A white paper on genomics in European Aquaculture Research

Report based on reviews by the AQUAFUNC and AQUGENOME projects, the expert opinions presented at the AQUAGENOME workshop in Bergen on September 20-21 2007 and discussions and conclusions of the AQUAGENOME consortium.

Edited by Patrick Prunet, Antonio Figueras, Kristina Sundell, Deborah Power, Brendan McAndrew, Yann Guiguen, Georgious Kotulas, Niklolai Mugue, Pierrick Haffray, Frank Nilsen, Anna Wargelius and Geir Lasse Taranger.

April 14 2009
1. Introduction

The development of technologies which enables studies of genomes, genes and their products (genomics) has in the last 20 years revolutionized the whole field of biology. The rapid advance in understanding and integrating the outcome of molecular studies with whole organism biology and the refinement of techniques in the field of genomics has recently generated important tools in for example agricultural research which includes aquaculture. Initiated about 15 years ago with the Human Genome project, the development of genomics has more recently reached fish and shellfish species.

In a broad sense, genomics not only cover the sciences of studying the entire DNA content of an organism but also the studies of the RNA (transcriptome) and the protein content (proteome). Genomics can be divided into structural genomics which study the structure, organization and evolution of the genome, and functional genomics which study expression and functions of the genomes. As genome functions are reflected in transcripts and proteins, genomics must also include the study of the transcriptome and the proteome.

The major benefits of genomic approaches are linked to the capability to drive a much broader and systematic analysis of the genome and its expression products. Such perspectives are profoundly modifying our ability to understand biological processes and to propose solutions to key problems in aquaculture production. As such, several research projects have started to take genomic approaches to solve industrially relevant problems.

The aquaculture Industry represents a particular challenge for the implementation of genomic resources. This is related to the structure of the industry which is composed of both large and small producers, its diversity both in terms of species and environment, lack of integration of the sector making implementation and uptake of new tools difficult, lack of standardization of production parameters which makes it high risk and economic sustainability issues.

In recent years major investments have been made to understand the biology of aquaculture species through genomic approaches. The approaches are diverse and cover a large variety of questions related to fish and shellfish aquaculture production, and encompass, issues related to safe and healthy food; optimal growth; selection for desirable traits; fish welfare in aquaculture production; development of vaccines against major diseases and monitoring the impact of aquaculture on the environment. However, the direct application of genomics in the aquaculture industry is still very limited, and genomics and genomic tools are still mainly a research tool for aquaculture.

The current white paper provides a short overview of genomic resources available for European aquaculture species based on an inventory developed within the AQUAGENOME project. It includes: 1) a short analysis of the state-of-the-art of genomic resources and their application to resolve key bottle-necks in European fish and shellfish farming; 2) suggestions for European research priorities in this area; 3) the need for further genomic resource and tool development; and 4) mechanisms for maintenance and access to resources and tools. The white paper is a summation of the outcome of the project AQUAFUNC (state-of-the-art reviews about application of functional genomics in Fp5 and Fp6 and a preliminary analysis of genomic resources in Europe); an in depth inventory of genomics tools (task 1) and a AQUAGENOME workshop in Bergen on September 20-21 2007 assembling around 60 leading scientist in the field. Finally, this information has been analysed and discussed by the AQUAGENOME partners and are summarized in this white paper. The background information on genomic tool and resources and state-of-the-art of their applications in European aquaculture research is publicly available on the AQUAGENOME web.

The following topics are covered in the white paper:

1. Genomics of host-pathogen interaction
2. Genomics in reproduction and breeding programmes
3. Genomics and nutrition impact on growth, development and product quality
4. Genomics and aquaculture-environment interactions
5. Genomic tools and bioinformatic resources
2. State of the art of selected research topics

2.1 Genomics of host-pathogen interaction

Introduction
European aquaculture is menaced by viral, bacterial and parasites (protozoan and ectoparasites) diseases. One of the factors with great influence on the outcome of the research on this field is the diversity of species in culture throughout Europe (huge geographical range) with a whole range of potential pathogens interacting with the commercially important species. The most efficient viral vaccines to date, such as DNA vaccines against the salmonid rhabdoviruses VHSV and IHNV are at the experimental level. A range of efficiently working vaccines has been developed, particularly for some bacterial diseases (e.g. *Vibrio salmonicida* in salmon, *Yersinia ruckeri* in rainbow trout). These could be regarded as potential models to develop functional/structural studies to understand host response against bacterial and viral infections. No commercial vaccines are available against parasite diseases, but ongoing projects using genomic approaches are exploring this possibility on salmon lice. As expected, vaccines are not totally successful, and some vaccines only work efficiently when injected intraperitoneally with strong oil-based adjuvant that often cause unwanted collateral negative effects in the vaccinated fish (e.g. strong and long lasting inflammatory responses, organ adhesions and in some cases skeletal deformities).

As an alternative complementary method of controlling disease immune-stimulants (β-glucans, levamisole) are added in commercial feed but there is a lack of genomic studies used to understand their role in disease resistance. Moreover, genomic approaches can have a large potential in studies on how both innate and acquired immune responses are modulated during critical life-stages and following exposure to various environmental factors and husbandry practices that may comprise immune function and disease resistance. Notably, such approaches can also provide a much better understanding on the links between physiologic stress and immune responses. There is a lack of correlation between laboratory and field experiments, and genomic approach may provide the basic knowledge that is needed for better prediction from laboratory to field situations, e.g. in terms of vaccine efficiency. In general the outcome of host-pathogen interactions (in terms of a naïve infectable host) is dependent on the balance between the host immune response and the virulence of the pathogen.

