CHAPTER 17

Marker-assisted selection in fish and shellfish breeding schemes

Victor Martinez

SUMMARY

The main goals of breeding programmes for fish and shellfish are to increase the profitability and sustainability of aquaculture. Traditionally, these have been carried out successfully using pedigree information by selecting individuals based on breeding values predicted for traits measured on candidates using an "animal model". This methodology assumes that phenotypes are explained by a large number of genes with small effects and random environmental deviations. However, information on individual genes with medium or large effects cannot be used in this manner. In selective breeding programmes using pedigree information, molecular markers have been used primarily for parentage assignment when tagging individual fish is difficult and to avoid causing common environmental effects from rearing families in separate tanks. The use of these techniques in such conventional breeding programmes is discussed in detail.

Exploiting the great biological diversity of many fish and shellfish species, different experimental designs may use either chromosomal manipulations or large family sizes to increase the likelihood of finding the loci affecting quantitative traits, the so-called QTL, by screening the segregation of molecular markers. Using information on identified loci in breeding schemes in aquaculture is expected to be cost-effective compared with traditional breeding methods only when the accuracy of predicting breeding values is rather low, e.g. for traits with low heritability such as disease resistance or carcass quality. One of the problems facing aquaculture is that some of the resources required to locate QTL accurately, such as dense linkage maps, are not yet available for the many species. Recently, however, information from expressed sequence tag (EST) databases has been used for developing molecular markers such as microsatellites and single nucleotide polymorphisms (SNPs). Marker-assisted selection (MAS) or genome-wide marker-assisted selection (G-MAS) using linkage disequilibrium within families or across populations are not widely used in aquaculture, but their application in actual breeding programmes is expected to be a fertile area of research. This chapter describes how genomic tools can be used jointly with pedigree-based breeding strategies and the potential and real value of molecular markers in fish and shellfish breeding schemes.

INTRODUCTION

The main goals of fish and shellfish breeding programmes are to increase the profitability and sustainability of production enterprises, while maintaining genetic variability in the cultured stock. Traditionally, selective breeding has targeted traits such as body weight that can be easily improved using mass selection. Relatively few studies have analysed other traits that are economically important. However, disease resistance and carcass quality are traits that are difficult to measure on candidates for selection, but have major effects on the production efficiency and profitability of many species in aquaculture.

When developing efficient breeding programmes, pedigree information is required to maximize effective population sizes and to use information from relatives to increase the accuracy of predicting breeding values for all traits included in the breeding objective. In most commercial applications, pedigree information is lacking; therefore, markers can be used to infer relatedness between individuals, with or without parental information. Several issues need to be considered on a case-by-case basis when applying such molecular information for increasing the profitability of breeding programmes in practice.

For traits that are difficult to measure on candidates for selection, prediction of breeding value has to rely on measurements on relatives. Under these circumstances, the accuracy of predicted breeding values (and thus, response) is lower than when records are obtained directly on candidates for selection. In addition, there is an increased probability of co-selecting relatives. It is especially for these traits that molecular markers that directly affect or are linked to quantitative trait loci (QTL) have been regarded as useful for marker-assisted selection (MAS) or gene-assisted selection (GAS) programmes.

This chapter begins by discussing the status of "conventional" breeding programmes, the challenges involved when starting such programmes for new species and the possibilities of incorporating marker information in "conventional" programmes. An outline is then provided of the molecular markers developed for aquaculture species and of their use for genetic analysis. The main features of designs for QTL mapping, including the use of chromosomal manipulations, are described, followed by a discussion of the prospects and challenges of GAS or MAS for disease or carcass traits. Finally, new genomic tools are considered briefly.

BREEDING PROGRAMMES AND RESPONSE TO SELECTION

Management of modern breeding programmes in aquaculture requires using pedigree information to carry out sound and efficient statistical evaluations (using best linear unbiased prediction [BLUP] methodology). This approach enables breeders to maximize genetic gain while limiting rates of inbreeding to acceptable levels (Meuwissen, 1997; Toro and Mäki-Tanila, 1999).

Most of the genetic improvement in fish and shellfish species to date has been made through the use of traditional selective breeding (reviewed by Hulata, 2001). Well-designed breeding programmes have shown substantial response to selection for body weight, e.g. Atlantic salmon, 10–14 percent (Gjøen and Bentsen, 1997). In rainbow trout, rates of genetic gain varied from 8 percent for indirect selection for body weight at sea (Kause *et al.*, 2005) to 13 percent for direct selection (Gjerde, 1986). The response to selection was about

10 percent for body weight in a breeding programme for coho salmon (CMG-IFOP) initially funded by FAO (Martinez and Hidalgo, unpublished data), and a similar response was obtained for this species in the United States of America (Hershberger *et al.,* 1990). Estimates for tilapia follow largely the same trend, with a response of about 10 percent (Ponzoni *et al.,* 2005). In common carp, responses to selection for body weight were inconsistent between up-selected and down-selected lines, although exhaustion of additive genetic variation for increased growth rate, genotype-by-environment interaction, or competition effects could not be ruled out (Moav and Wohlfarth, 1976). In oysters, asymmetrical response to selection for body weight was found (Toro *et al.,* 1995; Ward, English and McGoldrick, 2000).

Although responses to selection have not been well documented, significant estimates of genetic parameters have been obtained for carcass traits (Gjerde and Schaeffer, 1989; Kause *et al.,* 2002; Quinton, McMillan and Glebe, 2005) and disease resistance (Gjøen *et al.,* 1997; Henryon *et al.,* 2002, 2005). Rates of genetic gain are expected to be lower for these traits than for body weight because breeding value predictions rely solely on measurements from relatives.

Several breeding programmes have been initiated recently for new aquaculture species, such as mussels, scallops, *Artemia* and shrimp. The biology of these species poses interesting avenues for the design of conventional breeding programmes, taking into account factors such as self-fertilization, intrafamily competition, cannibalism, lack of methods for physical tagging, and mating preferences. For example, competition can affect the expression of quantitative traits due to co-variances among members of a group managed together in a pond or tank and, if not considered properly, this effect can seriously affect the rates of response to selection (Muir, 2005). However, this effect can be included explicitly in the model of analysis using the co-variance among members of a group, the so-called "associative effects" from other genotypes in the group. The theory of Griffing (1967) for BLUP evaluation was developed in the context of tree breeding, but deserves further investigation in the analysis of fish and shellfish breeding. This may be especially true for species taken recently from the wild or those that show cannibalistic behaviour.

Another recent example is the development of scallop breeding programmes. *Argopecten purpuratus* is a simultaneously hermaphroditic species. In the first breeding phase, the scallop liberates sperm, after which the eggs are expelled. To decrease the level of self-fertilization, it is customary to use only the last pulses of eggs. This system reduces rates of self-fertilization to 20 percent (A. Vergara, personal communication), but a residual proportion of eggs are still already fertilized with sperm from the same individual. As this process occurs within the reproductive tract, it is not possible to detect which individuals are selfed or outcrossed, although the rate of residual self-fertilization varies widely among families and produces biased estimates of heritability (Martinez and di Giovanni, 2006). Information from molecular markers can be of benefit under these circumstances (see below).

DNA MARKERS USED IN AQUACULTURE

Mutations in the genome create genetic variability (or polymorphism), which is reflected as allelic diversity of molecular markers. While genomic sequencing would greatly facilitate the development

of molecular markers, the many species in aquaculture would make this a costly task (Liu and Cordes, 2004). Hence, a variety of approaches have been taken to develop genetic markers for aquaculture species.

Dominantly-expressed markers have been used extensively in aquaculture studies. Amplified fragment length polymorphism (AFLP) markers (Vos *et al.,* 1995) provide a cost-effective alternative for species where DNA sequencing is not under way or when there are restricted resources for QTL mapping. Dominant AFLP markers are preferred over random amplified polymorphic DNA (RAPD) markers because they are more reproducible both in other lines or populations and in other laboratories (e.g. Nichols *et al.,* 2003), and they can generate hundreds of markers (a single polymerase chain reaction commonly generates over ten markers). Furthermore, heterozygotes can often be distinguished from homozygotes using the fluorescent band intensity (Piepho and Koch, 2000; Jansen *et al.,* 2001).

Microsatellite markers are simple sequence repeats (SSRs) arranged in tandem arrays scattered throughout the genome, both within known genes and in anonymous regions. Microsatellite markers are used increasingly in aquaculture species (reviewed by Liu and Cordes, 2004), due to their elevated polymorphic information content (PIC), co-dominant mode of expression, Mendelian inheritance, abundance and broad distribution throughout the genome (Wright and Bentzen, 1994). Microsatellites are generally Type II markers, which are associated with genomic regions that have not been annotated to known genes (O'Brien, 1991). Other molecular markers can be distinguished as Type I markers, which are linked to genes (of known function). Type I markers are more desirable because they are generally more conserved across evolutionarily distant organisms, enabling comparative genomics, assessment of genome evolution and candidate gene analysis.

Two procedures are used to generate microsatellite markers. The first uses a genomic library enriched with microsatellite-bearing sequences to generate clones that bear specific SSRs. These clones are then sequenced to identify microsatellite-bearing sequences and then to design primers to amplify the regions with specific SSR. Validation is required to study the level of polymorphism and the number of null alleles, and to identify any loci that are duplicates due to any recent evolutionary genome duplication event giving rise to multiple copies of loci in the haploid genome (Coulibaly *et al.,* 2005). This is done by screening a sample of individuals from the target population.

Many laboratories have been working on developing expressed sequence tags (ESTs) derived from complementary DNA (cDNA) libraries for a variety of fish and shellfish species (Panitz *et al.,* 2002; Rise *et al.,* 2004a; Hayes *et al.,* 2004; Rexroad *et al.,* 2005; A. Alcivar-Warren, personal communication). EST sequences can be used for marker development in species where the full genome is not currently being sequenced. The cDNA libraries are constructed using messenger RNA (mRNA) that was expressed in different tissues, such as kidney and gills. The expressed fragments of sequence data are not the full sequence of a known gene, but what was incorporated into a mature mRNA molecule.

In addition to the library-based method of marker development previously described, microsatellites can be developed from EST databases or from known gene sequences. As it is possible to connect the

TABLE 1
Recently published linkage maps for various fish and shellfish species used in aquaculture

Species	Number of markers	Marker type	Map length Female/Male	Male cM (Kosambi)	Female cM (Kosambi)	Reference
Atlantic salmon	473	AFLP	8.26:1	103	901	Moen *et al.*, 2004a
	54	Microsatellites				
	65	Microsatellites	3.92	np	np	Gilbey *et al.*, 2004
Rainbow trout	226	Microsatellites	-	4 590		Nichols *et al.*, 2003
	973	AFLP				
	4	Allozymes				
	72	VNTR				
	29	Known genes				
	12	Minisatellites				
	5	RAPDs				
	38	SINE*				
Oysters	115	Microsatellites	1.31:1	776	1 020	Houbert and Hedgecock, 2004
Sea bass	174	Microsatellites	1.6:1	567.4	905.9	Chistiakov *et al.*, 2005
Kuruma prawn	195	AFLP		1 780	1 026	Li *et al.*, 2003
Tilapia	525	Microsatellites	1:1	1 300		Lee *et al.*, 2005
	21	Genes				
Scallops	503	AFLP	1.27:1	2 468	3 130	Wang *et al.*, 2005
Common carp	110	Microsatellites	-	4 111		Sun and Liang, 2004
	105	Known genes				
	57	RAPDs				
Japanese flounder	111	Microsatellites	7.4:1	741.1	670.4	Coimbra *et al.*, 2003
	352	AFLP				
Channel catfish	313	Microsatellites	3.18:1	1 958[B]		Waldbieser *et al.*, 2001

[B] Sex-averaged
* Short interspersed elements
np = not published

function of the transcript of genes (from an EST sequence) with the presence of a microsatellite, these markers are Type I markers (O'Brien, 1991; Serapion *et al.*, 2004; Ng *et al.*, 2005). This strategy of developing microsatellite markers from known genes and ESTs has been used for common carp (Yue, Ho and Orban, 2004), rainbow trout (Rexroad *et al.*, 2005; Coulibaly *et al.*, 2005) and Atlantic salmon (Ng *et al.*, 2005; Vasemägi, Nilsson and Primmer, 2005).

In all these analyses, high levels of transferability between populations and species can be expected if the microsatellites are included in coding regions. Such transferability has been observed e.g. between Atlantic salmon and rainbow trout (Vasemägi *et al.*, 2005; Rexroad *et al.*, 2005), making these markers ideal for analyses of population genetics and comparative maps. For example, microsatellites derived from EST sequences have been used to study divergence of Atlantic salmon populations in salt, brackish and freshwater habitats (Vasemägi, Nilsson and Primmer, 2005).

Bioinformatic tools can be used for potential discovery of SNPs using DNA sequence alignment "*in silico*" (Marth *et al.*,

1999). Although it is possible to use base quality values to discern true allelic variations from sequencing errors, validation is a key step for true positive detection of SNPs (Marth *et al.*, 1999). This is generally carried out using a proportion of the SNPs detected in a sample of individuals from the target population. This strategy has been used recently for SNP detection using EST sequences from Atlantic salmon (Panitz *et al.*, 2002; Hayes *et al.*, 2004).

Linkage maps

A linkage map is an ordered collection of the genes and genetic markers occurring along the lengths of the chromosomes of a species, with distances between them estimated on the basis of the number of recombination events observed in the data. Genetic linkage maps have been published for rainbow trout (Young *et al.*, 1998; Sakamoto *et al.*, 2000; Nichols *et al.*, 2003), channel catfish (Waldbieser *et al.*, 2001), tilapias (Kocher *et al.*, 1998; Lee *et al.*, 2005) and Japanese flounder (Coimbra *et al.*, 2003). References to updated linkage maps of the major aquaculture species are given in Table 1. Dense linkage maps including a relatively large number of markers are under development.

Different patterns of recombination appear among regions of linkage groups in certain male maps, with markers clustered in centromeric regions, an extreme example being Atlantic salmon where recombination in males is greatly reduced (Moen *et al.*, 2004b). The molecular mechanisms responsible for the differences in recombination rates between sexes are not well understood, although studies on model organisms such as zebrafish, where genomic sequencing is currently under way, may help to clarify this (Singer *et al.*, 2002).

USING MARKERS TO AID CONVENTIONAL FISH AND SHELLFISH BREEDING PROGRAMMES

Molecular markers may be used in a number of ways to aid conventional breeding of fish and shellfish species, and some of these are described and exemplified below.

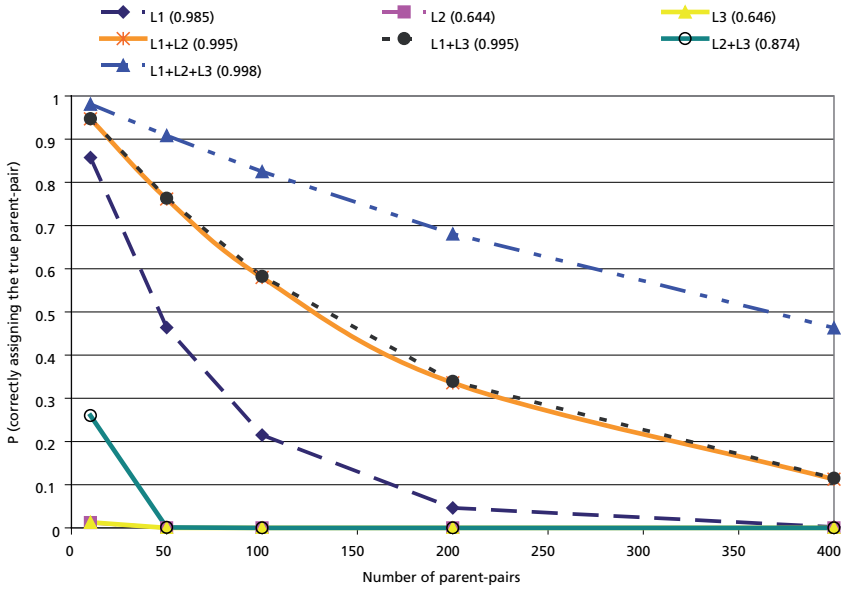
Parentage analysis

One of the main constraints facing effective breeding programmes for fish and shellfish is that newborn individuals are too small to be tagged individually. Application of the animal model approach (i.e. using a statistical genetic model to predict individual breeding values) requires tagging a constant number of individuals from each family with passive integrated transponders (PIT tags) when they become sufficiently large after a period of individual family rearing. However, this system of early management creates common environmental (i.e. tank) effects for full-sib families (Martinez, Neira and Gall, 1999). To address this issue, mixtures of equal-aged progeny from different families can be reared communally to preclude the development of such family-specific environmental effects, and genetic markers can be used subsequently to assign individuals to families after evaluation of individual performance (Doyle and Herbinger, 1994). Thus, the impact of early common environmental effects is considerably reduced if markers are used for parentage analysis when selecting individuals for early growth rate traits (Herbinger *et al.*, 1999; Norris, Bradley and Cunningham, 2000).

