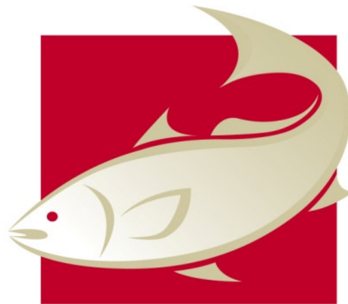
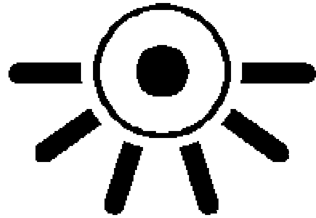
Benefit-cost analysis of marker assisted selection in Australian aquaculture species

FINAL REPORT

Dr Nick Robinson





Project No. 2008/904

Benefit-cost analysis of marker assisted selection in Australian aquaculture species

Project No. 2008/904

Dr Nick Robinson[1,2]

[1]Flinders University, GPO Box 2100, Adelaide, SA 5001.
[2]Nofima Marine, P.O. Box 5010, N-1432 Ås, Norway.

May 2009

Copyright Australian Seafood CRC, Flinders University and Nofima 2009.

This work is copyright. Except as permitted under the Copyright Act 1968 (Cth), no part of this publication may be reproduced by any process, electronic or otherwise, without the specific written permission of the copyright owners. Neither may information be stored electronically in any form whatsoever without such permission.

The Australian Seafood CRC is established and supported under the Australian Government's Cooperative Research Centres Programme. Other investors in the CRC are the Fisheries Research and Development Corporation, Seafood CRC company members, and supporting participants.

ISBN 978-0-9805789-2-8

Published by Flinders University, GPO Box 2100, Adelaide, SA 5001.

Preferred way to cite:

Robinson N. A. (2009). Benefit-cost analysis of marker assisted selection in Australian aquaculture species. Final report to the Australian Seafood CRC, Project No. 2008/904. Flinders University.

Table of Contents

1. NON-TECHNICAL SUMMARY ... 7

2. ACKNOWLEDGEMENTS ... 10

3. BACKGROUND ... 11

3.1 WHAT IS A MARKER? ... 11
3.2 HOW ARE MARKERS DETECTED? ... 11
3.3 GENETICS OF THE TRAIT ... 11
3.4 GENE EXPRESSION ... 13
3.5 RESOURCES NEED TO FIND MARKERS ... 14
3.6 NEW TECHNOLOGIES ... 14
3.7 TARGET AUDIENCE AND SCOPE OF REPORT ... 15

4. NEED ... 16

5. OBJECTIVES ... 16

6. METHODS ... 17

6.1 TRAITS TO TARGET FOR MAS ... 17
6.2 RESOURCES NEEDED AND CURRENT STATUS OF DEVELOPMENT ... 17
6.3 LIKELIHOOD OF SUCCESS OR FAILURE ... 17
6.4 COSTS OF MAS ... 17
6.5 BENEFITS FROM MAS ... 17
6.6 APPLICATION OF MAS ... 18

7. RESULTS/DISCUSSION ... 19

7.1 STRENGTHS, WEAKNESSES, OPPORTUNITIES, THREATS AND RISKS TO SELECTIVE BREEDING USING MARKER ASSISTED SELECTION 19
7.2 TRAITS TO TARGET FOR MAS ... 24
7.3 RESOURCES NEEDED AND CURRENT STATUS OF DEVELOPMENT ... 26
7.3.1 Development of DNA or cDNA libraries (Step 1) ... 27
7.3.2 Sequence of expressed genes or genomic DNA (Step 2) ... 27
7.3.3 Search sequence for genetic variation (Step 3) ... 28
7.3.4 Creation of tools for studying genetic variation (Step 4) ... 29
7.3.5 Development of mapping population resources (Step 6) and collection of trait data and DNA (Step 7) ... 30
7.4 LIKELIHOOD OF SUCCESS OR FAILURE ... 32
7.4.1 Influence of the species and the composition of it's genome ... 32
7.4.2 Technology used and resources available ... 33
7.4.3 Genetic basis for trait ... 33
7.4.4 Accuracy of information recorded and samples collected ... 34
7.4.5 Experimental design and method of analysis ... 34
7.4.6 Effective means for implementing MAS ... 35
7.5 BENEFIT-COST RATIO FOR MAS ... 36
7.5.1 Examples from the literature ... 36
7.5.2 Predicting the relative benefit-costs for species in the Australian Seafood CRC ... 38
7.5.3 Side benefits from MAS ... 40

7.6 APPLICATION, DEMONSTRATION AND VALIDATION OF MAS 40
7.6.1 Validation *40*
7.6.2 Application and demonstration *41*
7.6.3 Influence on future strategies *42*

8. BENEFITS AND ADOPTION 43

9. FURTHER DEVELOPMENT 43

9.1 WHAT ARE THE BIG OPPORTUNITIES FOR THE PARTNERS IN THE AUSTRALIAN SEAFOOD CRC? 43
9.2 OFFSET OF COSTS 44

10. PLANNED OUTCOMES 44

11. CONCLUSION 44

12. REFERENCES 47

13. APPENDIX 1. INTELLECTUAL PROPERTY 55

14. APPENDIX 2. STAFF 55

15. APPENDIX 3. DETAILED LITERATURE SEARCH HIGHLIGHTING AVAILABLE RESOURCES OF USE FOR DEVELOPING MAS 56

1. Non-technical summary

The development of marker assisted selection (MAS), its application using genomic breeding values or other means and selection using gene expression profiling are rapidly developing areas, and significant technological and methodological advances have been made in recent years. There is potential that these techniques could greatly speed the genetic progress in selective breeding programs and provide faster access and more benefit than conventional selective breeding techniques. But before embarking on research in this area, aquaculture industries need answers to a number of questions. What basic resources are available or needed for each species before research to find markers or to apply MAS, genomic estimated breeding values or gene expression profiles can be carried out? What are the likely cost-benefits for use of these technologies? Which traits would be best targeted when using these technologies? When should research be undertaken or what developments might trigger research in this area for each species? What has been the measured impact of the application of these technologies on other industries? This project relates to the research strategies developed by the CRC from projects that have scoped selective breeding for abalone (*Haliotis rubra and H. laevigata*), Yellowtail Kingfish (*Seriola lalandi*) and Barramundi (*Lates calcarifer)*. It also relates to the R&D strategies developed by many of the other aquaculture sectors in the CRC, eg. Pacific oyster and Sydney rock oyster (*Crassostrea gigas* and *Saccostrea glomerata*), Atlantic salmon (*Salmo salar*), black tiger prawn (*Penaeus monodon*), etc., which have included in the past, or currently include, projects to develop or apply marker assisted selection.

To answer these questions for the Seafood industries participating in the Australian Seafood CRC, this report

- Draws together an overview of trait priorities identified by these industries,
- Undertakes a literature review to determine the current state of knowledge and resources available for marker assisted selection,
- Explores which factors are likely to affect the success of the development and application of marker assisted selection,
- Estimates some of the costs and likely benefits involved in developing and from applying marker assisted selection,
- Provides an overview of current methods suggested for optimizing the discovery and application of markers, and,
- Reviews cases where the impact from the application of marker assisted selection has been demonstrated.

The report has identified:

1. *Basic resources that are available or needed for each species before research to find markers or to apply these technologies can be carried out.*

Because of the advances in technology, it is now relatively inexpensive to develop a very large set (10's of thousands) of useful (polymorphic- having multiple alleles within a population) tests. This large number of polymorphic tests can be quickly developed and families of animals efficiently genotyped to enable fine scale mapping and the identification of potentially useful marker tests for genes affecting economically important traits. These resources are currently only available for Atlantic salmon (*S. salar*), but initiatives are underway to develop resources for black tiger prawn (*P. monodon*) and are being planned for abalone (*Haliotis sp.*) and Pacific oyster (*C. gigas*).

Other important resources that are needed are trait data, pedigree data and DNA collected from individuals in large families (eg. several families each consisting of a few hundred full-siblings and parents). Ideally the families should be part of an existing selective breeding program so that any discoveries are directly applicable to the breeding population. Selective breeding programs are underway in Australia for *S. salar*, *C. gigas*, *S. glomerata*, *H. rubra* and *H. laevigata*, *P. monodon* and *F. merguiensis*. Relevant measurements and samples would need to be taken. A database containing this data and pedigree data with reference to stored samples would be needed.

Marker assisted selection also needs to be applied in an optimum way for the species and traits in question so that the benefit-cost ratio from the application of marker assisted selection is maximized. Some simulation is needed to predict how to apply marker information in an optimum way. This has been done for *H. rubra* and *H. laevigata,* and for some traits for *S. salar*.

All other software needed to develop resources, detect and apply MAS is freely available or can be purchased.

2. *Likely cost-benefits for the development and application of these technologies and traits that would be best targeted using the technologies.*

For traits like disease resistance, meat quality or feed conversion efficiency, for which improvement could have a highly beneficial economic impact in each aquaculture sector, and which are all difficult or inefficient to improve using selective breeding without MAS, the benefit-cost ratio for the development and application of MAS technologies is likely to be high. The benefit-cost ratio will be specific to the species, trait, set of markers and approach used. The highest benefits would come from application of MAS to the Atlantic salmon and prawn industries. The benefit derived depends on a number of factors including the value of the industry in Australia and generation interval of the species for marker assisted selection.

3. *When research should be undertaken and developments that might trigger research in this area for each species.*

A major prerequisite for the use of MAS is the existence of a selective breeding program for the species concerned in Australia. Also, the families developed for selective breeding would be the best resource to use for developing MAS technology. It therefore would be sensible to delay the development of MAS for some species (yellow tail kingfish, Barramundi and some prawn species) until selective breeding programs are established. The development of selective breeding programs is therefore one trigger point for beginning to develop MAS.

Although polymorphic tests are available for most of the species included in this study, the only comprehensive single nucleotide polymorphism resource that is available is for *S. salar*. Recent technology developments have lead to efficient means to characterise and screen the genome using tens-of-thousands of these types of polymorphisms. This has meant that we can quickly and efficiently identify some very useful markers. The development and availability of single nucleotide polymorphism-chips (SNP-chips) for the species other than *S. salar*, will be another important trigger point for beginning to develop MAS. It would be beneficial if Australian researchers participated in the development these SNP-chips (contributing DNA from our populations and contributing to the design of the SNP-chip). So another important trigger for beginning research might be the initiation of the development of these resources by researchers overseas.

Finally, MAS will be most beneficial for traits that are of high economic importance to our industries and traits that are otherwise difficult or slow to improve using selective breeding without markers or other means. Therefore, it is recommended that marker research only be undertaken for traits that meet these criteria (eg. disease resistance, meat quality or feed conversion efficiency traits). For disease resistance, we need an effective way of being able to measure an individual's level of resistance before we can find markers. Another trigger to the development of MAS might therefore be the characterization of disease causing organisms affecting production, and the development of mechanisms for their propagation and for controlled experimental challenge tests. A final trigger for the development of MAS might be the development of a significant market driven meat quality issue (an issue that cannot be easily or adequately addressed by manipulating nutrition, processing or grow-out environments).

4. *Measured impact of the application of these technologies on other industries.*

There is very little published work demonstrating the impact of the application of these technologies to other aquaculture or livestock industries. This is partly because, 1) there has been little application in the past, 2) it is difficult to get existing breeding companies to compare or benchmark the results of their breeding work, and 3) because there has not been enough time for the most recent developments in technology to be applied to aquaculture breeding programs and the impact of this application to be evaluated. However, two recent demonstrations exist.

The same marker for infectious pancreatic necrosis resistance in *S. salar* has been discovered and is being applied to selective breeding programs in Norway (Aquagen) and the United Kingdom (Landcatch Natural Selection). This marker has a very large effect on resistance to this disease with Aquagen reporting that animals that are homozygous QQ for the marker have a survival rate of 0.87 while qq homozygotes have survival rates of 0.51 under challenge tests. Aquagen have also shown that selection for IPN resistance by challenge testing has lead to a change in the frequency of the high resistance allele from 0.3 in 2005 to 0.5 in 2008.

Application of the use of "genomic estimated breeding values" to dairy cattle selective breeding programs in Australia, New Zealand and the United States have been shown to be 2-20% more reliable than parental average breeding values, and has allowed companies to market bull teams based on genomic estimated breeding values at 2 years of age, a strategy which doubles the rate of genetic improvement in milk composition.

In summary, recent developments in sequencing and genotyping technology, and in methods for the analysis of this information, have made it possible to more efficiently develop and apply MAS to benefit the participating industries in the Australian Seafood CRC. However, there are a number of important resources that need to be developed before this can be done, and we need to target traits and species where the potential benefit-cost ratio from MAS is greatest. There are few reports demonstrating the impact of the application of MAS. However, the new technologies should lead to much higher benefits than were possible before, breeding companies have sometimes been reluctant to divulge precisely what they do, and only a few marker discoveries are likely to have been effectively adopted by breeding companies around the world.

2. Acknowledgements

Thanks to Graham Mair Seafood CRC and Alex Safari Flinders University, who have all provided useful feedback and information for the report. Thanks also to all in the aquaculture industries who contributed information and provided helpful advice.

3. Background

Both a strain's performance in the face of disease, and its product eating quality, have large economic effects on aquaculture industries worldwide. Genetic variation for disease resistance and meat quality traits have been found, but such traits are often complex to understand and measure. Furthermore, these traits are costly to evaluate and slow to improve using traditional methods of selective breeding. A major limitation is the inability to directly test highly valuable potential broodstock.

Aquaculture researchers are therefore looking to develop and apply technologies such as marker assisted selection that have been useful for livestock and plant breeding.

3.1 What is a marker?

Markers are tests revealing variation in the sequence of DNA (ie. polymorphisms) that differ between individuals in a population and whose inheritance is associated with the inheritance of a trait of interest.

3.2 How are markers detected?

To find markers associated with traits of interest we need to look at the inheritance of many hundreds, if not thousands, of different polymorphisms (eg. single nucleotide polymorphisms or microsatellites, Fig. 1) spread throughout the genome and look for association (linkage with) genes affecting the trait of interest. We are then able to identify a region of chromosome containing loci affecting the trait (quantitative trait loci otherwise known as QTL) and markers, which when inherited lead to either a higher or lower trait value (Fig. 2).

3.3 Genetics of the trait

Most of the important economic traits (eg. growth rate, disease resistance, feed conversion efficiency, meat quality) are what are known as "quantitative" traits. That is, these traits vary in a continuous manner and are thought to be affected by many different genes of small and large effect. It has been shown that these types of traits are affected by a few genes which have a large effect on the trait, and by many genes with small effects on the trait (Hayes and Goddard, 2001). These types of traits have often been described as being controlled by a complex cascade or network of interacting genes. Key or major rate limiting steps in the gene pathways affecting the trait are controlled by genes that have a large effect on the trait.

There are some important traits (eg. the control of sexual maturation or reproduction) that are controlled by a few genes with major effects. An example is the gene KiSS-1 and its receptor GPR54 which have been shown to have a major role in controlling the development of the reproduction system in mice (eg. Funes et al., 2003) and has now been extensively characterised in teleost fish (Elizur, 2009) with GPR54 highly expressed in early pubertal development (eg. Nocillado et al., 2007). For such genes, that have been shown to have a large effect on similar traits in other species, the homologous gene in fish may play a similar role. In these instances it may be possible to use the product of the gene to affect the trait or to use the gene as a marker that explains a large proportion of the effect on the trait. But this report will focus on the most economically important traits in fish, most of which are probably influenced by many genes, and the benefits and costs associated with developing and applying marker tests for such traits.

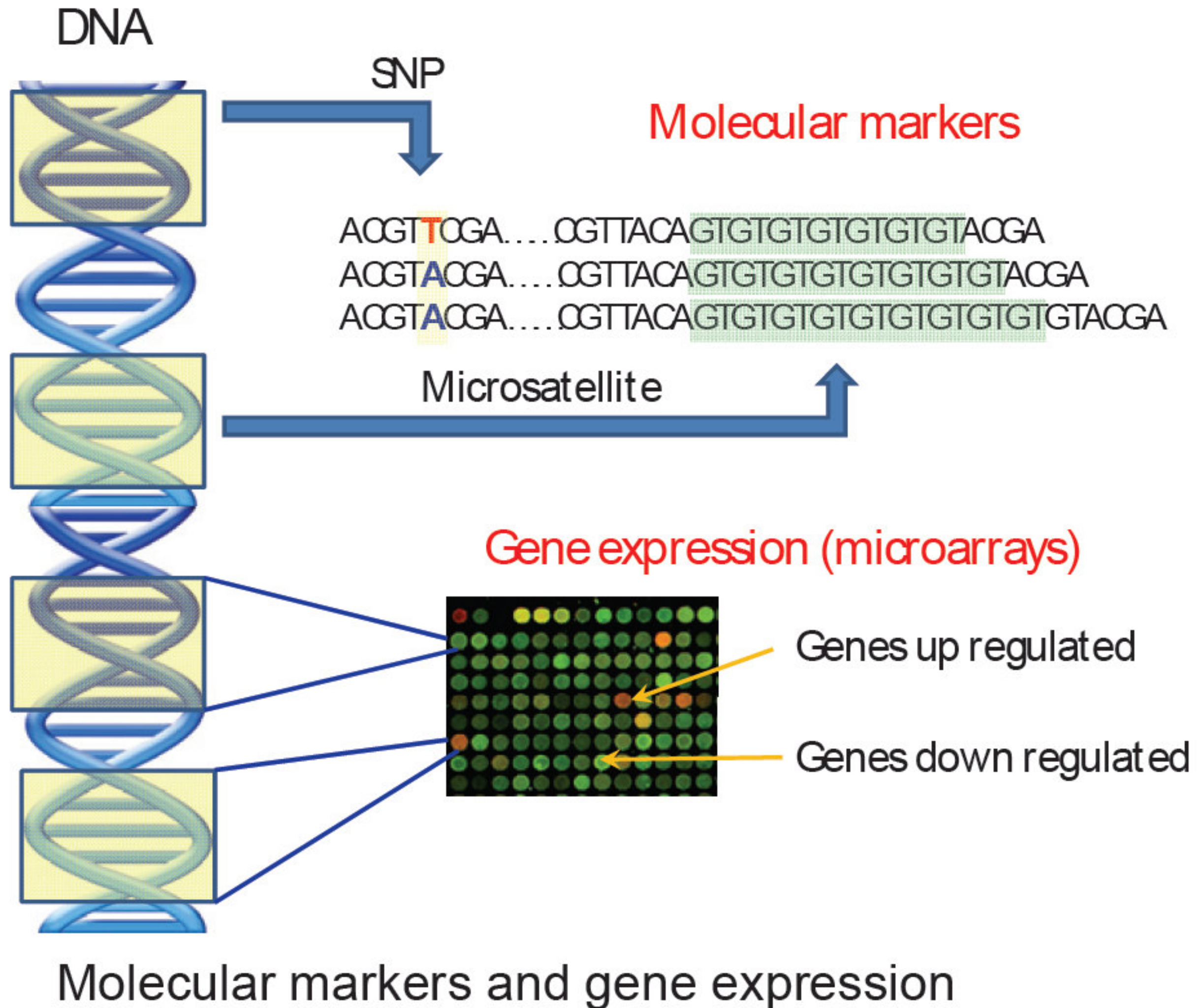


Fig. 1. Variation (polymorphism) in the sequence of DNA occurs between different individuals in a population, resulting in single nucleotide polymorphisms (SNPs) and differences in the number of repeated sequence in a stretch of tandemly repeated DNA (microsatellite locus). The regulation of how particular genes in the body are expressed (transcribed and translated into protein) under different states also differs between individuals. Gene expression can be measured for thousands of genes spotted on a glass slide (a microarray).