There is also a lack of functional genomics studies in which all the above factors have been taken into consideration. The involved mechanisms of gene regulation/antimicrobial components or other elements remain to be determined. It is necessary to point out that in invertebrates, of which Europe produces more than a million tons per year (mussels, clams and oysters) cannot be vaccinated. In this case the basis for resistance should be tackled from a genomics perspective. Recent genome sequencing activities implies that now it will be possible to identify gene and gene product homologues in these animals and to develop suitable molecular and immunological (antibodies) tools for studies of structure and function. This is also the case for early life-stages of fish were the immune system is not mature, and hence they cannot be vaccinated at these early stages.

State-of-the-Art
At the structural level there are several tools to study the genetic basis of disease resistance such as linkage maps (limited coverage depending on species), AFLP (Salmonids and some Mediterranean species), QTLs for traits in Atlantic salmon (resistance against viral, bacterial and ectoparasite infections), and radiation hybrid maps have been prepared for sea bass. However, in other species with commercial importance in Europe there are almost no genomic tools available for disease resistance at this level. At the functional level several microarrays are available for disease studies, and range of immune related genes have been identified and studied in the recent years.

Among others we can quote for salmonids TRAITS, the Kuopio array (Finland), INRA, and GRASP. The network Marine Genomics Resources will develop microarrays for some species in consortium with some other European partners. Recently, microarrays (based on ESTs obtained using experimentally challenged individuals), have been or have begun to be developed for carp, turbot (*Scophthalmus maximus*), sea bass (*Dicentrarchus labrax*), cod (*Gadus morhua*), halibut (*Hippoglossus hippoglossus*) and also for molluscs such as the mussel (*Mytilus galloprovincialis*), flat oyster (*Ostrea edulis* and *Crassostrea gigas*) and clam (*Ruditapes decussatus*). The activities in this field could be distributed in the following categories.
Key-issues for future research

**Immunology:**
- Improved understanding of the basic regulation of immune response in the species used (factors involved in both innate and acquired immunity)

**Host-pathogen interactions:**
- Increased understanding of the basic biology of the pathogen
- Increased knowledge on the host-pathogen interaction
- Development of diagnostic tools including genotyping of different pathogen strains and understanding mechanism of virulence

**Disease treatment:**
- Vaccine development and testing
- Disease resistance by selection, including mechanisms underlying genetic and phenotypic differences in disease resistance against various pathogens
- Disease treatments (immunostimulants/chemotherapeutants/antibiotics)

**Conclusions & suggested strategy**
Projects should attempt to reinforce interdisciplinary activity between microbiology/virology and immunology. They should focus on critical developmental stages and increase the knowledge base for immunological responses in the host. Better understanding of the immune system in aquaculture species will significantly improve vaccine development. Also genomic studies of the pathogen will yield better diagnostic tools and better understanding of the infectability of the pathogen. Strategies might also include QTL/SNP-studies on disease resistance based on functional genomics data of host–pathogen interactions. Fish health research can also take advantage of model species (e.g. knock down studies), comparative genomics and *in silico* data data mining (e.g. using zebrafish, medaka and fugu genome data). Other approaches such as monoclonal antibodies and cell lines are valuable tools to be combined with genomic tools, e.g. to model transcriptional networks. It is necessary to point out the crucial need of basic genomics studies in shellfish production, and specifically in disease studies since improvement in disease resistance based on the knowledge of its molecular basis might be the only potentially successful approach. It is also suggested that new projects should be established to integrate results from previous relevant EU funded projects, to explore further specific tools and results obtained in a coordinated manner at a European level.

---

2.2. Genomics in reproduction and breeding programmes

**Introduction**
Controlled reproduction and development of efficient selective breeding programs are two key-issues in the further development of a sustainable aquaculture industry in Europe.

Control of reproduction is the first step to move away from capture based aquaculture, and improved understanding and control of reproduction is essential for production of mono-sex stocks, puberty control, broodstock management, improved gamete quality and production of sterile animals (e.g. polyploids or novel ways of inducing sterility). Such knowledge is essential for reproductive isolation of aquaculture animals (e.g. to avoid genetic or ecological impact on wild stocks), improved welfare (e.g. to avoid maturation related disorders, diseases and mortality) and production of high quality offspring, e.g. by proper broodstock nutrition, environments and handling. Newly available molecular and genomic tools and resources greatly improves the possibility to understand the underlying mechanisms and impact of various endogenous and environmental factors that controls reproductive processes, and can also support development of practical protocols and solutions for reproductive control in the farming industry.

Selective breeding has a huge potential for improving production performance of aquaculture animals, such as improvements in growth, feed conversion efficiency, harvest quality, disease resistance and welfare under farming conditions, as well as later puberty, tolerance to novel feed ingredients and lowered mortality. The new molecular tools and resources that are becoming available can contribute to; implementing selective breeding programs in new species at a lower cost, more efficient selective breeding programs, and notably also to further our knowledge on potential negative side-effects of selective breeding, e.g. loss of broadness in disease resistance or selection of bold individuals with greater potential impact on wild stocks after escape.
State-of-the-art
A range of key-genes related to sex differentiation, puberty and sexual maturation have been identified in fish species such as rainbow trout, salmon and sea bass, and global transcriptome approaches has been applied to study these processes, e.g. in rainbow trout. A range of genetic markers are available in European aquaculture species for implementation in selective breeding programs, ranging from simple genotyping of broodstock and offspring to tools for marker assisted selection. Also a range of key genes are identified connected to breeding relevant processes such as early development, metamorphosis, muscle and skeletal growth, stress responses, immune function, metabolism and reproduction. Combining studies of the expression of such key-genes together with global genomic (QTL, SNP and microarray) approaches will provide novel opportunities to study genotype x environment interactions - and can e.g. contribute greatly to a deeper understanding of the actual changes that takes place during domestication and selective breeding in aquaculture animals.