The amount of marker data needed to achieve acceptable levels of correct parentage assignment depends on the number of loci, the number of alleles and the number of parent-pairs (sires and dams) available for reconstructing the pedigree (Jamieson

FIGURE 1

Predicted probability of assigning the correct sire-dam (parent-pair) for a given number of parent-pairs (*x* axis) using different combinations of three microsatellites (L1, L2 and L3) amplified in *Oncorhynchus kisutch*



Probabilities of exclusion included between parentheses in the legend are obtained using the deterministic method of Villanueva, Verspoor and Visscher, 2002.
Data obtained from Diagnotec SA.

and Taylor, 1997; Villanueva, Verspoor and Visscher, 2002). The information from the marker data available for each species can be studied using exclusion probabilities, which are then used to calculate the probability (PC) of correctly assigning the true parent-pair (sire and dam) to offspring that are genotyped (Villanueva, Verspoor and Visscher, 2002).

Figure 1 presents the results for three microsatellites and combinations of microsatellites to predict the probabilities of exclusion and PC. The allelic frequencies of the three microsatellites were calculated with a sample (n=100) from a coho salmon (*O. kisutch*) farm in southern Chile managed under commercial conditions. The analysis showed that the probability of assigning the true parent-pair depended greatly on the number of parent pairs available for parentage. Only for an unrealistically small number of ten sires and dams is there a high probability of assigning the correct parent-pair to offspring. For a breeding programme of 200 or 300 parent-pairs, PC decreased considerably. Therefore, in this example, more markers are needed for accurate pedigree reconstruction. Successful parentage assignment experiments typically have used six to eight microsatellite markers (Herbinger *et al.*, 1995; Garcia de Leon *et al.*, 1998; Norris, Bradley and Cunningham, 2000; Castro *et al.*, 2004). In practice, the presence of genotyping errors, null alleles, realized mutations and non-Mendelian segregation can seriously affect the efficiency of parentage assignment (Castro *et al.*, 2004). Parentage assignment in the con-

text of fish breeding is also discussed by Sonesson (this volume).

For most breeding programmes, physical tagging will prove efficient both in economic and biological terms to achieve acceptable rates of genetic gain, while minimizing rates of inbreeding. Genetic marker technology can still be costly for routine assignment of parentage, although these costs can be reduced using multiplex polymerase chain reaction (PCR) technology (Paterson, Piertney and Knox, 2004; Taris, Baron and Sharbel, 2005) in which more than one marker can be genotyped simultaneously in a single gel lane or capillary. This is especially the case when only DNA markers are used without physical tagging, as individuals must be retyped when records for multiple traits are included in the selection criteria (Gjerde, Villaneuva and Bentsen, 2002).

It is expected that rates of genetic gain for economic traits will not be affected significantly when common environmental effects are present. This is because, in many species of cultured salmonids, the common environmental effect decreases considerably, from about 20 percent for alevin weight to 5 percent for body weight at harvest, which is the trait with most impact on profit (Herbinger *et al.,* 1999; Henryon *et al.,* 2002; Kause *et al.,* 2005). Hence, common environmental effects should not decrease the rates of genetic gain for traits measured at harvest when physical tagging is used. Furthermore, multistage selection offers the possibility of first selecting individuals on a within-family basis directly from tanks (for traits influenced by common environmental effects), and then selecting at a second stage for traits measured at harvest (Martinez, 2006a). This alternative would either maintain rates of gain while decreasing the costs associated with tagging, or even increase rates of gain, when recording from tanks (within families) can be carried out relatively inexpensively (Martinez, 2006a).

The sample size (i.e. the numbers of individuals and markers required for reconstructing the pedigree of a population accurately) is a practical issue, as not all individuals in a population can be genotyped for all markers available. Such issues arise in species where physical tagging is not possible or not economically sound, as in nucleus populations without electronic tagging (e.g. when recovering a back-up population for nucleus breeding programmes) or when disease challenges (e.g. for infectious pancreatic necrotic virus [IPNV]) are carried out early in the life cycle (Martinez *et al.,* in preparation). Small sample sizes, together with sperm competition (Withler and Beacham, 1994), mating preference (as in *Artemia*; G. Gajardo, personal communication) and other biological factors after fertilization can increase the variance of family size, thereby decreasing the effective population size to unsustainable levels (Brown, Woolliams and McAndrew, 2005).

Another problem arises in practice when selection is carried out before genotyping with markers. In this case, BLUP of breeding values is likely to be biased because not all phenotypic information is used when predicting breeding values. The magnitude of re-ranking is dependent on the amount of information from a family within the selected group. In these instances, the mixed model equations need to be modified to account for such selected data (Morton and Howarth, 2005).

Establishing breeding programmes using molecular information

The choices made at the founding of a breeding programme have a critical

bearing on its ultimate success. Criteria for choosing individuals that will be founders should be essentially the same as those used when the selection response is optimized under restricted co-ancestry when pedigree information is available (Meuwissen, 1997; Toro and Mäki-Tanila, 1999). Thus, it is necessary to avoid matings between close relatives for managing existing quantitative genetic variation at the start of the programme. Experiments with the planktonic microcrustacean *Daphnia* spp. have shown that neutral genetic variation gives little indication of the levels of quantitative genetic variation available for selection (Pfrender *et al.,* 2000). However, increasing the population size at the beginning of the breeding programme will diminish the subsequent effect of random genetic drift, and therefore larger founding populations will have an increased likelihood of showing response to selection. Lack of adequate base populations is the main reason for the lack of selection response observed in some species of fish (Gjedrem, 2000).

The effective population size (N_e) required for setting up a breeding programme depends on the policy regarding risk management (Brown, Woolliams and McAndrew, 2005), but to prevent decline in fitness, some authors have recommended N_e values ranging from 31 to 250, which in terms of rates of inbreeding should be less than 2 percent (Meuwissen and Woolliams, 1994). Due to the large family sizes possible for many fish and shellfish species, breeding programmes that fail to control the genetic contributions of parents in every generation are expected to incur relatively high rates of inbreeding (Meuwissen, 1997). The situation is even more extreme when selection is based on a complex breeding objective that includes information from relatives and many traits jointly (Martinez, 2006b).

Fish within commercial production populations generally are not tagged individually and pedigree information is therefore lacking. Genetic markers allow the estimation of pairwise relatedness between individuals or sib-ship reconstruction even with unknown ancestors (Toro and Mäki-Tanila, 1999; Thomas and Hill, 2000; Toro, Barragán and Óvilo, 2002; Wang, 2004; Fernandez and Toro, 2006). There is a plethora of estimators for calculating pairwise relatedness (Queller and Goodnight, 1989; Lynch and Ritland, 1999). The efficiency of inferring pairwise relatedness using markers without parental information is affected by assuming known allele frequencies in the base population and unlinked loci in Hardy-Weinberg equilibrium. Furthermore, pair-wise methods can lead to inconsistent assignations between triplets of individuals because they use information from only two individuals at a time (Fernandez and Toro, 2006). In addition, it is difficult to set thresholds for claiming different types of relatedness in the data (Thomas and Hill, 2000; Norris, Bradley and Cunningham, 2000). On the other hand, sib-ship reconstruction methods do not attempt to calculate co-ancestry; rather, they attempt to reconstruct full- or half-sib or other family groups (Thomas and Hill, 2000; Emery, Boyle and Noble, 2001; Smith, Herbinger and Merry, 2001). Such reconstructions of full- or half-sib families or even other groups of relatives appear robust to lack of knowledge of base population allele frequencies (Thomas and Hill, 2000; Fernandez and Toro, 2006).

Marker information can be used to infer relatedness between individuals available as candidate broodstock to generate the first generation of offspring in the breeding programme, and thereby avoid mating among close relatives. This approach uses molec-

ular information to infer the genealogical pedigree. A simulation was conducted to reconstruct the pedigree of 100 potential candidates from ten full-sib families (with a Poisson family size equal to ten using six equally-frequent microsatellites, without parental genotypes (Martinez, 2006c). The posterior probability of either full [*P*(FS)] or half-sib [*P*(HS)] groups was obtained using the Bayesian model of Emery, Boyle and Noble (2001). In the simulation results, there was a tendency to overestimate relationships, with posterior probabilities over 0.5 when individuals were in fact unrelated. On the other hand, not all true full-sibs were assigned to the correct full-sib family with the greatest probability, and some true full-sib family members were reconstructed as half-sibs. On average (among ten replicates), the probability of mating related individuals was significantly smaller when information from molecular markers was used, compared with what was expected by chance (4.7 percent versus 18.1 percent, $p = 0.002$). The practical implication is that inbreeding in the progeny generation would average 5 percent when random mating is used and 1 percent when optimization using molecular information is used.

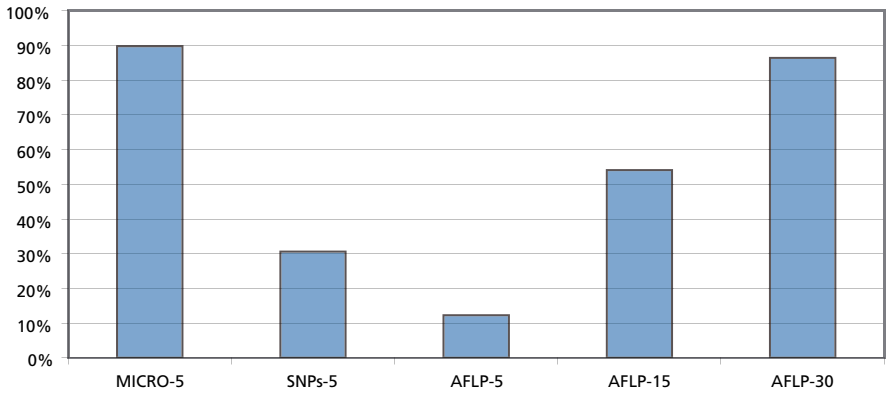
In practice, to perform mating in the base population, the relatedness inferred from molecular information does not need to be perfectly accurate, but it does require that relatedness is not underestimated greatly. Among the technical issues that arise when using marker data are that a pair of individuals could be misclassified as related when they are in fact unrelated (Type I error) or a pair may be wrongly classified as unrelated when the pair is in fact related (Type II error). Type II error is of greatest concern as this could result in related pairs being mated. This is because mating of individuals (males and females) as unrelated when in fact they are full sibs will increase true inbreeding in the population, while misclassification leading to unrelated individuals being assigned to a full-sib family would not increase the inbreeding in the progeny. The presence of mutations, null alleles or genotyping errors will underestimate the true relationships in the population and eventually increase the probability of mating true full-sibs (Butler *et al.,* 2004). Recently, Wang (2004) suggested a method for inferring relationships for marker data with a high error rate and mutation that can be used to address this issue. It should also be noted that studies dealing with estimation of heritability or prediction of breeding values with pedigrees reconstructed using molecular markers may be very inefficient when pedigrees are reconstructed with an increased rate of Type I errors (Mosseau, Ritland and Heath, 1998; Thomas, Pemberton and Hill, 2000).

Detecting self-fertilization in scallops

In scallops, a main drawback when implementing breeding programmes is the occurrence of self-fertilization, even when gametes from later spawning pulses are used for obtaining family material (Martinez and di Giovanni, 2006), i.e. a mixture of selfed and outcrossed individuals can be present even at later stages within a single family. Bias in estimating genetic parameters is expected due to this residual self-fertilization, which can occur with considerable frequency (average 20 percent) within particular families.

A simulation study was used to investigate to what extent markers with different information content can be used to discriminate between selfed and outcrossed individuals within a family (Figure 2). The results showed that microsatellites gave mean values of posterior probabilities greater

FIGURE 2
Identification of selfed individuals within families of scallops using different types of marker data[1]



Vertical bars represent the proportion of the 100 replicates in which the mean posterior probabilities of being selfed (for true selfed individuals) were greater than 0.95.

[1] MICRO-5: five microsatellites with six equally frequent alleles each. SNPs-5: five SNP markers with equal allele frequencies. AFLP-5, -15 or -30: 5, 15 or 30 AFLP markers. The design of the simulations of self-fertilization in scallops: the amount of self-fertilization was modelled using a truncated normal distribution which best fitted the empirical distribution of self-fertilization (Martinez and di Giovanni, 2006). A Bayesian model was used to infer mutually exclusive posterior probabilities of being either selfed or outbred (Anderson and Thompson, 2002). It was assumed that parental information was lacking, with unlinked markers and vague priors. Selfed individuals were regarded as having been detected correctly when the posterior probabilities of being selfed were greater than 0.95 (this criterion was determined empirically for operational reasons).

Source: V. Martinez, in preparation.

than 0.95 in about 90 percent of the families simulated (100 in total). Similar results were obtained with 30 AFLP markers, but these percentages were considerably reduced for smaller numbers of AFLPs or SNPs.

The information from these markers can be used to cull individuals, to construct a relationship matrix in which all unusual relationships are incorporated in analyses used for obtaining unbiased estimates of heritability and genetic correlations, and for estimating breeding values from real data sets (Martinez, 2006a).

IDENTIFYING QTL AND MAJOR GENES INFLUENCING COMPLEX QUANTITATIVE TRAITS

Molecular biology can greatly help the discovery of factors influencing the expression of quantitative traits. There are a number of ways in which this information can be used, the difference between them being the level of resolution with which these factors can be mapped. For example, loci with major effects on quantitative traits (QTL) are mapped by using markers to track inheritance of chromosomal regions in families or in inbred line crosses using the extent of linkage disequilibrium generated in the population. This approach gives a limited amount of mapping resolution. Fine mapping requires information from additional markers and individuals sampled across the outbred population and, while helping to narrow the confidence interval of the position of the QTL, this is only the starting point for identifying the polymorphisms in the performance-determining genes themselves. In practice, identification of genes influencing specific traits is achieved using

a combination of genetic mapping (linkage and fine mapping) to localize the QTL to a small region on the chromosome under analysis, and candidate gene or positional cloning approaches to identify the genes within the QTL region.

In some cases, sufficient biochemical or physiological information is available to investigate the association between the quantitative expression and the level of marker polymorphisms within specific genes. Nevertheless, this approach requires a great amount of detailed information in order to choose which gene explains the greatest effect and to have sufficient power to detect the association. This information is starting to appear in the aquaculture literature from multinational projects such as the Consortium of Genomic Resources for All Salmonids Project (cGRASP) (Ng *et al.,* 2005).

QTL mapping in fish using linkage disequilibrium: theoretical and practical considerations

Value of chromosomal manipulations

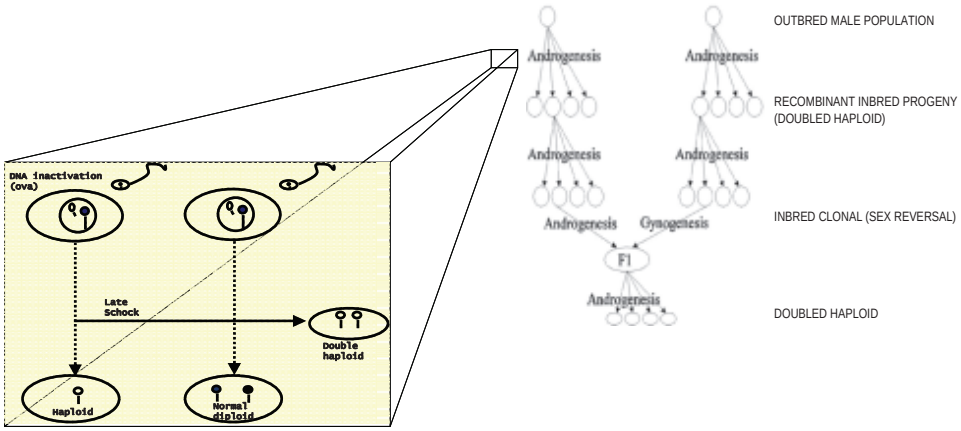
The great reproductive flexibility of fish enables different breeding designs to be implemented relatively easily. Completely homozygous fish can be produced in only one generation using chromosome set manipulations, without the many generations of inbreeding needed in other vertebrates. These manipulations enable doubling of the chromosomal complement of a haploid gamete (Young *et al.,* 1996; Corley-Smith, Lim and Bradhorst, 1996). Androgenetic double haploid individuals can be obtained by fertilizing eggs that were inactivated with gamma radiation, yielding haploid embryos containing only paternal chromosomes. Alternatively, gynogenetic double haploid individuals can be obtained by activating the development of eggs with ultraviolet-inactivated sperm, yielding haploid embryos containing only maternal chromosomes. In each case, diploidy is restored using methods that suppress the first mitotic division (Figure 3; Streisinger *et al.,* 1980; Corley-Smith, Lim and Bradhorst, 1996; Bijma, van Arendonk and Bovenhuis, 1997; Young *et al.*, 1998). The use of these reproductive manipulations to provide experimental populations for genetic analysis of complex quantitative traits has been well described (Bongers *et al.,* 1997; Robison, Wheeler and Thorgaard, 2001; Tanck *et al.,* 2001).