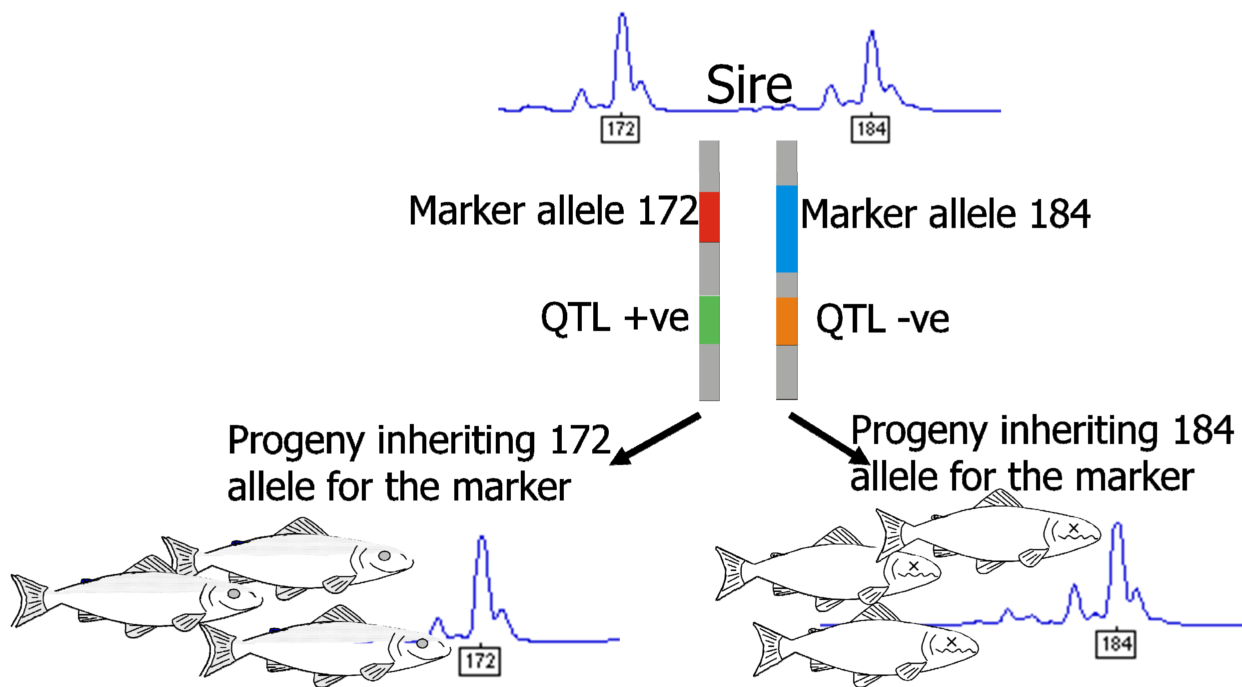


Fig. 2. Chromosome regions containing genes affecting the trait can be detected by genotyping progeny and looking for an association between the marker allele inherited from the parent and the inheritance of the trait.

3.4 Gene expression

Another type of variant that is only beginning to be used as a tool for marker assisted selection is gene expression analysis (Fig. 1). When a gene is "expressed" it is translated and transcribed into RNA and protein (amino acids). We can measure the amount of RNA that is produced by the gene in a given tissue at a given time (the genes expression level). Many thousands of genes, and their relative expression levels under two different states, can be measured on a single glass microscope slide (or microarray, Fig. 1). For instance, we might see which genes are turned on or off after infection of cells with a bacteria, and see across a number of animals a pattern of expression changes, or a gene expression profile, that is correlated with disease resistance or susceptibility (Fig. 3).

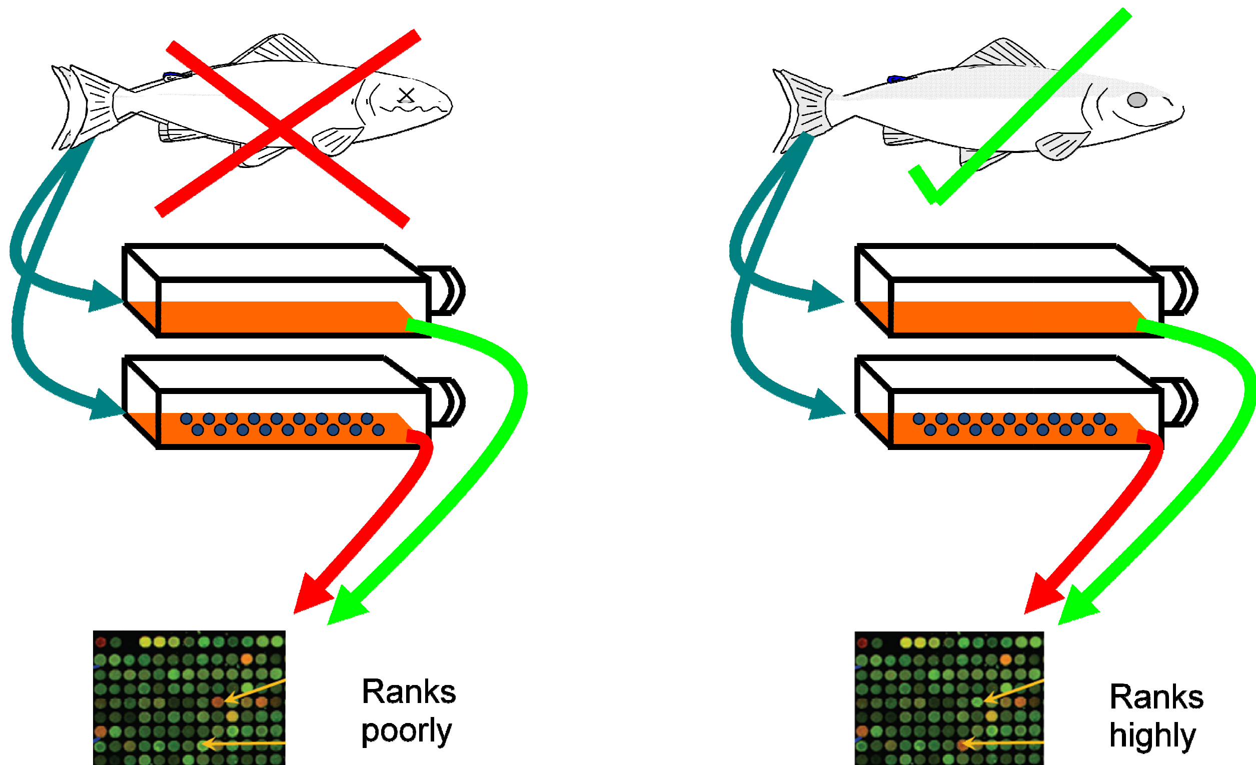


Fig. 3. Gene expression profiles may be associated with an animals performance, and may be a future form of marker assisted selection.

3.5 Resources need to find markers

To find linked polymorphic markers for genes affecting the trait, a number of resources need to be developed (such as polymorphic loci, families segregating for the trait etc). For the range of aquaculture species grown in Australia these resources are either fully developed, partially developed or non-existent. The value of disease resistance or meat quality differs between sectors, and sometimes between different markets within sectors. The industries in Australia also vary in production value and maturity and so the benefits that could be potentially realised from the use of marker assisted selection can be small, growing or large.

3.6 New technologies

Hetzel and Davis (2000) reviewed the integration of molecular genetic technology with traditional approaches for genetic improvement of aquaculture species. Since then advances in technology have dramatically improved the power of experiments to detect markers, the applicability of marker information and the costs of genome scans and marker assisted selection itself. A number of new technologies and methodologies have recently become available. Marker assisted selection (MAS), genomic breeding values (GEBVs) and selection using gene expression profiling (GEPS) are some variations that have recently arisen. There are now some good examples of the actual application of these technologies and the measured impact on different industries. There are also some new resources that are now available (eg. salmonid SNP-chips) that greatly reduce the cost and improve the

effectiveness of the application of MAS or GEBV's.

3.7 Target audience and scope of report

This project was developed out of needs that arose during the development of several of the Seafood CRC's selective breeding scoping initiatives and through the Seafood CRC theme business plan for "Breeding and Profit". This is a rapidly developing highly technical area. The industries are unsure about what benefits have been derived for other industries, basic resources that are needed and that are available for each species, when would be an appropriate time to invest, how much would MAS cost to develop and implement and what would be the likely benefits from MAS. This project therefore relates to existing breeding programs for Pacific oyster, Sydney rock oyster, prawns and Atlantic Salmon, to developing breeding programs for abalone and to potential new breeding programs for Barramundi and Yellowtail Kingfish.

The intention of the report is to present a concise benefit-cost analysis for the use of marker assisted selection for the industries participating in the Australian Seafood CRC. Although the focus is very much on "marker assisted selection" technologies, the report does also briefly describe some other "marker-like" technologies that have recently emerged (eg. gene expression profiling). The report will also be useful as background information for consideration by the Aquaculture Innovation Research Hub project in the Seafood CRC.

4. Need

The development of marker assisted selection, its application using genomic breeding values, or other means, and selection using gene expression profiling are rapidly developing areas, and significant technological and methodological advances have been made in recent years. There is potential that these techniques could greatly speed the genetic progress in selective breeding programs and provide faster access and more benefit than conventional selective breeding techniques. But before embarking on research in this area, aquaculture industries need answers to a number of questions. What basic resources are available or needed for each species before research to find markers or to apply these technologies can be carried out? What are the likely cost-benefits for the development and application of these technologies? Which traits would be best targeted using the technologies? When should research be undertaken or what developments might trigger research in this area for each species? What has been the measured impact of the application of these technologies on other industries? This project relates to the research strategies developed by the CRC from projects that have scoped selective breeding for abalone, Yellowtail Kingfish and Barramundi. It also relates to the R&D strategies developed by many of the aquaculture sectors in the CRC (eg. abalone, oyster, Atlantic salmon, prawn etc.) which have included in the past, or currently include, projects to develop markers or apply marker assisted selection.

5. Objectives

To provide a concise report that focuses on the relevance and benefit cost analysis of marker assisted selection (in its broader context) for the Abalone, Barramundi, Prawn, Tuna, Yellowtail Kingfish, Mulloway, Oyster and Salmon aquaculture industries.

6. Methods

6.1 Traits to target for MAS

Trait prioritisation exercises were already undertaken for many of the industries involved in the seafood CRC prior to this project. When trait prioritisation was unknown or unsure, a simple exercise was carried out over the phone or by email with representatives of the industry organisation. To assess a traits priority the following factors were considered:

i. the genetic effect on important traits (heritability)
ii. the economic effect of the trait on the industry
iii. practicality of measuring or selecting for improvement of the trait. Do individuals need to be stressed, killed or put at risk to measure the trait? Does measurement of the trait affect the speed of genetic improvement that can be achieved (for this or other traits)?
iv. demonstrability of the trait. How apparent will it be that the trait has been improved?
v. expense to measure the trait (either directly or indirectly without the use of markers)
vi. correlations with other traits
vii. what research has already been done on the trait

6.2 Resources needed and current status of development

A literature/internet review was carried out to determine the state of knowledge and resources. From this the resources needed to be developed for each species and industry in order to be able to discover markers and perform marker assisted selection were determined. This included resources such as microsatellite loci, single nucleotide polymorphisms (SNPs), SNP genotyping platforms (eg. SNP-chips), EST sequence or other resources for SNP discovery, linkage mapping families, linkage disequilibrium mapping pedigrees, test systems, tested animals, DNA samples, databases etc.

6.3 Likelihood of success or failure

A literature review was undertaken to review the development of new technologies and the progress in the application of MAS to genetic improvement for agriculture and aquaculture industries.

6.4 Costs of MAS

Companies and institutes were contacted and an estimate of the costs involved were determined. Note costs are changing rapidly (for the most part downwards). Costs were broken down into those associated with

i. development of polymorphic loci
ii. development of genotyping platforms
iii. development and collection of data from family material
iv. genome scan (genotyping costs)
v. analysis
vi. validation
vii. ongoing testing for marker assisted selection

6.5 Benefits from MAS

Estimates of the current and future industry production value were sourced from industry associations

participating in the CRC. From what was known/estimated about the heritability of the traits, an estimate of the genetic progress likely to be achieved with and without marker information was made for each species. This information was combined with that collected in point 1 above, and used to estimate the benefit likely to accrue both with and without the use of marker information.

6.6 Application of MAS

The literature review carried out to investigate point 6.3 was also used to describe alternative means that have been proposed/trialed for the application of MAS.

7. Results/Discussion

7.1 Strengths, weaknesses, opportunities, threats and risks to selective breeding using marker assisted selection

An overview of the strengths, weaknesses opportunities and threats to selective breeding entities from using marker assisted selection is shown in Table 1. This is a general overview and species specific cases will be discussed later in the report. From this analysis, threats and risks have been analysed further in Table 2. Major issues exist with the cost of quantitative trait loci (QTL) marker discovery, costs of application of MAS, lack of basic resources needed to discover markers, lack of breeding programs into which MAS could be implemented and misinformation/hype over the benefits from the use of marker information. Suggestions for managing these threats and risks are given (Table 2).

Major researchable constraints to the application of MAS are shown in Table 3. Wikipedia definitions of the terms used can be found by clicking on the hyperlinks in the document. Major restraints were associated with the development of resources (which are poorly developed for most aquaculture species), undertaking experiments to find and develop markers, research to determine how to apply marker information in the most efficient way possible and validation of the benefit from MAS (which has been poorly addressed in the past).

Table 1. General strengths weaknesses, opportunities and threats identified for marker assisted selection.

Strengths	Weaknesses
• **Industry perceptions (leading edge)** • **Potential advantage over competitors** • **Allows selection of best individuals (valuable potential broodstock cannot be directly tested for many traits eg. disease resistance or meat quality)** • **Allows early selection before trait manifests** • **Allows parentage to be determined (inbreeding avoidance)** • **Potential economic benefits for growers** • **Potential stimulus for expansion of industry** • **Potential stimulus for new entrants to industry** • **Helps meet increased demand for seafood and reduced volume of supply from wild fishery** • **Future product allows farms to produce more from use of same resources** • **Genetic improvement and benefits compound with each generation of selection** • **Many good researchers with experience in development of markers and MAS in Australia**	• **Mis-information and hype over potential for MAS** • **Expense to discover marker tests** • **Expense to apply marker assisted selection** • **Lack of available resources for discovery of marker tests** • **Lack of breeding programs into which marker assisted selection can be incorporated** • **Lack of basic genetic information (heritability of traits, correlations between traits) affects rate of genetic improvement** • **Size and limits (eg. regulations) on growth of industry and consequent restricted flow of benefits** • **Price farmers are willing to pay for genetically improved stock** • **Above points limit possible revenue stream and reduces attractiveness to investors**

Opportunities	Threats
• **Increase rate of improvement of disease resistance and meat quality traits** • **Break up association between negatively correlated traits (eg. improved growth rate with reduced disease resistance?)** • **Gene expression profiles allow early detection of environmental or nutritional stresses** • **Selective DNA pooling might be more efficient (microsatellite and SNP genome scans)** • **Combine research efforts (eg larger, fewer projects for QTL detection and for development of marker resources)** • **Combine genotyping/sequencing efforts to get a better service price (eg. join international sequencing consortiums)** • **Share genetic expertise across livestock and aquaculture industries (eg. national association)** • **New or niche markets (?)**	• **High AUS$ continues or value of aquaculture products slump** • **Marker information is false or inefficiently applied so that genetic improvement is poor and reputation of marker research is affected** • **Other countries develop technologies and restrict use by Australian industry (patent, in-house secret).**

Table 2. General risk analysis for selective breeding using marker assisted selection.

Threat/Risk	Management
Expense to discover marker tests **Expense to apply marker assisted selection** **Lack of available resources for discovery of marker tests** **→Lack "power" to develop marker tests**	Initiate or join in large collaborative international efforts where appropriate Use latest technology and pick best service provider
Size and limits (eg. regulations) on growth of industry and consequent restricted flow of benefits **Price farmers are willing to pay for genetically improved stock** **High AUS$ continues or value of product slumps**	Accept risk? As improvement is made to the stock that is grown the industry will become more resilient to this risk.
Lack of breeding programs into which marker assisted selection can be incorporated **Lack of basic genetic information (heritability of traits, correlations between traits) affects rate of genetic improvement**	Invest in development of basic selective breeding programs first or in overcoming reproductive or other barriers preventing family production
Mis-information and hype over potential for MAS **Marker information is false or inefficiently applied so that genetic improvement is poor, $'s wasted and reputation of marker research is affected**	Control expectations. Make industry aware of findings of this meta-analysis. Apply rigorous scientific scrutiny over marker proposals.
Other countries develop technologies and restrict use by Australian industry (patent, in-house secret)	Stay ahead, collaborate, stay informed of developments (optimised methods, research, protection)

Table 3. Key research needed in developing selective breeding using marker assisted selection. Costs are ball park estimates based on the budgets for recent projects.

Constraint/opportunity	Research/development needed	Cost of research	Notes
1. Development of tests for polymorphisms	1. Participation or access to ultra-high throughput sequence initiatives 2. Discovery of SNPs 3. Discovery of microsatellites	$30,000	To find QTL densely spaced markers are needed. SNPs are the cheapest, most reliable and most automated for high throughput genotyping. *Budget for transcriptome sequence to identify 100s of thousands of SNPs and some microsatellites (as per boxed example)*
2. Development of family resources for genome scan	Breed , tag (DNA fingerprint) and grow Measure trait(s) Take DNA samples Take RNA samples (if gene expression profiling)	$100,000	To find QTL, family resources are needed. The fish need to be well identified (so that parentage is certain). Appropriate measurements and samples need to be taken.
3. Develop genotyping or expression testing platform	Design and produce SNP chip. Design and produce microarray	Cost included in step 4	An inexpensive, high throughput, accurate and automated genotyping platform is essential
4. Undertake genome scan	Various methods exploiting LD, LDLA or LA	$350,000 (comprehensive)	Cost of genotyping families for the SNP markers developed and analysis to find QTL. *Budget for a comprehensive scan for immediate fine scale mapping of QTL (as per boxed example).*
5. Develop panel for MAS	Design a set of tests and develop a simple/cost effective way to genotype animals	$10,000	Need inexpensive and accurate ways to sample and test animals.
6. How to benefit most from implementation of marker assisted selection	Simulation studies, benefit-cost analysis, additional gain from use of MAS? Test new methods and use with selective breeding Weighting of marker/QTL information in EBV	$25,000	In order to maximize benefits from MAS it is necessary to know how they would most effectively be implemented into the selective breeding program
7. Validate and demonstrate utility of MAS	Apply selection with and without MAS Test genetic response with and without MAS	$50,000	Few published validations of the use of MAS. Need to test how well MAS performs and whether it meets expectations and predictions (*covers cost of implementing MAS, measurement and evaluation of impact*)

7.2 Traits to target for MAS

In general, similar prioritization considerations hold across all the species. These general considerations and some exceptions peculiar to particular species and industries are highlighted in Table 4. In deciding on traits that we should focus on in a breeding program, it is important to consider economic impact, genetic parameters and implications on the potential for genetic improvement (trait of focus and correlated traits), the practicality of measurement and selection, demonstrability (affecting uptake by growers) and the amount of research needed (could the trait be selected now). Most selective breeding programs start very simply (focusing on one or two important traits). Additional traits can be added later as the program develops and some traits, such as meat quality traits, can be monitored to check that no deleterious correlated side effects are encountered.