Key-issues for future research
Based on the above considerations a range of issues have been identified where genomic and molecular approaches are particularly well suited to solve problems related to European aquaculture sustainability within the domains of controlled reproduction and selective breeding. This includes use of such tools and resources in research to provide a better knowledge base and to find practical solutions:

Reproduction:
- Improved understanding of the mechanisms of sex control and differentiation with the aim to produce one sex populations.
- Knowledge for control of age at puberty and/or reproductive isolation/sterility to improve welfare, quality, production performance, and reduce environmental impact of spawning animals and escapees.
- Understand mechanisms controlling maturation, spawning and gamete quality to
- Development of methods in reproductive biotechnology to conserve sperm and eggs or fertilized embryos.

Genetics:
- Develop improved breeding programs which selects for i changing environments (temperatures, diseases, algal blooms, pollution…) and in response to new feeds and other constraint.
- Develop breeding programs which maintain genetic variability to ensure future development in breeding programmes: i.e. to ensure that variability in various traits such as survival, normal development, and broad disease resistance are maintained in farmed animals over time.
- Understand genetics of domestication by combined use of transcriptomal analysis and genetic markers or well defined sibling groups.
- Understand the genetic basis of reproductive disorders and improve reproductive performance in breeding programmes.
- Knowledge for minimising genetic impact of escapees or spawning of farmed animals – while optimising production performance including welfare, growth, quality, feed conversion, late puberty and high survival.

Conclusions & suggested strategy
High priority should be given to the application of the available tools for functional studies on key reproductive issues such as control of sex differentiation, puberty, spawning and gamete quality in a range of European aquaculture species - both in terms of improving the knowledge base and in development of applied solutions to industrial problems including reproductive isolation of farmed animals.

Research to support the implementation in selective breeding programmes and to reduce their costs should have high priority in genetic research. These approaches should be adapted to the status of the breeding programme for each species, the different breeding goals (e.g. growth, yield, flesh quality, feed efficiency and disease resistance), and take into consideration potential side-effects of breeding on normal development, welfare, disease transmission and diversity in disease resistance, as well as minimising risk of negative impact on wild stocks and natural ecosystems.

More genomic and molecular resources and tools should be developed to provide new knowledge and assist implementation of improved selective breeding programmes and methods for reproductive control and transfer of genetic progress in European aquaculture. These developments and approaches should focus on key-bottlenecks in reproductive control and genetic improvement, in making a solid knowledge base for sustainable aquaculture and in practical applications.
2.3. **Genomics and Nutrition impact on growth, development and product quality**

**Introduction**
A large body of scientific results exist on the nutritional requirements of fish to provide an optimal balance between growth/development, health and product quality. Less data are available for shellfish. These results have led to successful improvements of diets and feeding regimes for a wide range of aquaculture species. However, they do not provide a holistic understanding of the key metabolic and physiological processes that determine optimal growth, food utilisation, health, welfare and flesh quality. Moreover, there are many unresolved questions regarding nutritional requirements of early life-stages, including production of live-feeds, early weaning to formulated diets, feeding in extensive pond farming systems as well as shellfish feeding and rearing. Feed and feeding also has a decisive impact on product quality, including chemical composition, texture etc. There is also considerable pressure to develop diets using sustainable and economically practical alternative ingredients. Functional genomics offers a variety of common technological platforms for studying the multiple complex and interacting physiological pathways which determine optimal nutrition in fish and shellfish, and also how nutrition affects growth and flesh quality. It should be noted that knowledge about crustacean or mollusc feeding is limited, and the chapter focuses mainly on the impact of nutrition in finfish aquaculture.

**State-of-the-art**
Over twenty different aquaculture species are now cultured on a commercial aquaculture scale in Europe. Compared to the handful of terrestrial farm animals they represent a very diverse collection, phylogenetically and in terms of life history and nutritional requirements. A major bottleneck to diet formulation and feeding is the lack of basic understanding of metabolism in fish and shellfish. For example, carnivorous fish require a high lipid high protein diet and low carbohydrate and their metabolism is uniquely different to that of homeothermic farm animals. Although, the biochemistry of the key metabolic pathways is conserved, regulation and control differs radically between finfish and higher vertebrates. There is a need to discover which physiological pathways/mechanisms are important in determining health, energy homeostasis, feed conversion ratios and product quality in relation to diet. There are three critical points in the production cycle to consider. First, in the early stages there can be developmental and growth problems related to the quality of the gametes or limitations in diets for larval fish. Second, during on-growing, diets balanced for energy, bulk protein and lipid and micronutrients are critical. Finally, broodstock nutrition is critical for gamete quality, and offspring quality and production performance.

**Key-issues for future research**
Based on the above considerations a range of issues have been identified where genomic and molecular approaches are particularly well suited to solve problems related to nutrition growth and quality:

**Larval development and growth:**
- Improved understanding of first feeding of larvae
- Development of enriched live feed for fish larvae
- Increased understanding of the genetic basis of early growth
- Increased knowledge on the genetic basis of mal development regarding pigmentation, skeletal deformity, and muscle weakness.

**Alternative feed sources:**
- Increased understanding of replacement of fish oils and fish meals in finfish feed
- Improved understanding of how substitutions impacts on fish welfare
- Increased knowledge on how alternative feed may change the value, yield and quality of fish as human food
- To understand the genetic basis of possible physiological problems for the fish when fed alternative lipids, such as essential fatty acid deficiency, possible deficiency in other soluble lipid components (vitamins and cholesterol) and accumulation of persistent organic pollutants (POPS)
- To understand the genetic basis of possible physiological problems after protein replacement

**Broodstock nutrition:**
- To gain increased knowledge on how nutrition impact on broodstock health and welfare, as well as spawning and gamete quality in both fish and shellfish
• Application of molecular tool to investigate the underlying mechanism related to different feeding regimes, as well as the interactions with different rearing environments, genotypes and husbandry practises that affect sexual maturation and reproductive performance of the broodstock.