Double haploids from inbred line crosses

After a second round of uniparental reproduction (Figure 3), a collection of clonal lines can be obtained that collectively is likely to represent all the genetic variants from the base population (Bongers *et al.*, 1997). Crosses of sex-reversed double haploid individuals from lines that diverge for the traits of interest can produce F_1 lines in complete linkage disequilibrium. These F_1 populations can be used for further experimentation based on F_2 or backcross designs. Another round of androgenesis of F_1 individuals will produce a population of fully homozygous individuals. This design will have twice the power for detecting QTL as the standard F_2 design (Martinez, 2003). The standard deviation of QTL position estimates is halved for the double haploid design. This is due to an increase in the additive genetic variance, which is doubled for the double haploid design due to redistribution of the genotype frequencies in the progeny generation (Falconer and Mackay, 1996).

Informative double haploid populations of this sort have been utilized to perform QTL analysis for embryonic development rate in rainbow trout (Robison, Wheeler and Sundin, 2001; Martinez *et al.,* 2002;

FIGURE 3
Chromosomal manipulations in fish



ANDROGENESIS

In androgenesis, fertilization is carried out using radiation-inactivated ova, and a late shock is used to suppress first mitosis and thereby restore diploidy.
Adapted from Thorgaard and Allen, 1987.

Martinez *et al.*, 2005). At least four QTL of relatively large effect explain about 40 percent of the phenotypic variance of the mapping population and most of the 2.5 standard deviations of the difference between the original clonal lines used to generate the F_1 population (Robison, Wheeler and Thorgaard, 1999). Two linked QTL were in repulsion phase in the F_1 population, and were undetected in the analysis using composite interval mapping. This result was not surprising as evidence was accumulated among replicates of lines that were incubated at different temperatures (Robison, Wheeler and Sundin, 2001), and the Bayesian multiple QTL method incorporated all the available information of environmental co-variates in the analysis (Martinez *et al.*, 2005). Recently, these double haploid lines have been used for mapping QTL related to the number of pyloric caeca (Zimmerman *et al.*, 2005) and for confirming QTL influencing development rate (Sundin *et al.*, 2005).

When traits are associated and by taking into account the correlated structure of the data, multivariate estimation of QTL effects is expected to be more powerful than single trait analysis (Jiang and Zeng, 1995). Also, from a genetic standpoint, joint analysis provides the means for testing different hypotheses about the mode by which genes explained the genetic co-variation (Wu *et al.*, 1999). For example, after hypothesis testing (following Knott and Haley, 2000), a single pleiotropic QTL with opposite effects for development rate and length best explained the multivariate data (as detailed earlier by Martinez *et al.*, 2002b). This finding was also consistent with the negative correlation estimated with the data (Martinez *et al.*, 2002).

Double haploids in outbred populations

Martinez, Hill and Knott (2002) derived analytical formulae to predict the power of linkage analysis for interval mapping under three different mating designs in outbred

populations: full-sib mating, hierarchical mating, or double haploid designs. This analysis suggested that the use of double haploids appeared to be of benefit when detecting QTL, particularly when both the variance of the QTL and of the polygenic effects was small. Furthermore, given the relatively large size of full-sib families in fish, there appeared to be little advantage of hierarchical mating over full-sib mating designs for detecting QTL, the optimum family size depending on the size of the QTL and the population structure used for mapping (Martinez, Hill and Knott, 2002). The gain in power of the double haploid design comes from the increase in the variance of the Mendelian sampling term within families, which is effectively doubled for traits that are explained by additive effects (Falconer and Mackay, 1996).

As experimental settings constrain the total number of individuals genotyped, designs aimed at QTL mapping should include a small number of families of relatively large size in order to maximize the likelihood of detecting the QTL. This is because most of the information for mapping QTL uses linkage information that comes from within-family segregation (Muranty, 1996; Xu and Gessler, 1998). However, increasing power comes at the expense of reducing the accuracy of estimating the additive genetic variance for polygenic effects. A QTL mapping method has been developed for double haploids, which efficiently accommodates all the uncertainties that pertain to outbred populations, such as unknown linkage phases and differing levels of marker informativeness, using the identical-by-descent variance component method (see below; Martinez, 2003). Also, it is possible to combine double haploids and outbred relatives in the same family. Simulations of differing amounts of marker information and heritability for the QTL were used to compare the empirical power of the double haploid and full-sib designs. While the power of the full-sib design was lower than that for double haploids, QTL position estimates for double haploids had large confidence intervals (about 30 cM as compared with 40 cM for full-sibs; Martinez, 2003).

The double haploid design was used for mapping QTL for stress response in common carp using single marker analysis (Tanck *et al.*, 2001). The authors found only suggestive evidence for QTL, which is not surprising due to limited genome coverage for markers used in the analysis.

Published results have shown that double haploid lines are a useful resource for QTL detection studies. However, double haploid lines are difficult to develop due to the expression of deleterious recessive alleles (McCune *et al.*, 2002) and the low survival following shocks applied to restore diploidy to the haploid embryo. As the rate of male recombination is depressed, the precision of mapping QTL in androgenetic families is lower than that obtained using recombination events from females. Another practical matter is the labour needed for developing a clonal line, as at least two generations are required (Figure 3). This delay can be quite expensive and time-consuming for species with a long generation interval, such as salmon or trout (two to four years).

Aspects of QTL mapping in outbred populations of fish

Inbred line crosses are ideal for mapping QTL because they are expected to be completely informative for both markers and QTL, providing that the inbred lines are fixed for alternative alleles. Outbred populations are not completely informative for both QTL and markers; thus, experimental

power is expected to be lower than that for crosses between clonal lines. The power for detecting the QTL depends on allele frequencies, the probability of sampling an informative parent and family size.

Factors influencing the power of detecting QTL

Due to the large family sizes that can be obtained in many fish species, different mating designs using full-sib groups can be carried out for outbred populations. For example, full factorial designs may be used in which many males and females are mated to one another, and hierarchical designs may be applied in which each male is mated with multiple females, or each female with multiple males. For a given size of experiment, factorial and hierarchical designs have potentially a lower probability of sampling a heterozygous parent (because fewer sires and or dams are sampled overall), compared with the full-sib design in which each family has potentially two informative parents. For this reason, factorial and hierarchical designs can potentially give lower power compared with the simple full-sib design (Muranty, 1996; Martinez, Hill and Knott, 2002).

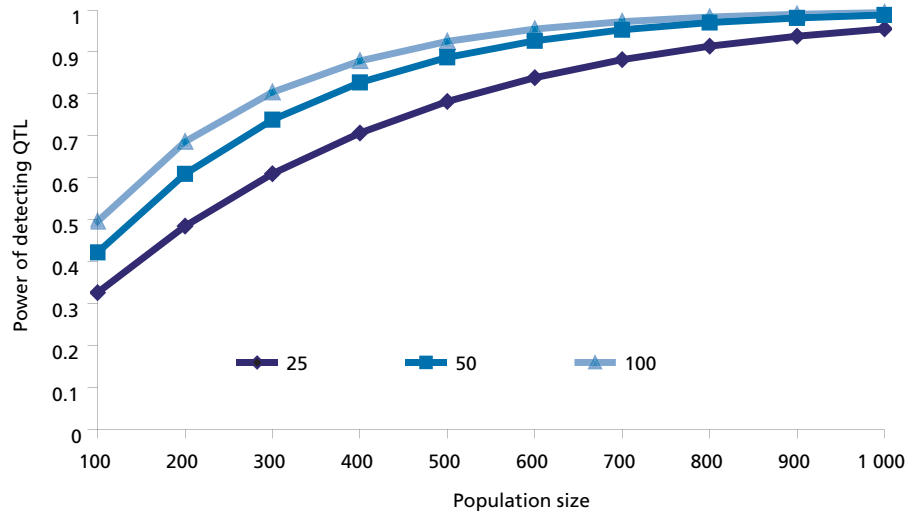
The optimum number of full-sib families sampled in the QTL mapping population depends on the intrinsic power of the experiment (i.e. size of the QTL effect and size of the population). As expected, large family sizes are needed for detecting QTL of small effects (Martinez, Hill and Knott, 2002). When the QTL explains 10 percent of the phenotypic variance, the optimum family size appears to be 50 individuals per family for a reasonably-sized QTL mapping experiment in outbred populations (Figure 4). Further increases in the number of individuals per family provide only a modest increase in power. Further, the same results used simulation models showing dominance and additive effects under the variance components method for mapping QTL (Martinez *et al.,* 2006a).

Methods of analysis

The method of choice when analysing data from outbred populations is the variance component method, in which QTL effects are included as random effects with a covariance proportional to the probability that relatives (e.g. full-sibs) share alleles identical by descent conditional on marker data (Xu and Atchley, 1995). This model is similar to the one used more generally for genetic evaluation of candidate fish for selection, but includes the random QTL effect.

A considerable proportion of the genetic variance for growth-related traits in fish populations has been explained by dominance (Rye and Mao, 1998; Pante, Gjerde and McMillan, 2001; Pante *et al.,* 2002). When mapping QTL using the random model, it is assumed that only additive effects are of importance and therefore only matrices of additive relationships conditional on marker data are fitted in the residual effect maximum likelihood procedure (George, Visscher and Haley, 2000; Pong-Wong *et al.,* 2001). However, the large family sizes in fish enable hypotheses for different modes of inheritance at the QTL to be tested using the within-family variance. While some authors have speculated that including dominance in the model will increase the power of detecting QTL (Liu, Jansen and Lin, 2002), others (Martinez, 2003; Martinez, 2006a) have shown that power to detect QTL was comparable between models including or not including dominance. This was particularly the case for the larger family sizes simulated and it was concluded that for most scenarios, the additive model was quite

FIGURE 4
Deterministic power calculation (following Martinez, Hill and Knott, 2002) for a QTL explaining 10 percent of the phenotypic variance for a variable total population size (x axis) and family size (25, 50 and 100)



robust for detecting QTL and there was little loss of information for detecting QTL when dominance is present but not used in the QTL mapping analysis.

QTL mapping in practice

To date, QTL mapping in fish using outbred populations has been carried out mostly with single marker analysis (microsatellites and AFLP markers), and using relatively sparse linkage maps when interval mapping is used. In tilapia, the F_2 design and a four-way cross between different species of *Oreochromis* have been used for detecting QTL affecting cold tolerance and body weight (Cnaani *et al.*, 2003; Moen *et al.*, 2004c). In outbred populations of salmonids, QTL that influence body weight have been mapped (Reid *et al.*, 2005 and references therein).

Studies seeking linkage of markers to traits amenable to MAS, such as disease resistance, have begun to appear in the literature over the past few years. For example, QTL for resistance have been mapped for infectious pancreatic necrosis virus (Ozaki *et al.*, 2001), infectious salmonid anemia (Moen *et al.*, 2004c), infectious haematopoietic necrosis (Rodriguez *et al.*, 2004; Khoo *et al.*, 2004), and stress and immune response (Cnaani *et al.*, 2004). Also, Somorjai, Danzmann and Ferguson (2003 and references therein) reported evidence of QTL for upper thermal tolerance in salmonids with differing effects in different species and genetic backgrounds.

From fine mapping to finding genes influencing complex traits

When the number of meioses in the genotyped pedigree is not sufficient for the linkage analysis to obtain a precise position for the QTL, there is a wide confidence interval around an estimated QTL position.

Fine mapping methods attempt to overcome this problem by quantifying the gametic phase or linkage disequilibrium (LD) present in an outbred population, i.e. across families. This method makes use of the number of generations as the appearance of a mutation and can produce extremely precise estimates of the QTL position (Pérez-Enciso *et al.,* 2003). The rationale behind using LD for mapping QTL is that when the population size is rather small, founders of the population would have only a limited number of haplotypes, and with very tightly linked loci there may not be sufficient time for recombination to break up the association between markers and the mutation affecting the quantitative trait.

LD mapping is carried out by calculating the probabilities that haplotypes shared by individuals are identical by descent from a common ancestor conditional on marker data (assuming t generations as the common ancestor and a certain N_e; Meuwissen and Goddard, 2001). The LD in the population depends on a number of population parameters such as the degree of admixture or stratification in the population and the actual level of association between the causal mutation and the polymorphisms. The correct determination of phases and of genotypes at the QTL is required for fine mapping purposes (Meuwissen and Goddard, 2001; Pérez-Enciso, 2003). For these reasons, a pure LD analysis is likely to result in a large number of false positives, i.e. falsely inferring association when there is no linkage.

Methods that incorporate the linkage information (within families) and LD jointly are preferred, because the likelihood of spurious association (i.e. LD without linkage) diminishes, making much better use of the whole data set (Meuwissen and Goddard, 2001, 2004; Pérez-Enciso, 2004). All of these methods, however, require a great deal of genotyping of tightly linked markers such as SNPs, which currently are not widely available for fine mapping in aquaculture species.

Using fine mapping techniques, the confidence interval for QTL position can be reduced considerably. However, to develop a direct test for a favourable polymorphism requires use of comparative mapping approaches with model species, such as zebrafish or *fugu*, to select the candidate genes that most likely affect the trait of interest. Otherwise, enrichment of markers in a specific region of the genome (to narrow further the most likely position of the polymorphism) following sequencing is needed to compare sequences between individuals that show different phenotypes or alternative QTL alleles.

Candidate gene analysis

It is tempting to invoke variation at genes with a known role in the physiology underlying a complex trait such as growth to explain phenotypic variability for the trait. These genes can be searched for polymorphisms (e.g. SNPs) and the variants then tested to determine whether they are correlated with the expression of the quantitative trait. This approach requires knowledge of the biology of the species, biochemical pathways and gene sequences in order to target variation at those specific genes. In aquaculture, most of this information is currently lacking, but it is expected that more genes will be incorporated in databases in the near future. The possibility exists to utilize data from highly studied model species, such as zebrafish or rainbow trout, in comparative bioinformatic approaches.

To date, this strategy has not proven particularly successful for explaining genetic

variation underlying complex (polygenic) traits. This is because although the biology of the trait and the genes most likely involved in the expression of the phenotype may be known, in complex traits many other genes may be involved in the metabolic pathway that are not obvious candidates. For example, in aquaculture species, candidate genes have been studied for growth-related traits using ten conserved gene sequences known to be related to the growth hormone axis (Tao and Boulding, 2003). In this study of Arctic charr, only a single SNP (of ten) from five of ten genes was found to be associated with growth rate.

Another example for disease resistance traits is the major histocompatibility complex (MHC). The genes of this complex encode highly polymorphic cell surface glycoproteins involved in specific immune responses and either specific alleles or heterozygotes at this complex were associated with resistance and susceptibility to *A. salmonicida* or infectious haematopoietic necrosis (IHN) virus (Langefors, Lohm and Grahn, 2001; Lohm *et al.,* 2002; Arkush *et al.,* 2002; Grimholt *et al.,* 2003; Bernatchez and Landry, 2003). Nevertheless, the background genome was quite important for explaining the difference in resistance between individuals within a family (Kjøglum, Grimholt and Larsen, 2005).

