i.e. identified genes or QTL) from these wild

ancestors into breeding populations is possible. However, one assumes that all other traits of the wild fish are unwanted in the breeding population, such that only the particular gene of interest should be introgressed, leaving the genome of the breeding population otherwise intact (Hospital and Charcosset, 1997). In white shrimp, for example, wild stocks have been found to have higher disease resistance but lower growth rates than stocks in culture. In this example then, only the genes for disease resistance should be introgressed (taking account of the possible effect of these genes on growth). The problem of actually implementing marker-assisted introgression in populations is to find the trait of interest in wild populations at a reasonable cost, and then to identify genes or marked QTL for the trait to be introgressed (Visscher, Haley and Thompson, 1992; Koudande *et al*., 2000). This is a costly and time-consuming process, especially for species with long generation intervals. Methods for simultaneous QTL mapping and introgression would be useful.

ACKNOWLEDGEMENTS

I am grateful to Bjarne Gjerde, Eric Hallerman and Thomas Moen for their help with the manuscript.

REFERENCES

Agresti, J.J., Seki, S., Cnaani, A., Poompuang, S., Hallerman, E.M., Umiel, N., Hulata, G., Gall, G.A.E. & May, B. 2000. Breeding new strains of tilapia: development of an artificial center of origin and linkage map based on AFLP and microsatellite loci. *Aquaculture* 185: 43–56.

Bernatchez, L. & Duchesne, P. 2000. Individual-based genotype analysis in studies of parentage and population assignment: how many loci, how many alleles? *Can. J. Fish Aquat. Sci.* 57: 1–12.

Castro, J., Bouza, C., Presa, P., Pino-Querido, A., Riaza, A., Ferreiro, I., Sánchez, L. & Martínez, P. 2004. Potential sources of error in parentage assessment of turbot (*Scophthalamus maximus*) using microsatellite loci. *Aquaculture* 242: 119–135.

Chistiakov, D.A., Hellemans, B., Haley, C., Law, A.S., Tsigenopoulos, C.S., Kotulas, G., Bertotto, D., Libertini, A. & Volckaert, F.A.M. 2005. A microsatellite linkage map of the European sea bass *Dicentrarchus labrax* L. *Genetics* 170: 1821–1826.

Chourrout, D. 1984. Pressure-induced retention of second polar body and suppression of first cleavage in rainbow trout: production of all-triploids, all-tetraploids, and heterozygous and homozygous gynogenetics. *Aquaculture* 36:111–126.

Cnaani, A., Hallerman, E.M., Ron, M., Weller, J.I., Indelman, M., Kashi, Y., Gall, G.A.E. & Hulata, G. 2003. Detection of a chromosomal region with two quantitative trait loci, affecting cold tolerance and fish size, in an F2 tilapia hybrid. *Aquaculture* 223: 117–128.

Cnaani, A., Zilberman, N., Tinman, S. & Hulata, G. 2004. Genome-scan analysis for quantitative traits in an F_2 tilapia hybrid. *Mol. Gen. Genomics* 272: 162–172.

Coimbra, M.R.M., Kabayashi, K., Koretsugu, S., Hasegawa, O., Ohara, E., Ozaki, S., Sakamoto, T., Naruse, K. & Okamoto, N. 2003. A genetic map of the Japanese flounder *Paralichthys olivaceus*. *Aquaculture* 220: 203–218.

Danzmann, R.G., Jackson, T.R. & Ferguson, M.M. 1999. Epistasis in allelic expression at upper temperature tolerance QTL in rainbow trout. *Aquaculture* 173: 45–58.

Dekkers, J.C.M. & Hospital, F. 2002. The use of molecular genetics in the improvement of agricultural populations. *Nature Revs. Genet.* 3: 22–32.

Dodds, K.G., Tate, M.L., McEwan, J.C. & Crawford, A.M. 1996. Exclusion probabilities for pedigree testing farm animals. *Theor. Appl. Genet*. 92: 966–975.