Conclusions & suggested strategy
In the area of nutrition, a priority is to try and understand the molecular basis of basic fish physiology, this could be achieved using microarrays and the currently developing new sequencing technologies to assay global transcriptome changes which will enable the discovery of affected metabolic pathways in response to specific nutritional regimes. It is considered that it should be possible to formulate an experimental approach based on comparing fish under varying nutrition regimes (for example fish grown on fish meal/oil versus fish on vegetable oil/meal) to address for example, energy homeostasis and metabolic control, disease resistance and larval development.

The application of genomic tools would enable global gene expression to be profiled under these conditions, thus allowing the formulation of testable physiological hypotheses. A further priority would be the advancement of QTL mapping technologies to enable the selection of fish and shellfish which are suited to particular diets or have a more efficient metabolic performance. A long term strategy could then be to implement this knowledge into breeding programs or alternatively this knowledge could be used adapt feed sources to the particular fish and shellfish nutritional requirements.

2.4 Genomics and aquaculture-environment interactions

Introduction
Unlike wild fish that can migrate to optimise their environment and live in large well buffered water volumes, aquaculture fish are confined and for this reason are probably subject to more stress. Environmental constraints therefore have a significant impact on the sustainability of aquaculture and influences production at all stages, hatchery, grow-on and broodstock management. Add to this the growing concern about the impact of aquaculture on the environment which may include, effects of chemicals released from aquaculture, genetic effects on natural populations etc. the extreme important of this issue and the potential gains from improved genomics tools is immediately evident. It is clear that considerations about genomics and the environment will inevitably cut across several different research domains and touch on inter-related issues. As a consequence of the broad range of issues covered by considerations about the environment, advances in genomics can be expected to have multiple impacts on a range of different domains which include but are not exclusive:
- Fish welfare and stress
- Adaptation and plasticity
- Xenobiotics and fish quality
- Monitoring of the impact of aquaculture on the environment (for example: Genetic contamination (gene flow between wild and aquaculture populations)

There are a number of different ways in which advances in fish and shellfish genomics may contribute to aspects associated with the environment. A number of key approaches may be envisaged which include development of “Genomic bio-indicators” which might be identified in the context of stress and fish quality, adaptation, response/detection of xenobiotics, and gene flow and impact on the environment. However, the lack of basic information about why and how fish respond to environmental challenges, including the stress response, and the impact of aquaculture on the environment makes the identification of robust bio-indicators difficult and the generation of such markers is clearly not possible in the short term. Unquestionably genomics will bring significant new basic information and potentially new perspectives which will contribute to improved monitoring and a better understanding of fish/environmental interactions.

State-of-the-art
Biological responses of fish exposed to environmental (biotic or abiotic) factors have been extensively studied at the level of tissue and blood mainly in situations in which physiological stress responses were observed. However, relatively few studies have examined the molecular mechanisms underlying the physiology of the stress response, and our understanding of the mechanisms underpinning the response is generally fragmented. New screening methods potentially provide a system-wide assessment of response which should not only generate overviews of the biological pathways and processes involved in the stress responses, but also provide the means to identify new candidate genes. Gene expression screening approaches using microarray technology has recently been applied to the study of stress responses in fish exposed to stressors such as temperature, hypoxia, confinement, xenobiotics or salinity. An overview of recent functional genomics analysis
of the stress responses in fish indicates that, although application of these methodologies is presently in its infancy, on-going research in this field should lead to considerable new knowledge among which will be indicators of stress with valuable diagnostic and/or predictive value. However, these data also confirm the complexity of the stress responses whatever the nature of the stressor. Both transcriptomic and proteomic information are required to generate an integrated picture of the mechanisms involved in these stress responses. If potential stress indicators identified through gene expressional surveys could be of interest for genetic selection is still a working hypothesis. However gene expression surveys have revealed many candidate factors/pathways which may contribute and will contribute to the understanding of complex traits underlying stress, health, adaptation and plasticity of fish. Gene expressional surveys studies have led to the identification of a transcriptome signature which could be used to assess exposure to xenobiotics and provide information about genes and gene networks implicated in the response to xenobiotics and underlying the physiological response. This should lead to the improved understanding at the molecular level of the response of fish to xenobiotics. Impact of fish farming activities for aquaculture management, stock conservation and environment safety have benefited from studies mainly focused on the genetic aspects. The occurrence of genetic contamination of wild stocks by aquaculture fish is well established for salmon but remains relatively poorly characterised and is virtually unstudied in other aquaculture species. Further, comparison between wild and domesticated fish lines has traditionally been exploited to relate genotype-phenotype and generate better knowledge about phenotype plasticity in response to the environment.

Key issues for future research
In the view of possible problems acquainted to the relationship between aqua cultured fish and the environment, as identified above, a broad range of key issues especially suited to be studied using genomic and molecular approaches have been identified. The use of genomic tools and resources are particularly timely for research that will provide fundamental scientific knowledge, a prerequisite to understand the complex interplay between fish and the environment. The key issues identified can be sub divided into three main areas:

The impact on fish welfare:
- Characterisation of “bad and good” welfare indicators in order to develop scientific based information on fish quality and the link between fish quality and welfare.
- Interaction between stress, husbandry conditions and other physiological functions to better understand the complexity of coping strategies employed by fish in response to abiotic and biotic environmental factors

The impact of genetic plasticity on the capacity to adapt:
- Development of genetic research on the plasticity and adaptation of aquacultured populations in order to understand how environmental changes affects new and old types of husbandry systems to meet sustainable objectives as well as climate changes can be faced by the fish populations.
- Identification of sources of variability (genetic or environmental) of individual responses for traits related to adaptation and domestication and carry out a genetic and functional molecular characterization of these traits

The impact of environmental xenobiotics:
- Establishment of screening methods to assess for the presence and effect of xenobiotics and their impact on aquaculture species using specific or generic molecular markers

The impact of aquaculture on the environment:
- Elucidation of the genetic structure of wild fish populations in association with functionality of the phenotype, including both genomic and physiological analysis, in order to be able to evaluate the impact of aquaculture escapees on wild population
- Comparison of natural and aqua cultured populations to understand the impact of domestication
- Monitoring tools for the impact of aquaculture production (chemicals, food, waste, pathogen) on natural fish and shellfish populations (health, disease, genetic).