Microarrays, gene expression and identification of candidate genes for QTL analysis

Microarray technology (Knudsen, 2002) enables the expression of thousands of genes to be studied simultaneously. Until now, this information has been used primarily for following gene expression in treatment and control experiments in many fields such as disease exposure and stress response. This information can be used to discover new sets of candidate genes, possibly with or without functional assignment that may be related to the quantitative trait of interest (Walsh and Henderson, 2004). Genes whose expression differs between treatments are likely to be *trans*-acting genes, i.e. their expression is regulated by other genes. Therefore, it seems likely that seeking polymorphisms within these genes may not yield information about factors that explain the phenotype, and there might be problems assigning the correct significance threshold (Pérez-Enciso *et al.,* 2003). Further, because many genes are part of metabolic pathways and do not act individually, the expression of a single gene may be insufficient to explain phenotypic differences between individuals. Only those genes that directly affect phenotypic expression (i.e. *cis*-acting genes) can be treated as candidate genes for subsequent use in MAS after studying polymorphisms in their sequences. In salmonids, a microarray made available from the Consortium for Genomics Research on all Salmonids Project (cGRASP) has been used to study gene expression in fish exposed or not exposed to *Piscicricketsia salmonis* (Rise *et al.,* 2004b), and microarrays in other fish and shellfish species are currently under development.

A gene expression pattern can itself be regarded as a quantitative trait. Here, the interest is in finding associations between different patterns of gene expression and marker loci. This analysis was coined as "genetical genomics" by Jansen and Nap (2001). As is usual in QTL mapping, the analysis attempted to dissect the transcriptional regulation of the entire transcriptome and to identify the effects of individual QTL affecting gene expression (the so-called eQTL; e.g. Hubner *et al.,* 2005). To

date, this analysis relies upon the use of segregating populations (of known origin) such as recombinant inbred lines (Carlborg *et al.,* 2005), and the analysis of outbred populations poses greater challenges (Pérez-Enciso, 2004). Still, aquaculture species can provide sufficient information due to the large family sizes needed to unravel complex regulatory gene networks. How all this information can be included in MAS programmes is yet unclear.

INCORPORATING MOLECULAR MARKERS INTO BREEDING PROGRAMMES FOR FISH AND SHELLFISH

General aspects of incorporating molecular information in breeding programmes

The response to selection ΔG is estimated as:

$$\Delta G = i\sigma_H r$$

where i = the intensity of selection, r = the correlation between the breeding objective and the selection criteria (i.e. accuracy), and σ_H = the additive genetic standard deviation for the breeding objective. As the major impact of incorporating information from molecular markers will be on accuracy estimates, improvement of the response to selection will be higher for traits that have relatively small accuracy than for traits of relatively large accuracy. Thus, breeding programmes for traits with low heritability and relatively few records per trait measured such as carcass and disease resistance are those most benefiting from incorporating marker information (Meuwissen, 2003).

The relative increase in accuracy depends on the amount of variation explained by markers, which in turn depends on the number of QTL identified and used in MAS or GAS schemes (Lande and Thompson, 1990). QTL experiments in other species have shown that the effects of marked genes have a leptokurtic distribution, with a small number of genes having large effects and polygenes (Hayes and Goddard, 2001), which is likely to be the case in aquaculture species (Martinez *et al.,* 2005). Hence, it is expected that more than a single marked gene will be needed for MAS schemes to be efficient.

Due to the biology of many fish and shellfish species, multistage selection will likely prove useful in MAS or GAS schemes. Basically, a first stage of selection can be applied for traits expressed early in the life cycle (e.g. body weight), and a second stage of selection will incorporate information from relatives plus marked QTL. Optimization will be needed to determine the intensity of selection that should be applied at each stage to maximize profit while reducing the cost and labour of keeping individuals until later stages (Martinez *et al.,* 2006b).

Health and carcass traits are difficult to select for in fish and shellfish because phenotypic records are obtained from relatives and not from candidates for selection (Gjoen and Bentsen, 1997). Sib or pedigree evaluation has many disadvantages in relation to the amount of genetic progress that can be realized within a selection programme using only pedigree information to predict breeding values using an animal model. First, selection accuracy using sib information is lower than when predicting breeding values based on an individual's own information (Falconer and Mackay, 1996). Second, there is no variation of estimated breeding value for polygenic effects. Thus, variation of Mendelian sampling effects within a family cannot be used and consequently there may be a limited scope for constraining rates of inbreeding

to acceptable levels when the number of families is relatively low.

To date, little has been published regarding the economic profits arising from the extra genetic gain obtained by MAS or GAS schemes in aquaculture or terrestrial species. Information of this nature is essential because the additional gains are dependent on the magnitude of the allelic effects and thus the marginal increase should offset the costs of applying the technology. This trade-off may be more important when a single marked QTL, rather than multiple marked QTL (and multiple traits), is targeted by selection.

Pleiotropic effects can be important if the polymorphisms under MAS or GAS also have negative effects on fitness or other traits of economic importance. For example, negative genetic correlations have been found for resistance to viral and bacterial diseases (Gjøen *et al.,* 1997; Henryon *et al.,* 2002, 2005), which may be a problem in practical breeding when the goal is to select fish resistant to a range of pathogens. For example, in natural and selected populations, MHC polymorphism is likely to be maintained by frequency-dependent selection (Langefors, Lohm and Grahn, 2001; Lohm *et al.,* 2002; Bernatchez and Landry, 2003), suggesting that selection favours rare alleles, but works against the same alleles at high frequency. Therefore, it seems likely that a MAS scheme using MHC information or QTL in LD with disease resistance should focus on maintaining polymorphism rather than on selecting for a particular combination of alleles.

MAS in populations in linkage equilibrium

When populations are in LE between markers and QTL, the information used for selection purposes is given by the Mendelian co-segregation of markers and QTL within each of the full-sib families in the population under selection. In practical terms, this means that co-ancestry conditional on marker information needs to be computed within a family for a given segment in the genome. In effect, the segregation of regions that individuals share as identical-by-descent ("more" or "less" than average) is being traced and, under such circumstances, the accuracy of predicting breeding values using marker information is mainly dependent on the proportion of the within-family variance due to the QTL (Ollivier, 1998).

The effect of family size on the relative accuracy of predicting breeding values (comparing MAS and BLUP) using marker information was studied in detail using simulations (Table 2; V. Martinez, unpublished data). Compared with the GAS schemes presented below, for LE-MAS to be efficient, large full-sib families are required for predicting breeding values for the QTL accurately. This is because breeding value prediction is carried out on a within-family basis; thus, large families are required to obtain breeding values for predicting QTL effects with reasonable accuracy. When individuals do not have records for the quantitative trait, the extra accuracy of MAS was highest for the largest family size simulated (50 individuals, 25 with records and 25 without records; the difference is equal to 7 percent). The accuracy of predicting breeding values was very similar in BLUP or MAS for individuals that have records for the trait in most of the scenarios simulated, suggesting that MAS is expected to be of little use under these circumstances (Villanueva, Pong-Wong and Woolliams, 2002).

The advantage of MAS will come both from increased accuracy and from

TABLE 2
Empirical correlation between predicted breeding values using molecular* and pedigree information (M+BLUP) or pedigree information (BLUP) and true breeding values

Scenario	Individuals with records	Family size (number of families)					
		10 (100)		20 (50)		50 (20)	
		M+ BLUP	BLUP	M+ BLUP	BLUP	M+ BLUP	BLUP
I	NO	0.47	0.45	0.55	0.52	0.64	0.57
	YES	0.60	0.60	0.65	0.64	0.70	0.65
II	NO	0.41	0.41	0.49	0.47	0.56	0.52
	YES	0.58	0.58	0.62	0.61	0.64	0.63

* Molecular information comprises a completely informative marker bracket of 10 cM around a QTL and all individuals genotyped for the markers. The matrix of identity-by-descent values was calculated using the deterministic method of Martinez (2003). The estimated values of h^2 using residual effect maximum likelihood for the polygenic and QTL effects were, on average, 0.13 and 0.09, respectively. The results are presented for different nuclear family sizes (number of families, between parentheses) and for candidates with or without phenotypic records. The population size was equal to 1 000, where 50 percent (Scenario I) or 25 percent (Scenario II) of the individuals within each full-sib family had records for the trait.
Source: Martinez, unpublished data.

increasing the realized selection intensity in sustainable breeding schemes with restricted rates of inbreeding. In sib-testing schemes, candidates without records can only be selected randomly within families because an estimate of the Mendelian sampling terms cannot be obtained. Markers provide an estimate of the QTL effects that segregate within a family, and therefore the realized selection differential (at the same rates of inbreeding) is expected to be greater than that obtained using standard sib/family testing.

All the benefits outlined above come at an expense. MAS using LD within families requires a great deal of genotyping and recording of phenotypes on relatives, due to the fact that the linkage phase between markers and QTL needs to be re-estimated in each generation. This is because LD between markers and the QTL is established only within families in each generation and not across the population. For this reason, it is not possible to predict breeding values for the QTL using molecular marker data without records when exploiting information from a single generation. Therefore, pre-selection using this approach is more difficult to apply in practice. This means that for disease resistance or carcass quality traits, challenge (measurement) will have to be carried out at every generation, in all the families available within the programme, as is always the case for conventional breeding programmes.

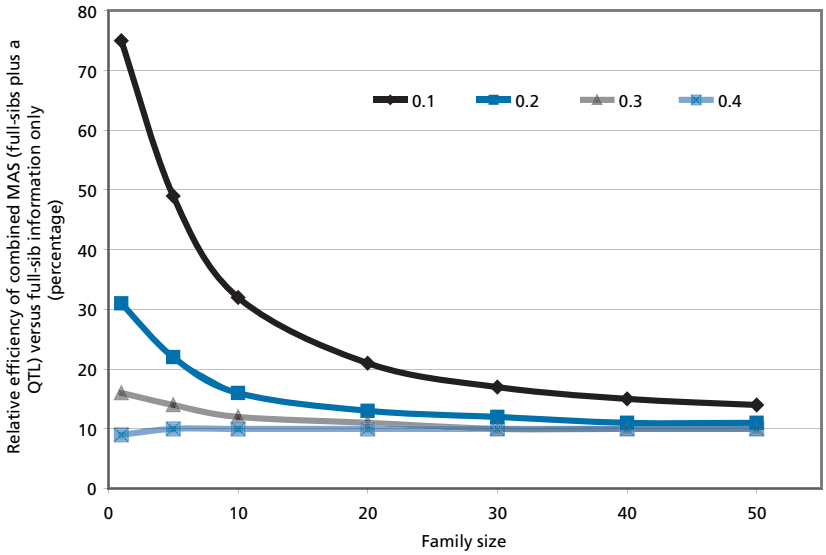
Due to the low resolution when mapping the QTL, it is likely that inaccurate estimates of position will lead to over-optimistic estimates of rates of genetic gain. In the simulations, it was assumed that the QTL position was known within the interval and the markers surrounding the QTL were completely informative. Thus, the increase in accuracy presented in Table 2 represents the upper bounds of accuracy estimates.

Utilizing direct test of genes in GAS schemes

The mean phenotype of the population for a quantitative trait can be modified by increasing the frequency of favourable alleles of genes influencing the trait. In the literature, greater genetic gain has been predicted for GAS schemes than for MAS schemes (using LE populations) at

FIGURE 5

Relative efficiency of combined GAS (for different family sizes [full-sibs]) and a known QTL, explaining 10 percent of the genetic variance) versus an index using information from full-sibs only for different values of the overall heritability (h^2)



The data were obtained from the ratio of square root of the variance of the indices. Selection index formulae were derived from Lande and Thompson (1990).

the same rate of inbreeding (Pong-Wong *et al.*, 2002). This is because the accuracy of predicting QTL effects using markers is always smaller than when the QTL effects are known, as in GAS schemes. In reality, it is likely that MAS will be carried out using information from many markers to predict the allelic effects of more than one QTL simultaneously whereas, in GAS schemes, only a limited number of polymorphisms are likely to be available. Therefore, on the whole, MAS schemes may yield greater genetic response because a greater proportion of the genetic variation is marked and used. Still, more marker genotyping is required for MAS schemes, which means that the additional proportion of the variance typed should pay for the increase in the cost of many markers typed simultaneously.

Due to the biology of many species in aquaculture, large family sizes can be used in a breeding programme. Following the deterministic model of Lande and Thompson (1990), Figure 5 describes the effect of family size and amount of polygenic variation on the relative efficiency of accuracy estimates for an index using different numbers of full-sibs measured for the trait, versus an index also including information on candidates for selection genotyped at loci targeted for GAS schemes (V. Martinez, unpublished results). For a single QTL explaining 10 percent of the genetic variance, when the heritability is relatively large, family size has a small impact on the accuracy. On the other hand, when the heritability of the trait is small, selection for a known QTL has a major impact on relative efficiency, particularly when the family size

is relatively small. Hence, this approach can be important for traits that are expensive or difficult to measure such as carcass quality, disease resistance or antibody response.

Given the research efforts carried out at diverse laboratories worldwide, it is likely that direct tests will be available in the near future for GAS schemes for different traits. With an increasing amount of data on ESTs, together with a greater understanding of the function of known genes in aquaculture species and new gene discovery, there is a possibility of more rapidly identifying and subsequently using polymorphisms that are within coding regions. However, the research effort required to develop tests for polymorphisms explaining allelic effects cannot be underestimated, and the factors influencing the profitability of GAS will include:

- the amount of variation explained by the test and the number of tests (genes) available for explaining the phenotype;
- the frequency of the favourable allele in, and the presence of the direct test (e.g. SNPs), for the target population;
- the interaction between the polymorphism and the background genome and possible pleiotropic effects on fitness;
- the trade-off between the marginal return given by the additional genetic gain obtained through the non-linear changes in the allele frequency of the favourable allele until fixation;
- fixed costs of implementing genotyping and patenting.

MAS in populations in LD

Using information from dense marker maps, it is possible to make use of LD between the markers and the beneficial mutations influencing the quantitative traits across the population. Under this scenario, there are two possible ways to use the LD in MAS programmes i.e. using information on a single haplotype effect in LD with the beneficial polymorphism across the population, or predicting the total genetic value using genome-wide, dense marker maps (genome-wide marker-assisted selection, or G-MAS) (Lande and Thompson, 1990; Meuwissen, Hayes and Goddard, 2001).

The effectiveness of each scenario is largely dependent on the actual magnitude of the effects associated with the polymorphism, either across the whole genome or at specific genes. It is likely that, in the near future, high-throughput SNP technology will make dense marker maps cost effective for selective breeding purposes in aquaculture. Thus, it can be expected that LD-MAS will be implemented over the whole genome, basically using markers to unravel the genetic architecture of quantitative traits. Information from multiple traits jointly and for multiple genes (and their interactions within and between loci) will be used, rather than first relying on mapping QTL in experimental populations and then implementing this information in MAS programmes. A profit analysis including multiple traits (e.g. to study undesirable pleiotropic effects on the breeding goal) will be needed on a case-by-case basis to determine whether the use of a single or multiple haplotypes simultaneously is most profitable and which method of LD-MAS better suits the population under selection.

Specific genes are not being evaluated when LD is used across the population; rather, haplotype effects on the phenotype are being estimated. As this is done on a single generation across the whole genome, it would be possible to use these haplotype effects for selecting candidates some generations after the initial estimation without relying on phenotypes (Meuwissen, Hayes and Goddard, 2001). Recombination will

erode the initial LD and therefore it is expected that accuracy of estimating the breeding value of many haplotypes will decay (Zhang and Smith, 1992), the extent of the erosion being dependent on several population parameters (Meuwissen, Hayes and Goddard, 2001). In practice, the response to selection obtained needs to be verified in each generation; thus, re-estimation can be used based on a random sample of individuals from the population.

One possible caveat is that by assuming a certain mode of gene action (i.e. only additive effects), there may in fact be a more complicated genetic architecture influencing quantitative traits. For example, when estimating dominance and epistasis with the same data, more haplotype effects need to be estimated. Therefore, it is likely that the accuracy of individual effects will decrease. Another potential complication that arises when the true model involves non-additive effects is that assignment of potential mates needs to be optimized to increase the mean phenotype of the population simultaneously through heterosis arising from combination of different QTL alleles. In the long term, the frequency of homozygotes that are identical-by-descent will increase within the population as a whole; consequently, methods are required to constrain the rates of inbreeding to obtain similar changes of the population mean across generations. Furthermore, expression of different combinations of alleles after selection will require re-estimation of between-haplotype effects in each generation.