Doyle, R.W. & Herbinger, C.M. 1994. The use of DNA fingerprinting for high-intensity within-family selection in fish breeding. *Proc. 5th Wrld. Congr. Genetics Appl. Livest. Prodn.* pp. 364–371. Guelph, Ontario. Department of Animal and Poultry Science, University of Guelph.

Estoup, A., Gharbi, K., SanCristobal, M., Chevalet, C., Haffray, P. & Guyomard, R. 1998 Parentage assignment using microsatellites in turbot (*Scophthalmus maximus*) and rainbow trout (*Oncorhynchus mykiss*) hatchery populations. *Can. J. Fish. Aquat. Sci.* 55: 715–725.

Falconer, D.S. & Mackay, T.F.C. 1996. *Introduction to quantitative genetics, 4th edition*. Harlow, UK, Longman Publishing.

FAO. 2003. FAOSTAT statistical database. (available at http://faostat.fao.org).

Fernando, R.L. & Grossman, M. 1989. Marker-assisted selection using best linear biased prediction. *Genet. Sel. Evol*. 21: 467–477.

Geldermann, H. 1975. Investigations on inheritance of quantitative characters in animals by gene markers. I. Methods. *Theor. Appl. Genet*. 46: 319–330.

Gilbey, A., Verspoor, E., McLay, A. & Houlihan, D. 2004. A microsatellite linkage map for Atlantic salmon (*Salmo salar*). *Anim. Genet*. 35: 98–105.

Gjedrem, T. & Olesen, I. 2005. Basic statistical parameters. *In* T. Gjedrem, ed. *Selection and breeding programs in aquaculture*, pp. 45–72. Doordrecht, Netherlands, Springer.

Gjerde, B., Mahapatra, K.D., Reddy, P.V.G.K., Saha, J.N., Jana, R.K. & Rye, M. 2003. Genetic and phenotypic parameters for growth in rohu (*Labeo rohita*) in mono- and polyculture production systems. In *Proc. 8th Internat. Symp. Genet. in Aquaculture*, p. 85. Puerto Varas, Chile.

Gjerde, B., Terjesen, B.F., Barr, Y., Lein, I. & Thorland, I. 2004. Genetic variation for juvenile growth and survival in Atlantic cod (*Gadus morhua*). *Aquaculture* 236: 167–177.

Grundy, B., Villanueva, B. & Woolliams, J.A. 1998. Dynamic selection procedures for constrained inbreeding and their consequences for pedigree development. *Genet. Res*. 72: 159–168.

Grundy B., Villanueva B. & Woolliams J. 2000. Dynamic selection for maximizing response with constrained inbreeding in schemes with overlapping generations. *Anim. Sci*. 70 373–382.

Hanset, R. 1975. Probabilité d'exclusion de paternité et de monozygotie, probabilité de similtude. Généralisation à N alleles co-dominants. *Ann. Med. Vet*. 119: 71–80.

Henderson, C.R. 1984. *Applications of linear models in animal breeding.* Guelph, Ontario, Canada,Guelph Univ. Press.

Herbinger, C.M., Doyle, R.W., Pitman, E.R., Paquet, D., Mesa, K.A., Morris, D.B., Wright, J.M. & Cook, D. 1995. DNA fingerprint-based analysis of paternal and maternal effects on offspring growth and survival in communally reared rainbow trout. *Aquaculture* 137: 245–256.

Hubert, S. & Hedgecock, D. 2004. Linkage maps of microsatellite DNA markers for the Pacific oyster *Crassostrea gigas*. *Genetics* 168: 351–362.

Hospital, F. & Charcosset, A. 1997. Marker-assisted introgression of quantitative trait loci. *Genetics* 147: 1469–1485.

