

# Genomics and Genome Sequencing: Benefits for Finfish Aquaculture

Nicole L. Quinn<sup>1</sup>, Alejandro P. Gutierrez<sup>1</sup>,  
Ben F. Koop<sup>2</sup> and William S. Davidson<sup>1</sup>

<sup>1</sup>*Department of Molecular Biology and Biochemistry,  
Simon Fraser University, Burnaby, British Columbia*

<sup>2</sup>*Department of Biology, University of Victoria,  
Victoria, British Columbia  
Canada*

## 1. Introduction

### 1.1 What is genomics?

The central dogma of molecular biology states that DNA is transcribed to RNA, which is translated into proteins (Crick, 1970), the molecules that facilitate all biological functions. A genome comprises all of an organism's DNA, or hereditary information. The field of genomics is the study of whole genomes. Genomics has been defined as "a branch of biotechnology concerned with applying the techniques of genetics and molecular biology to the genetic mapping and DNA sequencing of sets of genes or the complete genomes of selected organisms" (Mirriam-Webster Dictionary). Specifically, whereas genetics is the study of a single gene, or a few genes in isolation, genomics examines all of the genes, as well as the non-coding elements (i.e., regions that do not encode proteins or RNA components of the cell), within the DNA of a genome. Although the term genomics was first used in 1987, the field is relatively new and is growing rapidly as new technologies for exploring genomes emerge. The applications of genomics are vast, spanning realms such as medicine, industry and ecology, with implications for global issues including cancer diagnosis, prevention and treatment, alternative energy sources, agriculture, conservation and sustainable development. Furthermore, given that the fundamental basis of genomics—DNA—is common across all living organisms, genomics stands to bridge the gaps between these fields of study, with synergistic results as multi-disciplinary approaches to answering biological questions are developed.

### 1.2 Genomics in aquaculture: Phenotypic vs. genotypic selection

In traditional selective breeding practices, individuals showing desirable phenotypic, or visible characteristics are bred to produce new strains of plants or animals that exhibit features such as faster growth, greater size, increased overall robustness, or improved aesthetic appeal. Farmers, through thousands of years of following the mantra "breed the best to the best and hope for the best" (quotation attributed to American Thoroughbred and

Standardbred breeder, John E. Madden), have drastically changed natural populations, as exemplified by the now hundreds of breeds of the domestic dog, or differing cattle strains for dairy or beef production. However, these phenotype-based breeding tactics come with inherent drawbacks. Specifically, they are draining of resources and time, as they often rely on trial and error, they require vast numbers of individuals exhibiting extensive phenotypic variation, and often numerous generations of repetitive selective breeding are required to see an effect, which is especially difficult for species with long generation times and for which extensive parental investment is needed. Furthermore, often animals need to be sacrificed to detect or measure morphological characteristics such as flesh color and tissue quality, and these, as well as disease challenged (i.e., to determine immune function or disease resistance) animals cannot be bred. Finally, most characteristics of interest to aquaculture are complex traits - i.e., those that are governed by numerous genes and complex interactions between multiple pathways, or those for which morphological variation is based on environmental cues, and it is very difficult to select for multiple complex traits together based on phenotypic information alone. Thus, the amount of time and the costs associated with traditional, phenotypic approaches to selective breeding are large.

In contrast to phenotype-based selection, which directly selects for a given heritable trait, in genotype-based selection, or marker-assisted selection (MAS), traits are indirectly selected for based on variability at the DNA level. More specifically, the genomes of breeding populations are screened for multiple markers, or DNA tags, that are associated with genes of interest that work together to produce a desirable phenotype for a complex trait. This genetic screening can be accomplished using high-throughput genomics systems that incorporate tools such as polymerase chain reaction (PCR) or DNA sequence-based screening.

MAS has been used extensively for the genetic improvement of cultivated plant cultivars, such as wheat (Gupta et al., 2008), corn (Tuberosa et al., 2002) and soy (Kim et al., 2010), as well as cultured stocks (broodstocks) in agriculture species such as swine (Liu et al., 2007) and cattle (Veerkamp & Beerda 2007). As the availability of genomics resources has increased for aquatic species, more research is being done to improve broodstocks for many key aquaculture species worldwide. Specifically, genomics, and the use of genomic tools, enables one to examine the differences and similarities among organisms at the genotypic (DNA) level, as opposed to more traditional broad-scaled phenotype-based (appearance-based) approaches. This very fine-scaled perspective, looking at differences in genes (alleles) as well as variability at non-coding loci, or positions within the genome, means that complex traits that are driven by more than one gene or pathway can be broken down into their components. That is, differences at the individual, family, population, species and even the organism level can be assessed in very fine detail. Such insight into the genetic factors that drive complex traits can facilitate the development of effective and efficient breeding methods, which have far-reaching implications for the aquaculture industry.

## **2. Atlantic salmon: A model aquaculture species for MAS**

The holy grail of any genomics program for a species is a whole genome sequence that is well assembled and annotated. The advantages of this are many, and are discussed in further detail in subsequent sections, but are mainly centered around two things: 1) the

wealth of data that is produced, including the full gene repertoire with additional information such as gene location and copy number, and 2) the ability of the sequenced genome to act as a reference genome, both for the sequenced species itself (i.e., such that the genomes of additional individuals can be easily re-sequenced using the original as a reference for assembly), as well as to provide information for other, closely related species. Currently, however, even with the great advances in sequencing technology that have come to light in recent years, obtaining a whole genome sequence remains an extremely difficult, costly and time-consuming undertaking. This is particularly true for fish species simply due to the evolutionary age of fish and the more than 20,000 extant species (Nelson, 2006), factors that make the fish genomes diverse and complex and complicate sequencing. Indeed, only five fish genomes have been reported to date (medaka, *Oryzias latipes*; tiger pufferfish, *Takifugu rubripes*; green spotted pufferfish, *Tetraodon nigroviridis*; zebrafish, *Danio rerio* and stickleback, *Gasterosteus aculeatus*), although more are underway. Their sequences, as well as those for the sequenced genomes of other organisms are publically available within the Ensembl database ([www.ensembl.org](http://www.ensembl.org)). These fish species were chosen for their abilities to act as model sequences for genetics research, rather than for their utility for aquaculture. Specifically, the medaka and zebrafish genomes were sequenced to provide model organisms for studying developmental biology, while the stickleback genome serves as a model for studying adaptive evolution, and the two pufferfish represent the smallest known vertebrate genomes. Figure 1 illustrates the phylogenetic relationships among these species as well as some key aquaculture species, and shows that the full spectrum of teleosts is not represented by the genome sequences that are currently available.

When a whole genome sequence is not available, there are numerous genomics resources and tools that can be developed which, particularly when used in combination, can provide extensive insight into a genome and can be used for applications such as MAS, both for the species of interest, and for other closely-related species. In the following sections, we will use the example of Atlantic salmon (*Salmo salar*) and its genomics program to describe these tools, their development and utility for aquaculture. Atlantic salmon is a particularly good model fish species for genomics because it is a major aquaculture species, with approximately 1.5 million tonnes produced worldwide in 2009 (Food and Agriculture Organization [FAO] Fishery Statistic, 2009), and like other salmonids, is of substantial environmental, economic and social importance. In addition, there are extensive genomics resources for Atlantic salmon, which were developed using standard methods and approaches that are applicable to other genomes. Furthermore, a strong argument has been made for obtaining the full genome sequence for Atlantic salmon, a project that is currently in progress (Davidson et al., 2010). Specifically, aside from the merits of the species itself, there are no salmonid species yet sequenced, and thus Atlantic salmon can serve as a reference salmonid genome, providing extensive opportunities for cross-referencing, or comparative synteny analyses with other salmonids, particularly the Pacific salmon (*Oncorhynchus* sp.), rainbow trout (*Oncorhynchus mykiss*) and Arctic charr (*Salvelinus alpinus*).

Atlantic salmon also provides an example of some of the challenges that face fish genomics in general. The common ancestor of salmonids underwent a whole genome duplication event between 20 and 120 million years ago (Allendorf & Thorgaard, 1984; Ohno, 1970). Thus, whereas there are usually two copies of each gene within a genome, Atlantic salmon have four, and the duplicate copies are evolving into genes with new functions or non-coding DNA. The genome duplication also increased the size and the repetitiveness of the

genome. These characteristics, combined with the lack of a closely related guide sequence, mean that sequencing and assembling the Atlantic salmon genome are extremely challenging.