CONCLUSION

QTL mapping and MAS are not as well advanced in aquaculture species as in terrestrial plants and animals. However, the merger between genetics and genomics is expected to be a fertile area of research in the coming years due to the plethora of information that is currently being gathered by many laboratories around the world. It is through these research efforts that variations affecting complex traits in fish and shellfish species may be detected and used for increasing the usefulness of MAS schemes. In the final analysis, however, all these techniques must be cost-effective if they are to be profitable in actual breeding programmes.

ACKNOWLEDGEMENTS

I am indebted for comments from Mr Eric Hallerman who greatly helped clarify the manuscript. This chapter has been funded partially by: INNOVA-Corporación de Fomento de la Producción (CORFO) (05CT6 PP-10) on "Using functional genomics for understanding disease resistance in salmon" from the Government of Chile; Vicerrectoria de Investigacion y Creación, Universidad de Chile hatch project 03/03312; and Fondo de Cienciias y Tecnologia (FONDECYT) 1061190.

REFERENCES

Anderson, E.C. & Thompson, E.A. 2002. A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* 160: 1217–1229.

Arkush, K.D., Giese, A.R., Mendonca, H.L., McBride, A.M., Marty, G.D. & Hedrick, P.W. 2002. Resistance to three pathogens in the endangered winter-run Chinook salmon (*Oncorhynchus tshawytscha*): effects of inbreeding and major histocompatibility complex genotypes. *Can. J. Fisheries and Aquatic Sci.* 59: 966–975.

Bernatchez, L. & Landry, C. 2003. MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *J. Evolut. Biol*. 16: 363–377.

Bijma, P., van Arendonk, J.A.M. & Bovenhuis, J.H. 1997. Breeding value and variance component estimation from data containing inbred individuals: application to gynogenetic families in common carp.(*Cyprinus carpio* L.). *Genet.*145: 1243–1249.

Bongers, A.B., Bovenhuis, J.H., van Stokkom, A.C., Wiegertjes, G.F., Zandieh-Doulabi, B., Komen, J. & Ritcher, C.J.J. 1997. Distribution of genetic variance in gynogenetic or androgenetic families. *Aquaculture* 153: 225–238.

Brown, C.R., Woolliams, J.A. & McAndrew, B.J. 2005. Factors influencing effective population size in commercial populations of gilthead seabream, *Sparus aurata*. *Aquaculture* 247: 219–225.

Butler, K., Field, C., Herbinger, C.M. & Smith, B.R. 2004. Accuracy, efficiency and robustness of four algorithms allowing full sibship reconstruction from DNA marker data. *Mol. Ecol*. 13: 1589–1600.

Carlborg, Ö., De Koning, D.J., Manly, K.F., Chesler, E., Williams, R.W. & Haley, C.S. 2005. Methodological aspects of the genetic dissection of gene expression. *Bioinformatics* 21: 2383–2393.

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

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

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 F_2 tilapia hybrid. *Aquaculture* 223: 117–128.

Cnaani, A., Zilberman, N., Tinman, S., Hulata, G. & Ron, M. 2004 Genome-scan analysis for quantitative trait loci in an F_2 tilapia hybrid. *Mol. Genet. Genom*. 272: 62–172.

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

Corley-Smith, G.E., Lim, C.H.J. & Brandhorst, B.P. 1996. Production of androgenetic zebrafish (*Danio rerio*). *Genetics* 142: 1265–1276.

Coulibaly, I., Gharbi, K., Danzmann, R., Yao, J. & Rexroad, C.E. 2005. Characterization and comparison of microsatellites derived from repeat-enriched libraries and expressed sequence tags. *Anim. Genet*. 36: 309–315.

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. Livestock Prodn.*, Vol. 19, pp. 364–371. Guelph, Ontario.

Emery, W., Boyle, C. & Noble, F. 2001. Assignment of paternity groups without access to parental genotypes: multiple mating and development plasticity in squid. *Molec. Ecol*. 10: 1265–1278.

Falconer, D.S. & Mackay, T.F.C. 1996. *Introduction to quantitative genetics*, *4th edition.* Harrow, Essex, U.K, Longman.

Fernandez, J. & Toro. M. 2006. A new method to estimate relatedness using molecular markers. *Molec. Ecol*. (in press).

Garcia de Leon, F.J., Canonne, M., Quillet, E., Bonhomme, F. & Chatain, B. 1998. The application of microsatellite markers to breeding programs in the sea bass, *Dicentrarchus labrax*. *Aquaculture* 159: 303–316.

George, A., Visscher, P. & Haley, C.S. 2000. Mapping quantitative trait loci in complex pedigrees: a two step variance component approach. *Genetics* 156: 2081–2092.

Gilbey, J., Verspoor, E., McLay, A. & Houlihan, D. 2004. A microsatellite linkage map for Atlantic salmon (*Salmo salar*). *Anim. Genet*. 5: 98–105.

Gjedrem, T. 2000. Genetic improvement of cold-water fish species. *Aquaculture Res*. 31: 25–33.

Gjerde, B. 1986. Growth and reproduction in fish and shellfish. *Aquaculture* 57: 37–55.

Gjerde, B. & Schaeffer, L.R. 1989. Body traits in rainbow trout: II. Estimates of heritabilities and of phenotypic and genetic correlations. *Aquaculture* 80: 25–44.

Gjerde, B., Villanueva, B. & Bentsen, H.B. 2002. Opportunities and challenges in designing sustainable fish breeding programs. *Proc. 7th Wrld. Congr. Genetics Appl. Livestock Prodn.,* CD-ROM 06-01. Montpellier, France.

Gjøen, H. & Bentsen, H. 1997. Past present and future of genetic improvement in salmon aquaculture. *ICES J. Marine Sci*. 54: 1009–1014.

Gjøen, H.M., Refstie, T., Ulla, O. & Gjerde, B. 1997. Genetic correlations between survival of Atlantic salmon in challenge and field tests. *Aquaculture* 158: 277–288.

Griffing, B. 1967. Selection in reference to biological groups. I. Individual and group selection applied to populations of unordered groups. *Austr. J. Biol. Sci*. 20: 127–139.

Grimholt, U., Larsen, S., Nordmo, R., Midtlyng, P., Kjoeglum, S., Storset, A., Saebo, S. & Stet, R.J. 2003. MHC polymorphism and disease resistance in Atlantic salmon (*Salmo salar*); facing pathogens with single expressed major histocompatibility class I and class II loci. *Immunogen*. 55: 210–219.

Hayes, B. & Goddard, M.E. 2001. The distribution of the effects of genes affecting quantitative traita in livestock. *Genet. Sel. Evol*. 33: 209–229.

Hayes, B., Lærdahl, J., Lien, S., Adzhubei, A. & Høyheim, B. 2004. Large scale discovery of single nucleotide polymorphism (SNP) markers in Atlantic salmon (*Salmo salar*). AKVAFORSK, Institute of Aquaculture Research (available at www.akvaforsk.no/norsk/HentFil.asp?fid=41).

Henryon, M., Jokumsen, A., Berg, P., Lund, I., Pedersen, P.B., Olesen, N.J. & Slierendrecht, W.J. 2002. Genetic variation for growth rate, feed conversion efficiency and disease resistance exists within a farmed population of rainbow trout. *Aquaculture* 209: 59–76.

Henryon, M., Berg, P., Olesen, N.J., Kjaer, T.E., Slierendrecht, W.J., Jokumsen, A. & Lund. I. 2005. Selective breeding provides an approach to increase resistance of rainbow trout (*Onchorhynchus mykiss*) to the diseases, enteric redmouth disease, rainbow trout fry syndrome and viral haemorrhagic septicaemia. *Aquaculture* 250: 621–636.

Hershberger, W.K., Myers, J.M., Iwamoto, R.N., Mcauley, W.C. & Saxton, A.M. 1990. Genetic changes in the growth of coho salmon (*Oncorhynchus kisutch*) in marine net-pens, produced by ten years of selection. *Aquaculture* 85: 187–197.

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 of offspring growth and survival in communally reared rainbow trout. *Aquaculture* 137: 245–256.

Herbinger, C.M., O'Reilly, P.T., Doyle, R.W., Wright, J.M. & O'Flynn, F. 1999. Early growth performance of Atlantic salmon full-sib families reared in single family tanks versus in mixed family tanks. *Aquaculture* 173: 105–116

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

Hubner, N., Wallace, C.A., Zimdahl, H., Petretto, E., Schulz, H., Maciver, F., Mueller, M., Hummel, O., Monti, J., Zidek, V., Musilova, A., Kren, V., Causton, H., Game, L., Born, G., Schmidt. S., Muller, A., Cook, S.A., Kurtz, T.W., Whittaker, J., Pravenec, M. & Aitman, T.J. 2005. Integrated transcriptional profiling and linkage analysis for identification of genes underlying disease. *Nature Genet*. 37: 243–253.

Hulata, G. 2001. Genetic manipulations in aquaculture: a review of stock improvement by classical and modern technologies. *Genetica* 111: 155–173.

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

Jansen, R.C., Geerlings, H., Van Oeveren, A.J. & Van Schaik R.C. 2001. A comment on codominant scoring of AFLP markers. *Genetics* 158: 925–926.

Jansen, R.C. & Nap, J.P. 2001. Genetical genomics: the added value from segregation. *Trends Genet*. 17: 388–391.

Jiang, C. & Zeng, Z.B. 1995. Multiple trait analysis of genetic mapping for quantitative trait loci. *Genetics* 140: 1111–1127.

Kause, A., Ritola, O., Paananen, T., Ritola, O., Paananen, T., Mäntysaari, E. & Eskelinen, U. 2002. Coupling body weight and its composition: a quantitative genetic analysis in rainbow trout. *Aquaculture* 211: 65–79.

Kause, A., Ritola, O., Paananen, T., Wahlroos, H. & Mantysaari, E.A. 2005. Genetic trends in growth, sexual maturity and skeletal deformations and rate of inbreeding in a breeding programme for rainbow trout (*Oncorhynchus mykiss*). *Aquaculture* 247: 177–187.

Khoo, S.K., Ozaki, A., Nakamura, F., Arakawa, T., Ishimoto, S., Nickolov, R., Sakamoto, T., Akutsu, T., Mochizuki, M., Denda, I. & Okamoto, N. 2004. Identification of a novel chromosomal region associated with infectious hematopoietic necrosis (IHN) resistance in rainbow trout. *Fish Path*. 39: 95–102.

Kjøglum, S., Grimholt, U. & Larsen S. 2005. Non-MHC genetic and tank effects influence disease challenge tests in Atlantic salmon (*Salmo salar*). *Aquaculture* 250: 102–10.

Knott, S.A. & Haley, C.S. 2000. Multitrait least squares for quantitative trait loci detection. *Genetics* 156: 899–911.

Knudsen, S. 2002. *A biologist's guide to analysis of DNA microarray data*. New York, NY, USA, John Wiley and Sons.

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.

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

Langefors, J., Lohm, M. & Grahn, O. 2001. Association between major histocompatibility complex class IIB alleles and resistance to *Aeromonas salmonicida* in Atlantic salmon. *Proc. R. Soc. Lond. B* 268: 479–485.

Lee, B.Y., Lee, W.J., Streelman, J.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. & Lehnert, S. 2003. Genetic mapping of the kuruma prawn *Penaeus japonicus* using AFLP markers. *Aquaculture* 219: 143–156.

Liu, Y., Jansen, G.B. & Lin, C.Y. 2002. The covariance between relatives conditional on genetic markers. *Genet. Select. Evol*. 34: 657–667.

Liu, Z.J. & Cordes, J.F. 2004. DNA marker technologies and their applications in aquaculture genetics. *Aquaculture* 238: 1–37.

Lohm, J., Grahn, M., Langefors, A., Andersen, O., Storset, A. & von Schantz, T. 2002. Experimental evidence for major histocompatibility complex allele-specific resistance to a bacterial infection. *Proc.R. Soc. Lond. B* 269: 2029–2033.

Lynch, M. & Ritland, K. 1999. Estimation of pairwise relatedness with molecular markers. *Genetics* 152: 1753–1766.

Marth, G.T., Korf, I., Yandell, M.D., Yeh, R.T., Gu, Z., Zakeri, H., Stitziel, N.O., Hillier, L., Kwok, P.Y. & Gish, W.R. 1999. A general approach to single-nucleotide polymorphism discovery. *Nature Genet*. 23: 452–456.

Martinez, V. 2003. *Quantitative genetic analysis of fish populations using molecular markers*. University of Edinburgh, UK (Ph.D. thesis).

Martinez, V. 2006a. Using marker data in conventional breeding programmes: a case of study in selective breeding programs of scallops. In *Proc. Internat. Assocn. Genet. Aquaculture*, p 66. Montpellier, France.

Martinez, V. 2006b. Further insights of the variance component method to detect QTL in fish populations: relaxing the assumption of additive genetic effects. In *Proc. Internat. Assocn. Genet. Aquaculture,* p. 66. Montpellier, France.

Martinez, V. 2006c. Importance and implementation of molecular technologies in selective breeding programs for aquaculture species. In *Proc. 8th Wrld. Congr. Genet. Appl. Livest. Prodn*., Communication 09-01. Minas Gerais, Brazil.

Martinez, V. & di Giovanni. S. 2006. Breeding programmes of scallops: effect of self-fertilization when estimating genetic parameters. In *Proc. Internat. Assocn. Genet. Aquaculture*, p.65. Montpellier, France.

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

Martinez, V., Neira, R. & Gall, G.A.E. 1999. Estimation of genetic parameters from pedigreed populations: lessons from analysis of alevin weight in coho salmon (*O. kisutch*). *Aquaculture* 330: 22–30.

Martinez, V., Sillanpaa, M., Thorgaard, G., Robison, B., Woolliams, J. & Knott, S. 2002. Evidence of a pleiotropic QTL influencing components of early development in double haploid lines of rainbow trout. *Proc. 7th Wrld. Congr. Genet. Appl. Livestock Prodn*., CD ROM Communication 06-08. Montpellier, France.

Martinez, V., Thorgaard, G., Robison, B. & Sillanpää, M. 2005. Simultaneous consideration of environmental and genetic effects in QTL mapping: an application to early development in double haploid lines of rainbow trout. *Genet. Res*. 86: 209–221.

Martinez, V., Kause, A., Mäntysaari, E. & Mäki-Tanila, A. 2006a. The use of alternative breeding schemes to enhance genetic improvement in rainbow trout: I. One-stage selection. *Aquaculture* 254: 182–194.

Martinez, V., Kause, A., Mäntysaari, E. & Mäki-Tanila, A. 2006b. The use of alternative breeding schemes to enhance genetic improvement in rainbow trout: II. Two-stage selection. *Aquaculture* 254: 195–202.

McCune, A.R., Fuller, R.C., Aquilina, A.A., Dawley, R.M., Fadool, J.M., Houle, D., Travis, J. & Kondrashov, A.S. 2002. A low genomic number of recessive lethals in natural populations of bluefin killifish and zebrafish. *Science* 296: 2398–2401.

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

Meuwissen, T.H.E. 2003. Genomic selection the future of marker-assisted selection and animal breeding. In *Marker-assisted selection: a fast track to increase genetic gain in plant and animal breeding?* (available at www.fao.org/biotech/torino.htm).

Meuwissen, T.H.E. & Goddard, M.E. 2001. Prediction of identity by descent probabilities from marker-haplotypes. *Genet. Select. Evol*. 33: 605–634.

Meuwissen, T.H.E. & Goddard, M.E. 2004. Mapping multiple QTL using linkage disequilibrium and linkage analysis information and multitrait data. *Genet. Select. Evol.* 36: 261–279.

Meuwissen, T.H.E. & Woolliams, J.A. 1994. Effective sizes of livestock populations to prevent a decline in fitness. *Theor. Appl. Genet*. 89: 1019–1026.

Meuwissen, T.H.E., Hayes, B. & Goddard, M. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genet*. 157: 1819–1829.

Moav, R. & Wohlfarth, G. 1976. Two-way selection for growth rate in the common carp (*Cyprinus carpio L.*). *Genet.* 82: 83–101.

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

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

Morton, R. & Howarth, J.M. 2005. Estimation of breeding values using selected pedigree records. *Biomet.* 61: 422–432.

Mousseau, T.A., Ritland, K. & Heath, D.D. 1998. A novel method for estimating heritability using molecular markers. *Hered*. 80: 218–224.

Muir, W.M. 2005. Incorporation of competitive effects in forest tree or animal breeding programs. *Genet*. 170: 1247–1259.