Jackson, T.R., Ferguson, M.M., Danzmann, R.G., Fishback, A.G., Ihssen, P.E., O'Connell, M. & Crease, T.J. 1998. Identification of two QTL influencing upper temperature tolerance in three rainbow trout (*Onchorhynchus mykiss*) half-sib families. *Heredity* 80: 143–151.

Jackson, T.R., Martin-Robichaud, D.J. & Reith, M.E. 2003. Application of DNA markers to the management of Atlantic halibut (*Hippoglossus hippoglossus*) broodstock. *Aquaculture* 220: 245–259.

Jamieson, A. & Taylor, S.C.S. 1997. Comparisons of three probability formulae for parentage exclusion. *Anim. Genet*. 28: 397–400.

Kashi, Y., Hallerman, E. & Soller, M. 1990. Marker-assisted selection of candidate bulls for progeny testing programs. *Anim. Prod*. 51: 63–74.

Khoo, S.K., Ozaki, A., Nakamura, F., Arakawa, T., Ishimoto, S., Nikolov, R., Sakamoto, T., Akatsu, T., Mochuzuki, M., Denda, I. & Okamoto, N. 2004. Identification of a novel chromosomal region associated with infectious hematopoietic necrosis (IHN) resistance in rainbow trout, Oncorhynchus mykiss. *Fish Path*. 39: 95–101.

Kocher, T.D., Lee, W.J., Sobolewska, H., Penman, D. & McAndrew, B. 1998 A genetic linkage map of a cichlid fish, the tilapia (*Oreochromis niloticus*). *Genetics* 148: 1225–1232.

Koudandé, O.D., Iraqi, F., Thomson, P.C., Teale, A.J. & van Arendonk, J.A.M. 2000. Strategies to optimize marker-assisted introgression of multiple linked QTL. *Mamm. Genome* 11:145–150.

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

Leder, E.H., Danzmann, R.G. & Ferguson, M.M. 2006. The candidate gene, Clock, localizes to a strong spawning time quantitative trait locus region in rainbow trout. *J. Heredity* 97: 74–80.

Lee, B.-Y., Penman, D.J. & Kocher, T.D. 2003. Identification of the sex-determining region in tilapia (*Oreochromis niloticus*) using bulked segregant analysis. *Anim. Genet*. 34: 379–383.

Lee, B.-Y., Hulata, G. & Kocher, T.D. 2004. Two unlinked loci control the sex of blue tilapia (*Oreochromis aureus*). *Heredity* 92: 543–549.

Lee, B.-Y., Lee, W.J., Streelman, T., Carleton, K.L., Howe, A.E., Hulata, G., Slettan, A., Stern, J.E., Terai, Y. & Kocher, T.D. 2005. A second-generation genetic linkage map of tilapia (*Oreochromis* spp.). *Genetics* 170: 237–244.

Li, Y., Byrne, K., Miggiano, E., Whan, V., Moore, S., Keys, S., Crocos, P., Preston, N. & Lehnert, S. 2003. Genetic mapping of the kuruma prawn (*Penaeus japonicus*) using AFLP markers. *Aquaculture* 219: 143–156.

Li, L., Xiang, J., Liu, X., Zhang, Y., Dong, B. & Zhang, X. 2005. Construction of AFLP-based genetic linkage map for Zhikong scallop, *Chlamys farreri* Jones et Preston and mapping of sex-linked markers. *Aquaculture* 245: 63–73.

Liu, Z., Karsi, A., Li, P., Cao, D. & Dunham, R. 2003. An AFLP-based genetic linkage map of channel catfish (*Ictalurus punctatus*) constructed by using an interspecific hybrid resource family. *Genetics* 165: 687–694.

Mackinnon, M.J. & Georges, M.A.J. 1998. Marker-assisted preselection of young dairy sires prior to progeny-testing. *Livest. Prod. Sci*. 54: 229–250.

Martinez, V.A., Hill, W.G. & Knott, S.A. 2002. On the use of double haploids for detecting QTL in outbred populations. *Heredity* 88: 423–431.