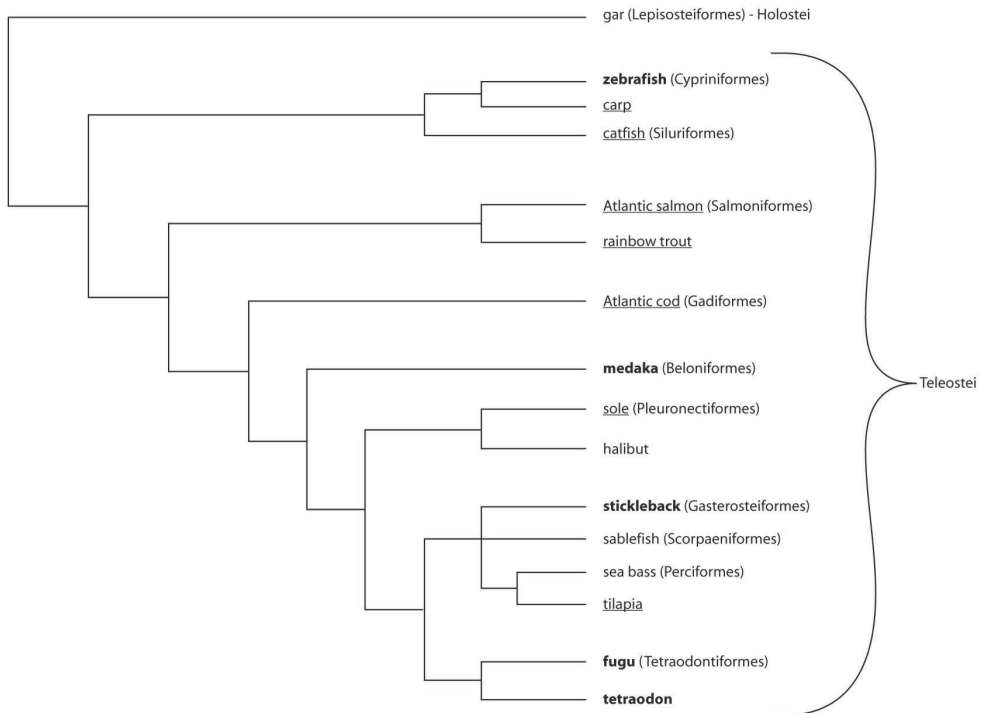


Fig. 1. Schematic representation of the phylogenetic relationships among fish species. Species in bold have publically available full genome sequences, while those that are underlined are currently being sequenced. Note that the full spectrum of teleosts is not represented by the genome sequences that are currently available. Species are listed by their common names with orders in parentheses. Gar is used as an outgroup.

### 3. Genomics resources

#### 3.1 ESTs

An expressed sequence tag (EST) is a sequenced section of a transcribed, or expressed gene. ESTs are developed by isolating the mRNA, i.e., the transcribed DNA, from a tissue, which is then reverse-transcribed into cDNA (complementary DNA). Thus, an EST library is the full repertoire—essentially an inventory—of expressed genes within a genome in a given tissue at a given time. More than 499,000 ESTs have been identified for Atlantic salmon, the sequences of which are available publicly in an online database (Rise et al., 2004; Koop et al., 2008) as well as within the NCBI database.

Perhaps the most common use of ESTs is to study gene expression (described below). Other uses of ESTs include the identification of molecular markers, or DNA tags, which are

described in detail in the subsequent section. EST-based markers have unique applications because they occur in expressed genes, as opposed to non-coding DNA. Thus, they can provide insight into allelic differences between individuals or populations, thereby facilitating an understanding of the relationship between phenotypes and particular genes.

### 3.2 Genetic markers

Genetic markers are short sequences of DNA that are associated with a specific location, or locus, on a chromosome. To be useful, a genetic marker must be sufficiently variable (polymorphic) among individuals or populations. Markers are valuable tools in molecular biology and they have many practical uses. For example, they are used to examine ancestry and to define discrete populations, as well as for forensic work, paternity testing and tissue typing for transplants given that the genotypes of the markers are inherited and thus an individual's combination of marker genotypes (commonly called a "DNA fingerprint") is unique.

#### 3.2.1 Microsatellite markers

Satellite markers are a group of DNA markers that are characterized by containing tandemly repeated sequences of nucleotides (e.g., ...ACACACAC...). Depending on their size, which is a function of the number of repeated bases in a repeat unit and the number of tandemly arrayed units, satellite markers are classed as microsatellites (short repeat unit) or minisatellites (long repeat unit). Microsatellite markers are heavily utilized in genomics. They act as DNA anchors or tags within a genome because: 1) like all molecular markers, their genotypes are heritable, 2) the repeat units make them more vulnerable to mutation than other types of DNA sequence (e.g., coding sequence), which means they are variable at all levels, from individual to species, and generally have several alleles, 3) their locations (loci) within the chromosome can be determined relative to one another, and 4) their small size (vs. minisatellites, which can be very long) facilitates ease of detection using PCR. The development of a suite of microsatellite markers is often one of the first steps in a genomics program for a species of interest.

The first microsatellites characterized for Atlantic salmon came from genomic libraries (McConnell et al. 1995; O'Reilly et al., 1998; Sanchez et al., 1996; Slettan et al., 1995). These were classified as "anonymous" or "Type II" markers as they were not associated with a particular gene. The mining of large EST databases allowed the identification of Type I microsatellites, which could be related directly to a gene (Anti Vasemagi et al., 2005; Ng et al., 2005). More recently, the sequencing of the ends of Bacterial Artificial Chromosome (BAC) clones has provided a rich source of microsatellites (Danzmann et al., 2008; Phillips et al., 2009). Microsatellites have been used extensively to examine the population structure of Atlantic salmon (e.g., King et al., 2001) as well as for assigning parentage and relatedness of individuals in a mixed family stock (Norris et al., 2000), which precludes the necessity of rearing families in individual tanks or using physical tags. Many of the microsatellite markers that have been developed for one species of salmonid can be used in others (e.g., McGowan et al., 2004; Presa & Guyomard, 1996; Scribner et al., 1996), which enables genomic comparisons among salmonids such as Atlantic salmon, brown trout (*Salmo trutta*), rainbow trout and Arctic charr (Danzmann et al., 2005, 2008; Woram et al., 2004). Most aquaculture species have a set of microsatellites now, and the list is too extensive to reference here.

### 3.2.2 Single Nucleotide Polymorphisms (SNPs)

A single nucleotide polymorphism (SNP) (pronounced *snip*) is a sequence change in a single nucleotide – A, T, C or G - that occurs with sufficient frequency in a population to serve as a genetic marker. SNPs can be within genes or within non-coding regions of the genome. Because they are usually bi-allelic (i.e., exhibit only two alleles, such as an A or a T at a given locus), they are easily identified and studied. The development of novel sequencing technologies has made it possible to obtain thousands SNP markers for genotyping (see Molecular Ecology Resources (2011) 11: Supplement 1, for several recent articles on this subject). Moreover, new genotyping technologies enable the simultaneous analysis of a great number of these markers, which allows the construction of high density genetic maps (Goddard & Hayes, 2009). These techniques are now being widely used in species of agricultural importance and have led to the development of numerous dense SNP arrays, which are tools that are used to screen entire genomes for the genotypes of known SNPs (Matukumalli et al., 2009; Muir et al., 2008; Ramos et al., 2009).

There have been several attempts to identify SNPs in individual genes in Atlantic salmon (Ryynanen & Primmer, 2006) and other salmonids (Smith et al., 2005), but these were small-scale efforts. It was not until large EST databases became available that an intensive search was undertaken for SNPs that could be tied to actual genes (Hayes et al., 2007). Re-sequencing BAC-end sequences (described below) was another approach taken to identify SNPs for Atlantic salmon (Lorenz et al., 2010), but this process was highly labor intensive relative to the information gained. A next generation sequencing strategy (Van Tassell et al., 2008) has also been used to identify SNPs in Atlantic salmon; however, the repetitive nature of the genome made it difficult to distinguish between putative SNPs arising from duplicated regions and those from a unique locus. Recently a custom SNP-array (iSelect by Illumina) containing approximately 6,500 SNP markers was developed for Atlantic salmon. This has been used to search for generic genetic differences between farmed and wild Atlantic salmon (Karlsson et al., 2011) and to assess the level of heterozygosity in Tasmanian Atlantic salmon (Dominik et al., 2010). In addition, the SNP-array was used to construct a relatively dense SNP-based linkage map (described below) (Lien et al., 2011). For other aquaculture species the availability of SNPs is still somewhat limited, with only a large number of SNPs being reported for catfish (*Ictalurus* sp.), (Wang et al., 2010) rainbow trout (Sanchez et al., 2009) and Atlantic cod (*Gadus morhua*) (Hubert et al., 2010; Moen et al., 2008b).