Muranty, H. 1996. Power of tests for quantitative trait loci detection using full-sib families in different schemes. *Hered*. 76: 156–165.

Ng, S.-H., Chang, S., Brown, A., Koop, B. & Davidson, W.S. 2005. Type I microsatellite markers from Atlantic salmon (*Salmo salar*) expressed sequence tags. *Mol. Ecol. Notes* 5: 762–766.

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 genetic 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'Brien, S.J. 1991. Molecular genome mapping: lessons and prospects. *Curr. Opin. Genet. Devel*. 1: 105–111.

Ollivier, L. 1998. The accuracy of marker-assisted selection for quantitative traits within populations in linkage equilibrium. *Genet.* 148: 1367–1372.

Ozaki, A., Sakamoto, T., Khoo, S., Nakamura, K., Coimbra, M.R., 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.

Panitz, F., Holm, L.-E., Høyheim, B., Hagen-Larsen, H., Taggart, J., Cairney, M., Davey, G.C., Powell, R. & Bendicen, C. 2002. EST sequencing in Atlantic salmon. *Proc. 7th Wrld. Congr. Genet. Appl. Livestock Prodn*. CD-ROM Communication No 06-14. Montpellier, France.

Pante, M.J.R, Gjerde, B. & McMillan, I. 2001. Effect of inbreeding on body weight at harvest in rainbow trout, *Oncorhynchus mykiss*. *Aquaculture* 192: 201–211.

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

Paterson, S., Piertney, S.B. & Knox, D. 2004. Characterization and PCR multiplexing of novel highly variable tetranucleotide Atlantic salmon (*Salmo salar L.*) microsatellites. *Mol. Ecol. Notes* 4: 160–162.

Pérez-Enciso, M. 2003. Fine mapping of complex trait genes combining pedigree and linkage disequilibrium information: a Bayesian unified framework. *Genet*. 163: 1497–1510.

Pérez-Enciso, M. 2004. In silico study of transcriptome genetic variation in outbred populations. *Genet.* 166: 547–554.

Pérez-Enciso, M., Toro, M.A., Tenenhaus, M. & Gianola. D. 2003. Combining gene expression and molecular marker information for mapping complex trait genes: a simulation study. *Genet*. 164: 1597–1606.

Pfrender, M.E., Spitze, K., Hicks, J., Morgan, K.L. & Lynch, L.M. 2000. Lack of concordance between genetic diversity estimates at the molecular and quantitative-trait levels. *Conserv. Genet*. 1: 263–269.

Piepho, H-P. & Koch, G. 2000. Codominant analysis of banding data from a dominant marker system by normal mixtures. *Genet*. 155: 1459–1468.

Pong-Wong, R., Villanueva, B. & Woolliams, J.A. 2002. Comparison of direct and marker-assisted selection with optimised contributions. In *Proc. 7th Wrld. Congr. Genet. Appl. Livestock Prodn*. CD-ROM Communication No. 33-02. Montpellier, France.

Pong-Wong, R., George, A.W., Woolliams, J.A. & Haley C.S. 2001 A simple and rapid method for calculating identity-by-descent matrices using multiple markers. *Genet. Select. Evol.* 33: 453–471.

Ponzoni, R.W., Hamzah, A., Tan, S. & Kamaruzzaman, N. 2005. Genetic parameters and response to selection for live weight in the GIFT strain of Nile tilapia (*Oreochromis niloticus*). *Aquaculture* 247: 203–210.

Queller, D.C. & Goodnight, K.F. 1989. Estimating relatedness using genetic markers. *Evol.* 43: 258–275.

Quinton, C., McMillan, I. & Glebe, B.D. 2005. Development of an Atlantic salmon (*Salmo salar*) genetic improvement program: Genetic parameters of harvest body weight and carcass quality traits estimated with animal models. *Aquaculture* 247: 211–217.

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*). *Hered.* 94: 166–172.

Rexroad, C.E., Rodriguez, M.F., Gharbi, K., Danzmann, R.G., DeKoning, J., Phillips, R. & Palti, Y. 2005. Comparative mapping of expressed sequence tags containing microsatellites in rainbow trout (*Oncorhynchus mykiss*). *BMC Genom*. 6: 54.

Rise, M.L., Von Schalburg, K.R., Brown, G.D., Mawer, M.A., Devlin, R.H., Kuipers, N., Busby, M., Beetz- Sargent, M., Alberto, R., Gibbs, A.R., Hunt, P., Shukin, R., Zeznik, J.A., Nelson, C., Jones, S.R., Smailus, D.E., Jones, S.J., Schein, J.E., Marra, M.A., Butterfield, Y.S., Stott, J.M., Ng, S.H., Davidson, W.S. & Koop, B.F. 2004a. Development and application of a salmonid EST database and cDNA microarray:data mining and interspecific hybridization characteristics. *Genome Res*. 14: 478–490.

Rise, M.L., Jones, S.R., Brown G.D., von Schalburg, K.R., Davidson, W.S. & Koop, B.F. 2004b. Microarray analyses identify molecular biomarkers of Atlantic salmon macrophage and hematopoietic kidney response to *Piscirickettsia salmonis* infection. *Physiol. Genom*. 20: 21–35.

Robison, B., Wheeler, P. & Sundin, K. 2001. Composite interval mapping reveals a QTL of major effect on embryonic development rate in rainbow trout (*Onchorhynchus mykiss*). *J. Hered*. 92: 16–22.

Robison, B., Wheeler, P. & Thorgaard, G. 1999. Variation in development rate among clonal lines of rainbow trout (*Oncorhynchus mykiss*). *Aquaculture* 173: 131–141.

Rodriguez, F.M., LaPatra, S., Williams, S., Famula, T. & May, B. 2004. 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* L.). *Livest. Prodn. Sci*. 57: 23–32.

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. *Genet*. 155: 1331–1345.

Serapion, J., Waldbieser, G.C., Wolters, W. & Wolters, Z.J. 2004. Development of Type I markers in channel catfish through intron sequencing. *Anim. Genet.* 35: 463–466.

Singer, A., Perlman, H., Yan, Y., Walker, C., Corley-Smith, G., Brandhorst, B. & Postlethwait, J. 2002. Sex-specific recombination rates in zebrafish (*Danio rerio*). *Genet*. 160: 649–657.

Smith, B.R., Herbinger, C.M. & Merry, H.R. 2001. Accurate partition of individuals into full-sib families from genetic data without parental information. *Genet*. 158: 1329–1338.

Somorjai, I.M.L., 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.

Streisinger, G.,Walker, C.H., Dower, N., Knauber, D. & Singer, F. 1980. Production of clones of homozygous diploid zebra fish (*Brachydanio rerio*). *Nature* 291: 293–296.

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 a development rate QTL in backcrosses of clonal rainbow trout, *Oncorhynchus mykiss*. *Aquaculture* 247: 75–83.

Tanck, M.W.T., Palstra, A.P., van de Weerd, M., Leffering, C.P., van der Poel, J.J., Bovenhuis, H. & Komen, J. 2001. Segregation of microsatellite alleles and residual heterozygosity at single loci in homozygous androgenetic common carp (*Cyprinus carpio* L.). *Genome* 44: 743–751.

Tao, W.J. & Boulding, E.G. 2003. Associations between single nucleotide polymorphisms in candidate genes and growth rate in Arctic charr (*Salvelinus alpinus* L.). *Hered.* 91: 60–69.

Taris, N., Baron, S. & Sharbel, T.F. 2005. A combined microsatellite multiplexing and boiling DNA extraction method for high-throughput parentage analyses in the Pacific oyster (*Crassostrea gigas*). *Aquaculture Res*. 36: 516–518.

Thomas, S. C., Pemberton, J.M. & Hill, W.G. 2000. Estimating variance components in natural populations using inferred relationships. *Hered*. 84: 427–436.

Thomas, S.C. & Hill, W.G. 2000. Estimating quantitative genetic parameters using sib-ships reconstructed from marker data. *Genet*. 155: 1961–1972.

Thorgaard, G.H. & Allen S.K. 1987. Chromosome manipulation and markers in fishery management. *In* N. Ryman & F.M. Utter, eds. *Population genetics and fishery management*, pp. 319–331. Seattle, WA, USA, University of Washington Press.

Toro, M. & Mäki-Tanila, A. 1999. Establishing a conservation scheme. *In* J.K. Oldenbroek, ed. *Gene banks and the conservation of farm animal genetic resources,* pp.75–90. Lelystad, Netherlands, DLO Institute of Animal Health.

Toro, M., Barragán, C. & Óvilo, C. 2002. Estimation of coancestry in Iberian pigs using molecular markers. *Conserv. Genet*. 3: 309–320.

Toro, J.E., Sanhueza, M.A., Winter, J.E., Aguila, P. & Vergara, A.M. 1995 Selection response and heritability estimates for growth in the Chilean oyster *Ostrea chilensis* (Philippi, 1845). *J. Shellfish Res*. 14: 87–92.

Vasemägi A., Nilsson, J. & Primmer, C.R. 2005. Expressed sequence tag (EST)-linked microsatellites as a source of gene associated polymorphisms for detecting signatures of divergent selection in Atlantic salmon (*Salmo salar* L.). *Mol. Biol. & Evol*. 22: 1067–1076.

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.

Villanueva, B., Pong-Wong, R. & Woolliams, J.A. 2002. Marker assisted selection with optimised contributions of the candidates to selection. *Genet. Select. & Evol*. 34: 679–703.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van De Lee, T., Hornes, M., Friters, 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*. *Genet.* 158: 727–734.

Walsh, B. & Henderson, D.J. 2004. Microarrays and beyond: What potential do current and future genomics tools have for breeders? *Anim. Sci*. 82: E292–E299.

Wang, J. 2004. Sibship reconstruction from genetic data with typing errors. *Genet.* 166: 1963–1979.

Wang, L., Song, L., Chang, Y. & Xu, W. 2005. A preliminary genetic map of Zhikong scallop (*Chlamys farreri,* Jones et Preston 1904). *Aquaculture Res*. 36: 643–653.

Ward, R D., English, L. & McGoldrick, J. 2000. Genetic improvement of the Pacific oyster, *Crassostrea gigas* (Thunberg) in Australia. *Aquaculture Res.* 31: 35–44.

Withler, R. & Beacham, T. 1984. Genetic consequences of the simultaneous or sequential addition of semen from multiple males during hatchery spawning of chinook salmon (*Oncorhynchus tshawytscha*). *Aquaculture* 126: 11–20.

Wright, J.M. & Bentzen, P. 1994. Microsatellites – genetic markers for the future. *Rev. Fish Biol. & Fisheries* 4: 384–388.

Wu, W.-R., Li, W.-M., Tang, D.-Z., Lu, H.-R. & Worland, A.J. 1999. Time-related mapping of quantitative trait loci underlying tiller number in rice. *Genet.* 151: 297–303.

Xu, S. & Atchley, W. 1995. A random model approach to interval mapping of quantitative trait loci. *Genet.* 141: 1189–1197.

Xu, S. & Gessler, D. 1998. Multipoint genetic mapping of quantitative trait loci using a variable number of sibs per family. *Genet. Res*. 71: 73–83.

Young, W.P., Wheeler, P.A., Fields, R.D. & Thorgaard, G.H. 1996. DNA fingerprinting confirms isogenicity of androgenetically-derived rainbow trout lines. *J. Hered.* 87: 77–81.

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

Yue, G.H., Ho., M.Y. & Orban, L. 2004. Microsatellites within genes and ESTs of common carp and their applicability in silver crucian carp. *Aquaculture* 234: 85–98.

Zhang, W. & Smith, C. 1992. Computer simulation of marker assisted selection utilizing linkage disequilibrium. *Theor. Appl. Genet*. 83: 813–820.

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. *Aquaculture* 247: 85–95.

SECTION VI

Selected issues relevant to applications of marker-assisted selection in developing countries

CHAPTER 18

Marker-assisted selection in crop and livestock improvement: how to strengthen national research capacity and international partnerships

Maurício Antônio Lopes

SUMMARY

It is generally recognized that marker-assisted selection (MAS) is a tool that breeders can use to accelerate the speed and precision of crop and livestock breeding in developing countries. However, its practical application has been more difficult than previously expected. Although advances in molecular marker technology have uncovered many possibilities for transferring genes into desired crops and livestock through MAS, more methodological development and better planning and implementation strategies will be needed for its successful and expeditious application to breeding programmes. Also, this technology should not be regarded as an end in itself, but as an interacting part of complex strategies and decision-making processes. An appropriate mix of technologies and capabilities together with effective approaches to networking must be viewed as key ingredients for its correct development and application to breeding programmes. This chapter describes some strategies to guide decisions about structures, methods and capacities that may contribute to enhancing the access and successful use of MAS in developing countries.

INTRODUCTION

The tremendous advances made in molecular marker techniques in the past two decades have led to increased understanding of the genetic basis of many agricultural traits in a variety of plant and animal species. The use of these techniques has also made it possible to accelerate the transfer of desirable traits among varieties and to introgress novel genes from related wild species.

DNA markers have many advantages over conventional approaches available to breeders. They are especially advantageous for traits that are otherwise difficult to tag, such as resistance to pathogens, insects and nematodes, tolerance to abiotic stresses and quality parameters. They offer great scope for improving the efficiency of conventional breeding by carrying out selection not directly on the trait of interest but on linked genomic regions. Additionally these markers are unaffected by environmental conditions and are detectable during all stages of growth (Mohan *et al*., 1997).

Molecular marker techniques have therefore moved beyond their early projected role as tools for identifying chromosomal segments and genes to uncovering many possibilities for easing the transfer of genes into desired cultivars and lines. MAS generated great enthusiasm as it was seen as a major breakthrough, promising to overcome many limitations of conventional breeding processes (FAO, 2003). However, despite advances in the theory of MAS, direct utilization of the information it provides for selecting superior individuals with complex traits is still very limited (Young, 1999; Ferreira, 2003). Nevertheless, there is still optimism about the contributions of MAS, which is now balanced by the realization that genetic improvement of quantitative traits using this tool may be more difficult than previously considered (FAO, 2003). In 1999, Young reviewed the development of MAS, analysing in detail its main drawbacks, many of which remain today. He concluded that because MAS technology was so challenging it should not be a reason for discouragement but, instead, reason for more ingenuity and better planning and execution.

Recent developments in high-throughput genotyping, single nucleotide polymorphism (SNP) and the integration of genomic technologies are advances that will play an important role in the development of MAS as an effective tool for sustainable conservation and increased use of crop genetic resources (Ferreira, 2006). However, research teams, funding agencies, commodity groups and the private sector will need to work together to develop MAS technology further and ensure that breeders have the best available tools. Also, the tools and strategies will need to go beyond markers themselves to include genome-based knowledge derived from model systems, high-throughput cost effective technology, as well as better technologies and strategies for handling large volumes of information.

The purpose of this chapter is to discuss the access to and utilization of MAS technology by breeding programmes, with special emphasis on strategies to help strengthen research capacity and partnerships in developing countries. Whenever possible, recommendations are presented to help guide decisions that may contribute to enhancing the access and successful use of MAS by national programmes.

PERCEPTIONS ABOUT THE USE OF MAS IN CROP AND LIVESTOCK IMPROVEMENT

As MAS is still an evolving technology, there are not many detailed studies available describing the state-of-the-art of its

application to breeding programmes. Also, there are very few prospective studies indicating future trends in the application of this technology. The FAO Biotechnology Forum hosted an e-mail conference on "Molecular marker-assisted selection as a potential tool for genetic improvement of crops, forest trees, livestock and fish in developing countries". This provided a comprehensive overview of the perceptions of scientists from different parts of the world about key aspects of the application of MAS to genetic improvement in developing countries (www.fao.org/biotech/logs/c10logs.htm).

As described in Chapter 21, this FAO conference was very inclusive, with a total of 627 people subscribing. Eight percent of these (52 people) submitted 85 messages, which were received from all major regions of the world, including Asia (33 percent), Europe (26 percent), Latin America and the Caribbean (14 percent), Africa (9 percent) Oceania (9 percent) and North America (8 percent). People from 26 different countries participated, with a total of 50 messages (59 percent) from developing countries and 35 messages (41 percent) from developed countries. Institutional representation was also ample, including national research institutes, centres belonging to the Consultative Group on International Agricultural Research (CGIAR), universities, consultants, farmer organizations, government agencies, non-governmental organizations (NGOs), etc. Although only 52 people out of 627 subscribers participated directly in the conference, the number is significant considering the broad representation, the high level of the (moderated) discussions and the number of relevant issues discussed (www.fao.org/biotech/logs/c10logs.htm).