Conclusions and suggested strategy
Understanding the interaction between aquaculture and the environment should have a significant impact on sustainability. Key contribution would arise from a variety of genomic approaches in several domains:
- Basic biological information on complex traits underlying the relationship between an organism, its environment and behavioral response (coping strategy) in relation with fish welfare and health.
- Knowledge base of genetic plasticity/variability in relation adaptation to challenging environmental factors in order to contribute to sustainability of aquaculture production and domestication of new species.
New screening methods of xenobiotics: development of genomic signatures correlated with presence of pollutants and health impacts with a long term goal to improve traceability of aquaculture products

Characterization of wild population and comparison with aquaculture populations in order have the knowledge base for robust evaluation of gene and pathogen flow and the level of domestication.

2.5 Genomic tools and bioinformatic resources

Introduction
A few important characteristics make that aquaculture species stand out from economically important terrestrial species regarding their genomic resources. First, the number of aquaculture species is much larger than the number of farmed terrestrial species (at least an order of magnitude). Hence, it is not realistic to foresee the production of all genomic resources for all species. This picture is complicated by the high evolutionary divergence in groups such as finfish (Teleostei) and shellfish (Mollusca and Crustacea). For instance, within the teleost fish, the Japanese pufferfish and the European eel have been separated for at least 350 million years. This makes the quality of the comparison of genomic information between species dependent on evolutionary distance. Finally some aquaculture species present an additional complication through whole genome duplications. Best known for this feature are the salmonids. These three characteristics should be taken into account when establishing new important and relevant genomic resources and tools to facilitate future research, including the development of comparative approaches.

State-of-the-art
An inventory of resources and tools important for aquaculture species have been established and been further updated within the Aquagenome project. All the detailed results are available within the Aquagenome Website (http://www.aquaculture-europe.org/, login: Aquagen, password: Aquagen). This inventory clearly shows that genomic resources are growing rapidly in many aquaculture species. All the species investigated have at least some genomic resources (see Table 1) but some species including rainbow trout, Atlantic salmon, sea bream, sea bass do already have quite important genomic resources. This importance also reflects the economic importance of these species and of the scientific community behind these resources. Some genomics resources including cDNA libraries, ESTs sequences, microarrays dedicated for gene expression analysis, genetic linkage maps, bioinformatics databases, BAC DNA libraries, microsatellite markers and SNP markers are quite frequent among the species investigated. However, a few resources are still lacking in all species including whole genome sequences, SNP microarrays and physical/genetic maps integration. Some of these gaps, like whole genome sequences are however currently underway (Seabass MPI genome sequence project, and the GENOFISK Cod genome project in Norway) or planed (Salmonids cGRASP project).

The most common resource available in nearly all species is cDNA libraries that are used to produce Expressed Sequence Tags (ESTs). By August 2007 a total of nearly 3.5 millions ESTs in fish and shellfish were available, but less than 1.5 millions EST were available in aquaculture species. Furthermore, the amount of EST sequence information publicly available varied extensively from one species to another with for instance 430 000 ESTs in Atlantic salmon to 2-3 000 in sea bream and turbot, Psetta maxima. Genetic maps are also a well developed genomic resource in numerous species, but even if they are available in most aquaculture species their quality that depends on the number of markers available, is highly variable from one species to another (data not collected now in the resource information gather within the Aquagenome Website). However, follow-up of all these resources is not systematic and the demand from the research community for an inventory, the curation and the exploitation of these resources is high.
Table 1. State of the art of the genomic resources available in important European aquaculture species. Non existing resources: open squares and coefficient 0; Resources ongoing or planned: blue squares and coefficient 1; Existing resources: black squares and coefficient 2.

<table>
<thead>
<tr>
<th>RESOURCES</th>
<th>SPECIES</th>
<th>Clam</th>
<th>Ostrea edulis</th>
<th>Mussel</th>
<th>Senegalese sole</th>
<th>Crassostrea gigas</th>
<th>Cod</th>
<th>Carp</th>
<th>Seabream</th>
<th>Turbot</th>
<th>Salmon</th>
<th>Seabass</th>
<th>Trout</th>
</tr>
</thead>
<tbody>
<tr>
<td>cDNA library, ESTs sequences</td>
<td>Crassostrea virg</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Microarray, gene expression</td>
<td>Clam</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Genetic, linkage map</td>
<td>Ostrea edulis</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Bioinformatics database</td>
<td>Mussel</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>BAC library</td>
<td>Senegalese sole</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>SNP marker</td>
<td>Crassostrea gigas</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Microsatellite marker</td>
<td>Cod</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Bioinformatics tool</td>
<td>Carp</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Bioinformatic service</td>
<td>Seabream</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>QTL detection</td>
<td>Turbot</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>QTL map</td>
<td>Salmon</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>AFLP marker</td>
<td>Seabass</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Candidate genes marker</td>
<td>Trout</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Physical map</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Special genetic background</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radiation Hybrid Panel</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Segregating families</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radiation hybrid map</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transgenic lines</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complete genome sequence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical/genetic maps integration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Microarray, genotyping (SNP)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Key issues for future research
To support and enable further research strengthen European aquaculture and ensure its sustainable development a range of resources and tool need to be developed and maintained in aquaculture species. The following items ranked by priority strongly need to be addressed by the community at a European level:

- Full genome sequences are available for a few model species (Japanese pufferfish, spotted green pufferfish, three-spined stickleback, medaka and zebrafish). Recently full genome sequencing of three aquaculture fish species (i.e., Channel catfish, European sea bass and Nile tilapia) has been launched. There is an important need for some additional full genome sequences as they will provide key resources for future functional and structural studies.
- Sequence annotation is an important bottleneck in the exploitation of the genomic information (e.g. ESTs, high throughput gene expression profiles) that is currently growing exponentially in many aquaculture species. Limitations mainly relate to the diversity of aquaculture species, their evolutionary distance and in some cases the presence of whole genome duplications.
• There is a need for maintenance of genomic in resources in aquaculture species (biological, genomic and bioinformatics), to prevent loss, duplications and maintain efficient dissemination of these resources.