Martinez, V., Thorgaard, G., Robison, B. & Sillanpää, M.J. 2005. An application of Bayesian QTL mapping to early development in double haploid lines of rainbow trout including environmental effects. *Genet. Res. Camb*. 86: 209–221.

Martyniuk, C.J., Perry, G.M.L., Mogahadam, H.K., Ferguson, M.M. & Danzmann, R.G. 2003. The genetic architecture of correlations among growth-related traits and male age at maturation in rainbow trout. *J. Fish Biol.* 63: 746–764.

May, B. & Johnson, J.E. 1990. Composite linkage map of salmonid fishes. *In* S.J. O'Brien, ed. *Genetic maps,* 5th edition, pp. 4.151–4.159. Cold Spring Harbor, NY, USA, Cold Spring Harbor Laboratory Press.

McConnell, S.K.J., Beynon, C., Leamon, J. & Skibinski, D.O.F. 2000. Microsatellite marker-based genetic linkage maps of *Oreochromis aureus* and *O. niloticus* (Cichlidae): extensive linkage group segment homologies revealed. *Anim. Genet*. 31: 214–216.

McDonald, G.J., Danzmann, R.G. & Ferguson, M.M. 2004. Relatedness determination in the absence of pedigree information in three cultured strains of rainbow trout (*Oncorhynchus mykiss*). *Aquaculture* 233: 65–78.

Meuwissen, T.H.E. 1997. Maximizing the response of selection with a predefined rate of inbreeding. *J. Anim. Sci*. 75: 934–940.

Meuwissen, T.H.E. & Sonesson, A.K. 1998. Maximizing the response of selection with a predefined rate of inbreeding – overlapping generations. *J. Anim. Sci.* 76: 2575–2583.

Meuwissen, T.H.E. & Sonesson, A.K. 2004. Genotype-assisted optimum contribution selection to maximise selection response over a specified time period. *Genet. Res*. 84: 109–116.

Meuwissen, T.H.E., Karlsen, A., Lien, S., Olsaker, I. & Goddard, M.E. 2002. Fine mapping of a quantitative trait locus for twinning rate using combined linkage and linkage disequilibrium mapping. *Genetics* 161: 373–379.

Moen, T., Hoyheim, B., Munck, H. & Gomez-Raya, L. 2004a. A linkage map of Atlantic salmon (*Salmo salar*) reveals an uncommonly large difference in recombination rate between the sexes. *Anim. Genet.* 35: 81–92.

Moen, T., Agresti, J.J., Cnaani, A., Moses, H., Famula, T.R., Hulata, G., Gall, G.A.E. & May, B. 2004b. A genome scan of a four-way tilapia cross supports the existence of a quantitative trait locus for cold tolerance on linkage group 23. *Aquaculture* 35: 893–904.

Moen, T., Fjalestad, K.T., Munck, H. & Gomez-Raya, L. 2004c. A multistage testing strategy for detection of quantitative trait loci affecting disease resistance in Atlantic salmon. *Genetics* 167: 851–858.

Moen, T., Hayes, B., Sonesson, A.K., Lien, S., Høyheim, B., Munck, H. & Meuwissen, T.H.E. 2006. Mapping of a quantitative trait loci for resistance against infectious salmon anemia in Atlantic salmon (*Salmo salar*): comparing survival models with binary models. *Genetics*. Submitted.

Moore, S.S., Whan, V., Davis, G.P., Byrne, K., Hetzel, D.J.S. & Preston, N. 1999. The development and application of genetic markers for the kuruma prawn (*Penaeus japonicus*). *Aquaculture* 173: 19–32.

Nakamura, K., Ozaki, A., Akutsu, T., Iwai., K., Sakamoto, T., Toshizaki, G. & Okamoto, N. 2001. Genetic mapping of the dominant albino locus in rainbow trout (*Oncorhynchus mykiss*). *Mol. Genet. Genom*. 265: 687–93.