To prepare this chapter, a detailed review was carried out of the conference results in an attempt to capture the main perceptions and concerns related to access to and utilization of MAS in developing countries. This analysis revealed a variety of ideas and creative suggestions to overcome the problems of MAS utilization. Although there is a risk of narrowing views on important issues discussed during the conference, four major perceptions were clear from the rich content of the discussions:

Perception 1. There is a need for development of priority-setting mechanisms and cost benefit analysis to guide informed decisions on how best to apply MAS and other technological innovations to crop and livestock breeding in developing countries.

Perception 2. MAS has to be understood as part of a complex process. Complementarities, mix of technologies, integration of capabilities and networking must always be viewed as key ingredients for its correct application in breeding programmes.

Perception 3. There is a need for an objective definition of public-private functions and responsibilities in relation to funding and development of technological innovation in developing countries. Public-private and north-south partnerships are essential to accelerate progress and effective application of MAS and other innovations to breeding programmes in developing countries.

Perception 4. Developing countries must focus on capacity building and human resource development oriented to shape effective strategies of technological innovation.

In the following sections, possible strategies and alternatives to deal with the challenges and opportunities indicated above are outlined, including the need

for objective priority-setting, development of partnerships, complementarities and capacity building for compatible human resource formation.

MAS AS PART OF A COMPLEX PROCESS – SETTING PRIORITIES AND TAKING ACTION

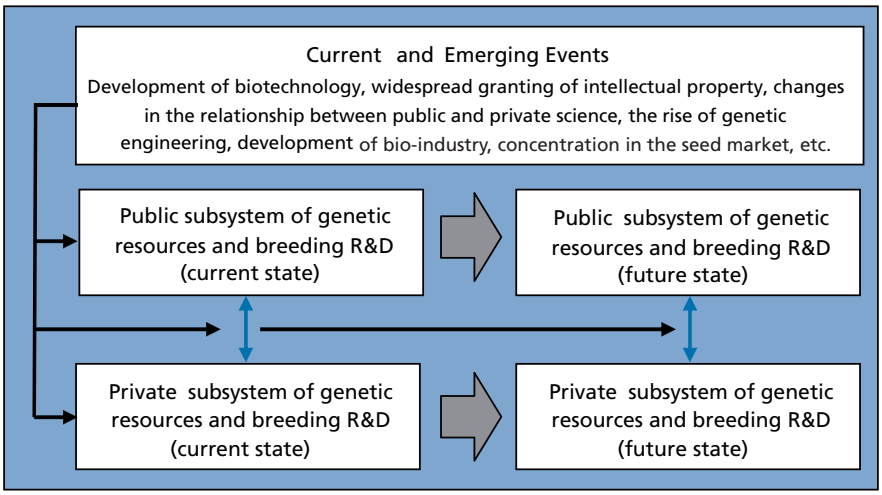
Before discussion of MAS as a technological alternative to increase the capacity of breeding programmes it is important to discuss and consider the future of the breeding process itself. Until recently, selection was based on observable phenotypes, without knowledge of the genetic architecture of the selected characteristics (Dekkers and Hospital, 2002). However, advances in molecular marker techniques and rapid advances in large-scale sequencing are creating new perspectives for exploiting the immense reservoir of polymorphism in genomes. Molecular genetic analysis of traits in plant and animal populations is leading to a better understanding of quantitative trait genetics. More recently, the discovery and scoring of single nucleotide polymorphisms (SNPs) using automated and high-throughput instrumentation are already providing the increased resolution needed to analyse sets of genes involved in complex quantitative traits (Altshuler *et al*., 2000; De La Vega *et al.*, 2002; Rafalsky, 2002, Lörz and Wenzel, 2005; Ferreira, 2006).

What impacts will all these developments have on breeding programmes? As anticipated by Stuber, Polacco and Senior, in 1999, "when genomics is added to future strategies for plant and animal breeders, the projected outcomes are mind-boggling. There is every reason to believe that the synergy of empirical breeding, MAS and genomics will truly produce a greater effect than the sum of the various individual actions." Despite the positive view of many who find technological development an open venue for enhancement or complete redesign of traditional breeding, there are many uncertainties about its future. The rise of genetic engineering and the bio-industry, and the widespread granting of intellectual property rights, followed by profound changes in the relationship between public and private science make it very difficult to anticipate future developments in both publicly funded breeding research and the commercial biotechnology industry.

Unfortunately, very little effort has been directed to thinking about the future of breeding, especially in developing countries (Castro *et al*., 2002, 2006). Many past and current events are changing the performance, the relationships and the space that public and private research organizations have in the market, raising the need for a deeper understanding of their unfolding impacts on the public activity of research (Price, 1999; Graff *et al*., 2003). The current scenario of changes and uncertainties has generated the necessity for strategic re-alignment of public research in many parts of the world. Therefore, research organizations need information that is not currently available about such changes and influences and their impact on the future of key activities, such as crop and livestock breeding. To obtain and to organize this information, prospective studies need to be developed on the present and future performance of breeding programmes and their related production systems.

The future configuration of breeding programmes depends on knowledge to guide strategic decisions about structures, methods and capacities in order to take advantage of new opportunities and technological niches. Foresight methodologies have been applied to this end, using systemic analysis of the past and

FIGURE 1
Conceptual framework of a prospective study on genetic resources and breeding R&D in Brazil



Dark arrows indicate the impact of current and emerging events on both public and private subsystems of R&D in genetic resources and breeding, at present and in the future, considering several alternative scenarios. Vertical arrows indicate the state of the relationship between the public and the private R&D subsystems as it is affected by current and emerging events.
Source: Castro *et al.,* 2002, 2006.

present performance of a research field, determining critical factors of performance (Linstone and Turoff, 1975; Castro, de Cobbe and Goedert, 1995; Castro, de Lima and Freitas Filho, 1998; Castro *et al*. 2002, 2006; Lima *et al*., 2000).

An innovative model of a prospective study was proposed and tested by Castro *et al*. (2002, 2006), based on the Brazilian national system of genetic resources and breeding. The effort started with the distinction between two component subsystems – public and private. The authors considered that the two subsystems admit two possible states or situations, current and future, after the effect of current and emerging events (Figure 1). Prospective efforts based on this framework can be very useful to guide diagnosis of national programmes, identifying the main determinants of current and past system performance that can be used to guide decisions about the configuration of genetic resources, breeding programmes and the associated seed industry.

This type of study can help identify changes in the system and in the corresponding technology market, analysing their current and future impacts, determining future opportunities and threats to the strategic positioning of research organizations in the technology market. There is also the perspective of developing possible alternative scenarios for the relationships between public and private research, and of these with the market, to guide the strategic positioning of public research. Results of this effort could indicate new opportunities and niches for public breeding programmes, as well as areas of extreme value where the public sector would have to acquire capacity in the future. Key decisions on investments in new technologies and processes applied

to genetic resources and breeding research, such as MAS, genomic tools, transgenic technology and others, are better taken if these results are available.

The results of this forward-looking approach developed in Brazil allowed the identification of some important trends that must be considered by managers in the process of adapting breeding efforts for the future (Castro *et al.*, 2002, 2005, 2006). Current and emerging events identified in the process will certainly affect the performance, methods, technological processes, portfolio of products and institutional relations in the public and private R&D sectors dedicated to plant breeding in Brazil. This complexity indicates that it is quite dangerous for developing countries, pressured by market evolution and rapid expansion of methods and technologies, to face the challenge of identifying priority areas for investment without a minimum prospective effort.

In summary, the ability to predict changes that might affect the performance of public and private R&D organizations is essential for decision-makers and managers to guide adjustments in the focus of these sectors in a timely manner, avoiding threats and promoting access to new tools and opportunities. Although the same prospective methodology may be applied to a wide range of countries, it is important to point out that situations differ drastically from country to country, thereby requiring examination of future configuration of a sector on a case-by-case basis.

MAS AS PART OF A COMPLEX PROCESS – BUILDING CAPACITIES, COMPLEMENTARITIES AND ENHANCING NETWORKING

MAS cannot be considered an end in itself or a tool detached from the complexities of breeding strategies. It has to be understood and analysed in the context of an interacting mix of tools and strategies that have to be targeted towards crop and livestock improvement in a coordinated manner. Independently of the outcome of any priority-setting effort, the need for an expanded networking approach to breeding and biotechnological research will always be an objective to be pursued. This need arises because networking and partnerships are essential to enable organizations to attain otherwise unattainable goals, add value to their products and processes and reduce costs. Also, the continuous demand for efficiency and relevance presses R&D programmes to move in the direction of cooperation and alignment of efforts.

One of the key problems limiting the use of MAS and other advanced technologies in developing countries is exactly the difficulty of building effective teams and networks. Unfortunately, very few developing countries have trained scientists and advanced facilities concentrated in one place or institution. Usually, these scarce resources are scattered over different places and institutions, and many times away or disconnected from the relevant breeding programmes. This is a serious drawback as the increasing interdependence of traditional and upstream disciplines makes it necessary to build and manage multidisciplinary teams consisting of breeders, agronomists, molecular biologists, biochemists, pathologists, entomologists, physiologists, soil scientists, statisticians, etc. – a goal always difficult to achieve. In addition to the challenge of working within team alignments and cooperation, there is the pressing need to develop ways to share capacities, infrastructure, materials and information among research teams located across a country, a region, or even continents.

The main problem in fostering collaboration and effective cooperation to achieve common goals seems to be the difficulty of recognizing that different teams and organizations have different general interests and norms. For this reason, competition usually prevails. While it has been well accepted that competition is one of the key forces that keep industry competitive and dynamic, this view is being challenged by the concept that many activities can benefit from a rational mix of competition and cooperation that leads to complementary products and expansion of possibilities through the formation of new relationships or even new modes of operation and management. Increasingly, the same is also true for R&D organizations, which can benefit from working with partners (competitors) whose abilities make their own more attractive in the eyes of clients (Brandenburger and Nalebuff, 1997). Also, faced with growing competition from industry and increasing pressures and demands, public R&D institutions must look at ways to do more with fewer resources. Collaboration through team nets and other networking strategies have the potential to reduce costs, add value and promote capacity to respond quickly to changes. Besides, with the new tools of information technology, collaboration with any part of the world is possible as this promotes information and other resource sharing without the need for geographical proximity (Lipnack and Stamps, 1993).

How should a R&D organization behave in a multifaceted relationship, when partners can be also competitors? Organizations that enter competitive collaboration must be aware that their partners may be out to disable them. This dilemma has been faced by a growing number of organizations, which rapidly understand that effectiveness will be more and more a product of recognizing and using interdependence. With networks and interdependent teams, cooperation must be designed in the name of mutual needs and with a clear sense of sharing risks to reach objectives that are common to all partners (Lopes, 2000).

In many parts of the world, including in developing countries like Brazil, competitive funding systems for agricultural R&D are assuming growing importance as new sources of funding and as drivers for cooperation among universities, R&D institutes and the private sector, in many cases allowing collaboration even among institutions that are traditional competitors (Lopes, 2000). Although the rules and procedures governing the competitive granting system indicate the need for partnership and the general mode of interaction, experience has shown that industry/university/R&D institutes cooperations succeed only if they are founded on trust and understanding and promise mutual benefits. Also, successful experiences have come from the clear recognition of objectives and well structured management with intense communication.

Two experiences are described below that rely heavily on cooperation and networking directed to effective application of advanced technologies, including MAS, to genetics and breeding. Both are excellent examples of strategies that promote effective partnerships and collaboration by researchers from different institutions, disciplines or countries working on specific high-priority projects.

The case of the CGIAR Generation Challenge Programme: an international R&D network in genetic resources, genomics and breeding

As the number of stakeholders in the agricultural decision-making process increases and

the agricultural research agenda expands, organizations must be able to respond to an increasingly diverse and complex portfolio of priorities by strengthening interactions within the system and developing links and partnerships with groups traditionally outside the system. Towards this end, the CGIAR has designed a strategy to nurture the definition of objective R&D agendas in key themes and to guide scientists and teams worldwide towards integrated, synergistic involvement and operation. This strategy became known as the "Global Challenge Programmes" (www.cgiar.org/impact/challenge/index.html).

The strategy of the Global Challenge Programmes recognizes both that the cost of conducting research is escalating and that the complexity of the science needed for agricultural research is increasing. Research in most fields requires not only specialized equipment and facilities but also highly trained technical support in diverse disciplines. Increasingly, multidisciplinary teams of scientists will be required to address the complex issues facing agriculture and, in many cases, the professional expertise needed may have to be accessed in different parts of the world.

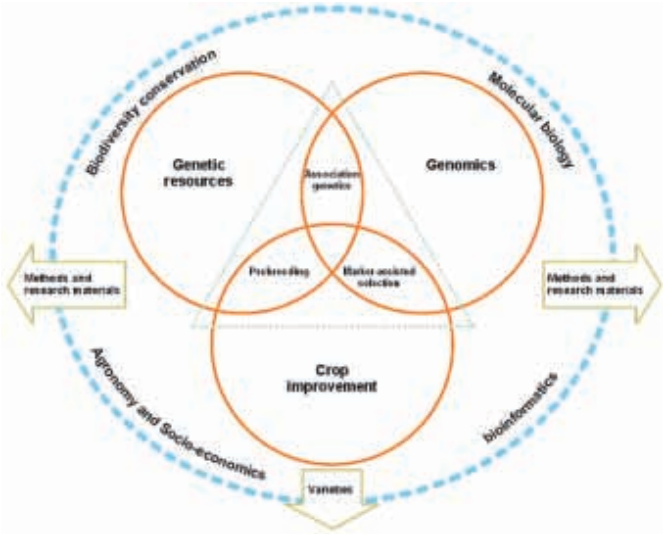
One such Challenge Programme, entitled "Unlocking Genetic Diversity in Crops for the Resource-Poor", also known as the "Generation Challenge Programme (GCP)" (CGIAR, 2003) is an international, multi-institutional, cross-disciplinary public platform for accessing and developing new genetic resources using advanced molecular technologies associated with conventional methods. Founded in July 2003 by the Executive Council of the CGIAR with start-up funding from the World Bank and the European Commission, the GCP has a membership of twenty-two public research institutions around the world, including nine CGIAR centres, four advanced research institutes and nine national agricultural research system institutions. Its budget in 2005 totalled at US$14 million (GCP, 2005).

This platform was designed to ensure that the advances of crop science and technology are applied to the specific problems and needs of resource-poor people who rely on agriculture for subsistence and their livelihoods. The GCP aims to "bridge that gap by using advances in molecular biology and harnessing the rich global stocks of crop genetic resources to create and provide a new generation of plants that meet these farmers' needs".

The concrete objective of the GCP is to access and develop genomic and genetic resources as enabling technologies and intermediate products for crop improvement programmes. It will not produce and release finished crop varieties for farmers, but develop new genetic resources and make the initial gene transfers to locally adapted germplasm, and then transfer the derived materials to crop improvement programmes, particularly those conducted in national agricultural research systems of developing countries, and to any other entities that have crop improvement goals, especially those dedicated to resource-poor farmers.

The GCP is, to date, the most comprehensive effort to cover, in a well structured and feasible manner, the complex interactions between genetic resources, genomics and breeding (Figure 2) in order to capture the benefits of the revolutions in biology and direct them to help solve some of the agricultural problems in the world's most difficult and marginal environments. It addresses its three key component parts in a separate but interconnected manner: (1) genetic resource collections provide the

FIGURE 2
Conceptual basis for the Generation Challenge Programme – Unlocking Genetic Diversity in Crops for the Resource-Poor



Reprinted by permission from the proposal for a CGIAR Challenge Programme (CGIAR, 2003).