Conclusions and suggested strategies

Whole genome sequences

The sequencing of whole genomes of selected aquaculture species is a ultimate addition to understand genome architecture and functioning. New sequencing technologies are emerging very quickly and facilitate already a much higher throughput at a lower cost than conventional Sanger sequencing technology, which is presently still essential for genome finishing. It is expected that the sequence of a single genome will cost much less in the near future. However target species should be chosen carefully; we propose the following guidelines to make a selection (list without priority):

✓ Phylogenetic relevance with regards to available sequenced genomes (new species should take phylogenetically unique positions).
✓ Economic significance (i.e., aquaculture species with a high economic value at the European scale).
✓ The availability of background genomic resources.
✓ Support from a significant European and international scientific community to exploit the data.

In parallel for less prioritized or newly emerging species there will be a need to develop physical maps to facilitate comparative approaches. The quality of the anchoring of these physical maps to full genome sequences will depend on the evolutionary distance. These maps are not necessarily classic linkage or radiation hybrid maps, but should be based on characterized / end-sequenced BAC libraries. Additionally, BAC-libraries are ideal carrier for the exchange of genetic material and genomic information between groups and resource centers.

Gene annotation

Gene annotation for aquaculture relevant species should be improved and standardized to characterize unique sets of sequences, to provide unified gene names and do functional annotation linked to model fish. This could be developed by a global approach that would cross-link genes (or ESTs) from between aquaculture species based on their phylogenetic affinity. A specific task within WP1 of the Aquagenome project will explore this approach to annotate teleost fish genes and ESTs. Other global approaches could complement and enrich this evolutionary annotation such as synteny relationships based either on comparisons of gene location on physical maps or the comparative analysis of full genome sequences.

Resource maintenance and sharing

There is need for an European action to maintain the existing and upcoming biological resources of aquaculture species. However, it is considered that no specific new database bioinformatics environments have to be developed for aquaculture species, but the implementation of already developed and existing databases has to be preferred (e.g. EMBL-EBI bioinformatics tools).

3. Strategic Recommendations

3.1 General considerations

The recent reviews of the state-of-the art by the AQUAFUNC and AQUAGENOME projects reveal that application of genomic tools already contributes substantially to the knowledge generation in several key aquaculture research areas such as host-pathogen interactions, reproduction, breeding, nutrition & growth, and aquaculture-environment interactions. By contrast, the implementation of genomic tools in applications in the industry is still in a very early stage and is limited to a few species and few examples.

The review reveal that substantial efforts have been carried out by a range of European research laboratories related to genomics in aquaculture species, and this have already involved considerable resources based on European and national funding. However, the review also illustrates a rather low level of interactions between these projects. It is also demonstrated that the large-scale nature of genomic approaches leads to very rich sources of information, but it is very demanding to fully exploit and analyse the vast amount of data that is generated. Hence, there is a potential for new European actions to more fully exploit these data.

Similarly to what has been observed in higher vertebrates, the major benefits of the genomic approaches are closely associated with the ability to carry out a much broader and systematic analysis of the genome and its expression products in close relation to a large set of biological processes. Genomic approaches can profoundly
extend our ability to understand biological processes, but the approaches are also very demanding both in terms of experimental set-ups and bioinformatic processing. As an example, currently genomic approaches such as micro array analysis serve mainly to generate hypotheses about affected processes and mechanisms in aquaculture species (i.e. to find groups of affected genes), but we foresee that these tools in the near future also should be used in hypothesis driven approaches in well defined experimental set-ups that will more fully exploit their potential.

Tools which scan the state of an organism’s transcriptome (e.g. cDNA microarrays and cDNA libraries) are now available for a range of European aquaculture species, and have recently been applied to study various processes such as stress, early development, reproduction, effect of diet and impact of vaccines or pathogens. These approaches have been successful in identifying factors involved in diverse processes, identify underlying mechanisms and/or affected metabolic pathways. Genome information has been also very important for the recent development of quantitative trait locus (QTL) in breeding programs for fish (e.g. in Atlantic salmon). Single nucleotide polymorphisms (SNP) analysis is presently being developed in order to assist efficient selection of attractive production traits. Successful characterization of such genetic information and the future development of genomic selection protocols will require very precise phenotype characterization of the traits for which functional genomic information will bring key information. The above mentioned techniques are also used in more basic science to understand the genetic changes that take place during domestication, as well as providing a broader basic understanding of the genetic basis of a range of phenotypic traits.

By contrast, the direct adoption of genomic techniques and knowledge has so far been relative slow in the European Aquaculture industry. This may in part be a reflection of the characteristic of the industry which includes: 1) its diversity both in terms of species, environment and size of producers; 2) the rather unstructured nature of the sector which makes uptake of new methods difficult; 3) limited domestication/genetic selection of aquaculture stocks, 4) the lack of standard production parameters which leads to variable performance and 5) issues related to economic sustainability. The diversity of species cultured in aquaculture and their very divergent physiology compared to the animal production sector does not favour models which try to transfer technology and approaches between sectors. Moreover, this high diversity of species in aquaculture and the increasing demand for diversification into new species makes it difficult to transfer molecular resources and tools between species, so that there is a continuous need for new resources and studies to understand their basic biology. In summary, 1) there is a need to intensify studies of established species for which generally there is a deficit of basic biological knowledge, and 2) studies are required for new species to ensure establishment of sustainable production methods. The high information content of well conducted biological trials analysed using a range of molecular resources represent a cost effective way of generating knowledge with high information content.