Nichols, K.M., Bartholomew, J. & Thorgaard, G.H. 2003. Mapping multiple genetic loci associated with *Ceratomyxa shasta* resistance in *Oncorhynchus mykiss*. *Dis. Aquat. Org*. 56: 145–154.

Nichols, K.M., Wheeler, P.A. & Thorgaard, G.H. 2004. Quantitative trait loci analyses for meristic traits in *Oncorhynchus mykiss*. *Env. Biol. Fish*. 69: 317–331.

Nichols, K.M., Robison, B.D., Wheeler, P.A. & Thorgaard, G.H. 2000. Quantitative trait loci (QTL) associated with developmental rate in clonal *Oncorhynchus mykiss* strains. *Aquaculture* 209:233.

Nichols, K.M., Young, W.P., Danzmann, R.G., Robison, B.D., Rexroad, C., Noakes, M., Phillips, R.B., Bentzen, P., Spies, I., Knudsen, K., Allendorf, F.W., Cunningham, B.M., Brunelli, J., Zhang, H., Ristow, S., Drew, R., Brown, K.H., Wheeler, P.A. & Thorgaard, G.H. 2003. A consolidated linkage map for rainbow trout (*Oncorhynchus mykiss*). *Anim. Genet*. 34: 102–115.

Norris, A.T., Bradley, D.G. & Cunningham, E.P. 2000. Parentage and relatedness determination in farmed Atlantic salmon (*Salmo salar*) using microsatellite markers. *Aquaculture* 182: 73–83.

O'Malley, K.G., Sakamoto, T., Danzmann, R.G. & Ferguson, M.M. 2003. Quantitative trait loci for spawning date and body weight in rainbow trout: testing for conserved effects across ancestrally duplicated chromosomes. *J. Hered.* 94: 273–84.

Ozaki, A., Sakamoto, T., Khoo, S., Nakamura, K., Coimbra, M.R.M., Akutsu, T. & Okamoto, N. 2001. Quantitative trait loci (QTLs) associated with resistance/susceptibility to infectious pancreatic necrosis virus (IPNV) in rainbow trout (*Oncorhynchus mykiss*). *Mol. Genet. Genom*. 265: 23–31.

Palti, Y., Shirak, A., Cnaani, A., Hulata, G., Avtalion, R.R. & Ron, M. 2002. Detection of genes with deleterious alleles in an inbred line of tilapia (*Oreochromis aureus*). *Aquaculture* 206: 151–164.

Pante, M.J.R., Gjerde, B., McMillan, I. & Misztal, I. 2002. Estimation of additive and dominance genetic variances for body weight at harvest in rainbow trout, *Oncorhynchus mykiss*. *Aquaculture* 204: 383–392.

Parsons, J.E. & Thorgaard, G.H. 1984. Induced androgenesis in rainbow trout. *J. Exp. Zool*. 231: 407–412.

Pérez, F., Erazo, C., Zhinaula, M., Volckaert, F. & Calderon, J. 2004. A sex-specific linkage map of the white shrimp *Penaeus* (*Litopenaeus*) *vannamei* based on AFLP markers. *Aquaculture* 242: 105–118.

Perez-Enriquez, R., Takagi, M. & Taniguchi, N. 1999. Genetic variability and pedigree tracing of a hatchery-reared stock of red sea bream (*Pagrus major*) used for stock enhancement, based on microsatellite DNA markers. *Aquaculture* 173: 413–423.

Perry, G.M., Ferguson, M.M. & Danzmann, R.G. 2003. Effects of genetic sex and genomic background on epistasis in rainbow trout (*Oncorhynchus mykiss*). *Genetica* 119: 35–50.

Perry, G.M., Danzmann, R.G., Ferguson, M.M. & Gibson, J.P. 2001. Quantitative trait loci for upper thermal tolerance in outbred strains of rainbow trout (*Oncorhynchus mykiss*). *Heredity* 86: 333–341.