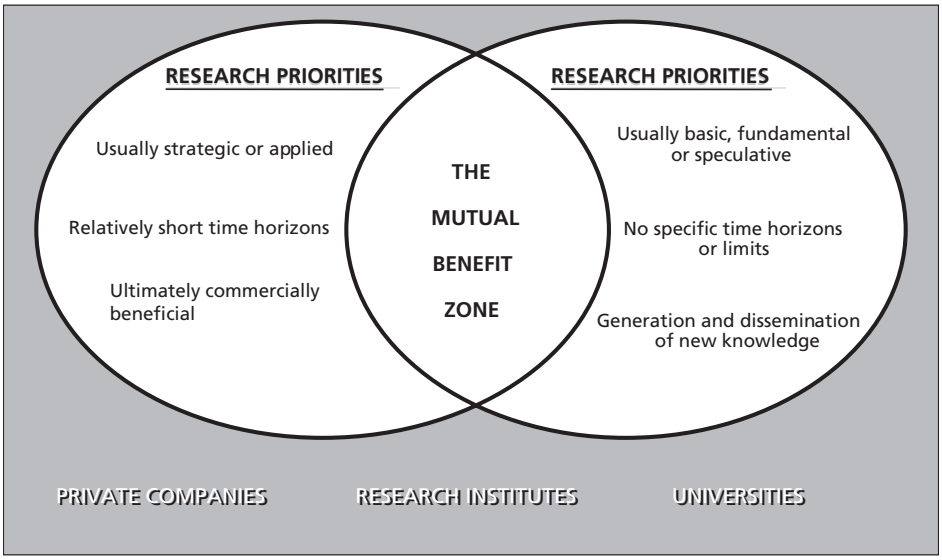
raw materials; (2) genomic science provides the means to exploit genetic resources (i.e. identify new alleles); and (3) crop improvement applies traditional and modern methods of gene/allele transfer into functional crop varieties (CGIAR, 2003).

The GCP is therefore an ambitious initiative to put into action a complex mix of tools, capacities, concepts and strategies. It is organized and managed to direct these resources towards the pursuit of goals that are not attainable through the disciplinary and isolated modes of operation that unfortunately prevail in the international agricultural R&D arena. As such, it is possibly the best structured international effort for development, adaptation and promotion of effective (and inclusive) access and use of tools such as MAS.

As part of its complex strategy, the GCP will define protocols for more efficient gene transfer including molecular markers that are closely linked to the genes for the desired trait, rapid tests for phenotype recognition, and genetic transformation of new genes into locally adapted genetic materials, such as improved varieties and landraces. All of these strategies depend on the adaptation and development of marker technology and marker-assisted procedures, hopefully helping to consolidate a networking approach to breeding and biotechnological research with effective impact, especially on resource-poor countries.

Research activities commenced in January 2005 with the first round of competitive research grants awarded for 17 three-year projects of approximately US$1 million each. In early 2005 a new round of commissioned grants was started, which served as the basis of the GCP platform of tools and technologies for genetic studies and applications. In total, the GCP initiated

FIGURE 3
Research priorities and partnerships among private companies, research institutes and universities



Source: adapted from Laider (1998).

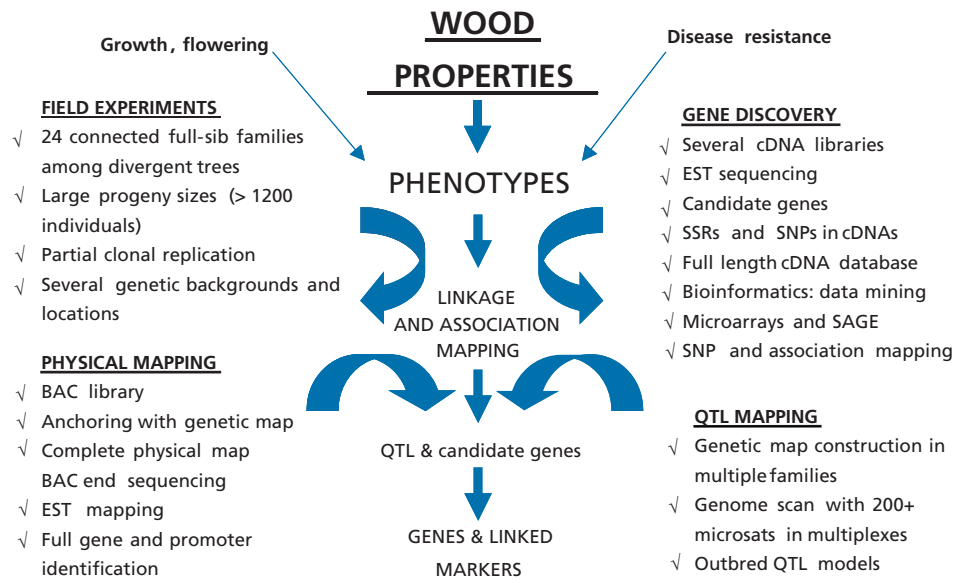
67 competitive and commissioned research projects and capacity-building activities in 2005. The 2005 Annual Report and 2006 Work Plan summarizes research progress and capacity building achievements in 2005 and presents an overview of the competitive and commissioned research portfolio and the capacity-building and delivery activities for 2006 (GCP, 2005, 2006).

The Case of the Genolyptus Programme in Brazil: a public/private network in genomics and molecular breeding of *Eucalyptus*

"The challenge and the opportunity for publicly supported agricultural research are not in duplicating the private sector's research agenda, but in building unique public/private partnerships or perhaps even jointly supported consortia for agricultural research" (CAST, 1994). Increasingly, agricultural research will be conducted through partnerships among private companies, public research institutes and universities (Figure 3). In forming such alliances, these organizations must recognize that developing productive relationships involves non-competitive dialogue and understanding of each others' abilities and limitations. Partnerships will flourish only if founded on trust and understanding and if differences in drivers and objectives are recognized and accommodated in initiatives with a real perspective of mutual benefits (Lopes, 2000).

An example of a successful public/private partnership with clear understanding of partners' abilities and limitations and clear definition of responsibilities and benefits to be pursued is the Genolyptus Network in Brazil (Grattapaglia, 2003). This R&D network was created to establish a foundation for a genome wide understanding of the molecular basis of wood formation in

FIGURE 4
The Genolyptus project – from phenotypes to genotypes in an integrated way



Source: Grattapaglia, 2003.

Eucalyptus, coupled with the translation of knowledge into improved tree breeding technologies.

This programme relies heavily on the development of aligned R&D efforts in genetic resources, genomics and molecular breeding (Figure 4). It mobilizes capacities and infrastructure in constructing physical maps, developing expressed sequence tag (EST) databases, generating a database of expression profiling of genes that control key traits and developing methods for MAS for traits of high heritability in wood formation. Also, the network develops a capacity-building and training programme for professionals in universities and forestry companies, targeting the integration of genetics, genomics and breeding efforts of *Eucalyptus*.

The rationale of the network is based on the understanding that, even with the more powerful tools allowing a much more global and integrated view of genetic processes, genomics will only succeed in contributing to the development of improved eucalypt if it is deeply interconnected with intensive fieldwork and creative breeding. The Genolyptus project therefore differs from other plant genome initiatives in the intensity, refinements and scope of the effort devoted to field experiments to generate the diversity of phenotypes necessary to study gene function. Quantitative trait loci (QTL) detection, the development of SNP haplotypes for association mapping and physical mapping will link the phenotypes to genes that control processes of wood formation that define industrial level traits (Grattapaglia, 2003).

A key feature of the Genolyptus network is its pre-competitive nature. The research programme was designed collaboratively with no immediate intention of marketing its results, even although its

planned outputs will eventually lead to the creation of many new products and processes of commercial value. Thus, the activities during its first phase are designed to resolve basic, common technological problems – a sufficient reason to mobilize several private companies (that are competitors in the market) and public organizations. After the first phase of six years, the network will have generated a consolidated base of knowledge and tools that will promote the development of specific interest projects, either in partnership or individually, according to specific business strategies and market targets.

Also, team organization and management are based on modern tools and concepts, involving a competent, highly respected scientist with talent to lead network operations, a steering committee and a technical committee for adequate planning, decision-making and follow-up, as well as contract models and negotiation strategies appropriate to the complexity of the network. Intellectual property rights provisions are based on access limited to participants, with all genetic materials and patents produced being co-owned by the 20 participating institutions. Scientific publications are highly encouraged.

As in the Generation Challenge Programme, the Genolyptus network is an original initiative to integrate and align a complex mix of tools, capacities, concepts and strategies. The ability to mobilize such a wide range of organizations, including 12 private companies operating in a highly competitive market space, indicates that the network design was successful, while its pre-competitive nature, organization and management strategy allowed the definition of a "zone of mutual benefits" (Figure 3), facilitating the pursuit of goals that are not attainable through isolated research. The Genolyptus network is therefore an excellent example of the feasibility of developing a structured public/private effort for integrated and effective use of advanced tools such as MAS.

CONCLUSIONS

- Although advances in molecular marker technology have uncovered many possibilities for easing the transfer of genes into desired crops and livestock through MAS, there is still limited recorded impact of these technologies in breeding programmes.
- It is generally recognized that genetic improvement of complex traits using MAS is more difficult than previously considered. Therefore, more methodology development, better planning and implementation strategies will be needed for its successful and rapid application to breeding programmes.
- The future configuration of breeding programmes is dependent on knowledge to guide strategic decisions about structures, methods and capacities that take advantage of new opportunities and technological niches. Unfortunately, there are very few efforts directed at thinking about the future of breeding programmes, especially in less developed countries. Research organizations need information, which is not currently available, about changes and influences and their impact in the future on key activities such as crop and livestock breeding. To acquire and organize this information, prospective studies on the present and future performance of breeding programmes and their related activities will have to be developed.
- Priority-setting strategies, together with cost–benefit analysis are necessary to guide informed decisions on how best

to apply MAS and other advanced technologies to crop and livestock breeding in developing countries.

- MAS has to be understood and undertaken as part of a complex process. Complementarities and a mix of technologies and capabilities, together with effective approaches to networking, must be viewed as key ingredients for its appropriate development and application to breeding programmes.
- One of the key problems limiting the use of MAS and other advanced technologies in developing countries is the difficulty of building effective teams and networks. Approaches to networking and partnerships are key to enabling organizations to attain new goals with less cost and to adding more value to their products and processes. Also, the demand for efficiency and relevance presses R&D programmes to move in the direction of cooperation and alignment of efforts.
- The present and future challenges and opportunities for agricultural research organizations are to build public/private partnerships or new types of consortia dedicated to innovation. In forming such alliances, these organizations must recognize that developing productive relationships involves non-competitive dialogue and understanding of each others' abilities and limitations. In order to survive and flourish, partnerships have to be sustained on trust and understanding.
- Developing countries must focus on training to build and shape capacities and effective strategies to support research in advanced biology applied to breeding. Also, new management strategies are needed to deal with the complex nature of modern breeding programmes.

REFERENCES

Altshuler, D., Pollara, V.J., Cowles, C.R., Van Etten, W.J., Baldwin, J., Linton, L. & Lander, E.S. 2000. An SNP map of the human genome generated by reduced representation shotgun sequencing. *Nature* 407: 513–516.

Brandenburger, A.J. & Nalebuff, B.J. 1997. Co-opetition, New York, USA, Doubleday.

CAST (Council for Agricultural Science and Technology). 1994. Challenges confronting agricultural research at land grant universities. Issue Paper 5. Ames, IA, USA.

Castro, A.M.G., de Cobbe, R.V. & Goedert, W.J. 1995. Manual de prospecção de demandas para o SNPA. Brasília, Brazil, Embrapa.

Castro, A.M.G., Lima, S.M.V. & Freitas Filho, A. 1998. Manual de Capacitación en Análisis de Cadenas Productivas. Brasília, Brazil, Embrapa.

Castro, A.M.G., Lima, S.M.V., Lopes, M.A. & Martins, M.A.A.G. 2002. Estratégia de P&D para o Melhoramento Genético em uma Época de Turbulência. *Anais do XXII simpósio de gestão da inovação tecnológica,* p.17. Salvador, BA, Brasil.

Castro, A.M.G., Lopes, M.A., Lima, S.M.V. & Bresciani, J.C. 2005. Trends and technological strategy in the Brazilian seed sector (Cenários do Setor de Sementes e Estratégia Tecnológica*). Revista de Política Agrícola Brasília* 3: 58–72.

Castro, A.M.G., Lima, S.M.V., Lopes, M.A., Machado, M.S. & Martins, M.A.G. 2006. O Futuro do Melhoramento Genético Vegetal no Brasil – Impactos da Biotecnologia e das Leis de Proteção de Conhecimento (The future of plant breeding in Brazil – impacts of biotechnology and intellectual property legislation). Brasília, DF, Embrapa Informação Tecnológica.

CGIAR (Consultative Group on International Agricultural Research). 2003. *Unlocking Genetic Diversity in Crops for the Resource-poor: A Proposal for a CGIAR Challenge Program* (available at www.generationcp.org/sccv10/sccv10_upload/CP_proposal.pdf).

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

De La Vega F.M., Dailey D., Ziegle J., Williams J., Madden D. & Gilbert D.A. 2002. New generation pharmacogenomic tools: a SNP linkage disequilibrium map, validated SNP assay resource, and high-throughput instrumentation system for large-scale genetic studies. *Biotechniques* Suppl 32: S48–S54.

FAO. 2003. Molecular marker-assisted selection as a potential tool for genetic improvement of crops, forest trees, livestock and fish in developing countries. *Background document to Conference 9 of the FAO Biotechnology Forum* (available at www.fao.org/biotech/C10doc.htm).

Ferreira, M.E. 2003. Melhoramento Genético do Arroz: Impactos da Genômica. *In* A. Borém, M. Giúdice & T. Sediyama, eds. *Melhoramento genômico.* Universidade Federal de Viçosa,Viçosa, Brazil.

Ferreira, M.E. 2006. Molecular analysis of gene banks for sustainable conservation and increased use of crop genetic resources. *In* J. Ruane & A. Sonnino, eds. *The role of biotechnology in exploring and protecting agricultural genetic resources*. FAO, Rome (available at www.fao.org/docrep/009/a0399e/a0399e00.htm).

GCP (Generation Challenge Programme). 2005. *Generation Challenge Programme research highlights 2005.* Mexico, DF. CIMMYT (available at http://www.generationcp.org/sccv10/sccv10_upload/Highlightslittle.pdf).

GCP. 2006. *Generation Challenge Program Annual Report and Year Three Workplan.* Mexico, DF, (available at www.generationcp.org/comm/manual/doc_100.pdf).

Graff, G., Cullen, S., Bradford, K., Zilberman, D. & Bennet, A. 2003. The public-private structure of intellectual property ownership in agricultural biotechnology. *Nature Biotech*. 21: 989–995.

Grattapaglia, D. 2003. Genolyptus. *In*: A. Borém, M. Giúdice, & T. Sediyama, eds. *Melhoramento genômico,* pp. 51–71. Universidade Federal de Viçosa,Viçosa, Brazil.

Laider, D.A. 1998. *Research partnerships between industry & universities. A guide to better practice*. Runcorn, UK, ICI Technology.

Lima, S.M.V., Castro, A.M.G., de Mengo, O. & Medina, M. 2000. *Entorno y Enfoque Estratégico en I & D: Proyecto Nuevo Paradigma.* Artigo submetido ao XXI Simpósio Internacional de Gestão Tecnológica, FIA/ USP/PACTO.

Linstone, H.A. & Turoff, M. 1975. *The Delphi method: techniques and applications.* Readings, MA, USA, Addison Wesley.

Lipnack, J. & Stamps, J. 1993. *The age of the network.* New York, USA, Oliver Wright Publications.

Lopes, M.A. 2000. Cooperation and competition in competitive grants: is "coopetition" being achieved? *In Proc. of the Competitive Grants in the New Millenium: A Global Workshop for Designers and Practitioners,* pp. 219–230. Brasília, Brazil, Embrapa/IDB/The World Bank.

Lörz, H. & Wenzel, G. 2005. Molecular marker systems in plant breeding and crop improvement. *Biotech. in Agric. & Forestry* 55: XXII. Berlin, Germany, Springer.

Mohan, M., Nair, S., Bhagwat, A., Krishna, T.G. & Yano, M. 1997. Genome mapping, molecular markers and marker-assisted selection in crop plants. *Mol. Breeding* 3: 87–103.

Price, S.C. 1999. Public and private plant breeding. *Nature Biotech*. 17: 938.

Rafalski, J.A. 2002. Novel genetic mapping tools in plants: SNPs and LD-based approaches. *Plant Sci.* 162: 329–333,

Stuber, C.W., Polacco, M. & Senior, M.L. 1999. Synergy of empirical breeding, marker-assisted selection, and genomics to Increase crop yield potential. *Crop Sci*. 39: 1571–1583.

Young, N.D. 1999. A cautiously optimistic vision for marker-assisted breeding. *Mol. Breeding* 5: 505–510.