The development and use of MAS is one way to tackle the traits highlighted in Table 4 that are:

1. Of high economic importance
2. Low heritability
3. Show unfavourable correlations with other traits
4. Difficult, expensive or impractical to directly measure or otherwise manipulate

In general, the choice of traits for selective breeding is influenced by the economic benefit likely to be derived from genetic improvement and ability to demonstrate improvement in the trait to growers (salability). There are three other important factors to consider. The trait should be practical to measure, have high potential for genetic improvement (ie. high heritability and show favourable correlations with other traits of importance) and require little research. If a trait scores well with regard to these considerations, then the trait should be simple, practical and effective to select in the absence of markers. If a trait is of high economic importance and highly demonstrable, but difficult to measure and select, lowly heritable, negatively correlated with improvement of other important traits and/or requires substantial research to be able to measure and select for it in an accurate way, there may be some benefit in considering research for the development of MAS.

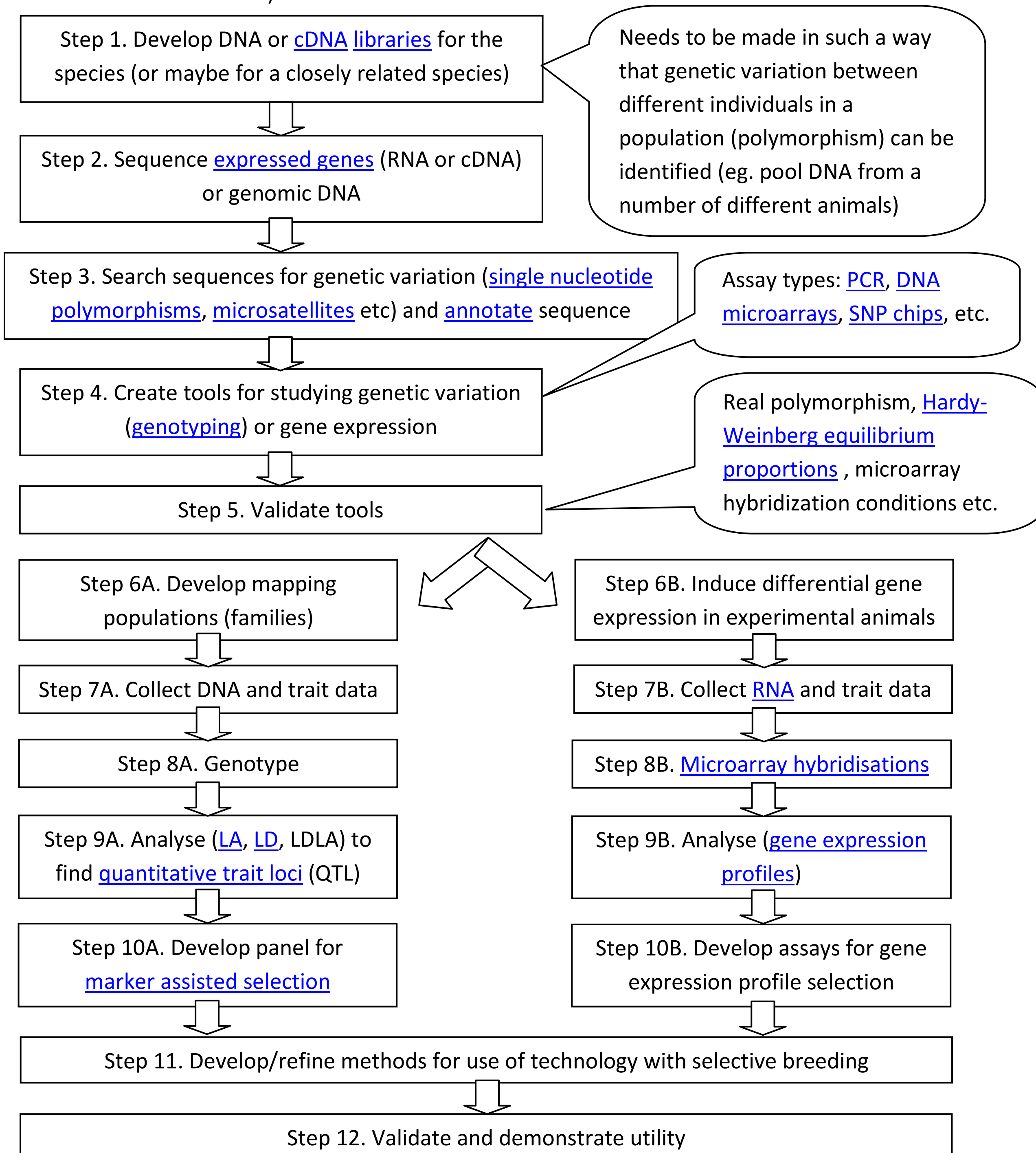
Simulation models have been used to predict which traits and conditions would be the best targeted for marker assisted selection (eg. Hayes et al., 2007a; Sonesson, 2007). Simulation models predict that the efficiency of MAS over more traditional approaches for selective breeding should be greater under the circumstances shown in Table 4.

Table 4. Traits and species that would benefit most from MAS.

Consideration	**Type of trait advantaged**
Unable to pick the best individuals within families without MAS	Disease resistance, meat quality, feed conversion efficiency
Measurement of trait when selected is poorly correlated with trait at harvest	Eg. Abalone growth rate at harvest
Unable to “break up” negative correlation between traits without MAS	Eg. Prawn growth rate and resistance to WSSV
Low heritability	Disease resistance, meat quality
Few breeding candidates available	Eg. Barramundi
Uneven male to female mating ratio	Eg. Kingfish & Barramundi
QTL explains a lot of variation for trait	Eg. Some disease resistance QTL
Other factors affecting benefit from MAS	
Multiple markers linked to QTL	Eg 4 instead of 2 or 1
Frequency of positive QTL allele	

7.3 Resources needed and current status of development

There are several necessary steps that need to be taken to develop a system for marker assisted selection (steps 6B to 10B describe the steps for developing gene expression profile assays (see overview in section 3.4):



7.3.1 Development of DNA or cDNA libraries (Step 1)

DNA or cDNA libraries can be quickly and cheaply produced (eg. a biotech company may take around 1 month to make a normalized and/or subtracted library at a cost of <$10,000). Many services producing libraries are available, or the library might be made by a research provider in Australia.

Box 1. mRNA/cDNA protocol

Illumina has their own protocol for this, however it does not include normalisation http://www.illumina.com/pages.ilmn?ID=291

A common problem appears to be getting too much ribosomal RNA sequence, however it appears that mRNA prep using oligo-dT-25 DynaBeads is an excellent solution. It is possible to purchase the mRNA-seq kit from Illumina themselves (they use the Dynabeads kit apparently). There is a thread on seqanswers with a brief discussion about this http://seqanswers.com/forums/showthread.php?t=1026&highlight=ribosomal

7.3.2 Sequence of expressed genes or genomic DNA (Step 2)

New ultra high throughput and inexpensive methods are available. There are few research providers or companies with the latest ultra high throughput technology, and the technology is changing at an astonishing rate. It is common now to pay for a company or specialist group to provide the sequencing service. In Australia the Australian Genome Research Facility (AGRF) and AgGenomics are examples of service providers. There are also some excellent service providers in Asia (eg. Genotypic in Bangalore India). Examples of the latest technology include SNP discovery using the Illumina Genome Analyser (Solexa, see box below). For some species there are already good DNA/cDNA library resources and many hundreds of thousands of expressed sequence tags that are freely available (eg. Atlantic salmon, Table 5 & Appendix 3). But for most species in aquaculture, there has been no single coordinated effort to develop these resources. It would best if the sequencing effort was part of an international collaboration (ie. competitors could duplicate this effort easily in any case, and costs can be shared).

Box 2. Overview of example of one of the latest strategies for SNP discovery (Matvienko et al., 2008).

1. RNA is purified and rRNA is removed leaving mRNA (eg. dynabead preparation)
2. cDNA library preparation (eg. Clontech SMART kit)
3. Library normalization to reduce representation of highly abundant transcripts (eg. Evrogen trimmer kit)
4. Illumina Genome Analyser (Solexa) paired end sequencing (“massively parallel sequencing-by-synthesis” approach)

The above process could be completed in one week for around AU$ 20- 30,000 and would result in around 105 million reads or around 15 billion bases of sequence. Assuming 50,000 transcribed genes, average size 1kb (50MB = 1X coverage), the latest generation Solexa would yield approximately 200X coverage from a single flow cell (billions of bases of high quality sequence per run). This deep coverage should allow easy discrimination of true SNPs from sequencing errors, and should yield hundreds of thousands of SNPs .

Box 3.
Commercial software for handling EST data
Eg. SeqMan Pro by LaserGene–A complete package with Sequencing Manager, Sequence Assembly, Contig Management, and SNP Discovery

Open source software for handling EST data
http://estpiper.cgb.indiana.edu/ ESTpiper (Tang et al., 2009)
http://www.ebi.ac.uk/~guy/estate/ various
http://seqanswers.com/forums/showthread.php?t=43 various
http://www.ebi.ac.uk/~zerbino/velvet/ Velvet is a de novo genomic assembler specially designed for short read sequencing technologies, such as Solexa or 454. Need about 20-25X coverage and paired reads. Developed by Daniel Zerbino and Ewan Birney at the European Bioinformatics Institute (EMBL-EBI).

Knowing the genes, annotation and comparative mapping
An advantage with sequencing expressed genes for SNP or microsatellite discovery is that you also obtain some information about the type/function of the gene in which the SNP or microsatellite lies. Microsatellites or SNPs in ESTs also provide an opportunity to assemble comparative maps across species. For instance Rexroad et al (2005) has proposed using expressed sequence tags containing microsatellites to compare rainbow trout with salmon genomes.

7.3.3 Search sequence for genetic variation (Step 3)

Box 4. Methods and software for SNP detection
Numerous software and methodology is now available. Here are a few examples:
Deco-tiling (Garvin and Gharrett, 2007)
PreAssemble (Adzhubei et al., 2006)
SNP detection exploiting multiple sources of redundancy (Hayes et al., 2007c)
Single nucleotide discovery in duplicated genomes (Ryynanen and Primmer, 2006)
Targetted locus approach to find SNPs (Sprowles et al., 2006)
SNPIDENTIFIER software (Gorbach et al., 2009)

Box 5. SNP discovery analysis pipelines for handling multiple open source software options
Despite the attractiveness of an integrated software package for SNP discovery, the cost of purchasing such a package is large (eg. 10,000 euros). Open source programs are a better option. Using larger sequence lengths, Ben Hayes developed an analysis pipeline linking the Phred, Phrap and PolyBayes programs. With the smaller sequence lengths for the ultra-high throughput sequencing, this pipeline is being updated to replace Phred, Phrap and PolyBayes programs with:

- Velvet: For *de novo* assembly. http://www.ebi.ac.uk/~zerbino/velvet/
- Gigabayes (new generation of PolyBayes): Short read SNP detection software. http://bioinformatics.bc.edu/marthlab/GigaBayes
- Eagleview: Sequence and SNP viewing program. http://bioinformatics.bc.edu/marthlab/EagleView

7.3.4 Creation of tools for studying genetic variation (Step 4)

Again this is a rapidly developing field of technology. The technology chosen depends on

1. Type and number of markers (eg. SNPs),
2. number of samples (eg. individuals in families) and
3. cost.

The latest trend is to develop and use SNPs for QTL detection. This is because

1. It is becoming much cheaper to genotype individuals using SNPs
2. Little manual checking is required (save time and labour costs)
3. Extremely high throughput is available
4. There are many more SNP markers in the genome than microsatellites, giving denser coverage and greater power of detection

The technology platforms now being used for thorough genome scans (eg. those aiming to genotype 1,000-10,000 SNPs over 1,000-10,000 samples in an aquaculture species) are those supplied by Illumina (Golden Gate or Custom Pannels) or Affymetrix (10K or custom SNPs). Affymetrix now only create custom SNP-chips for species where a very high overall production of chips is likely to be required (eg. human or model species chips).

The latest high throughput SNP genome scan technologies (eg. using llumina SNP chips) are capable of genotyping 10’s of thousands of single nucleotide polymorphisms throughout the genome cheaply and quickly (see boxed example below). Like the sequencing and SNP detection it would be best if the creation of SNP genotyping tools was part of an international consortium. The cost of producing SNP chips is highly dependent on the number of print runs (ie. number of users, number of animals to be genotyped). Also, by including samples from the population you intend to genotype, and by influencing which SNPs are printed on the SNP-chip, you can greatly affect the value of the SNP-chip for your intended purpose.

For instance, SNP chips are available for Atlantic salmon through such an international consortium (led by Norway). CSIRO is currently evaluating the use of these SNP chips for its research with Atlantic salmon. CSIRO have found that only around 3,000 of the 15,000 salmon SNPs are polymorphic in the Australian population of Atlantic salmon (Nick Elliot pers. com.). This is likely to mainly be due to selection bias with polymorphism tending to be highest in the population which you are using to search for the SNP (as is the case for polymorphism detection in microsatellite loci, Ellegren et al., 1997). The salmon SNPs were developed using sequence from Canada and Norway. Use of the SNP-chips so far have revealed that more than 5,800 of the SNPs are polymorphic in some selectively bred Norwegian Atlantic salmon populations.

The new sequencing technologies allow the samples that are sequenced to be tagged or separated, so that SNPs that are polymorphic within multiple populations could now be identified and printed on the chips. SNP chips may also soon be available for *Penaeus monodon* (Norwegian-Indian collaborative project underway), for *Crassostrea gigas* (Chinese initiative to begin in 2009) and for *Haliotis sp*. (discussions to be had at the International Society for Genetics in Aquaculture Conference in Bangkok June 2009). If Australia wishes to make use of these resources, and to make the resources valuable for detecting markers in our populations of animals, we need to contribute, collaborate and become part of these initiatives.

After the genome scan has been performed, and has identified some markers associated with genes affecting traits of importance, a different type of technology can be used to genotype animals for routine marker assisted selection. In performing marker assisted selection, less than 100 markers over more than 1000 samples might be used. The options for routine MAS testing include technologies such as Pyrosequencing, Luminex, TaqMan or SNPlex.

Box 6. Example of a genome scan allowing immediate fine-scale mapping of the QTL
An Illumina SNP-chip (Infinium iSelect SNP-chip) is developed for the species with up to 60,000 SNPs. This gives a SNP density of approximately 1 SNP every 15kb, capturing linkage disequilibrium and enabling immediate fine scale mapping of QTL. 1000 animals (50 families each with 20 individuals) are tested and sampled for the genome scan. Selective genotyping is performed (best 20% of animals and worst 20% of animals for a trait such as disease resistance (time until death with challenge). That is 500 animals are tested for the 60,000 SNPs. The cost of the genotyping are decreasing rapidly, but last year the costs for this genome scan (excluding DNA extractions) would have been somewhere between $350,000-$400,000 (ie. around 1 cent per genotype).

This Illumina technology is well described by (Fan and al., 2003) and by the University of California Davis Genome Centre at
http://www.genomecenter.ucdavis.edu/dna_technologies/illumina.html

7.3.5 Development of mapping population resources (Step 6) and collection of trait data and DNA (Step 7)

To find markers you need to collect data and DNA from large families. This is best done by using the same families, data and DNA that are collected to establish a selective breeding program for the species. Normally, several full-sibling families each consisting of a few hundred siblings and their parents are used. Data should be carefully collected and stored in a database. See sections 7.4.4 and 7.4.5 for further discussion.

Table 5. Summary of state of resources for species cultured in Australia (and related species for comparison where useful). Colours highlight large scale SNP developments under discussion (yellow), underway (red) or completed (green).

Species	Microsats	SNP s	EST s	Linkage map	Reported QTL discovery or expression profile difference
Haliotis rubra	≈ 150			Yes	Growth rate
Haliotis laevigata	≈90				
Haliotis midae		≈20	≈110		
Haliotis discus	≈180	≈12. New SNP development?	≈841	Yes	Growth related traits, stress tolerance
Haliotis asinina			≈232		
Crassostrea gigas	≈150	≈290. New SNP development?	≈30,000	Yes	Summer mortality (gene expression)
Crassostrea virginica	≈80	≈6,500	≈15,000	Yes	Disease resistance (Dermo/summer mortality-resistance)
Penaeus monodon	>1,000	Under develpmnt	≈40,000	Yes	WSSV resistance
Litpenaeus vannamei	≈500	≈500-1000. New SNP development?	≈180,000	Yes	Genes whose expression is modulated by taura syndrome virus
Penaeus japonicas (kuruma)	≈20		≈4,000	AFLP	Weight and length AFLP
Fenneropenaeus chinensis & merguiensis	≈2000		≈10,443	Yes	WSSV resistance
Seriola lalandi	≈250			Yes	
Salmo salar	≈2000	≈3,000	500,000	Yes 138 mic 304 snp 473aflp	Infectious salmon anaemia, infectious pancreatic necrosis ,body weight condition factor, age of sexual maturation, gyrodactylus., salmon lice resistance.
Onchorynchus mykiss	Lots of resources				
Lates calcarifer	≈350		8,655	Yes	Growth related traits

7.4 Likelihood of success or failure

The likelihood of the success or failure of MAS is going to depend on the particulars of the trait, the power of the QTL scan, the genetic variation explained by the QTL linked to the markers, the frequency of the QTL alleles in the population and many of the other factors already discussed in the sections above. A summary of these factors is shown in Table 6.

Table 6. Factors likely to contribute to the success or failure of the use of MAS

Consideration	**Detail**
Species and its genome	Duplication, linkage disequilibrium, ability to produce pedigreed families
Technology used & resources available	Coverage, scoring accuracy
Genetic basis for trait	Number of genes explaining trait, size of effect of gene variant on trait
Accuracy of information recorded and samples collected	Pedigree relationships, trait performance, sample collection and processing
Experimental design and method of analysis	Multistage design, cost, power to detect real effects
Validation	Ensure association is real. Large scale independent data sets for replication. Meta-analysis. Winners curse!
Effective means of implementation	Ongoing breeding program, genomic EBV's, predictions from simulation

7.4.1 Influence of the species and the composition of it's genome

The extent of linkage disequilibrium, and differences in recombination rates between different genomes can have a large effect on the power to detect quantitative trait loci and on the effect of MAS. For instance, Atlantic salmon have been found to have the largest difference in recombination rate between male and female gametes of any vertebrate species (Moen et al., 2004a). However, this difference can be exploited to improve the likelihood of finding QTL and to speed MAS and its application to genetic improvement in this species. For instance, an initial genome scan using a few markers per linkage group and analyzing segregation from male parents could be used to give a quick and efficient indication of the linkage groups containing QTL (Moen et al., 2004a). Simulation can be used to determine how to best exploit these situations, as has been done for Atlantic salmon (Hayes et al., 2006).

Duplication of the Atlantic salmon and rainbow trout genome also affects the likelihood of success or failure for finding QTL. For rainbow trout 11 homeologous duplicated regions (linkage groups) have been detected (Guyomard et al., 2006). Gene orders are conserved between the duplicated groups and each duplicated group is located on a single chromosome arm. After a major duplication event, the duplicated

genes tend to gradually diverge. Polymorphism in such regions may become fixed over time and result in two diverging gene functions. When such polymorphism is detected, all individuals in the population may appear as heterozygous for the polymorphism. Such polymorphisms do not segregate in meiosis and are not useful for mapping purposes. The effects of duplication can also reduce the success of finding polymorphism that differs between different individuals. So although genome duplication could be an important source of polymorphism that contributes to trait effects and changing functional roles of different genes (eg. Myostatin and gene duplication in teleosts, Østbye et al., 2007), it tends to make the identification of markers more difficult and costly.

7.4.2 Technology used and resources available

The marker density and marker type used to scan the genome for QTL has a large influence on the likelihood of success or failure. Marker density affects the likelihood of detecting QTL, and once detected, affects the ability and accuracy of MAS. This is because the denser the set of markers that is used, the higher the chance that a marker will occur in close proximity to the QTL. The size and structure of the families used for mapping the trait also affects the chances of success of detecting QTL affecting the trait.