Poompuang, S. & Hallerman, E.M. 1997. Toward detection of quantitative trait loci and marker-assisted selection in fish. *Rev. Fisheries Sci*. 5: 253–277.

Reid, D.P., Szanto, A., Glebe, B., Danzmann, R.G. & Ferguson, M.M. 2005. QTL for body weight and condition factor in Atlantic salmon (*Salmo salar*): comparative analysis with rainbow trout (*Oncorhynchus mykiss*) and Arctic charr (*Salvelinus alpinus*). *Heredity* 94: 166–172.

Robison, B.D., Wheeler, P.A., Sundin, K., Sikka, P. & Thorgaard, G.H. 2001. Composite interval mapping reveals a major locus influencing embryonic development rate in rainbow trout (*Oncorhynchus mykiss*). *J. Hered*. 92: 16–22.

Rodriguez, M.F., LaPatra, S., Williams, S., Famula, T. & May, B. 2005. Genetic markers associated with resistance to infectious hematopoietic necrosis in rainbow and steelhead trout (*Oncorhynchus mykiss*) backcrosses. *Aquaculture* 241: 93–115.

Rye, M. & Mao, I.L. 1998. Nonadditive genetic effects and inbreeding depression for body weight in Atlantic salmon (*Salmo salar*). *Livest. Prodn. Sci*. 57: 15–22.

Sakamoto, T., Danzmann, R.G., Okamoto, N., Ferguson, M.M. & Ihssen, P.E. 1999. Linkage analysis of quantitative trait loci associated with spawning time in rainbow trout (*Oncorhynchus mykiss*). *Aquaculture* 173: 33–43.

Sakamoto, T., Danzmann, R.G., Gharbi, K., Howard, P., Ozaki, A., Khoo, S.K., Woram, R.A., Okamoto, N., Ferguson, M.M., Holm, L.-E., Guyomard, R. & Hoyheim, B. 2000. A microsatellite linkage map of rainbow trout (*Oncorhynchus mykiss*) characterized by large sex-specific differences in recombination rates. *Genetics* 155: 1331–1345.

SanCristobal, M. & Chevalet, C. 1997. Error tolerant parent identification from a finite set of individuals. *Genet. Res. Camb*. 70: 53–62.

Smith, C. & Smith, D.B. 1993. The need for close linkages in marker-assisted selection for economic merit in livestock. *Anim. Breed. Abstr*. 61: 197–204.

Solberg, C., Saugen, E., Swenson, L.P., Bruun, L. & Isaksson, T. 2003. Determination of fat in live farmed Atlantic salmon using non-invasive NIR techniques. *J. Sci. Food Agr*. 83: 692–696.

Somorjai, I.M., Danzmann, R.G. & Ferguson, M.M. 2003. Distribution of temperature tolerance quantitative trait loci in Arctic charr (*Salvelinus alpinus*) and inferred homologies in rainbow trout (*Oncorhynchus mykiss*). *Genetics* 165: 1443–1456.

Sonesson, A.K. 2005. A combination of walk-back and optimum contribution selection for fish – a simulation study. *Genet. Sel. Evol.* 37: 587–599.

Sonesson, A.K. 2006. Marker-assisted selection in breeding schemes for aquaculture species. *Genet. Sel. Evol*. (in press*).*

Sonesson, A.K., Gjerde, B. & Meuwissen, T.H.E. 2005. Truncation selection for BLUP-EBV and phenotypic values in fish breeding schemes. *Aquaculture* 243: 61–68.

Stoss, J. 1983. Fish gamete preservation and spermatozoan physiology. *In* W.S. Hoar, D.J. Randall, & E.M. Donaldson, eds. *Fish physiology, vol. 9, part B*. pp. 305–350. New York, NY, USA, Academic Press.

Streelman, J.T., Albertson, R.C. & Kocher, T.D. 2003. Genome mapping of the orange blotch colour pattern in cichlid fishes. *Molec. Ecol*. 12: 2465–2471.