### 3.3 Linkage maps

One of the most useful applications of genetic markers is the generation of a genetic map, also known as a linkage map, of an organism's genome. Specifically, markers that are located physically close to one another on a chromosome, or 'linked', tend to be inherited together. By determining the frequency that alleles at different genetic loci are inherited together (i.e., the degree of linkage between markers) one can estimate the recombination distances that separate markers (i.e., the relative positions of markers within a genome). With sufficient markers (i.e., a "dense" linkage map), the number of linkage groups should correspond to the number of different chromosomes comprising a genome. Linkage maps are extremely useful tools in genomics and form the foundation for QTL mapping (described below).

The first linkage maps for Atlantic salmon were constructed using relatively few microsatellite markers (50–60) (Gilbey et al., 2004; Moen et al., 2004). The number of microsatellite markers was expanded to approximately 2,000 thanks to BAC-end sequencing, and a relatively dense genetic map covering the 29 pairs of chromosomes in the European strain of Atlantic salmon was built (Danzmann et al., 2008). Additional maps were constructed using SNPs derived from ESTs (Moen et al., 2008a) and BAC-end sequences (Lorenz et al., 2010), and these were incorporated into a dense SNP map comprising ~5,600 SNPs (Lien et al., 2011). In terms of other salmonids, linkage maps have been constructed for rainbow trout (Guyomard et al., 2006; Nichols et al., 2003; Rexroad et al., 2005; Sakamoto et al., 2000), brown trout (Gharbi et al., 2006), coho salmon (*Oncorhynchus kisutch*) (McClelland & Naish, 2008), and Arctic charr (Woram et al., 2004). Other aquaculture species of interest for which linkage maps have been developed include catfish (Kucuktas et al., 2009; Waldbieser et al., 2001), Japanese flounder (Coimbra et al., 2003) European sea bass (Chistiakov et al., 2005), tilapia (Lee et al., 2005), common carp (Sun & Liang, 2004), olive flounder (Kang et al., 2008), Atlantic cod (Hubert et al., 2010; Moen et al., 2009b) and the list is expanding.

### 3.4 Bacterial Artificial Chromosome (BAC) libraries

#### 3.4.1 BACs and BAC libraries

A bacterial artificial chromosome (BAC) clone is comprised of a special plasmid and a large segment of genomic DNA from the species of interest. Specifically, the entire genome of a selected individual is cut into pieces of approximately 150,000–300,000 base pairs, each of which is inserted into the plasmid, thereby generating thousands of closed, circular “hybrid DNA constructs”. Thus, a BAC library refers to all of the BACs, which in combination comprise the entire genome, for a representative individual of a species. Each BAC is then inserted (“transformed”) into live *E. coli* cells; thus, the foreign (non-bacterial) DNA can be replicated as if it were a bacterial chromosome using the cellular machinery found within the bacterial cell. Ideally there should be redundancy of the segments of the genome represented in a BAC library such that overlaps of individual clones occur. The ability to identify which BACs share a segment of the genome enables a network called a physical map to be constructed. Therefore, the physical map represents the genome, and has traditionally been used to assist with sequence assembly following whole genome sequencing. BAC libraries also facilitate the identification of target sequences or regions, as the library can be “screened” for genetic markers or other sequence-specific probes. An Atlantic salmon BAC library comprising approximately 300,000 clones and providing an 18-fold coverage of the genome is available (Thorsen et al., 2005), and it was used to generate a physical map (Ng et al., 2005). BAC libraries have been constructed for several aquaculture species including rainbow trout (Palti et al., 2009), catfish (Xu et al., 2007), common carp (Li et al., 2011b) and barramundi (*Lates calcarifer*) (Xia et al., 2010).

#### 3.4.2 BAC-end sequencing

Once a BAC library has been generated, typically the next step is to select a subset of BACs for end sequencing. Because the sequence of the plasmid vector within the BAC is known, the bacterial sequences at the junctions of the genomic DNA and vector DNA can act as sequence primers. Thereby, the first 500–1,000 nucleotides (based on Sanger sequencing) of

the genomic DNA inserts—i.e., the BAC-ends—can be determined. End-sequencing a large subset of BACs can provide extensive insight into the full genome sequence. For example, the 207,869 BAC-end sequences for the Atlantic salmon BAC library cover approximately 3.5% of the whole genome sequence. This glance into a genome is very powerful, as it can provide information about the complexity of the genome (i.e., repeat content), and the BAC-end sequences can be a source of molecular markers (as described above for Atlantic salmon (Danzmann et al., 2008; Lorenz et al., 2010)). Finally, the BAC-end sequences can be used for comparative synteny analyses by aligning them against other, fully sequenced genomes, which can provide insight into the gene content of the BACs, thus providing a partial, putative annotation of segments of the genome (Sarropoulou et al., 2007). This is particularly powerful if the two genomes being compared are closely related, as one can act as a reference genome (e.g., common carp and zebrafish, which are both Cypriniformes) (Xu et al., 2011). Even when there is no obvious reference genome there is value in this exercise as it often suggests syntenic regions among phylogenetically distantly related species, and can lead to candidate genes being suggested for specific traits (see below for an example with ISA resistance in the MAS section). The Atlantic salmon BAC end sequences have been made available in a publicly accessible searchable database ([www.asalbase.org](http://www.asalbase.org)) along with comparative genomic information for four of the five published fish genomes. Similar work is ongoing with aquaculture species such as catfish (Liu et al., 2009), rainbow trout (Genet et al., 2011) and sea bream (Kuhl et al., 2011).

### 3.4.3 Integration of physical and linkage maps and chromosomes

Once both a physical map (i.e., a BAC library is assembled into overlapping, contiguous BACs), and a linkage map (generated using the relationships between variable markers) exist for a species, the next step is to integrate the two maps. This procedure is relatively straight forward if the linkage map contains many markers from BAC-ends, as illustrated by the integration of the Atlantic salmon physical and linkage maps (Danzmann et al., 2008). However, this is seldom the case and considerable effort is normally required to carry out this process (e.g., see preliminary data for rainbow trout) (Palti et al., 2011). The integration of the linkage and physical maps then allows BACs containing particular markers to be selected for fluorescent *in situ* hybridization on chromosomes so that the each linkage group can be assigned to a particular chromosome. This information is now available for several salmonid species including Atlantic salmon (Phillips et al., 2009) and rainbow trout (Phillips et al., 2006).

### 3.5 Quantitative trait loci (QTL)

Generally, the characteristics that are desirable for selection are complex, or 'quantitative', which means that they are not simply governed by a single gene, but rather are controlled by a suite of genes and regulatory factors that interact to produce a phenotypic characteristic. The resulting phenotype usually falls somewhere on a spectrum. This is not always the case; traits such as genetic disorders can be determined by a single gene (e.g., cystic fibrosis and Duchenne muscular dystrophy), but most phenotypic traits associated with health, growth, meat quality and the capacity for utilizing new and ecologically sustainable food items are complex traits that are strongly influenced by the environment, and thus show a range of phenotypes.



A deep biological understanding of phenotypic patterns, ranging from intracellular molecular signatures to whole individual physiological, morphological and behavioral characteristics of genes of commercial importance is synonymous with an understanding of the functioning of the underlying gene regulatory networks and how these are molded by the environment. This calls for the identification of genes and characterization of their functional repertoires in various systemic settings. Specifically, the first step in understanding the mechanisms that govern complex traits is to determine the heritability of the trait – i.e., the extent to which it is determined genetically and by the environment. Heritability estimates vary tremendously depending on the population, the experimental design and the researchers. Garcia de Leaniz et al. provide the best available review of heritability estimates for Atlantic salmon (Garcia de Leaniz et al., 2007). What is clear is that the genes do play a role in the variation detected in the majority of economically important production traits.

The two most common approaches to identify genes associated with complex traits are to examine candidate genes, based on previous knowledge of a biochemical process, or by mapping quantitative trait loci (QTL). We will not discuss the candidate gene approach here, but rather concentrate on QTL analysis. QTL are regions of the genome that contain, or are linked to, genes that contribute to a particular trait (Lynch & Walsh, 1998). Note that the term "QTL" does not refer to a particular gene, but rather its putative location in the genome. This association, or linkage, is determined using the genetic markers within a linkage map, and how alleles at these markers co-segregate with the trait in question. Thus, finding functional markers to assist in MAS is largely dependent on the variability and polymorphic level of the species. In most agricultural examples, there are highly inbred strains that make MAS usable. When a species is highly variable, such as the case with fish and most marine species, the QTL markers have to be very close to the causal variant to be usable in MAS breeding studies. The more variable the species, the closer the marker has to be to be used in selective breeding. Given sufficiently high linkage, the markers that define QTL can feed directly into marker assisted selection (MAS). The number, positions and the magnitude of the QTL affecting a trait are determined by statistical associations between marker alleles or genotypes and particular trait phenotypes. QTL analysis, therefore, relies heavily on robust, dense linkage maps as well as large families that are accurately pedigreed and carefully described with respect to phenotype. Indeed, a limiting factor for effective QTL analysis is the difficulty in characterizing "phenomes", the full set of phenotypes of an individual (Houle et al., 2010).