Human disease research is driving the development of new genotyping technologies and techniques, and has opened up new possibilities for the creation of marker resources. For example, Affymetrix and Illumina produce human SNP chips with 500 thousand or 1.1 million SNP tests respectively. A 500K SNP chip was recently used to screen pooled human DNA samples to detect several loci associated with general cognitive ability (Butcher et al., 2008). Each of the SNPs accounted for less than 0.4% of the variance for this trait, but predictions of the trait value made using the combined 6 SNP set correlated 0.11 ($P<0.00000003$) with the actual measured trait value. Similarly a 100K SNP set was used to detect several genes for reading disability and ability using pooled DNA from a sample of over 5,000 children (Meaburn et al., 2008). These same technology platforms are now being used for genotyping some fish species.

A "complete genome sequence" for the species or related species can also help, particularly when it comes to doing finer mapping or identifying which genes lay in a particular chromosome region. Fully sequenced commercial fish species include *Sparus aurata*, *Dicentrarchus labrax* and *Oreochromis spp*. Genome information is being linked together using comparative genomics across model and non-model teleost species (Sarropoulou et al., 2008). Other complete genome sequences also underway include the Atlantic salmon genome and the Atlantic cod genome (both due for completion later this year).

Knowledge about key genes involved in particular processes, such as disease resistance, and the study of genetic variation in these genes can also influence the likelihood of success or failure. For instance, the major histocompatibility complex (MHC) has been found to be associated with disease resistance in Atlantic salmon (Bridle et al., 2006; Chen et al., 2004; Grimholt et al., 2003; Johnson et al., 2008; Jørgensen et al., 2007a; Jørgensen et al., 2006; Jørgensen et al., 2007b; Kjøglum et al., 2005; Kjøglum et al., 2006; Lukacs et al., 2007; Mjaaland et al., 2005; Morrison et al., 2006; Ozaki et al., 2007).

7.4.3 Genetic basis for trait

Heritability. The higher the traits heritability, the greater the likelihood that you will find a QTL affecting the trait. As mentioned before though, if the heritability is low, MAS may be advantaged over

conventional selective breeding.

Size of QTL effect. Also, the more genetic variation your set of markers is able to explain for the trait, the more useful those markers are likely to be for MAS and the higher the likelihood of success. The amount of genetic variation the markers can explain depends on, 1) the size of the effect on the trait of the genes linked to the markers, and, 2) the proximity of the markers to the genes (extent of recombination between the marker and gene).

Frequency of favourable/unfavourable alleles. If there is a mutation that has a strong effect, but it is found in very few individuals in the population, then you either need luck and a large number of families to find it, or you need to choose families you suspect are segregating for the allele or the trait

7.4.4 Accuracy of information recorded and samples collected

If accurate records are not maintained, or if errors are made (swapped identifications etc), the ability to find useful markers will be compromised. Where manual entry of data occurs, mistakes will always be made. There are many tools available now for automating the collection and entry of data (eg weight, length, pit tag identification, colour, fat etc). It is also important to have a reliable database system for keeping track of the data, organizing data for analysis, detecting errors etc.

In some instances there is a need to develop techniques and technology in order to be able to simply measure the trait.
Eg 1. Disease resistance. The most common way to measure disease resistance is by challenging animals to the disease and measuring either whether they die or how many days they survive over the course of the experiment. However, these experiments require a high degree of biosecurity and extensive tank infrastructure, etc to perform the tests. There are currently no suitable facilities available in Australia for performing such tests on the scale needed for a selective breeding program. Also, for each disease, research is needed to determine how to propagate the virulent strain of the disease in culture, the route of infection for the challenge test (which needs to mimic the natural route of infection as closely as possible) and the dose.
Eg. 2 Measuring feed conversion efficiency for individual fish. There is currently no suitable (efficient) technology available for measuring feed conversion efficiency for large numbers of individual fish.

Another issue is that there is often little correspondence between what is measured (eg. experimental challenge test where the route of infection might be by injection etc) and the actual desired trait (eg. survival to the actual disease outbreak in a sea cage where the route of infection might be by ingestion). Therefore when we select families that have a high probability of surviving the challenge test conditions, are they the same families that will have higher rates of survival to a disease outbreak at sea? At least with marker tests it should be possible to easily check whether the survivors at sea have the best marker conformation!

7.4.5 Experimental design and method of analysis

Critical factors in designing an experiment to give a high likelihood of finding useful markers are:

Family structure. Need a number of full-(preferably) or half-sibling families. This is one of the most important factors influencing the power of QTL experiments and can be a major constraint for species in aquaculture. There are some practical and biological constraints to producing this resource for some species. Some of the species do not have selective breeding programs underway, so families are not

available. Tagging is difficult for young fish and resources needed to separate families are costly (ie. it may be more efficient to use DNA pedigreeing in some instances).

General design of experiment. Genome scans can be performed in stages in order to save some costs. For instance, the first stage might involve selective DNA pooling where DNA from the best performers is pooled and compared to DNA from the worst performers. Selective DNA pooling is proving to be an efficient and effective way of detecting QTL with microsatellite (Baranski et al., 2007; Mariasegaram et al., 2007) and SNP genome scans (eg. Butcher et al., 2008; Meaburn et al., 2005; Meaburn et al., 2008). Selective DNA pooling could substantially reduce the genotyping costs. However, the same pools will not be effective for finding QTL for multiple traits. Ie. each pool is made of individuals that are extreme for a particular trait. As a second stage after selective DNA pooling, individuals might be genotyped to confirm any QTL associations that are detected (ie. genotyping all of the individuals with the few markers that are associated with QTL in stage 1). In this staged tpe of experimental design, genotyping costs can be saved.

Computer simulation studies can be used to determine how the design of the experiment should be adjusted so that the power of the experiment to detect true markers for QTL can be improved (Baro et al., 2001; Coppieters et al., 1999; Hayes et al., 2005b; Hayes et al., 2006).

Error checking and analysis. Genotype data should be checked for the occurrence of scoring error, null alleles, non-Mendelian inheritance and other anomalies that might lead to false positive or negative effects, and the data set should be edited to remove these anomalies before analysis. There are numerous open source software packages available for detection of QTL. It is important to choose an analysis package that is suited to the design of the experiment undertaken.

7.4.6 Effective means for implementing MAS

Simulation and prediction of economic benefit from MAS
Once markers associated with a trait are detected we need effective ways of integrating marker information into the selective breeding program. Because it is expensive and impractical to run replicate selective breeding programs to test different methods, computer simulation is an effective tool to develop or refine methods for application of MAS and other technologies with selective breeding. Simulation has been used to predict whether it would be beneficial to use gene expression profile information with selective breeding (Robinson and Hayes, 2008), in predicting the outcome of using walkback selection or marker assisted selection with optimum contribution selection (Sonesson, 2005; Villanueva et al., 2002), in predicting the outcome from the general implementation of discovered markers (Gomez-Raya and Klemetsdal, 1999; Hayes et al., 2007a), demonstrating the use of marker information for selection index prediction of genetic response and inbreeding (Dekkers, 2007) and in predicting the power of using many thousands of SNPs on a chip to estimate “genomic EBVs” (Bennewitz et al., 2009; Nielsen et al., 2009) (see boxed examples below). To maximize the likelihood of the successful implementation of MAS, simulation studies are important tools that should be used for determining how MAS should be integrated into the breeding program.

Box 7. Eg. Simulating the use of MAS for growth rate in abalone

A simulation model assuming MAS using 5 marker genes explaining 50% of the genetic variance for growth rate in *H. rubra* abalone (as detected by Baranski et al., 2007) was used to predict the benefit-cost ratio for the use of MAS under different circumstances (Hayes et al., 2007a). The study found that the lower accuracy of selecting breeding candidates on early growth rate rather than late growth rate was more than compensated by the ability to perform 2 rounds of selection instead of one when there were more than 5 progeny per family tested. The largest advantage from the use of marker assisted best linear unbiased prediction (MBLUP) over best linear unbiased prediction (BLUP) occurred when the genetic correlation between early growth and late growth rate was lowest. MBLUP was predicted to be generally more advantageous for disease resistance than for growth rate (although no markers are currently available for disease resistance in abalone). Selective genotyping reduced the extra response from MAS by 2% and gave a large reduction in the number and cost of genotyping. Profits were predicted to be maximized if 75 progeny per family are genotyped. At this point a balance was reached between the benefit from MAS relative to non-MAS and the increased cost of genotyping with higher numbers of progeny.

Box 8. Eg. General comparison of MAS to non-MAS using simulation

Sonesson (2007) has also compared within-family MAS to non-MAS for traits recorded on sibs of the breeding candidates (eg. disease resistance or meat quality traits). Sibs were tested for the trait and genotyped to establish genetic marker effects on the trait. Again, the use of MBLUP and BLUP breeding values were compared. All family members had identical BLUP breeding values in the non-MAS schemes, but MBLUP breeding values differed between family members making within-family selection possible for MBLUP. MAS gave up to twice the genetic gain of non-MAS. The efficiency of MAS was somewhat reduced with,

- higher heritabilities
- frequency of the positive allele <0.5
- less candidates
- more even male:female mating ratios (less half sib family relationships)
- QTL explaining less of the total variation for the trait
- Less markers linked to the QTL (eg. 2 instead of 4)

Rates of inbreeding have also been predicted to be lower for MAS than non-MAS selection because fewer full-sibs would be selected by MAS (Sonesson, 2007).

7.5 Benefit-cost ratio for MAS

7.5.1 Examples from the literature

Factors influencing the benefit-cost ratio for the development and use of MAS are shown in Table 7. Precise prediction of the benefit-cost ratio for all the different traits and species involved in the Seafood CRC is beyond the scope of this review, but some estimates of costs and industry parameters have been used in an effort to determine where the development and use of MAS might have a relatively high benefit-cost ratio (Table 8).

Table 7. Factors influencing the benefit-cost ratio for development and use of MAS

Costs	**Benefits**
Develop breeding program. Find variation. Scan genome. Validate tests. Application of testing	Economic benefit based on industry uptake & earnings, trait value, rate of improvement (generation interval, proportion of variation explained, practicalities etc). Profile of breeding company in face of competition. Stimulation to expand production.
Potential cost offsets & alternative costs	**Potential add-on benefits**
Government innovation research grants and bursaries. Collaborate in international effort. Alternatives to MAS might be inaccurate, low benefit, expensive, difficult and/or negatively affect image.	Export improved seedstock. Attract additional investors/partners.

A ball park estimate of the benefits and costs that might be associated with the future development and use of marker assisted selection is presented in Table 8. Very few genetic resources have been developed for most of the species cultured by participants in the Seafood CRC (with the exception of Atlantic salmon). It was assumed that full funding would be required to cover all costs (ie. potential cost offsets due to the factors listed in Table 7 above, such as collaboration in an international effort to develop SNP-chips, have been ignored). Benefits from MAS are highly dependent on how much variation in the trait is explained by the set of markers, how the markers are incorporated into a selective breeding program, how the selective breeding program is run and the particular economics of the industry. The benefits presented in Table 8 are therefore only roughly indicative based on what has been predicted and realized for other species.

Few studies have tried to predict the benefit-cost ratio from using marker assisted selection (Gomez-Raya and Klemetsdal, 1999; Hayes et al., 2007a) or for selection using gene expression profiling (Robinson and Hayes, 2008). For abalone Hayes et al. (2007a) predicted that the benefit-cost ratio after 1 generation of selective breeding using published markers to select for improved growth rate could be between 3:1 - 4:1. As a result of implementing MAS for improved growth rate the abalone industry could gain additional profit of around $ 0.785 million per year after year 4, but predicted benefit-cost ratios were similar with and without MAS for selection. It was also predicted in this study that use of marker assisted selection to improve disease resistance would give greater advantages than marker assisted selection for growth rate. Using gene expression profiles to select for improved disease resistance was predicted to result in benefit-cost ratios of less than 5:1 after 1 generation of selective breeding of Atlantic salmon in Norway (Robinson and Hayes, 2008). But benefits from selective breeding accumulate with every generation of selection. Salmon selective breeding programs using gene expression profiling could achieve 0.29 Euro added value per kg of fish produced after 10 generations compared to 0.23 Euro added value per kg of fish from selection based on challenge test data alone.

In summary, from the cases that have been reported in the literature, the economic benefits predicted from using MAS or gene expression profiling are substantially higher than those predicted from using selection without MAS or gene expression profiling. It may therefore be more profitable to use MAS in some circumstances. However,

1. none of these studies have considered the cost of developing marker technology,

2. because of the rapid technology advances made over the last year, it is likely that the markers or profiles that are developed in the future will be much more powerful than those that have been modeled so far, and
3. both the cost of developing, and applying the technologies, are decreasing every year.

7.5.2 Predicting the relative benefit-costs for species in the Australian Seafood CRC

Hayes et al (2007a) have predicted that extra gains from MAS for growth rate in abalone could be of the order of 15% and that this could result in 3% additional income for the abalone industry above that expected from selective breeding without MAS. If we assume that 12% improvement per generation is possible for a trait like growth rate without MAS, then based on Hayes et al's predictions, MAS might add an extra 1% improvement or more to the trait. Based on these figures an approximation of relative benefit possible accruing over 10 years from the application of MAS to the abalone, Barramundi, kingfish, oyster, salmon and prawn aquaculture industries was predicted (Table 8). The 10 year benefit was calculated assuming a lead time for the implementation of MAS of 3 years. The benefit from the 1% genetic improvement attributable to MAS on growth rate each generation was assumed to be captured as 1% improvement in the value of the production of the industry every year (using latest industry value estimates as a baseline) less some increased feed costs (the average price per kg of feed varies between industries, but for the calculation we assumed feed costs of $2.50/kg and feed conversion ratio of 2).

The main factors affecting benefit from MAS would be the trait targeted for MAS, production volume of the industry, farm gate price per kg and generation interval for MAS. The main factors affecting cost are associated with the development of necessary genetic resources and ongoing sampling, testing and analysis needed to apply the MAS each generation. The trait targeted for MAS is going to have a very large influence on the benefit. An approximation is made for growth rate (Table 8), but for traits like disease resistance the benefit will depend on the frequency of disease outbreaks, economic impact of outbreaks (deaths, reduced production), treatment costs etc.

Table 8. Approximation of additional costs and benefits that would be associated with the development and implementation of MAS for growth rate for industries participating in the Seafood CRC over 10 years (discount rate of zero assumed).

Species	Steps needed (from table 5)	Costs of development (from table 3)	Accumulated cost of implementation	Total accumulated cost	Approximate accumulated benefit (MAS over no MAS)	Recommendation
Abalone	i, ii, iii, iv, v, vii	$0.54 million	$0.15 million	$0.69 million	$2.4 million	Develop selective breeding programs. Become minor participant in developing SNP resources.
Barramundi	i, ii, iii, iv, v, vi, vii	$0.57 million	$0.1 million	$0.67 million	$1.1 million	Develop selective breeding program
Kingfish	i, ii, iii, iv, v, vi, vii	$0.57 million	$0.2 million	$0.77 million	$3.2 million	Develop selective breeding program
Oyster	i, ii, iii, iv, v, vi, vii	$0.57million	$0.2 million	$0.77 million	$2.3 million	Become minor participant in developing SNP resource (*C. gigas*).
Salmon	iv, v, vii	$0.41 million	$0.15 million	$0.56 million	$11.5 million	Undertake marker development, especially for AGD resistance or meat quality traits.
Prawn	i, ii, iii, iv, v, vi, vii	$0.57 million	$0.35 million	$0.92 million	$6.7 million	Become minor participant in developing SNP resource (*P. monodon*). Develop selective breeding programs.

Steps considered above:

i. development of tests for polymorphism (preferably SNPs)
ii. development of family material and collection of data/samples
iii. development of genotyping platforms (preferably SNP-chips)
iv. genome scan (comprehensive scale for immediate fine scale mapping, may necessitate research to develop methodology to measure trait eg. experimental challenge tests for disease resistance)
v. develop panel for MAS
vi. determine how to benefit most from the use of MAS
vii. validation and demonstration of utility (includes cost of implementing MAS)

Salmon and prawn are likely to accrue the largest benefits over 10 years because these are relatively larger industries and, in the case of prawn, the generation interval for MAS is potentially short.

7.5.3 Side benefits from MAS

Traceability

There are a few other benefits that can be derived from the use of marker assisted selection at little additional cost. Traceability is an important issue in all food industries now and strict rules are enforced by the European Union and other trading partners with Australia so that disease or contaminants can be traced back and so that product can be efficiently recalled from the market. It is also important sometimes to verify that labels accurately reflect what is being sold (eg. to test claims over the source of product). DNA samples can be taken and tested from product at virtually any point in the production chain. The same resources as are used to detect markers and perform marker assisted selection (SNP or microsatellite loci) could be used to ensure traceability (Hayes et al., 2005a).

Pedigree records

Accurate pedigree records are of great importance for making genetic improvement in any selective breeding program. Blocks of linked SNPs can be used as highly polymorphic genetic markers for parentage analysis in much the same way as microsatellite loci (Jones et al., 2009). Problems could be traced all the way back to the source family generated within the selective breeding program. Substituted imported product could be detected.

Restricting inbreeding

Markers can also be used to help optimize the genetic gain in the selective breeding program while restricting the rate of inbreeding (eg. Villanueva et al., 2004).

7.6 Application, demonstration and validation of MAS

7.6.1 Validation

Use of new SNP-chip genotyping platforms in human genome scans are now revealing a number of promising associations between variants and human traits. However, for clinical use, and for incorporation and demonstration of utility with selective breeding, such associations need to be thoroughly validated (see Ioannidis et al., 2009 for a review of this subject). Here in this report I only consider the use of marker assisted selection. Identifying the causative genes and mutations affecting quantitative traits requires a good deal of extra research and costs.

For gene expression information in particular, experiments typically collect expression information for many thousands of genes using a very small number of samples for comparison. This is mainly due to the high costs of microarray hybridization. This small number of samples results in a very high likelihood of detecting false positive effects. The validation of this type of data is best done with cross validation utilizing different subsets of animals (eg. Robinson et al., 2008).

For validating genome wide associations, large scale exact replication is necessary. There are a number of ways this can be achieved in practice. One way is to begin a “single-blind” trial where the researcher predicts the trait value or breeding value using the marker information, and the correlation between the predicted and actual phenotype is then tested. This trial must be done using a completely independent data set to that which was used to detect the marker associations in the first place. Alternatively, half the data could be set aside and used to perform the genome scan and find marker associations, and the other half used to validate the associations that are found. Another way to achieve validation is to swap or merge data with researchers conducting similar experiments on the same or a related species

overseas (eg. this might be possible for some salmon, prawn, oyster, kingfish, Barramundi genome scans). The ability to do this depends on whether the data is compatible and can be merged. Finally, multistage genome scan design, such as those using an initial stage in which selective DNA pools are genotyped first to detect associations (eg. Baranski et al., 2008; Mariasegaram et al., 2002), followed by individual genotyping of additional animals for validation, should also be considered. This approach has been shown to be cost effective for microsatellite genome scans and is now being trialed for SNP genome scans in salmon as well (using SNP-chips, CIGENE Norway).

Meta-analysis of data from multiple studies is another way of providing validation (eg. MacLeod et al., 2003). For instance, there have been many studies that have looked for associations with disease resistance in salmonids (eg. Gilbey et al., 2006; Houston et al., 2008; Houston et al., 2006; Johnson et al., 2008; Jorgensen et al., 2008; Khoo et al., 2004; Moen et al., 2004b; Moen et al., 2007; Salte et al., 1993). A meta-analysis could be performed to see if particular chromosomal regions contain markers with consistent associations with disease resistance traits. If enough information was available from such a meta-analysis it might be possible to stage the genome scan to target regions were associations are consistently found first. Such an approach has been clarify where genes affecting milk composition traits might be found in the dairy cattle genome (MacLeod et al., 2003).