A substantial part of the European genomic research for aquaculture has been funded by EU research programs. However, EU-funded project are normally developed for a limited period of time (usually 3 years) and without clear policy in term of project exchanges and maintenance of genomic resources and databases after the end of the project. This does not favour cross-talk between projects and various European research laboratories, both at the level of approaches, methodologies and the level of new data for meta-analysis. Considering the large investment of the European Commission on the development of these genomic approaches, there is certainly large potential in further developing such cross-talk to explore the full potential of the genomic tools and resources that are available and currently being further developed. Measures should therefore be found to secure the afterlife and availability of genomic resources and tools after the completion of individual projects and to more fully explore results derived from individual projects developing and using genomic resources.

3.2. Specific recommendations on the application of genomic tools and resources in aquaculture research

Fish health
Genomics, have a vast potential in studies on host-pathogen interaction, e.g. by the ability to sequence whole genomes of both host and pathogen, to understand the immunological spectrum (innate and acquired immunity) in the host and the way of infection by the pathogen. In addition transcriptomic surveys can explain how the relation between host and pathogen change with life-stages, environments and treatments.

Reproduction
High priority should be given to the application of the available tools for functional studies on key reproductive issues such as control of sex differentiation, puberty, spawning and gamete quality in a range of European species.
aquaculture species - both in terms of improving the knowledge base and in development of applied solutions to industrial problems, including reproductive isolation of farmed animals.

Breeding
Research to support the implementation in selective breeding programmes and to reduce their costs should have high priority in genetic research. These approaches should be adapted to the status of the breeding programme for each species, the different breeding goals (e.g. growth, yield, flesh quality, feed efficiency and disease resistance), and take into consideration potential side-effects of breeding on normal development, welfare, disease transmission and diversity in disease resistance, as well as minimising risk of negative impact on wild stocks and natural ecosystems.

Nutrition and growth
In the area of nutrition and growth, and particularly in trying to understand basic fish physiology, global gene expression surveys (microarray, cDNA libraries) are considered to provide a very promising tool. This would include the effects of varying nutrition regimes on energy homeostasis and metabolic control, disease resistance, fish welfare, larval development, muscle and skeletal development and growth and broodstock performance. Identification of QTLs linked to nutrition and metabolism should be a priority as it will offer the means to select strains of fish with improved performance or adaptability to modified diets.

Aquaculture-environment interactions
High priority should be given to the use of genomic approaches to provide a better understanding of the complex interactions between aquaculture and the environment. Such development will have major impacts on various aspects of aquaculture sustainability including fish welfare and health, product quality and safety, genetic impact on wild stocks. Key contributions for these topics should include genomic studies on basic biological information on complex traits related to physiological and behavioural responses to environment in relation to welfare, genetic impact of aquaculture species on wild stocks, basic information on genetic of plasticity and variability in relation to adaptation to challenging and fluctuating environment, new screening methods for effects of xenobiotics with the long term goal to improve traceability of aquaculture products.

3.1 Resource development and maintenance

Whole genome sequences
The sequencing of whole genomes of selected aquaculture species is an ultimate addition to understand genome architecture and functioning. New sequencing technologies are emerging very quickly and facilitate already a much higher throughput at a lower cost than conventional Sanger sequencing technology, which is presently still essential for genome finishing. It is expected that the sequence of a single genome will cost less than 100,000 € in the near future. However target species should be chosen carefully, and we propose the following guidelines to make a selection (list without priority):

- The availability of background genomic resources.
- Economic significance (i.e., aquaculture species with a high economic value at the European scale).
- Support from a significant European and international scientific community to exploit the data.
- Genetic structure
- Phylogenetic relevance with regards to available sequenced genomes (new species should take phylogenetically unique positions).

As example of a whole genome sequencing project which is prioritized is Atlantic salmon. Much work has already been done in this species with about 400000 EST’s available which will facilitate the assembly. In addition salmon is an important economic species. This species is also genetically interesting since it has undergone a recent genome duplication, which makes it partly tetraploid. Not much is known about genetic adjustment to genome duplication but by using this species many of the basic questions regarding gene silencing and evolution could be answered. In addition Atlantic salmon could fill out phylogenetically as a model for primitive teleost.

In parallel for less prioritized or newly emerging species there will be a need to develop physical maps to facilitate comparative approaches. The quality of the anchoring of these physical maps to full genome sequences will depend on the evolutionary distance. These maps are not necessarily classic linkage or radiation hybrid maps, but should be based on characterized / end-sequenced BAC libraries. Additionally, BAC-libraries are ideal carrier for the exchange of genetic material and genomic information between groups and resource centres.
**Gene annotation**

Gene annotation for aquaculture relevant species should be improved and standardized to characterize unique sets of sequences, to provide unified gene names and do functional annotation linked to model fish. This could be developed by a global approach that would cross-link genes (or ESTs) from between aquaculture species based on their phylogenetic affinity. A specific task within WP1 of the Aquagenome project has explored this approach to annotate teleost fish genes and ESTs. Other global approaches could complement and enrich this evolutionary annotation such as synteny relationships based either on comparisons of gene location on physical maps or the comparative analysis of full genome sequences. No specific new database bioinformatics environments needs to be developed for aquaculture species, but the implementation of already developed and existing databases has to be preferred (e.g., EMBL-EBI bioinformatics tools).