Growth is one of the most economically important traits for aquaculture species. It is no surprise therefore, that several QTL have been identified for body weight and condition factor in Atlantic salmon (Baranski et al., 2010; Boulding et al., 2008; Moghadam et al., 2007; Reid et al., 2005) rainbow trout (Haidle et al., 2008; Martyniuk et al., 2003; O'Malley et al., 2003; Wringe et al., 2010), coho salmon (McClelland & Naish, 2010; O'Malley et al., 2010) and Arctic charr (Moghadam et al., 2007; Kuttner et al., 2011). QTL for several other traits in salmonids have been mapped, including upper temperature tolerance (Jackson et al., 1998; Perry et al., 2001; Somorjai et al., 2003), spawning time (O'Malley et al., 2003), developmental rate (Nichols et al., 2008), and resistance to pathogens (Gheyas et al., 2010; Gilbey et al., 2006; Houston et al. 2008, 2009, 2010; Jones et al., 2002; Moen et al. 2007, 2009a). Given that many of the microsatellite markers derived from one salmonid species amplify the DNA from

other salmonid species, it has been relatively straightforward to carry out comparative analyses of the rainbow trout, Arctic charr and Atlantic salmon genomes (Danzmann et al., 2005).

Most of the QTL studies described above were based on a limited number of microsatellite markers; thus, many regions of the genome were sparsely represented. Therefore, it was not possible to obtain precise and complete information about the number and locations of most of these QTL, and to identify specific genes and the underlying genetic variation, or to obtain markers with sufficiently robust associations with the QTL to warrant their integration into MAS programs. Cost has been the main factor limiting the number of microsatellites that could reasonably be used. However, with the advent of SNP arrays, this problem has largely been overcome for agriculture animals such as cattle, pigs and chicken (Groenen et al., 2009; Matukumalli et al., 2009; Ramos et al., 2009) and great progress is being made for some aquaculture species such as Atlantic salmon (Kent M et al., in preparation).

### 3.6 Expression profiling (microarrays and qPCR)

The concept of gene expression was introduced above in the discussion on expressed sequence tags, or ESTs and EST libraries, which are assemblages of all small sequences of all the expressed genes in a tissue at a particular time point. The term “expression profiling” refers to the examination of the genes expressed in an individual, or group of individuals, under specific conditions. This is based on the fact that gene expression varies with genotype (i.e., the suite of genes and their alleles that are present within an individual’s genome) and environment, which cues the expression of particular genes as well as the extent of expression. Expression profiling studies generally compare groups (treatment groups) of individuals that differ in some way, such as genetic background, or environmental surroundings. Thereby, one can examine which genes are relatively over- and under-expressed across treatment groups using statistical analyses, thus gaining insight into the genes that govern particular traits that differ between the subjects.

There are two approaches to studying gene expression that are now standard: real-time PCR (also called quantitative real time PCR or qPCR, or sometimes abbreviated qrtPCR), and microarray analysis. The former is used to examine the expression of a single gene, or a few genes at a time. The technique follows the general principles of regular PCR, in which DNA (in this case, having been reverse-transcribed from RNA, which is present in amounts that are directly proportional to the extent of gene expression) is amplified from a template strand, with the key difference that the amount of product is measured in real time. This is accomplished in one of two ways: 1) the DNA components in the reaction mixture are tagged with dyes that fluoresce when integrated into the DNA molecule, and the level of fluorescence (i.e., corresponding to the amount of product) is measured after each PCR cycle, or 2) sequence-specific probes are similarly tagged, and fluoresce when hybridized to the complementary DNA target, thus providing a direct measurement of the amount of target present after each PCR cycle. The benefits of qPCR are that the process is gene-specific and the results are relatively accurate and reproducible. The cons are that only a small number of genes within a genome can be examined at once, and that sequence-specific PCR primers must be designed, which is dependent on knowing the sequence of the gene *a priori*. This is further complicated when examining gene families or duplicated genes, as it is often

not feasible to design primers specific for each gene, a problem of particular relevance when there is no genome sequence for the study organism.

Whereas qPCR targets single genes at a time, DNA microarrays are specifically designed to examine the gene expression at thousands of genes at once. A microarray (also referred to as a “chip”) is comprised of thousands of single-stranded DNA spots, or “features” attached in known locations to a surface, such as a glass slide, with each spot corresponding to a gene. To analyze gene expression using a microarray, cDNA (again, reverse-transcribed from RNA) that has been fluorescently tagged is washed over the slide, and complementary strands will hybridize to corresponding spots on the microarray. Thus, when the level of fluorescence is measured, the intensity of each spot will correspond to extent of the expression of that particular gene. Statistical tests are then used to determine the relative expression of each gene on the microarray for the individual/tissue/environmental condition being tested. This method is extremely powerful given the number of genes that can be assessed at once. For example, several microarrays have been designed using the Atlantic salmon EST library, including an initial 3.7K (i.e., 3700 spots) array (Rise et al., 2004), followed by 16K (von Schalburg et al., 2005) and a 32K cDNA (Koop et al., 2008) arrays, as well as two recent 44K oligonucleotide arrays. These resources, which are collectively referred to as the GRASP (Genomic Research on All Salmon Project) arrays, have been used extensively for gene expression studies throughout the salmonid genomics world. For example, recent studies have used the GRASP arrays to look for genetic markers of immune responses in Atlantic salmon (LeBlanc et al., 2010), as well as of thermal stress in rainbow trout (Lewis et al., 2010) and Arctic charr (Quinn et al., 2011a, 2011b), and to predict spawning failure in wild sockeye populations (Miller et al., 2011).

Despite their utility, however, microarrays (particularly those based on cDNA clones) have inherent drawbacks, such that they are resource intensive and highly susceptible to technical variations as well as statistical pitfalls (Martin et al., 2007). The latter may result in Type I errors (false positive results), a particular concern when using a large microarray, as the false-positive rate increases as the number of spots increases. Thus, often microarray data are confirmed, or validated, by selecting a few genes to test using qPCR. Another challenge is the potential for cross-hybridization between similar transcripts, which is a problem of particular concern for duplicated genomes, or for large gene families, such as that of the hemoglobin genes, which contains multiple members of high sequence similarity (Quinn et al., 2010). This can mute out evidence of differential expression between individual transcripts, and instead, produce the general trends for all highly similar genes on the microarray. In addition, this phenomenon may cause information to be lost if the elevated expression of a small number of family members is spread among all similar genes on the microarray such that no spots meet the statistical filters assigned. The use of transcript-specific oligonucleotide arrays may overcome some of these challenges. For these reasons, profiling using microarrays is sometimes described as an exploratory tool, for which the data must be interpreted with a caution, or which must be followed up or correlated with other types of analysis (Lewis et al., 2010; Quinn et al., 2011a).

### **3.7 Bioinformatics and genomics databases**

The advancement of genomics technologies has resulted in the generation of vast amounts of biological data. This has fueled the need for computational systems and techniques to

manage such data as well as for data analysis, the development of which has led to the advent of the field of bioinformatics. Specifically, bioinformatics refers to the application of computer science and information technology to biology to increase the understanding of biological processes. Bioinformatics and genomics now go hand in hand; indeed, often the bottlenecks for obtaining and interpreting genomics results occur in the bioinformatics steps, as high throughput genomics technologies produce more data than the available computational tools can handle. Individuals specializing both in biology and computer science are becoming increasingly in demand. For genomics, the general goals of bioinformatics include the development of computational tools for data management (e.g., databases and interactive online tools), data manipulation (map assembly, sequence/genome assembly, protein structure alignment) and data mining (e.g., pattern recognition, genome annotation, sequence comparisons, predictions of gene expression and genome-wide association studies). Given that the fundamental goals of genomics research are shared across all genomes, standard computational tools can often be used for numerous genomes, requiring only minor tweaking and manipulating to accommodate the specific resources for the species of interest.