Some validation of QTL effects has been carried out for Atlantic salmon disease resistance QTL. A QTL for infectious salmon anaemia (ISA) has been validated (confirmed in a larger new dataset from the same population and a more precise chromosome location determined) (Moen et al., 2007). This QTL has been shown to explain 6%-9% of the phenotypic variation for resistance to the disease. Because the heritability of ISA resistance is moderate (0.19) this QTL explains approximately 32-47% of the additive genetic variance for the trait. The markers were found not to be in population wide association to ISA resistance, so fine mapping would be needed to identify markers in linkage disequilibrium.

Independent validation of a QTL affecting infectious pancreatic necrosis in Atlantic salmon has come from two independent and coincidental genome scans for QTL affecting the disease (http://www.aquagen.no/filestore/QTL-presentation.pdf and Houston et al., 2008)

7.6.2 Application and demonstration

There is little information about the application, realized genetic response or realized economic benefit from the use of marker assisted selection in aquaculture species. This could be for three reasons,

1. There may have been little impact or failed examples from the application of MAS. These failures are unlikely to have been reported.
2. Breeding companies using MAS might see it as too risky to compare their results with competing companies that don't use MAS, and may wish to maintain details about the use of MAS as an in-house secret. The use of new technologies can provide a modern and progressive image for a breeding company. Some may wish to appear to be quick to adopt new technologies, but may actually not use the technologies at all. These companies may benefit greatly from the publicity and image associated with the technologies, while spending little on implementation.
3. There may be few markers of any great value discovered using the technologies developed until recently, little adoption by breeding companies or validation by researchers, and only the most recent discoveries are now just beginning to be incorporated into MAS schemes.

Markers for a single gene affecting infectious pancreatic necrosis (IPN) resistance are now being

used by Aquagen in Norway for selection of Atlantic salmon (http://www.aquagen.no/filestore/QTL-presentation.pdf). The markers explain 80% of the genetic variation in IPN resistance and have the same effect on the resistance of fry as post-smolt. A population level association has been detected (ie. so that the same association and linkage phase with the marker exists across all the families used by Aquagen). Aquagen have found that animals that are homozygous QQ for the marker have a survival rate of 0.87 while qq homozygotes have survival rates of 0.51. They have also shown that selection for IPN resistance by challenge testing has led to a change in the frequency of the high resistance allele from 0.3 in 2005 to 0.5 in 2008. Benefits from the use of this set of markers are,

- Improved accuracy of selection (in combination with phenotypic measurements)
- Possibility for within family selection
- Reduced challenge testing (ie. use marker as a substitute) to improve animal welfare image and reduce costs
- Selection of animals for the production line as well as the selection line (reduced incidence of disease)

There are some good recent examples of the application and demonstration of MAS for livestock selective breeding. Genomic EBVs are now commonly used for selection of dairy cattle (Hayes et al., manuscript). A genomic EBV is calculated as the sum of the effects of dense genetic markers, or haplotypes of the markers, across the entire genome. Genomic EBV's potentially capture all the QTL that contribute to variation in a trait. When used for dairy cattle selection in the USA, New Zealand and Australia, the reliabilities achieved have been 2-20% greater than those of parental average breeding values (the current criteria used to select bull calves for progeny test teams) (Hayes et al., manuscript). This means that breeding companies can market bull teams based on GEBVs at 2 years of age, a strategy that would double the rate of genetic gain in milk composition.

The use of genomic EBVs with selective breeding in aquaculture has high potential when dense genetic markers (eg. SNP-chip resources) are available. The first of these resources for an aquaculture species (Atlantic salmon) only became available last year. Plans for the development of a number of new SNP-chip resources in prawn, oyster and abalone are only now being formulated.

7.6.3 Influence on future strategies

It will be several years before the impact from the application of these latest technological developments (ultra-high thoughput sequencing and genotyping, improved analysis capacity, new ways of utilizing genomic information etc) can be evaluated and demonstrated. But there is no doubt that over the last year there have been very significant developments that should result in much more effective and applicable tools for marker assisted selection, and that it will be several years before the impact of these developments can be demonstrated (in terms of improved genetic response and economic benefit). There are strong indications that the use of these technologies will provide strong benefits for disease resistance, product quality and production efficiency. If the aquaculture industries in Australia choose to sit on the sidelines for too long our competitors in Asia and elsewhere could be reaping the benefits from these new technologies before we even get started. If our competitors are able to obtain a significant head-start (eg. 1-2 generations of implementation of MAS), and given the nature of the step-wise improvement from selective breeding, it would be very difficult for the Australian industries to ever catch up. Therefore we need to at least stay in touch with the latest developments and to be smart about the areas we choose to invest our limited project money into.

8. Benefits and adoption

In this section I am refering to the benefits and adoption of this report for the Australian Seafood Indstries, not to the benefits from the use of MAS per se.

The industries that would benefit would be abalone, Barramundi, prawns, tuna, Yellowtail Kingfish, mulloway, oyster and salmon aquaculture. All sectors of these industries would benefit, but the main benefit would be to growers, hatcheries and breeding organisations.

The benefit of this project will be that each sector will gain an understanding of:

1) Traits that would most benefit from the use of marker assisted selection.
2) The current status of development of genetic resources (polymorphic loci, tested/pedigreed families etc) and what additional resources would be needed before research to discover markers for the species could be undertaken
3) Past examples of success and failure in the application of marker assisted selection.
4) Costs involved to develop necessary resources, discover markers and undertake ongoing testing for marker assisted selection.
5) Benefits likely to accrue from the use of marker assisted selection and how this would compare to other means of selection for the traits of interest.
6) Knowledge of alternative means by which marker assisted selection would be applied to genetic improvement.

By gaining an understanding of these points above, each industry would benefit immediately by being able to prioritise and plan the timing and budget for research, and, if appropriate to their sector, by taking advantage of the latest developments in this area. The ultimate benefit to each industry will be that the industry organisations and selective breeding companies will be able to make informed decisions that maximise the benefit-costs that are derived from genetic improvement.

9. Further development

To realize additional benefits from genetic improvement using marker assisted selection each industry would need to make a substantial investment and commitment to establish research projects. Section 11 (Conclusion) summarises under what circumstances we would be likely to receive a high benefit-cost ratio and what are the most important trigger points for industries participating in the Australian Seafood CRC embarking on the development and application of marker assisted selection.

9.1 What are the big opportunities for the partners in the Australian Seafood CRC?

There may be an opportunity for the Seafood CRC to collaborate (by contributing some funding and DNA) to international efforts that are beginning for the development of SNP-chip resources. For instance, talks have been held about the joint development of genomic resources for *P. monodon*. A joint Norwegian-Indian *P. monodon* SNP-chip initiative is already underway. Chinese groups are looking to begin the genome sequence for *C. gigas*. Involvement in these initiatives allows us to influence the development of these new resources. By contributing DNA, the polymorphisms that are found and

included on the SNP chip will be more likely to be of use for Australian researchers in the future. The costs can be shared between the different countries involved in producing the resource.

There is also an opporunity to undertake a more coordinated world wide effort for genomics research. This type of approach could open up funding opportunities through International Linkage type grants. A colaborative approach would also allow the sharing of resources, data and skills, so that better value for research dollars could be achieved. For example, a joint project on the genomics disease resistance against external parasites such as amoebic gill disease (Australia), sea lice and gyrodactylus (Norway) in Atlantic salmon might identify common QTL or gene pathways.

There may be an opportunity to use genomics to help with early identification of problems in production that might arise further down the value chain. Again this is an area where we would benefit from international colaborations. Nofima last year completed a project characterising the gene expression response to vegetable matter in feeds and changes in feed rations. In France a similar study has been completed this year (Panserat et al.). Some of these nutritional responses have been shown to be linked closely to stress response, disease resistance and fillet quality. Monitoring using genomics in the "whole of chain" could be a way of providing an "early warning" so that management practices can be adjusted to meet the needs further down the value chain.

9.2 Offset of costs

Many of the costs that are highlighted in this report could be offset by government grants, industry levies, FRDC, research institute contributions and the Seafood CRC. The CRC will give priority to research that aims to alleviate constraints affecting a number of its partner industries. Unfortunately, the resources required for MAS are species specific and it is therefore difficult to envisage how resources could be developed as part of a larger research project across two or more different sectors. However it might be possible to come up with common solutions for common problems in applying the use of MAS across the different sectors.

The implementation of MAS requires that a breeding program is underway. If our breeding programs are profitable and competitive entities then it is likely that they themselves will look invest in research projects that will give them an advantage in the market place (such as the development of MAS for some traits). This is what has happened in Norway and the United Kingdom.

10. Planned outcomes

The principal outcome from this scoping study will be an Australian aquaculture industry timing and targeting its research effort in order to utilize the worlds latest advances in technology and available resources to give advantages in terms of a higher rate of genetic improvement, reduced rate of inbreeding, improved disease resistance, improved efficiency and improved profitability. The outputs of this project can be used to make informed decisions on how to proceed with the development of MAS in order to get the best value for our aquaculture industries in Australia.

11. Conclusion

This analysis was undertaken in order to answer a number of questions posed by the Seafood industries participating in the Australian Seafood CRC regarding the development and use of marker assisted

selection technologies. The technologies, methods and consequent power and potential impact of marker assisted selection are rapidly evolving on the back of developments made in human medicine and genetics. Over the last two years some major initiatives to develop resources that will be of use for the development of marker assisted selection for seafood species have been underway. The conclusions from this study are:

1. *Basic resources that are available or needed for each species before research to find markers or to apply these technologies can be carried out.*

Because of the advances in technology, it is now relatively inexpensive to develop a very large set of polymorphic tests (10's of thousands). This large number of polymorphic tests can be quickly developed and families of animals efficiently genotyped to enable fine scale mapping and the identification of potentially useful marker tests for genes affecting economically important traits. These resources are currently only available for Atlantic salmon (*S. salar*), but initiatives are underway to develop resources for black tiger prawn (*P. monodon*) and are being planned for abalone (*Haliotis sp.*) and Pacific oyster (*C. gigas*).

Another important resource that is needed is trait data, pedigree data and DNA collected from individuals in large families (eg. several families each consisting of a few hundred full-siblings and parents). Ideally the families should be part of an existing selective breeding program so that any discoveries are directly applicable to the breeding population. Selective breeding programs are underway in Australia for S*. salar*, *C. gigas*, *S. glomerata*, *H. rubra* and *H. laevigata* and *P. monodon*. Relevant measurements and samples would need to be taken. A database containing this data and pedigree data with reference to stored samples would be needed.

Marker assisted selection also needs to be applied in an optimum way for the species and traits in question so that the benefit-cost ratio from the application of marker assisted selection is maximized. Some simulation is needed to predict how to apply marker information in an optimum way. This has been done for *H. rubra* and *H. laevigata,* and for some traits for *S. salar*.

All other software needed to develop resources, detect and apply MAS is freely available or can be purchased.

2. *Likely cost-benefits for the development and application of these technologies and traits that would be best targeted using the technologies.*

For traits like disease resistance, meat quality or feed conversion efficiency, for which improvement could have a highly beneficial economic impact in each aquaculture sector, and which are all difficult or inefficient to improve using selective breeding without MAS, the benefit-cost ratio for the development and application of MAS technologies is likely to be high. The benefit-cost ratio will be specific to the species, trait, set of markers and approach used. The highest benefits would come from application of MAS to the Atlantic salmon and prawn industries. The benefit derived depends on a number of factors including the value of the industry in Australia and generation interval of the species for marker assisted selection.

3. *When research should be undertaken and developments that might trigger research in this area for each species.*

A major prerequisite for the use of MAS is the existence of a selective breeding program for the species concerned in Australia. Also, the families developed for selective breeding would be the best resource to use for developing MAS technology. It therefore would be sensible to delay the development of MAS for some species (yellow tail kingfish, Barramundi and some prawn species) until selective breeding programs are established. The development of selective breeding programs is therefore one trigger point for beginning to develop MAS.

Although polymorphic tests are available for most of the species included in this study, the only comprehensive single nucleotide polymorphism resource that is available is for *S. salar*. Recent technology developments have lead to efficient means to characterise and screen the genome using tens-of-thousands of these types of polymorphisms. This has meant that we can quickly and efficiently identify some very useful markers. The development and availability of single nucleotide polymorphism-chips (SNP-chips) for the species other than *S. salar*, will be another important trigger point for beginning to develop MAS. It would be beneficial if Australian researchers participated in the development these SNP-chips (contributing DNA from our populations and contributing to the design of the SNP-chip). So another important trigger for beginning research might be the initiation of the development of these resources by researchers overseas.

Finally, MAS will be most beneficial for traits that are of high economic importance to our industries and traits that are otherwise difficult or slow to improve using selective breeding without markers or other means. Therefore, it is recommended that marker research only be undertaken for traits that meet these criteria (eg. disease resistance, meat quality or feed conversion efficiency traits). For disease resistance, we need an effective way of being able to measure an individual's level of resistance before we can find markers. Another trigger to the development of MAS might therefore be the characterization of disease causing organisms affecting production, and the development of mechanisms for their propagation and for controlled experimental challenge tests. A final trigger for the development of MAS might be the development of a significant market driven meat quality issue (an issue that cannot be easily or adequately addressed by manipulating nutrition, processing or grow-out environments).

4. *Measured impact of the application of these technologies on other industries.*

There is very little published work demonstrating the impact of the application of these technologies to other aquaculture or livestock industries. This is partly because, 1) there has been little application in the past, 2) it is difficult to get existing breeding companies to compare or benchmark the results of their breeding work, and 3) because there has not been enough time for the most recent developments in technology to be applied to aquaculture breeding programs and the impact of this application to be evaluated. However, two recent demonstrations exist.

The same marker for infectious pancreatic necrosis resistance in *S. salar* has been discovered and is being applied to selective breeding programs in Norway (Aquagen) and the United Kingdom (Landcatch Natural Selection). This marker has a very large effect on resistance to this disease with Aquagen reporting that animals that are homozygous QQ for the marker have a survival rate of 0.87 while qq homozygotes have survival rates of 0.51. Aquagen have also shown that selection for IPN resistance by challenge testing has lead to a change in the frequency of the high resistance allele from 0.3 in 2005 to 0.5 in 2008.

Application of the use of “genomic estimated breeding values” to dairy cattle selective breeding programs in Australia, New Zealand and the United States have been shown to be 2-20% more

reliable than parental average breeding values, and has allowed companies to market bull teams based on genomic estimated breeding values at 2 years of age, a strategy which doubles the rate of genetic improvement in milk composition.

12. References

Adzhubei, A.A., Laerdahl, J.K., Vlasova, A.V., 2006. preAssemble: a tool for automatic sequencer trace data processing. Bmc Bioinformatics 7.

Adzhubei, A.A., Vlasova, A.V., Hagen-Larsen, H., Ruden, T.A., Laerdahl, J.K., Hoyheim, B., 2007. Annotated Expressed Sequence Tags (ESTs) from pre-smolt Atlantic salmon (Salmo salar) in a searchable data resource. Bmc Genomics 8.

Alcivar-Warren, A., Meehan-Meola, D., Park, S.W., Xu, Z., Delaney, M., Zuniga, G., 2007a. PrawnMap: A low-density, microsatellite-based linkage map of the pacific whiteleg prawn, Litopenaeus vannamei: Identification of sex-linked markers in linkage group 4. Journal of Shellfish Research 26, 1259-1277.

Alcivar-Warren, A., Song, L., Meehan-Meola, D., Xu, Z., Xiang, J.H., Warren, W., 2007b. Characterization and mapping of expressed sequence tags isolated from a subtracted cDNA library of Litopenaeus vannamei injected with white spot syndrome virus. Journal of Shellfish Research 26, 1247-1258.

Alcivar-Warren, A., Meehan-Meola, D., Wang, Y.P., Guo, X.M., Zhou, L.H., Xiang, L.H., Moss, S., Arce, S., Warren, W., Xu, Z.K., Bell, K., 2006. Isolation and mapping of telomeric pentanucleotide (TAACC)(n) repeats of the pacific whiteleg prawn, Penaeus vannamei, using fluorescence in situ hybridization. Mar. Biotechnol. 8, 467-480.

Aoki, T., Hirono, I., Tassanakajo, A., 2005. Expressed sequence tag analysis and construction of cDNA microarray of penaeid prawn, Marsupenaeus japonicus and Penaeus monodon. Aquaculture 247, 4-4.

Araneda, C., Neira, R., Iturra, P., 2005. Identification of a dominant SCAR marker associated with colour traits in Coho salmon (Oncorhynchus kisutch). Aquaculture 247, 67-73.

Baranski, M., Loughnan, S., Austin, C.M., Robinson, N., 2006a. A microsatellite linkage map of the blacklip abalone, *Haliotis rubra*. Anim. Genet. 37, 563-570.

Baranski, M., Rourke, M., Loughnan, S., Austin, C., Robinson, N., 2006b. Isolation and characterization of 125 microsatellite DNA markers in the blacklip abalone, *Haliotis rubra*. Mol. Ecol. Notes 37, 563-570.

Baranski, M., Robinson, N., Loughnan, S., Rourke, M., Hayes, B., Austin, C., 2007. Detection of QTL for growth rate in abalone using selective DNA pooling. Aquaculture 272, S242-S243.

Baranski, M., Rourke, M., Loughnan, S., Hayes, B., Austin, C., Robinson, N., 2008. Detection of QTL for growth rate in the blacklip abalone (*Haliotis rubra* Leach) using selective DNA pooling
Anim. Genet. 39, 606-614.

Baro, J.A., Carleos, C., Corral, N., Lopez, T., Canon, J., 2001. Power analysis of QTL detection in half-sib families using selective DNA pooling. Genet. Sel. Evol. 33, 231-247.

Barroso, R.M., Wheeler, P.A., LaPatra, S.E., Drew, R.E., Thorgaard, G.H., 2008. QTL for IHNV resistance and growth identified in a rainbow (Oncorhynchus mykiss) x Yellowstone cutthroat (Oncorhynchus clarki bouvieri) trout cross. Aquaculture 277, 156-163.

Bennewitz, J., Solberg, T., Meuwissen, T.H.E., 2009. Genomic breeding value estimation using nonparametric additive regression models. Genet. Sel. Evol. 41, 20.

Bester, A.E., Roodt-Wilding, R., Whitaker, H.A., 2008. Discovery and evaluation of single nucleotide polymorphisms (SNPs) for Haliotis midae: a targeted EST approach. Anim. Genet. 39, 321-324.

Boulding, E.G., Culling, M., Glebe, B., Berg, P.R., Lien, S., Moen, T., 2008. Conservation genomics of Atlantic salmon: SNPs associated with QTLs for adaptive traits in parr from four trans-Atlantic backcrosses. Heredity 101, 381-391.

Bridle, A.R., Morrison, R.N., Nowak, B.F., 2006. The expression of immune-regulatory genes in rainbow trout, Oncorhynchus mykiss, during amoebic gill disease (AGD). Fish Shellfish Immunol. 20, 346-364.

Butcher, L.M., Davis, O.S.P., Craig, I.W., Plomin, R., 2008. Genome-wide quantitative trait locus association scan of general cognitive ability using pooled DNA and 500K single nucleotide polymorphism microarrays. Genes Brain and Behavior 7, 435-446.

Castano, C., Yao, J., Rexroad III, C.E., 2008. Single Nucleotide Polymorphism Discovery in Rainbow Trout. . Meeting Abstract.