Sun, X. & Liang, L. 2004. A genetic linkage map of common carp (*Cyprinus carpio* L.) and mapping of a locus associated with cold tolerance. *Aquaculture* 238: 165–172.

Sundin, K., Brown, K.H., Drew, R.E., Nichols, K.M., Wheeler, P.A. & Thorgaard, G.H. 2005. Genetic analysis of development rate QTL in backcrosses of clonal rainbow trout. *Aquaculture* 247: 75–83.

Vandeputte, M., Kocour, M., Mauger, S., Dupont-Nivet, M., De Guerry, D., Rodina, M., Gela, D., Vallod, D., Chevassus, B. & Linhart, O. 2004. Heritability estimates for growth-related traits using microsatellite parentage assignment in juvenile common carp (*Cyprinus carpio* L.). *Aquaculture* 235: 233–236.

Villanueva, B., Verspoor, E. & Visscher, P.M. 2002. Parental assignment in fish using microsatellite genetic markers with finite numbers of parents and offspring. *Anim. Genet*. 33: 33–41.

Visscher, P.M., Haley, C.S. & Thomson, R. 1992. Marker-assisted introgression in backcross breeding programs. *Genetics* 144: 1923–1932.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., vanden Lee, T., Hornes, M., Fritters, A., Pot, J., Paleman, J., Kuiper, M. & Zabeau, M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucl. Acids Res*. 23: 4407–4414.

Waldbieser, G.C., Bosworth, B.G., Nonneman, D.J. & Wolters, W.R. 2001. A microsatellite-based genetic linkage map for channel catfish (*Ictalurus punctatus*). *Genetics* 158: 727–734.

Wang, S., Bao, Z., Pan, J., Zhang, L., Yao, B., Zhan, A., Bi, K. & Zhang, Q. 2004. AFLP linkage map of an intraspecific cross in *Chlamys farreri*. *J. Shellfish Res*. 23: 491–499.

Weir, B.S. 1996. *Genetic data analysis, 2nd edition*. Sunderland, MA, USA., Sinauer Associates, Inc.

Wilson, K., Li, Y., Whan, V., Lehnert, S., Byrne, K., Moore, S., Pongsomboon, S., Tassanakajon, A., Rosenberg, G., Ballment, E., Fayazi, A., Swan, J., Kenway, M. & Benzie, J. 2002. Genetic mapping of the black tiger shrimp (*Penaeus monodon*) with amplified fragment length polymorphism. *Aquaculture* 204: 297–309.

Woram, R.A., McGowan, C., Stout, J.A., Gharbi, K., Ferguson, M.M., Hoyheim, B., Sakamoto, J., Davidson, W., Rexroad, C. & Danzmann, R.G. 2004. A genetic linkage map for Arctic char (*Salvelinus alpinus*): evidence for higher recombination rates and segregation distortion in hybrid versus pure strain mapping parents. *Genome* 47: 304–315.

Young, W.P., Wheeler, P.A., Coryell, V.H., Keim, P. & Thorgaard, G.H. 1998. A detailed map of rainbow trout produced using double haploids. *Genetics* 148: 839–850.

Yu, Z.N. & Guo, X.M. 2003. Genetic linkage map of the eastern oyster (*Crassostrea virginica* Gmelin). *Biol. Bull.* 204: 327–338.

Zimmerman, A.M., Evenhuis, J.P., Thorgaard, G.H. & Ristow, S.S. 2004. A single major chromosomal region controls natural killer cell-like activity in rainbow trout. *Immunogen.* 55: 825–35.

Zimmerman, A.M., Wheeler, P.A., Ristow, S.S. & Thorgaard, G.H. 2005. Composite interval mapping reveals three QTL associated with pyloric caeca number in rainbow trout, *Oncorhynchus mykiss*. *Aquaculture* 247: 85–95.