The concept of “open data”, or public accessibility is generally adopted for all genomics projects, and most genomics data are posted online in publically accessible databases. Indeed, it is a requirement for publication in most peer-reviewed academic journals that sequences as well as gene expression data be made available online before any results can be considered for publication within the journal. Besides preventing redundant work and enabling research groups to build off of the work of others, open access data allows the potential for cross-species comparisons, which is particularly true among the Salmonidae. That is, online, public databases such as those managed by NCBI (National Center for Biotechnology Information), Ensembl (Ensembl Genome Browser), and GEO (Gene Expression Omnibus) provide an opportunity to make predictions about the genes in closely related species for which there are few genomic resources available (e.g., Arctic charr). Furthermore, for most species for which there is a large genomics program, a species-specific database will be generated. For example, the Atlantic salmon database, publically available at [www.asalbase.org](http://www.asalbase.org), comprises both the linkage and physical maps, as well as all of the BAC-end sequences and marker information. Also included are tools that enable comparative synteny analyses between Atlantic salmon and the other sequenced species. The next step for Asalbase will be integrating the sequence scaffolds for the whole genome sequence. Finally, the last decade saw many research labs pack away their pipettes and centrifuges and replace them with large computer networks and powerful servers, as the wealth of genomics data available online now allows entire experiments, such as comparative synteny analyses, evolutionary studies and even gene expression meta-analyses to be conducted “*in silico*”, or with computers alone. This not only greatly improves the speed of research, but also reduces the cost and number of live specimens, or research animals required.

## 4. MAS in aquaculture

### 4.1 Examples of QTL-based markers for aquaculture

Over the last decade breeders have begun to include molecular genetics in their strategies for genetically improving salmonids for aquaculture production efficiency. Markers that

exhibit strong associations with QTL that play a substantial role in a desirable trait can be integrated into MAS programs. That is, rather than using trial and error to cross individuals that exhibit phenotypic traits of interest, breeders can purposely select individuals that exhibit the marker associated with the desirable QTL. This fine-scale, targeted approach to artificial selection has the potential to be substantially more precise, efficient and effective than traditional, phenotype based approaches.

One example of an economically important trait for which there is a QTL is Infectious Salmon Anemia (ISA) resistance in Atlantic salmon. Previous studies determined that there is a genetic component to ISA resistance in Atlantic salmon (Gjøen & Bentsen 1997, Ødegård et al., 2007) and a QTL for this trait was mapped (Moen et al., 2004). A more recent study confirmed this QTL and showed that it explains 6% of the phenotypic variation and 32-47% of the additive genetic variation (Moen et al., 2007). A comparative genomics analysis of the most likely QTL region suggested several candidate genes for ISA resistance (Li et al., 2011a). These genes remain to be validated based on the association of polymorphisms within them and the trait. However, it is anticipated that a combination of fine mapping and genomic sequence information will identify the causative allele that confers resistance and that this will be incorporated into a MAS breeding program.

Another excellent example of the potential of MAS involves resistance to a pathogen that causes Infectious Pancreatic Necrosis (IPN) in Atlantic salmon. A major QTL for IPN resistance in Atlantic salmon has been identified and confirmed (Gheyas et al., 2010; Houston et al., 2008; Moen et al., 2009a). A four-marker haplotype has been identified that strongly segregates with IPN resistance in Norwegian Atlantic salmon broodstock, providing a solid framework for linkage-based MAS in this population (Moen et al., 2009a).

These results support the prediction that selective breeding for resistance ISA and IPN (caused by viruses) as well as furunculosis (caused by a bacterium) should be possible given the heritabilities for these traits (Kjoglum et al., 2008). However, the speed at which selective breeding for these traits can be performed efficiently and effectively will depend on the ability to identify the underlying genetic basis for the resistance, and this requires a reference sequence for Atlantic salmon. Other traits of interest for QTL identification for MAS in aquaculture include tissue quality, the ability to thrive on commercial diets (i.e., increased vegetable oil content, pathway analysis for vaccine development, and ability/suitability for high-density rearing conditions. Overall, integrating MAS into broodstock development programs can provide an effective and efficient, science-based means of minimizing animal stress while maximizing growth, health and sustainability

#### **4.2 Challenges facing MAS for aquaculture**

Currently, given the lack of fully sequenced genomes for aquaculture species, genomics-based selection, or MAS is fully dependent on a having a suite of well-developed genomics tools for the genome of the species of interest. In addition to the cost associated with the development of these tools, the complex nature of fish genomes combined with the large number of diverse fish species of interest for aquaculture complicates this goal. However, the availability of fully sequenced reference genomes for a few key species groups would go a long way to fill in the major under-represented gaps in the fish phylogeny. As illustrated in Figure 1, there is no reference genome from a closely related species for either the

Salmoniformes or the Gadiformes. However, work on the genomes of Atlantic salmon (Davidson et al., 2010) and Atlantic cod (Johansen et al., 2009) is in progress, and the results are eagerly anticipated by the aquaculture community that breeds these and related species.

## 5. Role of whole genome sequencing (WGS) in genomics

### 5.1 Benefits of whole genome sequences

The availability of a complete genome sequence for Atlantic salmon will have a major impact on all sectors of the international salmonid community. For the aquaculture industry it will provide a complete suite of genetic markers for the identification of the genes and alleles responsible for production traits. Companies will be able to develop tailored broodstock using nucleotide or allele assisted selection rather than the more general marker assisted selection. In conjunction with traditional breeding practices, this approach promises rapid gains that will make companies who embrace this technology more competitive.

### 5.2 Methods and approaches to whole genome sequencing

The conventional approach to sequencing whole genomes is to break the genome into small parts, and sequence each part using Sanger sequencing (reviewed in Hutchison, 2007). Although this approach remains the gold standard for sequence and assembly quality, limitations with respect to cost, labor-intensiveness and speed have fueled the demand for new approaches to DNA sequencing. In recent years, several novel high-throughput sequencing platforms, or 'next generation' sequencing approaches have entered the market (for a review, see Metzker, 2010). Most of these are targeted to the goal of re-sequencing an entire human genome for <\$1,000, and their capabilities with respect to sequencing whole genomes *de novo* (i.e., for the first time, and without a reference sequence for comparative assembly), remain unknown. This is complicated by the diverse, highly repetitive and overall complex nature of the fish genomes given that, these technologies are currently limited by their sequence read lengths and difficulties with specific types of repetitive sequences. Generally, the next generation sequencing platforms are capable of producing vast amounts of sequence data in very short periods of time for significantly lower cost than traditional Sanger sequencing, but the bottleneck lies in developing the computational tools to handle the data, and ultimately, to assemble the short sequence reads into contiguous (gap-free) stretches of DNA sequence when no reference genome is available (Quinn et al., 2008).

Aside from their potential contribution to whole genome sequencing, next generation sequencing technologies have countless implications for genomics. One of the most common uses for next generation sequencing in human genomics is re-sequencing, i.e., sequencing the genomes of individuals for which there is a full reference sequence. This could be applied to aquaculture to look for allelic variations between individuals of the same species that show differences in characteristics of interest (e.g., disease resistance or growth). Similarly, next generation sequencing has been used for "transcriptome sequencing" (also referred to as "transcriptomics") i.e., sequencing the entire repertoire of transcribed RNA of a species. This is a form of expression profiling, as the relative number of transcripts of genes can be compared against one another, thus providing insight into the expression

levels of the genes at a given time. This approach is being used much more extensively as the cost of next generation sequencing technologies decreases. Finally, next generation sequencing provides an efficient and effective means identifying molecular markers, including microsatellites, SNPs (Chan, 2009) and RAD (Restriction site Associated DNA) markers, which are markers (e.g., SNPs) contained within RAD tags, or the sequences immediately flanking specific DNA sites that are targeted by restriction enzymes. The use of next generation sequencing technologies to sequence RAD tags to identify RAD markers is referred to as RADseq, and its popularity is increasing quickly given its utility for identifying individual-specific markers, and thus to generate an indexed marker library for an individual organism. Therefore, as the capabilities and availability of next generation technologies continue to advance, we are sure to see drastic changes in the approaches to genomics centered around high-throughput, next generation technologies as well as bioinformatics-based analyses. These advances will provide vast opportunities for integrating genomics into aquaculture breeding programs.

## 6. Conclusion

Genomics has become an increasingly powerful field with far-reaching implications from medicine to pharmacology to industry to ecology. When applied to aquaculture, genomics has vast potential to improve factors such as production, efficiency, animal welfare, sustainability, reducing inbreeding, and increasing the profitability of the industry. The current genomics tools for fish, and specifically for aquaculture species, have already enabled drastic improvements to characteristics such as growth and disease resistance, and show great promise for future developments; however, their utility remains limited by the lack of whole genome sequences for a few key species, which could provide reference sequences for other aquaculture species. As these sequences, particularly those of Atlantic salmon, rainbow trout and Atlantic cod, become available over the next several years, which will certainly correspond to further advances in technology and computational tools, the potential for genomics to facilitate major developments in aquaculture is enormous.