Castano, C., Smith, T.P., Wiedmann, R.T., Yao, J., Rexroad III, C.E., 2009. Single nucleotide polymorphism discovery in rainbow trout using reduced representation libraries. . Meeting Abstract.

Cesar, J., Zhao, B., Yang, J., 2007. Analysis of expressed sequenced tags from abdominal muscle cDNA library of the pacific white prawn Litopenaeus vannamei. J. Anim. Sci. 85, 481-481.

Cesar, J.R., Zhao, B., Yang, J., 2008. Analysis of expressed sequence tags from abdominal muscle cDNA library of the pacific white prawn Litopenaeus vannamei. Animal 2, 1377-1383.
Chatziplis, D., Batargias, C., Tsigenopoulos, C.S., Magoulas, A., Kollias, S., Kotoulas, G., Volckaert, F.A.M., Haley, C.S., 2007. Mapping quantitative trait loci in European sea bass (Dicentrarchus labrax): The BASSMAP pilot study. Aquaculture 272, S172-S182.
Chen, S., Xu, M., Hu, S., Li, L., 2004. Analysis of immune-relevant genes expressed in red sea bream (Chrysophrys major) spleen. Aquaculture 240, 115-130.
Cheng, P.H., Liu, X., Zhang, G.F., Deng, Y.W., 2006. Heat-shock protein70 gene expression in four hatchery Pacific Abalone Haliotis discus hannai Ino populations using for marker-assisted selection. Aquac. Res. 37, 1290-1296.
Coppieters, W., Kvasz, A., Arranz, J.J., Grisart, B., Riquet, J., Farnir, F., Georges, M., 1999. The great-grand-daughter design: a simple strategy to increase the power of a grand-daughter design for QTL mapping. Genetical Research 74, 189-199.
Danzmann, R.G., Jackson, T.R., Ferguson, M.M., 1999. Epistasis in allelic expression at upper temperature tolerance QTL in rainbow trout. Aquaculture 173, 45-58.
Davis, G.P., Hetzel, D.J.S., 2000. Integrating molecular genetic technology with traditional approaches for genetic improvement in aquaculture species. Aquac. Res. 31, 3-10.
de Freitas, P.D., Pinheiro, A.P., Silva, T.B., Manoel Galetti, P., 2007. In silico analysis of polymorphic microsatellites in penaeid prawn and construction of a free-access database. Genetics and Molecular Biology 30, 1194-1197.
Dekkers, J.C.M., 2007. Prediction of response to marker-assisted and genomic selection using selection index theory. J. Anim. Breed. Genet. 124, 331-341.
Dhar, A.K., Meehan-Meola, D., Breland, V., Carr, W.H., Lotzz, J.M., Alcivar-Warren, A., 2007. Linkage mapping of developmentally expressed cDNAs containing AT-rich elements in prawn (Litopenaeus vannamei) - Differential expression after challenge with Taura Syndrome Virus (TSV). Journal of Shellfish Research 26, 1217-1224.
Dong, S.R., Kong, J., Meng, X.H., Zhang, Q.W., Zhang, T.S., Wang, R.C., 2008. Microsatellite DNA markers associated with resistance to WSSV in Penaeus (Fenneropenaeus) chinensis. Aquaculture 282, 138-141.
Drew, R.E., Schwabl, H., Wheeler, P.A., Thorgaard, G.H., 2007. Detection of QTL influencing cortisol levels in rainbow trout (Oncorhynchus mykiss). Aquaculture 272, S183-S194.
Du, Z., Gorbach, D., Jaramillo, G., Rothschild, M.F., 2009. SNP Discovery For Genetic Map Construction In The Pacific White Prawn, Plant & Animal Genomes XVII Conference, San Diego, CA.
Elfstrom, C.M., Smith, C.T., Seeb, J.E., 2006. Thirty-two single nucleotide polymorphism markers for high-throughput genotyping of sockeye salmon. Mol. Ecol. Notes 6, 1255-1259.
Elfstrom, C.M., Smith, C.T., Seeb, L.W., 2007. Thirty-eight single nucleotide polymorphism markers for high-throughput genotyping of chum salmon. Mol. Ecol. Notes 7, 1211-1215.
Elizur, A., 2009. The KiSS1/GPR54 system in fish. Peptides 30, 164-170.
Ellegren, H., Moore, S., Robinson, N., Byrne, K., Ward, W., Sheldon, B.C., 1997. Microsatellite evolution - A reciprocal study of repeat lengths at homologous loci in cattle and sheep. Mol. Biol. Evol. 14, 854-860.
Evans, B., White, R.W.G., Elliott, N.G., 2000. Characterization of microsatellite loci in the Australian Blacklip abalone (Haliotis rubra, Leach). Mol. Ecol. 9, 1183-1184.
Evans, B., Conod, N., Elliott, N.G., 2001. Evaluation of microsatellite primer conservation in abalone. Journal of Shellfish Research 20, 1065-1070.
Fan, J.B., al., e., 2003. Highly parallel SNP genotyping, Cold Spring Harbour Symposia on Quantitative Biology. Cold Spring Harbour Laboratory Press.
Fleury, E., Daniel, J.Y., Boulo, V., Favrel, P., Mazurais, D., Boudry, P., Moal, J., Huvet, A., 2009. Identification Of Genes Associated With Summer Mortality In The Pacific Oyster Crassostrea gigas Using cDNA Microarrays, Plant & Animal Genomes XVII Conference, San Diego CA.
Funes, S., Hedrick, J.A., Vassileva, G., Markowitz, L., Abbondanzo, S., Golovko, A., Yang, S.J., Monsma, F.J., Gustafson, E.L., 2003. The KiSS-1 receptor GPR54 is essential for the development of the murine reproductive system. Biochem. Biophys. Res. Commun. 312, 1357-1363.
Garcia, D.K., Alcivar-Warren, A., 2007. Characterization of 35 new microsatellite genetic markers for the pacific whiteleg prawn, Litopenaeus vannamei: Their usefulness for studying genetic diversity of wild and cultured stocks, tracing pedigree in breeding programs, and linkage mapping. Journal of Shellfish Research 26, 1203-1216.
García, J.C., Reyes, A., Salazar, M., Granja, C.B., 2009. Differential gene expression in White Spot Syndrome Virus (WSSV)-infected naïve and previously challenged Pacific white prawn Penaeus (Litopenaeus) vannamei. Aquaculture 289, 253-258.
Garvin, M.R., Gharrett, A.J., 2007. DEco-TILLING: an inexpensive method for single nucleotide polymorphism discovery that reduces ascertainment bias. Mol. Ecol. Notes 7, 735-746.
Gharbi, K., Glover, K., Stone, L., MacDonald, E., Matthews, L., Grimholt, U., Stear, M., 2009. Genetic dissection of MHC-associated susceptibility to Lepeophtheirus salmonis in Atlantic salmon. BMC Genetics 10, 20.

Gharbi, K., Gautier, A., Danzmann, R.G., Gharbi, S., Sakamoto, T., Hoyheim, B., Taggart, J.B., Cairney, M., Powell, R., Krieg, F., Okamoto, N., Ferguson, M.M., Holm, L.E., Guyomard, R., 2006. A linkage map for brown trout (Salmo trutta): Chromosome homeologies and comparative genome organization with other salmonid fish. Genetics 172, 2405-2419.

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

Gilbey, J., Verspoor, E., Mo, T.A., Sterud, E., Olstad, K., Hytterd, S., Jones, C., Noble, L., 2006. Identification of genetic markers associated with *Gyrodactylus salaris* resistance in Atlantic salmon *Salmo salar*. Dis. Aquat. Org. 71, 119-129.

Goetz, F.W., McCauley, L., Goetz, G.W., Norberg, B., 2006. Using global genome approaches to address problems in cod mariculture. Ices Journal of Marine Science 63, 393-399.

Gomez-Raya, L., Klemetsdal, G., 1999. Two-stage selection strategies utilizing marker-quantitative trait locus information and individual performance. J. Anim. Sci. 77, 2008-2018.

Gorbach, D.M., Hu, Z.L., Du, Z.Q., Rothschild, M.F., 2009. SNP discovery in Litopenaeus vannamei with a new computational pipeline. Anim. Genet. 40, 106-109.

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

Gross, P.S., Bartlett, T.C., Browdy, C.L., Chapman, R.W., Warr, G.W., 2001. Immune gene discovery by expressed sequence tag analysis of hemocytes and hepatopancreas in the Pacific White Prawn, Litopenaeus vannamei, and the Atlantic White Prawn, L-setiferus. Dev. Comp. Immunol. 25, 565-577.

Gueguen, Y., Cadoret, J.P., Flament, D., Barreau-Roumiguiere, C., Girardot, A.L., Garnier, J., Hoareau, A., Bachere, E., Escoubas, J.M., 2003. Immune gene discovery by expressed sequence tags generated from hemocytes of the bacteria-challenged oyster, Crassostrea gigas. Gene 303, 139-145.

Guyomard, R., Mauger, S., Tabet-Canale, K., Martineau, S., Genet, C., Krieg, F., Quillet, E., 2006. A Type I and Type II microsatellite linkage map of Rainbow trout (Oncorhynchus mykiss) with presumptive coverage of all chromosome arms. Bmc Genomics 7.

Haidle, L., Janssen, J., Gharbi, K., Danzmann, R.G., Ferguson, M.M., 2007. Identifying regions of the genome responsible for early maturation and associated traits in rainbow trout (Oncorhynchus mykiss). Aquaculture 272, S265-S266.

Haidle, L., Janssen, J.E., Gharbi, K., Moghadam, H.K., Ferguson, M.M., Danzmann, R.G., 2008. Determination of quantitative trait loci (QTL) for early maturation in rainbow trout (Oncorhynchus mykiss). Mar. Biotechnol. 10, 579-592.

Hayes, B., Goddard, M.E., 2001. The distribution of the effects of genes affecting quantitative traits in livestock. Genet. Sel. Evol. 33, 209-229.

Hayes, B., Sonesson, A.K., Gjerde, B., 2005a. Evaluation of three strategies using DNA markers for traceability in aquaculture species. Aquaculture 250, 70-81.

Hayes, B., Robinson, N., Baranski, M., Rourke, M., 2005b. Power of QTL detection in abalone using selective DNA pooling. Aquaculture 247, 18-18.

Hayes, B., Baranski, M., Goddard, M.E., Robinson, N., 2007a. Optimisation of marker assisted selection for abalone breeding programs. Aquaculture 265, 61-69.

Hayes, B., Laerdahl, J.K., Lien, S., Moen, T., Berg, P., Hindar, K., Davidson, W.S., Koop, B.F., Adzhubei, A., Hoyheim, B., 2007b. An extensive resource of single nucleotide polymorphism markers associated with Atlantic salmon (Salmo salar) expressed sequences. Aquaculture 265, 82-90.

Hayes, B.J., Gjuvsland, A., Omholt, S., 2006. Power of QTL mapping experiments in commercial Atlantic salmon populations, exploiting linkage and linkage disequilibrium and effect of limited recombination in males. Heredity 97, 19-26.

Hayes, B.J., Bowman, P.J., Chamberlain, A.J., Goddard, M.E., manuscript. Genomic selection in dairy cattle: progress and challenges.

Hayes, B.J., Nilsen, K., Berg, P.R., Grindflek, E., Lien, S., 2007c. SNP detection exploiting multiple sources of redundancy in large EST collections improves validation rates. Bioinformatics 23, 1692-1693.

He, N.H., Liu, H.P., Xu, X., 2004. Identification of genes involved in the response of haemocytes of Penaeus japonicus by suppression subtractive hybridization (SSH) following microbial challenge. Fish Shellfish Immunol. 17, 121-128.

He, N.H., Qin, Q.W., Xu, X., 2005. Differential profile of genes expressed in hemocytes of White Spot Syndrome Virus-resistant prawn (Penaeus japonicus) by combining suppression subtractive hybridization and differential hybridization. Antiviral Research 66, 39-45.

Houston, R.D., Haley, C.S., Hamilton, A., Guyt, D.R., Tinch, A.E., Taggart, J.B., McAndrew, B.J., Bishop, S.C., 2008. Major quantitative trait loci affect resistance to infectious pancreatic necrosis in Atlantic salmon (*Salmo salar*). Genetics 178, 1109-1115.

Houston, R.D., Guy, D.R., Hamilton, A., Ralph, J., Spreckley, N., Taggart, J.B., McAndrew, B.J., Haley, C.S., Bishop, S.C., 2006. Detection of major QTL affecting resistance to Infectious Pancreatic Necrosis (IPN) in a commercial atlantic salmon

population, Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Minas Gerais, Brazil, 13-18 August, 2006. Instituto Prociencia, Minas Gerais, Brazil, pp. 15-02.

Huang, B., Hanna, P.J., 1998. Identification of three polymorphic microsatellite loci in blacklip abalone, Haliotis rubra (Leach), and detection in other abalone species. Journal of Shellfish Research 17, 795-799.

Hubert, S., Hedgecock, D., 2004. Linkage maps of microsatellite DNA markers for the pacific oyster Crassostrea gigas. Genetics 168, 351-362.

Ioannidis, J.P.A., Thomas, G., Daly, M.J., 2009. Validating, augmenting and refining genome-wide association signals. Nat Rev Genet 10, 318-329.

Jackson, D.J., Degnan, B.M., 2006. Expressed sequence tag analysis of genes expressed during development of the tropical abalone Haliotis asinina. Journal of Shellfish Research 25, 225-231.

Jenny, M.J., Ringwood, A.H., Lacy, E.R., Lewitus, A.J., Kempton, J.W., Gross, P.S., Warr, G.W., Chapman, R.W., 2002. Potential indicators of stress response identified by expressed sequence tag analysis of hemocytes and embryos from the American oyster, Crassostrea virginica. Mar. Biotechnol. 4, 81-93.

Johnson, N.A., Vallejo, R.L., Silverstein, J.T., Welch, T.J., Wiens, G.D., Hallerman, E.M., Palti, Y., 2008. Suggestive association of major histocompatibility IB genetic markers with resistance to bacterial cold water disease in rainbow trout (Oncorhynchus mykiss). Mar. Biotechnol. 10, 429-437.

Jones, B., Walsh, D., Werner, L., Fiumera, A., 2009. Using blocks of linked single nucleotide polymorphisms as highly polymorphic genetic markers for parentage analysis. Molecular Ecology Resources 9, 487-497.

Jorgensen, S.M., Afanasyev, S., Krasnov, A., 2008. Gene expression analyses in Atlantic salmon challenged with infectious salmon anemia virus reveal differences between individuals with early, intermediate and late mortality. Bmc Genomics 9.

Jørgensen, S.M., Grimholt, U., Gjøen, T., 2007a. Cloning and expression analysis of an Atlantic salmon (Salmo salar L.) tapasin gene. Dev. Comp. Immunol. 31, 708-719.

Jørgensen, S.M., Lyng-Syvertsen, B., Lukacs, M., Grimholt, U., Gjøen, T., 2006. Expression of MHC class I pathway genes in response to infectious salmon anaemia virus in Atlantic salmon (Salmo salar L.) cells. Fish Shellfish Immunol. 21, 548-560.

Jørgensen, S.M., Hetland, D.L., Press, C.M., Grimholt, U., Gjøen, T., 2007b. Effect of early infectious salmon anaemia virus (ISAV) infection on expression of MHC pathway genes and type I and II interferon in Atlantic salmon (Salmo salar L.) tissues. Fish Shellfish Immunol. 23, 576-588.

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 Oncorhynchus mykiss. Fish Pathol. 39, 95-101.

Kjøglum, S., Grimholt, U., Larsen, S., 2005. Non-MHC genetic and tank effects influence disease challenge tests in Atlantic salmon (Salmo salar). Aquaculture 250, 102-109.

Kjøglum, S., Larsen, S., Bakke, H.G., Grimholt, U., 2006. How specific MHC class I and class II combinations affect disease resistance against infectious salmon anaemia in Atlantic salmon (Salmo salar). Fish Shellfish Immunol. 21, 431-441.

Kong, J., Gao, H.A., 2005. Analysis of tandem repeats in the genome of Chinese prawn Fenneropenaeus chinensis. Chinese Science Bulletin 50, 1462-1469.

Koop, B., Von Schalburg, K., Leong, J., Walker, N., Lieph, R., Cooper, G., Robb, A., Beetz-Sargent, M., Holt, R.A., Moore, R., Brahmbhatt, S., Rosner, J., Mcgowan, C., Rexroad Iii, C.E., Davidson, W., 2008. A salmonid EST genomic study: genes, duplications, phylogeny and microarrays. . . Biomed Central (BMC) Genomics 9, 545.

Koyama, T., Asakawa, S., Katagiri, T., Shimizu, A., Shimizu, N., Fagutao, F.F., Rapeepat, M., Kondo, H., Aoki, T., Hirono, I., 2009. The Presence Of Highly Repetitive Large Fragments In The Kuruma Prawn Genome, Plant & Animal Genomes XVII Conference, San Diego, CA.

Kuntal Mukherjee, N.M., 2009. A Microsatellite DNA Marker Developed for Identifying Disease-resistant Population of Giant Black Tiger Prawn, <i>Penaeus monodon</i>. Journal of the World Aquaculture Society 40, 274-280.

Lallias, D., Lapegue, S., Haley, C., Heurtebise, S., Boudry, P., Beaumont, A., 2007a. A preliminary genetic linkage map for the European flat oyster Ostrea edulis. Aquaculture 272, S281-S282.

Lallias, D., Beaumont, A.R., Haley, C.S., Boudry, P., Heurtebise, S., Lapegue, S., 2007b. A first-generation genetic linkage map of the European flat oyster Ostrea edulis (L.) based on AFLP and microsatellite markers. Anim. Genet. 38, 560-568.

Lehnert, S.A., Wilson, K.J., Byrne, K., Moore, S.S., 1999. Tissue-specific expressed sequence tags from the black tiger prawn *Penaeus monodon*. Mar. Biotechnol. 1, 465-476.

Leu, J.H., Chang, C.C., Wu, J.L., Hsu, C.W., Hirono, I., Aoki, T., Juan, H.F., Lo, C.F., Kou, G.H., Huang, H.C., 2007. Comparative analysis of differentially expressed genes in normal and white spot syndrome virus infected Penaeus monodon. Bmc Genomics 8.

Li, L., Guo, X.M., 2004. AFLP-based genetic linkage maps of the Pacific oyster Crassostrea gigas Thunberg. Mar. Biotechnol. 6, 26-36.

Li, Y.T., Dierens, L., Byrne, K., Miggiano, E., Lehnert, S., Preston, N., Lyons, R., 2006a. QTL detection of production traits for the Kuruma prawn Penaeus japonicus (Bate) using AFLP markers. Aquaculture 258, 198-210.

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

Li, Z.X., Li, J., Wang, Q.Y., He, Y.Y., Liu, P., 2006b. AFLP-based genetic linkage map of marine prawn Penaeus (Fenneropenaeus) chinensis. Aquaculture 261, 463-472.

Liu, X.D., Liu, X., Zhang, G.F., 2007. Identification of quantitative trait loci for growth-related traits in the Pacific abalone Haliotis discus hannai Ino. Aquac. Res. 38, 789-797.

Liu, X.D., Liu, X., Guo, X.M., Gao, Q.K., Zhao, H.G., Zhang, G.F., 2006. A preliminary genetic linkage map of the pacific abalone Haliotis discus hannai Ino. Mar. Biotechnol. 8, 386-397.

Lukacs, M.F., Harstad, H., Grimholt, U., Beetz-Sargent, M., Cooper, G.A., Reid, L., Bakke, H.G., Phillips, R.B., Miller, K.M., Davidson, W.S., Koop, B.F., 2007. Genomic organization of duplicated major histocompatibility complex class I regions in Atlantic salmon (Salmo salar). Bmc Genomics 8.