**Resource maintenance and sharing**

There is need for a European action to ensure that existing and new biological resources for aquaculture species are maintained and remain available for use by the wider scientific community. This could be achieved by assigning support and promoting networks to link research groups to maintain and share the biological resources (e.g. cDNA and BAC libraries, Radiation Hybrid panels, cDNA and BAC clones, mapping panels, QTL families), this process could be a combination of physical and virtual centres. Moreover, database maintenance and exploitation is important due to the considerable increase in sequence, mapping, and gene expression datasets in numerous aquatic farmed species. Data management at this scale requires sufficient resources and close cooperation between the involved parties, hence such topics may also be suited for new European actions.
4 Contributors

List of working group members that contributed to the report from the Bergen meeting, and the main basis of the white paper.

Working group 1 - Genomics of host-pathogen interaction
Antonio Figueras, CSIC, Spain (Moderator)
Simon Mackensie, Universitat Autònoma de Barcelona, Spain
Carlor Didcock, Genesis Faraday, UK
Ross Houston, Roselin Institute, UK
Cathrine Jones, University of Aberdeen, UK
Sven M Jørgensen, AKVAFORSK, Norway
Beatriz Novoa, CSIC, Spain
Elena Sarrapoulo, HCMR, Greece
Kristin Thorud, Ministry of Fisheries and Coastal Affairs, Norway
Rasmus Skern, Institute of Marine Research, Norway

Working group 2 - Genomics to study ecotoxicological, environmental and husbandry effects on stress, welfare and product safety & and environmental impact
Patrick Prunet, INRA, France
Bjørn O Kvamme, Institute of Marine Research, Norway
Pål Olsvik, NIFES, Norway
Marco Saroglia, University of Insubria, Italy
Per-Ove Thörnquist, Uppsala University, Sweden
Lluis Tort, Universitat Autònoma de Barcelona Spain
Geir Dable, Institute of Marine Research, Norway
Knut Jørstad, Institute of Marine Research, Norway

Working group 3 - Genomics in reproduction and breeding programs
Pierrick Haffray, SYSAAF, France (Moderator)
Geir Lasse Taranger, Institute of Marine Research, Norway (Moderator)
Pierre Boundry, IFREMER, France
Jean-Jacques Lareyre, INRA, France
Edwige Quillet, INRA, France
Thomas Moen, AKVAFORSK, Norway
Eva Andersson, Institute of Marine Research, Norway
Jon Vidar Helvik, University of Bergen, Norway
Luca Bargelloni, University of Padova, Italy
Sigbjørn Lien, CIGENE, Norway

Working group 4 - Genomics of Growth and Nutrition
Brendan McAndrew, University of Stirling, UK (Moderator)
Sam Martin, University of Aberdeen, UK
Ulla Nordgarden, Institute of Marine Research, Norway
Michael Leaver, University of Stirling, UK
Genciana Terova, University of Insubria, Italy
Nigel Finn, University of Bergen, Norway
Anna Wargelius, Institute of Marine Research, Norway
Anita Sagstad, NIFES, Norway

Working group 5 Genomic tools and bioinformatic resources
Yann Guiguen, INRA, France (Moderator)
Nikolai Mugue, VNIRO, Russia (Moderator)
Nilsen Frank, University of Bergen, Norway (Moderator)
Filip Volekaert, Katholieke Universiteit Leuven, Belgium
Laurence Amilhat, INRA, France
Carlos Fernandez, University of Santiago de Compostela, Spain
Alexei Krasnov, AKVAFORSK, Norway
Richard Reinhardt, MPI-MG, Germany
Vilas Roman, University of Santiago de Compostela, Spain
Appendix

Glossary

Genomics The study of an organism's entire genome.

Transcriptomics Also termed expression profiling and examines the expression level of mRNAs in a given cell population, often using high-throughput techniques based on DNA microarray technology.

Proteonomics Large-scale study of proteins, particularly their structures and functions.

Microarray Description of tool: A DNA microarray is a multiplex technology used in molecular biology and in medicine. It consists of an arrayed series of thousands of microscopic spots of DNA each containing picomoles of a specific DNA sequence. This can be a short section of a gene or other DNA elements that are used as probes to hybridize a cDNA or cRNA sample (called target). Probe-target hybridization is usually detected and quantified by fluorescence-based detection of fluorophore-labeled targets to determine relative abundance of nucleic acid sequences in the target. Use of tool: DNA microarrays can be used to detect DNA (as in comparative genomic hybridization), or detect RNA (most commonly as cDNA after reverse transcription) that may or may not be translated into proteins. The process of measuring gene expression via cDNA is called expression analysis or expression profiling.

ESTs Expressed Sequence Tag is a short sub-sequence of a transcribed cDNA sequence. They may be used to identify gene transcripts, and are instrumental in gene discovery and gene sequence determination

cDNA library A collection of cloned cDNA (complementary DNA) fragments produced from fully transcribed mRNA.

BAC Bacterial Artificial Chromosome (BAC) is a DNA construct, based on a fertility plasmid (or F-plasmid), used for transforming and cloning in bacteria. These types of plasmid are unique due to the amount of DNA material they can contain between 150-700 thousand basepairs. The large content of DNA makes them suitable for use in genome projects especially in the assembly process. SNP single nucleotide polymorphism (pronounced snip) is a DNA sequence variation occurring when a single nucleotide - A, T, C, or G - in the genome (or other shared sequence) differs between members of a species (or between paired chromosomes in an individual).

AFLP Amplified fragment length polymorphism, a highly sensitive tool used in molecular biology to detect DNA polymorphisms

QTL Quantitative trait loci are stretches of DNA that are closely linked to the genes that underlie the trait in question

Radiation hyb. mapping Is a somatic cell technique for ordering genetic loci along a chromosome and estimating physical distances between adjacent loci

Microsatellites Also termed Simple Sequence Repeats (SSRs), are polymorphic loci present in DNA that consist of repeating units of 1-6 base pairs in length. They are typically neutral, co-dominant and are used as molecular markers which have wide-ranging applications in the field of genetics, including kinship and population studies.