## 7. Glossary of terms

**Alleles:** Alternative forms (two or more) of a gene, located at a specific position on a chromosome.

**Annotation:** Identification of the genomic position of intron-exon boundaries, regulatory sequences, repeats, gene names and protein products within a DNA sequence.

**Bioinformatics:** refers to the application of computer science and information technology to developing computationally intensive techniques to increase the understanding of biological processes.

**Complex trait:** Trait influenced by more than one genetic and/or environmental factor.

**cDNA:** Complementary DNA, which is DNA synthesized from a mature mRNA template by reverse transcription.

**DNA sequence:** Linear representation of a DNA strand by the presence of the four nucleotide bases (A, T, C, G).

**Ensembl Genome Browser:** A joint project between EMBL-EBI and the Wellcome Trust Sanger Institute to develop a software system that produces and maintains automatic annotation on selected eukaryotic genomes ([www.ensembl.org](http://www.ensembl.org)).

**EST:** Expressed sequence tag, which is a short sequence obtained from one shot sequencing of cDNA, corresponding to a fragment (~500 bp) of an expressed gene.

**Gene:** Name given to stretches of DNA that code for a specific protein for a specific characteristic or function. It represents the heredity unit in living organisms.

**Gene Expression Omnibus (GEO):** A database repository of high-throughput gene expression data and hybridization arrays, chips and microarrays.

**Generation time:** The average interval between the birth of an individual and the birth of its offspring.

**Genome:** Is the entirety of an organism's hereditary information. It is encoded either in DNA or RNA (some viruses).

**Genomics:** Discipline in genetics concerning the study of the genomes of organisms.

**Genotype:** The total set of alleles possessed by an organism that determines a specific characteristic or trait. It comprises the entire complex of genes inherited from both parents.

**Heritability:** Proportion of phenotypic variation in a population that is due to genetic variation between individuals.

**High-throughput:** Processes usually performed via increased levels of automation and robotics. In sequencing: involves the application of rapid sequencing technology at the scale of whole genomes.

**Locus:** (pl. loci) Specific location of a gene or DNA sequence on a chromosome.

**Marker-assisted selection (MAS):** The use of DNA markers linked to traits of interest to assist in the selection of individuals for breeding purposes.

**Molecular marker:** Specific fragments of DNA that can be identified within the whole genome. These can be associated with the position of a particular gene or the inheritance of a particular characteristic.

**Molecular pathway:** Series of molecular processes that are connected by their intermediates such that the products of one process may trigger or participate in the other.

**Morphology:** The visible form and structure of living organisms.

**mRNA:** Molecule of RNA encoding a specific protein product. mRNA is transcribed from a DNA template in the cell nucleus then moves to the cytoplasm where it is translated by the ribosomes.

**NCBI:** The National Center for Biotechnology Information, which houses the world's biggest sequence datasets in GenBank and an index of biomedical research articles in PubMed.

**Non-coding:** Components of an organism's genome (DNA sequences) that do not encode for protein sequences.



**Phenotype:** Observable physical or biochemical characteristics of an organism, as determined by both genetic makeup and environmental influences.

**Polymerase Chain Reaction (PCR):** Scientific technique designed to amplify a DNA fragment across several orders of magnitude, generating thousands to millions of copies of a particular DNA sequence

**Quantitative trait locus (QTL):** Stretch of DNA containing or linked to genes that underlie a quantitative trait of interest.

**RAD:** Restriction site Associated DNA. RAD tags are the sequences that immediately flank a restriction enzyme site. RAD markers refer to sequence variations identified within RAD tags that can be used as molecular markers (e.g., SNPs).

**Reference genome:** Nucleic acid sequence database, assembled into a whole genome and representative of a species' genetic code. Typically used as a guide on which new genomes are built.

**Restriction enzyme:** An enzyme that cuts DNA at a specific recognition site/sequence, referred to as a restriction site.

**RNA:** Ribonucleic acid. Single stranded molecule transcribed from one strand of DNA. Unlike DNA, the sugar in RNA is ribose and one of the four bases, T (thymine) is replaced by U (uracil). There are several types of RNA mainly involved in protein synthesis (mRNA, tRNA, rRNA).

**Sequence assembly:** Refers to aligning and merging sequence fragments into a longer sequence.

**Synten:** Physical co-localization of genetic loci on the same chromosome within an individual or species.

**Transcription:** Process involved in the synthesis of mRNA from a DNA template, catalyzed by RNA polymerase.

**Translation:** Process in which the messenger RNA (mRNA) produced by transcription is decoded by the ribosome to produce a specific amino acid chain, or polypeptide, that will later fold into an active protein.

**Whole genome sequence:** The complete DNA sequence of the genome of an organism.

## 8. References

- Allendorf, F.W. & Thorgaard, G.H. (1984). Tetraploidy and the Evolution of Salmonid Fishes. in *Evolutionary Genetics of Fishes*, Plenum Press, pp. 55, ISBN 978-0306415203, New York, USA.
- Anti Vasemagi, Jan Nilsson, Craig R. Primmer (2005). Seventy-five EST-linked Atlantic Salmon (*Salmo salar* L.) Microsatellite Markers and their Cross-amplification in Five Salmonid Species. *Molecular Ecology Notes*, Vol. 5, No. 2, pp. 282-288, ISSN 1471-8286.

- Baranski, M., Moen, T. & Vage, D.I. (2010). Mapping of Quantitative Trait Loci for Flesh Colour and Growth Traits in Atlantic Salmon (*Salmo salar*). *Genetics, Selection, Evolution*, Vol. 42, No. 1, pp. 17. ISSN 1297-9686.
- 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*, Vol. 101, No. 4, pp. 381-391, ISSN 0018-067X.
- Chan, E.Y. (2009). Next-Generation Sequencing Methods: Impact of Sequencing Accuracy on SNP Discovery. In *Single Nucleotide Polymorphisms*, Humana Press, pp. 95-111, ISBN 978-1-60327-411-1.
- Chistiakov, D.A., Hellemans, B., Haley, C.S., Law, A.S., Tsigenopoulos, C.S., Kotoulas, G., Bertotto, D., Libertini, A. & Volckaert, F.A. (2005). A Microsatellite Linkage Map of the European Sea Bass *Dicentrarchus labrax* L. *Genetics*, Vol. 170, No. 4, pp. 1821-1826, ISSN 0016-6731.
- Coimbra, M.R.M., Kobayashi, K., Koretsugu, S., Hasegawa, O., Ohara, E., Ozaki, A., Sakamoto, T., Naruse, K. & Okamoto, N. (2003). A genetic Linkage Map of the Japanese Flounder, *Paralichthys olivaceus*. *Aquaculture*, Vol. 220, No. 1-4, pp. 203, ISSN 0044-8486.
- Crick, F. (1970). Central Dogma of Molecular Biology. *Nature*, Vol. 227, No. 5258, pp. 561-563, ISSN 0028-0836.
- Danzmann, R.G., Cairney, M., Davidson, W.S., Ferguson, M.M., Gharbi, K., Guyomard, R., Holm, L.E., Leder, E., Okamoto, N., Ozaki, A., Rexroad, C.E., 3rd, Sakamoto, T., Taggart, J.B. & Woram, R.A. (2005). A Comparative Analysis of the Rainbow Trout Genome with 2 Other Species of Fish (Arctic charr and Atlantic salmon) Within the Tetraploid Derivative Salmonidae Family (subfamily: Salmoninae). *Genome*, Vol. 48, No. 6, pp. 1037-1051, ISSN 0831-2796.
- Danzmann, R.G., Davidson, E.A., Ferguson, M.M., Gharbi, K., Koop, B.F., Hoyheim, B., Lien, S., Lubieniecki, K.P., Moghadam, H.K., Park, J., Phillips, R.B. & Davidson, W.S. (2008). Distribution of Ancestral Proto-Actinopterygian Chromosome Arms Within the Genomes of 4R-derivative Salmonid Fishes (Rainbow trout and Atlantic salmon). *BMC Genomics*, Vol. 9, No.1, ISSN 1471-2164.
- Davidson, W.S., Koop, B.F., Jones, S.J., Iturra, P., Vidal, R., Maass, A., Jonassen, I., Lien, S. & Omholt, S.W. (2010). Sequencing the Genome of the Atlantic Salmon (*Salmo salar*). *Genome Biology*, Vol. 11, No. 9, pp. 403, ISSN 1465-6906.
- Dominik, S., Henshall, J.M., Kube, P.D., King, H., Lien, S., Kent, M.P. & Elliott, N.G. (2010). Evaluation of an Atlantic Salmon SNP Chip as a Genomic Tool for the Application in a Tasmanian Atlantic Salmon (*Salmo salar*) Breeding Population. *Aquaculture*, Vol. 308, Suppl. 1, pp. S56, ISSN 0044-8486.
- Garcia de Leaniz, C., Fleming, I.A., Einum, S., Verspoor, E., Jordan, W.C., Consuegra, S., Aubin-Horth, N., Lajus, D., Letcher, B.H., Youngson, A.F., Webb, J.H., Vållstad, L.A., Villanueva, B., Ferguson, A. & Quinn, T.P. (2007). A Critical Review of Adaptive Genetic Variation in Atlantic Salmon: Implications for Conservation. *Biological Reviews*, Vol. 82, No. 2, pp. 173-211, ISSN 1469-185X.
- Genet, C., Dehais, P., Palti, Y., Gao, G., Gavory, F., Wincker, P., Quillet, E. & Boussaha, M. (2011). Analysis of BAC-end Sequences in Rainbow Trout: Content