Lyons, R.E., Dierens, L.M., Tan, S.H., Preston, N.P., Li, Y., 2007. Characterization of AFLP markers associated with growth in the kuruma prawn, Marsupenaeus japonicus, and identification of a candidate gene. Mar. Biotechnol. 9, 712-721.

MacLeod, I., Robinson, N.A., Goddard, M.E., 2003. A concensus map of quantitative trait loci QTL in dairy cattle., Proceedings of the 15th Australian Association for Animal Breeding and Genetics Conference, pp. 22-26.

Maneeruttanarungroj, C., Pongsomboon, S., Wuthisuthimethavee, S., Klinbunga, S., Wilson, K.J., Swan, J., Li, Y., Whan, V., Chu, K.H., Li, C.P., Tong, J., Glenn, K., Rothschild, M., Jerry, D., Tassanakajon, A., 2006. Development of polymorphic expressed sequence tag-derived microsatellites for the extension of the genetic linkage map of the black tiger prawn (*Penaeus monodon*). Anim. Genet. 37, 363-368.

Mariasegaram, M., Robinson, N.A., Goddard, M.E., 2007. Empirical evaluation of selective DNA pooling to map QTL in dairy cattle using a half-sib design by comparison to individual genotyping and interval mapping. Genet. Sel. Evol. 39, 267-283.

Mariasegaram, M., Robinson, N., Balasingham, T., Chamberlain, A., Lade, J., McPartlan, H., Goddard, M., 2002. Mapping genes affecting milk composition in Australian dairy cattle using selective DNA pooling, 7th World Congress on Genetics Applied to Livestock Production, Montpellier France.

Martin, S.A., Caplice, N.C., Davey, G.C., Powell, R., 2002. EST-based identification of genes expressed in the liver of adult Atlantic salmon (Salmo salar). Biochem. Biophys. Res. Commun. 293, PII S0006-0291X(0002)00263-00262.

Matvienko, M., Kozik, A., Michelmore, R., 2008. Evaluation Of Technologies For Constructing And Sequencing cDNA Libraries Plant & Animal Genomes XVI Conference, San Diego, CA.

McClelland, E.K., Naish, K.A., 2008. A genetic linkage map for coho salmon (Oncorhynchus kisutch). Anim. Genet. 39, 169-179.

Meaburn, E., Butcher, L., Liu, L., Fernandes, C., Hansen, V., Al-Chalabi, A., Plomin, R., Craiig, I., Schalkwyk, L., 2005. Genotyping DNA pools on microarrays: Tackling the QTL problem of large samples and large numbers of SNPs. BMC Genomics 6, 52.

Meaburn, E.L., Harlaar, N., Craig, I.W., Schalkwyk, L.C., Plomin, R., 2008. Quantitative trait locus association scan of early reading disability and ability using pooled DNA and 100K SNP microarrays in a sample of 5760 children. Molecular Psychiatry 13, 729-740.

Meehan, D., Xu, Z.K., Zuniga, G., Alcivar-Warren, A., 2003. High frequency and large number of polymorphic Microsatellites in cultured prawn, Penaeus (Litopenaeus) vannamei [Crustacea : Decapoda]. Mar. Biotechnol. 5, 311-330.

Mjaaland, S., Markussen, T., Sindre, H., Kjøglum, S., Dannevig, B.H., Larsen, S., Grimholt, U., 2005. Susceptibility and immune responses following experimental infection of MHC compatible Atlantic salmon (Salmo salar L.) with different infectious salmon anaemia virus isolates. Archives of Virology 150, 2195-2216.

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

Moen, T., Sonesson, A.K., Hayes, B., Lien, S., Munck, H., Meuwissen, T.H.E., 2007. Mapping of a quantitative trait locus for resistance against infectious salmon anaemia in Atlantic salmon (Salmo Salar): comparing survival analysis with analysis on affected/resistant data. Bmc Genetics 8.

Moen, T., Hayes, B., Baranski, M., Berg, P.R., Kjøglum, S., Koop, B.F., Davidson, W.S., Omholt, S.W., Lien, S., 2008. A linkage map of the Atlantic salmon (Salmo salar) based on EST-derived SNP markers. Bmc Genomics 9.

Moghadam, H.K., Poissant, J., Fotherby, H., Haidle, L., Ferguson, M.M., Danzmann, R.G., 2007. Quantitative trait loci for body weight, condition factor and age at sexual maturation in Arctic charr (Salvelinus alpinus): comparative analysis with rainbow trout (Oncorhynchus mykiss) and Atlantic salmon (Salmo salar). Molecular Genetics and Genomics 277, 647-661.

Monwadee Wonglapsuwan, A.P.W.C., 2009. Dynamic changes in gene expression during vitellogenic stages of the white prawn: <i>Fenneropenaeus merguiensis</i> de Man. Aquac. Res. 40, 633-643.
Moore, S.S., Whan, V., Davis, G.P., Byrne, K., Hetzel, D.J.S., Preston, N., 1999. The development and application of genetic markers for the Kuruma prawn Penaeus japonicus. Aquaculture 173, 19-32.
Morrison, R.N., Koppang, E.O., Hordvik, I., Nowak, B.F., 2006. MHC class II+ cells in the gills of Atlantic salmon (Salmo salar L.) affected by amoebic gill disease. Vet. Immunol. Immunopathol. 109, 297-303.
Munasinghe, H., Kang, H.S., Lee, J., 2006. Analysis of digestive gland expressed sequence tag library from the disk abalone, Haliotis discus discus. Journal of the World Aquaculture Society 37, 96-106.
Ng, S.H.S., Chang, A., Brown, G.D., Koop, B.F., 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., Edo, A.F., Wheeler, P.A., Thorgaard, G.H., 2008. The genetic basis of smoltification-related traits in Oncorhynchus mykissi. Genetics 179, 1559-1575.
Nichols, K.M., Broman, K.W., Sundin, K., Young, J.M., Wheeler, P.A., Thorgaard, G.H., 2007. Quantitative trait loci x maternal cytoplasmic environment interaction for development rate in Oncorhynchus myhiss. Genetics 175, 335-347.
Nichols, K.M., Young, W.P., Danzmann, R.G., Robison, B.D., Rexroad, C., Noakes, M., Phillips, R.B., Bentzen, P., Spies, I., Knudsen, K., Allendorf, F.W., Cunningham, B.M., Brunelli, J., Zhang, H., Ristow, S., Drew, R., Brown, K.H., Wheeler, P.A., Thorgaard, G.H., 2003. A consolidated linkage map for rainbow trout (Oncorhynchus mykiss). Anim. Genet. 34, 102-115.
Nielsen, H.M., Sonesson, A.K., Yazdi, H., Meuwissen, T.H.E., 2009. Comparison of accuracy of genome-wide and BLUP breeding value estimates in sib based aquaculture breeding schemes. Aquaculture 289, 259-264.
Nocillado, J.N., Levavi-Sivan, B., Carrick, F., Elizur, A., 2007. Temporal expression of G-protein-coupled receptor 54 (GPR54) gonadotropin-releasing hormones (GnRH), and dopamine receptor D2 (drd2) in pubertal female grey mullet, Mugil cephalus. Gen. Comp. Endocrinol. 150, 278-287.
Ohara, E., Nishimura, T., Nagakura, Y., Sakamoto, T., Mushiake, K., Okamoto, N., 2005. Genetic linkage maps of two yellowtails (Seriola quinqueradiata and Seriola lalandi). Aquaculture 244, 41-48.
Østbye, T.K.K., Wetten, O.F., Tooming-Klunderud, A., Jakobsen, K.S., Yafe, A., Etzioni, S., Moen, T., Andersen, Ø., 2007. Myostatin (MSTN) gene duplications in Atlantic salmon (Salmo salar): Evidence for different selective pressure on teleost MSTN-1 and -2. Gene 403, 159-169.
Ozaki, A., Sakamoto, T., Khoo, S., Nakamura, K., Coimbra, M.R.M., Akutsu, T., Okamoto, N., 2001. Quantitative trait loci (QTLs) associated with resistance/susceptibility to infectious pancreatic necrosis virus (IPNV) in rainbow trout (Oncorhynchus mykiss). Molecular Genetics and Genomics 265, 23-31.
Ozaki, A., Khoo, S.K., Yoshiura, Y., Ototake, M., Sakamoto, T., Dijkstra, J.M., Okamoto, N., 2007. Identification of additional quantitative trait loci (OTL) responsible for susceptibility to infectious pancreatic necrosis virus in rainbow trout. Fish Pathol. 42, 131-140.
Ozakil, A., Khoo, S.K., Yoshiura, Y., Ototake, M., Sakamoto, T., Dijkstra, J.M., Okamoto, N., 2007. Identification of additional quantitative trait loci (OTL) responsible for susceptibility to infectious pancreatic necrosis virus in rainbow trout. Fish Pathol. 42, 131-140.
Pan, D., He, N.H., Yang, Z.Y., Liu, H.P., Xu, X., 2005. Differential gene expression profile in hepatopancreas of WSSV-resistant prawn (Penaeus japonicus) by suppression subtractive hybridization. Dev. Comp. Immunol. 29, 103-112.
Panserat, S., Hortopan, G.A., Plagnes-Juan, E., Kolditz, C., Lansard, M., Skiba-Cassy, S., Esquerré, D., Geurden, I., Médale, F., Kaushik, S., Corraze, G., Differential gene expression after total replacement of dietary fish meal and fish oil by plant products in rainbow trout (Oncorhynchus mykiss) liver. Aquaculture In Press, Accepted Manuscript.
Perez, F., Erazo, C., Zhinaula, M., Volckaert, F., Calderon, J., 2004. A sex-specific linkage map of the white prawn Penaeus (Litopenaeus) vannamei based on AFLP markers. Aquaculture 242, 105-118.
Perez, F., Ortiz, J., Zhinaula, M., Gonzabay, C., Calderon, J., Volckaert, F., 2005. Development of EST-SSR markers by data mining in three species of prawn: Litopenaeus vannamei, Litopenaeus stylirostris, and Trachypenaeus birdy. Mar. Biotechnol. 7, 554-569.
Perry, G.M.L., Ferguson, M.M., Sakamoto, T., Danzmann, R.G., 2005. Sex-linked quantitative trait loci for thermotolerance and length in the rainbow trout. J. Hered. 96, 97-107.
Pongsomboon, S., Wongpanya, R., Tang, S., Chalorsrikul, A., Tassanakajon, A., 2008. Abundantly expressed transcripts in the lymphoid organ of the black tiger prawn, Penaeus monodon, and their implication in immune function. Fish Shellfish Immunol. 25, 485-493.
Preechaphol, R., Leelatanawit, R., Sittikankeaw, K., Klinbunga, S., Khamnamtong, B., Puanglarp, N., Menasveta, P., 2007. Expressed sequence tag analysis for identification and characterization of sex-related genes in the giant tiger prawn Penaeus monodon. Journal of Biochemistry and Molecular Biology 40, 501-510.
Qi, H.G., Liu, X., Zhang, G.F., 2008. Characterization of 12 single nucleotide polymorphisms (SNPs) in Pacific abalone, Haliotis discus hannai. Molecular Ecology Resources 8, 974-976.

Quilang, J., Wang, S.L., Li, P., Abernathy, J., Peatman, E., Wang, Y.P., Wang, L.L., Shi, Y.H., Wallace, R., Guo, X.M., Liu, Z.J., 2007. Generation and analysis of ESTs from the eastern oyster, Crassostrea virginica Gmelin and identification of microsatellite and SNP markers. Bmc Genomics 8.

Ramachandra, R.K., Salem, M., Gahr, S., Rexroad, C.E., Yao, J., 2008. Cloning and characterization of microRNAs from rainbow trout (Oncorhynchus mykiss): Their expression during early embryonic development. Bmc Developmental Biology 8.

Reece, K.S., Ribeiro, W.L., Gaffney, P.M., Carnegie, R.B., Allen, S.K., 2004. Microsatellite marker development and analysis in the eastern oyster (Crassostrea virginica): Confirmation of null alleles and non-Mendelian segregation ratios. J. Hered. 95, 346-352.

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

Rexroad, C.E., Palti, Y., 2003. Development of ninety-seven polymorphic microsatellite markers for rainbow trout. Trans. Am. Fish. Soc. 132, 1214-1221.

Rexroad, C.E., Palti, Y., Gahr, S.A., Vallejo, R.L., 2008. A second generation genetic map for rainbow trout (Oncorhynchus mykiss). Bmc Genetics 9.

Rexroad, C.E., Rodriguez, M.F., Coulibaly, I., 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 Genomics 6.

Robinson, N., Hayes, B., 2008. Modelling the use of gene expression profiles with selective breeding for improved disease resistance in Atlantic salmon (Salmo salar). Aquaculture 285, 38-46.

Robinson, N., Goddard, M., Hayes, B., 2008. Use of gene expression data for predicting continuous phenotypes for animal production and breeding. Animal 2, 1413-1420.

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

Rodriguez, M.F., LaPatra, S., Williams, S., Famula, T., May, B., 2004. Genetic markers associated with resistance to infectious hematopoietic necrosis in rainbow and steelhead trout (Oncorhynchus mykiss) backcrosses. Aquaculture 241, 93-115.

Rojtinnakorn, J., Hirono, I., Itami, T., Takahashi, Y., Aoki, T., 2002. Gene expression in haemocytes of kuruma prawn, Penaeus japonicus, in response to infection with WSSV by EST approach. Fish Shellfish Immunol. 13, 69-83.

Ryynanen, H.J., Primmer, C.R., 2006. Single nucleotide polymorphism (SNP) discovery in duplicated genomes: intron-primed exon-crossing (IPEC) as a strategy for avoiding amplification of duplicated loci in Atlantic salmon (Salmo salar) and other salmonid fishes. Bmc Genomics 7.

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

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

Salem, M., Kenney, P.B., Rexroad, C.E., Yao, J., 2008. Development of a 37 k high-density oligonucleotide microarray: a new tool for functional genome research in rainbow trout. J. Fish Biol. 72, 2187-2206.

Salem, M., Xiao, C., Jonah, T., Rexroad III, C.E., Yao, J., 2009. MicroRNA repertoire for functional genome research in rainbow trout (Oncorhynchus mykiss). . Meeting Abstract.

Salte, R., Gjoen, H.M., Norberg, K., Gjedrem, T., 1993. Plasma-Protein Levels as Potential Marker Traits for Resistance to Furunculosis. J. Fish Dis. 16, 561-568.

Sarropoulou, E., Nousdili, D., Magoulas, A., Kotoulas, G., 2008. Linking the genomes of nonmodel teleosts through comparative genomics. Mar. Biotechnol. 10, 227-233.

Sarropoulou, E., Sepulcre, P., Poisa-Beiro, L., Mulero, V., Meseguer, J., Figueras, A., Novoa, B., Terzoglou, V., Reinhardt, R., Magoulas, A., Kotoulas, G., 2009. Profiling of infection specific mRNA transcripts of the European Seabass Dicentrarchus labrax. BMC Genomics 10, 157.

Saski, C.A., Hilderman, R., Chapman, R.W., Benzie, J.A.H., 2009. Advancing Prawn Genomics, Plant & Animal Genomes XVII Conference, San Diego CA.

Sauvage, C., Bierne, N., Lapegue, S., Boudry, P., 2007. Single Nucleotide polymorphisms and their relationship to codon usage bias in the Pacific oyster Crassostrea gigas. Gene 406, 13-22.

Schalburg, K.R., Rise, M.L., Cooper, G.A., Brown, G.D., Gibbs, A.R., Nelson, C.C., Davidson, W.S., Koop, B.F., 2005. Fish and chips: various methodologies demonstrate utility of a 16,006-gene salmonid microarray. BMC Genomics 6, (15 September 2005).

Schiotz, B.L., Jorgensen, S.M., Rexroad, C., Gjoen, T., Krasnov, A., 2008. Transcriptomic analysis of responses to infectious salmon anemia virus infection in macrophage-like cells. Virus Research 136, 65-74.

Sekino, M., Hara, M., 2007. Linkage maps for the Pacific abalone (genus Haliotis) based on microsatellite DNA markers. Genetics 175, 945-958.

Sekino, M., Kobayashi, T., Hara, M., 2006. Segregation and linkage analysis of 75 novel microsatellite DNA markers in pair crosses of Japanese abalone (Haliotis discus hannai) using the 5 '-tailed primer method. Mar. Biotechnol. 8, 453-466.

Slettan, A., Olsaker, I., Lie, O., 1997. Segregation studies and linkage analysis of Atlantic salmon microsatellites using haploid genetics. Heredity 78, 620-627.

Smith, C.T., Park, L., Vandoornik, D., Seeb, W., Seeb, E., 2006. Characterization of 19 single nucleotide polymorphism markers for coho salmon. Mol. Ecol. Notes 6, 715-720.

Smith, C.T., Baker, J., Park, L., Seeb, L.W., Elfstrom, C., Abe, S., Seeb, J.E., 2005. Characterization of 13 single nucleotide polymorphism markers for chum salmon. Mol. Ecol. Notes 5, 259-262.

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

Sonesson, A.K., 2007. Within-family marker-assisted selection for aquaculture species. Genet. Sel. Evol. 39, 301-317.

Sprowles, A.E., Stephens, M.R., Clipperton, N.W., May, B.P., 2006. Fishing for SNPs: A targeted locus approach for single nucleotide polymorphism discovery in rainbow trout. Trans. Am. Fish. Soc. 135, 1698-1721.

Sun, Z.N., Liu, P., Li, J., Meng, X.H., Zhang, X.M., 2008. Construction of a genetic linkage map in Fenneropenaeus chinensis (Osbeck) using RAPD and SSR markers. Hydrobiologia 596, 133-141.

Supungul, P., Klinbunga, S., Pichyangkura, R., Hirono, I., Aoki, T., Tassanakajon, A., 2004. Antimicrobial peptides discovered in the black tiger prawn *Penaeus monodon* using the EST approach. Dis. Aquat. Org. 61, 123-135.

Taggart, J.B., Bron, J.E., Martin, S.A.M., Seear, P.J., Hoyheim, B., Talbot, R., Carmichael, S.N., Villeneuve, L.A.N., Sweeney, G.E., Houlihan, D.F., Secombes, C.J., Tocher, D.R., Teale, A.J., 2008. A description of the origins, design and performance of the TRAITS-SGP Atlantic salmon Salmo salar L. cDNA microarray. J. Fish Biol. 72, 2071-2094.

Tang, Z., Choi, J.-H., Hemmerich, C., Sarangi, A., Colbourne, J., Dong, Q., 2009. ESTPiper - a web-based analysis pipeline for expressed sequence tags. BMC Genomics 10, 174.

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

Tassanakajon, A., Klinbunga, S., Paunglarp, N., Rimphanitchayakit, V., Udomkit, A., Jitrapakdee, S., Sritunyalucksana, K., Phongdara, A., Pongsomboon, S., Supungul, P., Tang, S., Kuphanumart, K., Pichyangkura, R., Lursinsap, C., 2006. *Penaeus monodon* gene discovery project: the generation of an EST collection and establishment of a database. Gene 384, 104-112.

Thorsen, J., Zhu, B.L., Frengen, E., Osoegawa, K., de Jong, P.J., Koop, B.F., Davidson, W.S., Hoyheim, B., 2005. A highly redundant BAC library of Atlantic salmon (Salmo salar): an important tool for salmon projects. Bmc Genomics 6.

Tian, Y., Kong, J., Wang, W.J., 2008. Construction of AFLP-based genetic linkage maps for the Chinese prawn Fenneropaeneus chinensis. Chinese Science Bulletin 53, 1205-1216.

Tonteri, A., Vasemagi, A., Lumme, J., Primmer, C.R., 2008. Use of differential expression data for identification of novel immune relevant expressed sequence tag-linked microsatellite markers in Atlantic salmon (Salmo salar L.). Molecular Ecology Resources 8, 1486-1490.