- Characterization and Assessment of Synteny Between Trout and Other Fish Genomes. *BMC Genomics*, Vol. 12, No. 1, ISSN 1471-2164.
- "genomics" Mirriam-Webster.com, Mirriam-Webster, 2011. Web. 27 June 2011.
- 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*, Vol. 172, No. 4, pp. 2405-2419, ISSN 0016-6731.
- Gheyas, A.A., Haley, C.S., Guy, D.R., Hamilton, A., Tinch, A.E., Mota-Velasco, J.C. & Woolliams, J.A. (2010). Effect of a Major QTL Affecting IPN Resistance on Production Traits in Atlantic Salmon. *Animal Genetics*, Vol. 41, No. 6, pp. 666-668, ISSN 1365-2052.
- Gilbey, J., Verspoor, E., McLay, A. & Houlihan, D. (2004). A Microsatellite Linkage Map for Atlantic Salmon (*Salmo salar*). *Animal Genetics*, Vol. 35, No. 2, pp. 98-105, ISSN 1365-2052.
- Gilbey, J., Verspoor, E., Mo, T.A., Sterud, E., Olstad, K., Hytterod, S., Jones, C. & Noble, L. (2006). Identification of Genetic Markers Associated with *Gyrodactylus salaris* Resistance in Atlantic Salmon *Salmo salar*. *Diseases of Aquatic Organisms*, Vol. 71, No. 2, pp. 119-129, ISSN 0177-5103.
- Gjøen, H.M. & Bentsen, H.B. (1997). Past, Present, and Future of Genetic Improvement in Salmon Aquaculture. *ICES Journal of Marine Science: Journal du Conseil*, Vol. 54, No. 6, pp. 1009-1014, ISSN 1095-9289.
- Goddard, M.E. & Hayes, B.J. (2009). Mapping Genes for Complex Traits in Domestic Animals and Their use in Breeding Programmes. *Nature Reviews. Genetics*, Vol. 10, No. 6, pp. 381-391, ISSN 1471-0056.
- Groenen, M.A., Wahlberg, P., Foglio, M., Cheng, H.H., Megens, H.J., Croijmans, R.P., Besnier, F., Lathrop, M., Muir, W.M., Wong, G.K., Gut, I. & Andersson, L. (2009). A High-density SNP-based Linkage Map of the Chicken Genome Reveals Sequence Features Correlated with Recombination Rate. *Genome Research*, Vol. 19, No. 3, pp. 510-519, ISSN 1088-9051.
- Gupta, P.K., Mir, R.R., Mohan, A. & Kumar, J. (2008). Wheat Genomics: Present Status and Future Prospects. *International Journal of Plant Genomics*, Vol. 2008, pp.36, ISSN 1687-5389.
- 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*, Vol. 7, No. 1, ISSN 1471-2164.
- Haidle, L., Janssen, J., Gharbi, K., Moghadam, H., Ferguson, M. & Danzmann, R. (2008). Determination of Quantitative Trait Loci (QTL) for Early Maturation in Rainbow Trout (*Oncorhynchus mykiss*). *Marine Biotechnology*, Vol. 10, No. 5, pp. 579-592, ISSN 1436-2228.
- Hayes, B.J., Nilsen, K., Berg, P.R., Grindflek, E. & Lien, S. (2007). SNP Detection Exploiting Multiple Sources of Redundancy in Large EST Collections Improves Validation Rates. *Bioinformatics*, Vol. 23, No. 13, pp. 1692-1693, ISSN 1367-4803.

- Houle, D., Govindaraju, D.R. & Omholt, S. (2010). Phenomics: The Next Challenge. *Nature Reviews. Genetics*, Vol. 11, No. 12, pp. 855-866, ISSN 1471-0056.
- Houston, R.D., Bishop, S.C., Hamilton, A., Guy, D.R., Tinch, A.E., Taggart, J.B., Derayat, A., McAndrew, B.J. & Haley, C.S. (2009). Detection of QTL Affecting Harvest Traits in a Commercial Atlantic Salmon Population. *Animal Genetics*, Vol. 40, No. 5, pp. 753-755, ISSN 1365-2052.
- Houston, R.D., Gheyas, A., Hamilton, A., Guy, D.R., Tinch, A.E., Taggart, J.B., McAndrew, B.J., Haley, C.S. & Bishop, S.C. (2008). Detection and Confirmation of a Major QTL Affecting Resistance to Infectious Pancreatic Necrosis (IPN) in Atlantic Salmon (*Salmo salar*). *Developments in Biologicals*, Vol. 132, pp. 199-204, ISSN 1424-6074.
- Houston, R.D., Haley, C.S., Hamilton, A., Guy, D.R., Mota-Velasco, J.C., Gheyas, A.A., Tinch, A.E., Taggart, J.B., Bron, J.E., Starkey, W.G., McAndrew, B.J., Verner-Jeffreys, D.W., Paley, R.K., Rimmer, G.S., Tew, I.J. & Bishop, S.C. (2010). The susceptibility of Atlantic Salmon Fry to Freshwater I pancreatic Necrosis is Largely Explained by a Major QTL. *Heredity*, Vol. 105, No. 3, pp. 318-327, ISSN 0018-067X.
- Houston, R.D., Haley, C.S., Hamilton, A., Guy, 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*, Vol. 178, No. 2, pp. 1109-1115, ISSN 0016-6731.
- Hubert, S., Higgins, B., Borza, T. & Bowman, S. (2010). Development of a SNP Resource and a Genetic Linkage Map for Atlantic Cod (*Gadus morhua*). *BMC Genomics*, Vol. 11, No. 1, ISSN 1471-2164.
- Hutchison, C.A. 3rd (2007). DNA Sequencing: Bench to Bedside and Beyond. *Nucleic Acids Research*, Vol. 35, No. 18, pp. 6227-6237, ISSN 1362-4962.
- Johansen, S.D., Coucheron, D.H., Andreassen, M., Karlsen, B.O., Furmanek, T., Jorgensen, T.E., Emblem, A., Breines, R., Nordeide, J.T., Moum, T., Nederbragt, A.J., Stenseth, N.C. & Jakobsen, K.S. (2009). Large-scale Sequence Analyses of Atlantic Cod. *New Biotechnology*, Vol. 25, No. 5, pp. 263-271, ISSN 1871-6784.
- Jones, C.S., Lockyer, A.E., Verspoor, E., Secombes, C.J. & Noble, L.R. (2002). Towards Selective Breeding of Atlantic Salmon for Sea Louse Resistance: Approaches to Identify Trait Markers. *Pest Management Science*, Vol. 58, No. 6, pp. 559-568, ISSN 1526-4998.
- Kang, J.H., Kim, W.J. & Lee, W.J. (2008). Genetic Linkage Map of Olive Flounder, *Paralichthys olivaceus*. *International Journal of Biological Sciences*, Vol. 4, No. 3, pp. 143-149, ISSN 1449-2288.
- Karlsson, S., Moen, T., Lien, S., Glover, K.A. & Hindar, K. (2011). Generic Genetic Differences Between Farmed and Wild Atlantic Salmon Identified from a 7K SNP-chip. *Molecular Ecology Resources*, Vol. 11, Suppl. 1, pp. 247-253, ISSN 1755-098X.
- Kim, D.H., Kim, K.H., Van, K., Kim, M.Y. & Lee, S.H. (2010). Fine Mapping of a Resistance Gene to Bacterial Leaf Pustule in Soybean. *Theoretical and Applied Genetics. Theoretische und Angewandte Genetik*, Vol. 120, No. 7, pp. 1443-1450, ISSN 0040-5752.
- King, T.L., Kalinowski, S.T., Schill, W.B., Spidle, A.P. & Lubinski, B.A. (2001). Population Structure of Atlantic Salmon (*Salmo salar* L.): A Range-wide Perspective from