Vallejo, R.L., Wiens, G.D., Welch, T.J., Rexroad, C., Palti, Y., 2009. Evidence of major genes for resistance to bacterial cold-water disease in rainbow trout using mixed inheritance multiple-threshold models and Bayesian segregation analysis. . Gordon Research Conference Proceedings.

Vasemagi, A., Nilsson, J., Primmer, C.R., 2005. Expressed sequence tag-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., Pong-Wong, R., Woolliams, J.A., 2002. Marker assisted selection with optimised contributions of the candidates to selection. Genet. Sel. Evol. 34, 679-703.

Villanueva, B., Dekkers, J.C.M., Woolliams, J.A., Settar, P., 2004. Maximizing genetic gain over multiple generations with quantitative trait locus selection and control of inbreeding. J. Anim. Sci. 82, 1305-1314.

Wang, C.M., Lo, L.C., Zhu, Z.Y., Yue, G.H., 2006. A genome scan for quantitative trait loci affecting growth-related traits in an F1 family of Asian seabass (*Lates calcarifer*). BMC Genomics 7.

Wang, C.M., Lo, L.C., Feng, F., Zhu, Z.Y., Yue, G.H., 2008a. Identification and verification of QTL associated with growth traits in two genetic backgrounds of Barramundi (*Lates calcarifer*). Anim. Genet. 39, 34-39.

Wang, C.M., Zhu, Z.Y., Lo, L.C., Feng, F., Lin, G., Yang, W.T., Li, J., Yue, G.H., 2007. A microsatellite linkage map of Barramundi, Lates calcarifer. Genetics 175, 907-915.

Wang, C.M., Lo, L.C., Feng, F., Gong, P., Li, J., Zhu, Z.Y., Lin, G., Yue, G.H., 2008b. Construction of a BAC library and mapping BAC clones to the linkage map of Barramundi, Lates calcarifer. Bmc Genomics 9.

Wang, H.X., Li, F.H., Xiang, H.H., 2005. Polymorphic EST-SSR markers and their mode of inheritance in Fenneropenaeus chinensis. Aquaculture 249, 107-114.

Wang, Y., Ren, R., Yu, Z., 2008c. Bioinformatic mining of EST-SSR loci in the Pacific oyster, Crassostrea gigas. Anim. Genet. 39, 287-289.
Wang, Y.P., Guo, X.M., 2007. Development and characterization of EST-SSR markers in the eastern oyster Crassostrea virginica. Mar. Biotechnol. 9, 500-511.
Whankaew, S., Tangphatsornruang, S., Triwitayakorn, K., 2008. Development of simple sequence repeat markers from expressed sequence tags of the black tiger prawn (Penaeus monodon). Molecular Ecology Resources 8, 1494-1496.
Wilson, K., Li, Y.T., Whan, V., Lehnert, S., Byrne, K., Moore, S., Pongsomboon, S., Tassanakajon, A., Rosenberg, G., Ballment, E., Fayazi, Z., Swan, J., Kenway, M., Benzie, J., 2002. Genetic mapping of the black tiger prawn Penaeus monodon with amplified fragment length polymorphism. Aquaculture 204, 297-309.
Wynne, J.W., O'Sullivan, M.G., Cook, M.T., Stone, G., Nowak, B.F., Lovell, D.R., Elliott, N.G., 2008. Transcriptome analyses of amoebic gill disease-affected Atlantic salmon (Salmo salar) tissues reveal localized host gene suppression. Mar. Biotechnol. 10, 388-403.
Xu, Y., Zhu, Z.Y., Lo, L.C., Wang, C.M., Lin, G., Feng, F., Yue, G.H., 2006. Characterization of two parvalbumin genes and their association with growth traits in Asian seabass (*Lates calcarifer*). Anim. Genet. 37, 266-268.
You, E., Huang, S., Liu, K., Chen, M., Fann, S., Yu, H., 2009. Construction Of Genetic Linkage Maps Of Tiger Prawn (Penaeus monodon) Based On Microsatellite And AFLP Markers, Plant & Animal Genomes XVII Conference, San Diego CA.
Yu, H., Li, Q., 2007. EST-SSR markers from the Pacific oyster, Crassostrea gigas. Mol. Ecol. Notes 7, 860-862.
Yu, H., Li, Q., 2008. Exploiting EST databases for the development and characterization of EST-SSRs in the pacific oyster (Crassostrea gigas). J. Hered. 99, 208-214.
Yu, Z.N., Guo, X.M., 2003. Genetic linkage map of the eastern oyster Crassostrea virginica Gmelin. Biol. Bull. 204, 327-338.
Yu, Z.N., Guo, X.M., 2006. Identification and mapping of disease-resistance QTLs in the eastern oyster, Crassostrea virginica Gmelin. Aquaculture 254, 160-170.
Yue, G.H., Zhu, Z.Y., Wang, C.M., Xia, J.H., 2009. A simple and efficient method for isolating polymorphic microsatellites from cDNA. BMC Genomics 10, 125.
Zeng, D.G., Chen, X.H., Li, Y.M., Peng, M., Ma, N., Jiang, W.M., Yang, C.L., Li, M., 2008. ANALYSIS OF HSP70 IN LITOPENAEUS VANNAMEI AND DETECTION OF SNPS. Journal of Crustacean Biology 28, 727-730.
Zhang, L.S., Yang, C.J., Zhang, Y., Li, L., Zhang, X.M., Zhang, Q.L., Xiang, J.H., 2007. A genetic linkage map of Pacific white prawn (Litopenaeus vannamei): sex-linked microsatellite markers and high recombination rates. Genetica 131, 37-49.
Zhao, Y.M., Li, Q., 2007. Characterization of expressed sequence tag-derived microsatellites from the kuruma prawn, Marsupenaeus japonicus. Mol. Ecol. Notes 7, 1248-1250.
Zhao, Z.-Y., Yin, Z.-X., Weng, S.-P., Guan, H.-J., Li, S.-D., Xing, K., Chan, S.-M., He, J.-G., 2007. Profiling of differentially expressed genes in hepatopancreas of white spot syndrome virus-resistant prawn (Litopenaeus vannamei) by suppression subtractive hybridisation. Fish Shellfish Immunol. 22, 520-534.
Zhu, Z.Y., Wang, C.M., Lo, L.C., Feng, F., Lin, G., Yue, G.H., 2006. Isolation, characterization, and linkage analyses of 74 novel microsatellites in Barramundi (Lates calcarifer). Genome 49, 969-976.

13. Appendix 1. Intellectual property

All of the information generated from this scoping study should be made publically available to the industry.

14. Appendix 2. Staff

Principal Investigator Dr Nick Robinson. Other research investigators, Graham Mair and Alex Safari,.

15. Appendix 3. Detailed literature search highlighting available resources of use for developing MAS

State of microsatellite, SNP and EST development/usage (does not account for reports of microsatellite loci or SNP reported in a few or single loci). Reports of large scale SNP discovery/usage in the Web of Knowledge Database and those in the NCBI database at 11 March 2009 were captured, but many SNPs reported were unconfirmed (ie. may be sequencing or computational errors, not real SNPs) and some reports may have been missed.

Species	Microsatellite number and reference	SNP number and reference	EST number/BAC library and reference	Linkage map	Reported QTL discovery
Haliotis rubra	126 (Baranski et al., 2006b) 24 (Evans et al., 2000; Evans et al., 2001; Huang and Hanna, 1998)			Microsatellite linkage map (Baranski et al., 2006a)	Growth rate (Baranski et al., 2007; Baranski et al., 2008)
Haliotis laevigata	89 *H. rubra* markers cross amplify (Baranski et al., 2006b) 3 *H. rubra* markers cross amplify (Evans et al., 2001)				
Haliotis midae		20 (Bester et al., 2008)	110 (Bester et al., 2008)		
Haliotis discus	75 (Sekino et al., 2006)	12 (Qi et al., 2008)	841 (Munasinghe et al., 2006)	180 microsatellites (Sekino and Hara, 2007) Female 119 and male 94 polymorphic microsatellite RAPD and AFLP markers (Liu et al., 2006)	Growth related traits (Liu et al., 2007) Stress tolerance and heat shock protein (Hsp70) expression (Cheng et al., 2006)
Haliotis asinina			232 (Jackson and Degnan, 2006)		
Crassostrea gigas	147 identified 16 developed from ESTs (Yu and Li, 2008) 10 from ESTs (Yu and Li, 2007) 17 (Wang et al., 2008c)	290 (Sauvage et al., 2007)	30,000 NCBI (13/3/2009) 9,000 unigenes (Fleury et al., 2009) 4,300 ESTs at www.marine genomics.org. 4053 (genbank ??, Yu and Li, 2008) 5132 (genbank June 30 2007 Wang et al., 2008c) 1,142 in 2003 and 1,260 in 2009 (http://www.ifremer.fr/GigasBase Gueguen et al., 2003)	88 microsatellite male and 86 female map (Hubert and Hedgecock, 2004) 119 AFLP female and 96 male map (Li and Guo, 2004)	Summer mortality genes whose expression modulated in resistant vs suceptable animals (Fleury et al., 2009)
Crassostrea virginica	71 amplified from ESTs, 53 polymorphic (Wang and Guo, 2007) 35 from ESTs (Quilang et al., 2007) 18 (Reece et al., 2004)	6,533 putative (Quilang et al., 2007)	15,000 NCBI (13/3/2009) 5,542 (Quilang et al., 2007) 998 embryo & hemocyte (Jenny et al., 2002)	114 AFLP, microsatellite and type I male and 84 female markers (Yu and Guo, 2003)	12 disease resistance (Dermo/summer mortality-resistance) (Yu and Guo, 2006)

Species	Microsatellite number and reference	SNP number and reference	EST number/BAC library and reference	Linkage map	Reported QTL discovery
Ostrea edulis				246 AFLPs and 20 microsatellites (Lallias et al., 2007a; Lallias et al., 2007b)	
Penaeus monodon	997 from 10,000 ESTs (Maneeruttanarungroj et al., 2006) 5 confirmed polymorphic from ESTs (Whankaew et al., 2008) Summarise 300 polymorphic in 12 prawn species (http://www.prawn.ufscar.br de Freitas et al., 2007)		26,170 NCBI (13/3/2009) 150,000 BACs (Saski et al., 2009) 7,809 (genbank ?? Whankaew et al., 2008) 625 Lymphoid (Pongsomboon et al., 2008) 15,981 normal & WSSV infected (Leu et al., 2007) 1,051 ovary (Preechaphol et al., 2007) 10,100 in 2004 and 40,001 in 2009 (http://pmonodon.biotec.or.th and Maneeruttanarungroj et al., 2006; Tassanakajon et al., 2006) ? (Aoki et al., 2005) 1,062 hemocytes (Supungul et al., 2004)	148 microsatellites & 134 AFLPs (You et al., 2009) 144 new markers added including 36 microsateelite containing ESTsAFLPs, SNPs etc (Maneeruttanarungroj et al., 2006) 673 AFLP loci (Wilson et al., 2002)	WSSV resistance (Kuntal Mukherjee, 2009)
Litpenaeus vannamei	35 (Garcia and Alcivar-Warren, 2007) 10 from ESTs (Alcivar-Warren et al., 2007b) 83 genomic and 17 from ESTs (Alcivar-Warren et al., 2007a) 112 from ESTs (multiple species 69% cross amplify, Perez et al., 2005) 136 with 93 polymorphic (Meehan et al., 2003)	1,012 (Du et al., 2009) 504, 44% confirmed (Gorbach et al., 2009) 5 Hsp70 (Zeng et al., 2008)	157000 NCBI (13/3/2009) 150,000 BACs (Saski et al., 2009) 176,000 ESTs at www.marine genomics.org. Prawn microarray available with nearly 22K prawn unigenes (in duplicate) along with suitable controls and landing lights (late Paul Gross). Control spots. Arrays available for purchase at https://earray.chem.agilent.com/earray/PublishDesignLogin.do?eArrayAction=showPreviewForLogin&publishdesignid=PD410656091. 25,937 (genbank ??, Gorbach et al., 2009) 197 (Cesar et al., 2007; Cesar et al., 2008) 268 immune (Gross et al., 2001) 151 various (Lehnert et al., 1999) Many immune system ESTs from subtractive hybridization (García et al., 2009; Zhao et al., 2007) 224 (Alcivar-Warren et al., 2007b)	108 AFLP primer pairs and 30 microsatellites (2071 markers total) (Zhang et al., 2007) 83 microsatellites, 17 ESTs and pentanucleotide repeats (Alcivar-Warren et al., 2007a; Alcivar-Warren et al., 2006) 394 AFLP loci (Perez et al., 2004)	Geneses whose expression is modulated by taura syndrome virus (Dhar et al., 2007)

Species	Microsatellite number and reference	SNP number and reference	EST number/BAC library and reference	Linkage map	Reported QTL discovery
Penaeus japonicas (kuruma)	13 from ESTs (Zhao and Li, 2007)		50,000 BACs (Koyama et al., 2009) 3668 Genbank 2006 (Zhao and Li, 2007) ? (Aoki et al., 2005) 635 normal and WSSV infected haemocytes (Rojtinnakorn et al., 2002) Many immune system ESTs from subtractive hybridization (He et al., 2004; He et al., 2005; Pan et al., 2005)	401 AFLPs (Lyons et al., 2007) 217 AFLP (Li et al., 2003) 129 AFLP (Moore et al., 1999)	Weight and length AFLP (Li et al., 2006a; Lyons et al., 2007)
Fenneropenaeus chinensis & merguiensis	1714 identified ?? validated (Kong and Gao, 2005) 229 from ESTs (Wang et al., 2005)		11 merguiensis NCBI (13/3/2009) 10,443 (Wang et al., 2005) (Monwadee Wonglapsuwan, 2009)	44 microsatellites 460 RAPDs (Sun et al., 2008) 532 AFLPs (Tian et al., 2008) 472 AFLPs 197 marker male map and 194 marker female map (Li et al., 2006b)	WSSV (Dong et al., 2008)
Seriola lalandi	217 identified and validated (Ohara et al., 2005)			217 microsatellites mapped (Ohara et al., 2005)	
Salmo salar	16 confirmed from ESTs (Tonteri et al., 2008) 95 from ESTs (Vasemagi et al., 2005) 1975 identified from GRASP and 61 polymorphic ESTs (Ng et al., 2005) 5 (Slettan et al., 1997)	304 (Moen et al., 2008) 129 (Boulding et al., 2008) 2507 putative (Hayes et al., 2007b) 856 cross validated putative (Hayes et al., 2007c)	494,561 NCBI (13/3/2009) 17K microarray (Wynne et al., 2008) 16K microarray (Taggart et al., 2008) 4K immune (Tonteri et al., 2008) 58K- 15 tissues (Adzhubei et al., 2007) 00K Norwegian & Canadian (Hayes et al., 2007b) 240K (Hayes et al., 2007c) 300K ESTs with 16K on microarray (Schalburg et al., 2005) 733 liver (Martin et al., 2002) 300K ESTS in a database containing ESTs, assemblies, consensus sequences, open reading frames, gene predictions and putative annotation (Koop et al., 2008) cDNA microarray of 32,000 features was created (Koop et al., 2008) BAC library consisting of 300K highly redundant 187kb (average insert size) clones from the Norwegian aquaculture strain giving 18 fold coverage (Thorsen et al., 2005)	EST derived SNP linkage map consisting of 138 microsatellite markers and 304 SNPs located within genes (Moen et al., 2008) Microsatellite linkage map consisting of 54 microsatellites and 473 amplified fragment length polymorphism (AFLP) markers (Moen et al., 2004a) 50 microsatellites and 14 unlinked markers (Gilbey et al., 2004) 5 microsatellites (Slettan et al., 1997)	Infectious salmon anaemia (Moen et al., 2004b; Moen et al., 2007) and differential expression of 306 genes during infection of a macrophage-like Atlantic salmon kidney (ASK) cell line (Jorgensen et al., 2008; Schiotz et al., 2008) Infectious pancreatic necrosis (Houston et al., 2008) Body weight condition factor and age of sexual maturation (Moghadam et al., 2007) Gyrodactylus (Gilbey et al., 2006) Body weight and condition factor (Reid et al., 2005) Salmon lice resistance (Gharbi et al., 2009)
Salmo trutta				288 microsatellites, 13 allozymes (Gharbi et al., 2006)	

Species	Microsatellite number and reference	SNP number and reference	EST number/BAC library and reference	Linkage map	Reported QTL discovery
Onchorynchus keta (chum salmon)		83 (Elfstrom et al., 2006; Elfstrom et al., 2007; Smith et al., 2005)			
Onchorynchus kisutch (coho)		39 (Smith et al., 2006)		148 AFLPs 133 microsatellite (McClelland and Naish, 2008)	Colour traits (Araneda et al., 2005)
char					Body weight condition factor and age of sexual maturation (Moghadam et al., 2007; Reid et al., 2005) Growth rate (Tao and Boulding, 2003)
Onchorynchus mykiss (rainbow trout)	89 from ESTs (Rexroad et al., 2005) 97 microsatellites (Rexroad and Palti, 2003)	30,000 putative (Castano et al., 2009) 90 (Castano et al., 2008)	37K oligonucleotide microarray (Salem et al., 2008) MicroRNAs (Ramachandra et al., 2008; Salem et al., 2009) And many more see PAG XVII conference	191 microsatellites and (Sakamoto et al., 2000) 1,124 microsatellites (Rexroad et al., 2008) 30 ESTs (Rexroad et al., 2005) >900 microsatellites in double haploid gynogenetic lines (Guyomard et al., 2006) 973 AFLPs, 226 microsatellites, 72 VNTR, 38 SINE, 29 known genes, 12 minisatellites, 5 RAPDS and 4 allozymes mapped using androgenic double haploid progeny and clonal lines (Nichols et al., 2003)	Body weight condition factor and age of sexual maturation (Moghadam et al., 2007; Reid et al., 2005) Early maturation/ spawning/smoltification time (Haidle et al., 2007; Haidle et al., 2008; Nichols et al., 2008; Nichols et al., 2007; Sakamoto et al., 1999) Infectious pancreatic necrosis (Ozaki et al., 2001; Ozakil et al., 2007) Thermo tolerance (Danzmann et al., 1999; Perry et al., 2005) Infectious hematopoietic necrosis (Barroso et al., 2008; Khoo et al., 2004; Rodriguez et al., 2004) Bacterial cold water disease (Johnson et al., 2008; Vallejo et al., 2009) Cortisol levels (Drew et al., 2007) Embryonic development rate (Robison et al., 2001)
Gadus (Atlantic cod)			1,361 ovary (Goetz et al., 2006)		

Species	Microsatellite number and reference	SNP number and reference	EST number/BAC library and reference	Linkage map	Reported QTL discovery
Lates calcarifer	240 (Wang et al., 2007) 74 (Zhu et al., 2006) 41 (Yue et al., 2009)		8,655 NCBI (13/3/2009)	240 microsatellites mapped (Wang et al., 2007) 86 BAC library clones (out of 49,000 clones) mapped to linkage map and additional 62 microsatellites (Wang et al., 2008b) 50 microsatellites assigned to linkage groups (Zhu et al., 2006)	Growth related traits (Wang et al., 2006; Wang et al., 2008a) Growth traits and two parvalbumin genes (Xu et al., 2006)
European sea bass			9,605 sequences yielding 3,075 unique sequences from range of tissues from animals infected with *V. anguillarum* and Nodavirus (Sarropoulou et al., 2009)		(Chatziplis et al., 2007) 106 differentially expressed genes common to infecton with *V. anguillarum* and Nodavirus (Sarropoulou et al., 2009)