- Microsatellite DNA Variation. *Molecular Ecology*, Vol. 10, No. 4, pp. 807-821, ISSN 0962-1083.
- Kjoglum, S., Larsen, S., Bakke, H.G. & Grimholt, U. (2008). The Effect of Specific MHC Class I and Class II Combinations on Resistance to Furunculosis in Atlantic Salmon (*Salmo salar*). *Scandinavian Journal of Immunology*, Vol. 67, No. 2, pp. 160-168, ISSN 0300-9475.
- Koop, B.F., von Schalburg, K.R., Leong, J., Walker, N., Lieph, R., Cooper, G.A., Robb, A., Beetz-Sargent, M., Holt, R.A., Moore, R., Brahmhatt, S., Rosner, J., Rexroad, C.E., 3rd, McGowan, C.R. & Davidson, W.S. (2008). A Salmonid EST Genomic Study: Genes, Duplications, Phylogeny and Microarrays. *BMC Genomics*, Vol. 9, No. 1, ISSN 1471-2164.
- Kucuktas, H., Wang, S., Li, P., He, C., Xu, P., Sha, Z., Liu, H., Jiang, Y., Baoprasertkul, P., Somridhivej, B., Wang, Y., Abernathy, J., Guo, X., Liu, L., Muir, W. & Liu, Z. (2009). Construction of Genetic Linkage Maps and Comparative Genome Analysis of Catfish Using Gene-associated Markers. *Genetics*, Vol. 181, No. 4, pp. 1649-1660, ISSN 0016-6731.
- Kuhl, H., Sarropoulou, E., Tine, M., Kotoulas, G., Magoulas, A. & Reinhardt, R. (2011). A Comparative BAC Map for the Gilthead Sea Bream (*Sparus aurata* L.). *Journal of Biomedicine & Biotechnology*, Vol. 2011, 7 pp, ISSN 1110-7243
- Kuttner, E., Moghadam, H.K., Skulason, S., Danzmann, R.G. & Ferguson, M.M. (2011). Genetic Architecture of Body Weight, Condition Factor and Age of Sexual Maturation in Icelandic Arctic Charr (*Salvelinus alpinus*). *Molecular Genetics and Genomics*, Vol. 286, No. 1, pp. 67-79, ISSN 1617-4623.
- LeBlanc, F., Laflamme, M. & Gagne, N. (2010). Genetic Markers of the Immune Response of Atlantic Salmon (*Salmo salar*) to Infectious Salmon Anemia Virus (ISAV). *Fish & Shellfish Immunology*, Vol. 29, No. 2, pp. 217-232, ISSN 1050-4648.
- Lee, B.Y., Lee, W.J., Strelman, J.T., Carleton, K.L., Howe, A.E., Hulata, G., Slettan, A., Stern, J.E., Terai, Y. & Kocher, T.D. (2005). A Second-generation Genetic Linkage Map of Tilapia (*Oreochromis spp.*). *Genetics*, Vol. 170, No. 1, pp. 237-244, ISSN 0016-6731.
- Lewis, J.M., Hori, T.S., Rise, M.L., Walsh, P.J. & Currie, S. (2010). Transcriptome Responses to Heat Stress in the Nucleated Red Blood Cells of the Rainbow Trout (*Oncorhynchus mykiss*). *Physiological Genomics*, Vol. 42, No. 3, pp. 361-373, ISSN 1094-8341.
- Li, J., Boroevich, K.A., Koop, B.F. & Davidson, W.S. (2011a). Comparative Genomics Identifies Candidate Genes for Infectious Salmon Anemia (ISA) Resistance in Atlantic Salmon (*Salmo salar*). *Marine Biotechnology*, Vol. 13, No. 2, pp. 232-241, ISSN 1436-2228.
- Li, Y., Xu, P., Zhao, Z., Wang, J., Zhang, Y. & Sun, X.W. (2011b). Construction and Characterization of the BAC Library for Common Carp *Cyprinus Carpio* L. And Establishment of Microsynteny with Zebrafish *Danio Rerio*. *Marine Biotechnology*, Vol. 13, No. 4, pp. 706-712, ISSN 1436-2228.
- Lien, S., Gidskehaug, L., Moen, T., Hayes, B.J., Berg, P.R., Davidson, W.S., Omholt, S.W., & Kent, M.P. (2011). A dense SNP-based linkage map for Atlantic salmon (*Salmo salar*) reveals extended chromosome homeologies and striking difference in sex-specific recombination rates. *BMC Genomics* (in press).

- Liu, G., Jennen, D.G., Tholen, E., Juengst, H., Kleinwachter, T., Holker, M., Tesfaye, D., Un, G., Schreinemachers, H.J., Murani, E., Ponsuksili, S., Kim, J.J., Schellander, K. & Wimmers, K. (2007). A Genome Scan Reveals QTL for Growth, Fatness, Leanness and Meat Quality in a Duroc-Pietrain Resource Population. *Animal Genetics*, Vol. 38, No. 3, pp. 241-252, ISSN 1365-2052.
- Liu, H., Jiang, Y., Wang, S., Ninwichian, P., Somridhivej, B., Xu, P., Abernathy, J., Kucuktas, H. & Liu, Z. (2009). Comparative Analysis of Catfish BAC end Sequences with the Zebrafish Genome. *BMC Genomics*, Vol. 10, No. 1, ISSN 1471-2164.
- Lorenz, S., Brenna-Hansen, S., Moen, T., Roseth, A., Davidson, W.S., Omholt, S.W. & Lien, S. (2010). BAC-based Upgrading and Physical Integration of a Genetic SNP Map in Atlantic Salmon. *Animal Genetics*, Vol. 41, No. 1, pp. 48-54, ISSN 1365-2052.
- Lynch, M. & Walsh, B. (1998), *Genetics and Analysis of Quantitative Traits*, Sinauer Associates Inc., ISBN 0-87893-481-2, Sunderland, USA.
- Martin, L.J., Woo, J.G., Avery, C.L., Chen, H.S., North, K.E., Au, K., Broet, P., Dalmasso, C., Guedj, M., Holmans, P., Huang, B., Kuo, P.H., Lam, A.C., Li, H., Manning, A., Nikolov, I., Sinha, R., Shi, J., Song, K., Tabangin, M., Tang, R. & Yamada, R. (2007). Multiple Testing in the Genomics Era: Findings From Genetic Analysis Workshop 15, Group 15. *Genetic Epidemiology*, Vol. 31, Suppl. 1, pp. S124-131, ISSN 0741-0395.
- Martyniuk, C.J., Perry, G.M.L., Mogahadam, H.K., Ferguson, M.M. & Danzmann, R.G. (2003). The Genetic Architecture of Correlations Among Growth-related Traits and Male Age at Maturation in Rainbow Trout. *Journal of Fish Biology*, Vol. 63, No. 3, pp. 746-764, ISSN 1095-8649.
- Matukumalli, L.K., Lawley, C.T., Schnabel, R.D., Taylor, J.F., Allan, M.F., Heaton, M.P., O'Connell, J., Moore, S.S., Smith, T.P., Sonstegard, T.S. & Van Tassell, C.P. (2009). Development and Characterization of a High Density SNP Genotyping Assay for Cattle. *PloS One*, Vol. 4, No. 4, pp. e5350, ISSN 1932-6203.
- McClelland, E.K. & Naish, K.A. (2010). Quantitative Trait Locus Analysis of Hatch Timing, Weight, Length and Growth Rate in Coho Salmon, *Oncorhynchus kisutch*. *Heredity*, Vol. 105, No. 6, pp. 562-573, ISSN 0018-067X.
- McClelland, E.K. & Naish, K.A. (2008). A Genetic Linkage Map for Coho Salmon (*Oncorhynchus kisutch*). *Animal Genetics*, Vol. 39, No. 2, pp. 169-179, ISSN 1365-2052.
- McConnell, S.K., O'Reilly, P., Hamilton, L., Wright, J.M. & Bentzen, P. (1995). Polymorphic Microsatellite Loci from Atlantic Salmon (*Salmo salar*): Genetic Differentiation of North American and European Populations. *Canadian Journal of Fisheries and Aquatic Sciences*, Vol. 52, No. 9, pp. 1863-1872, ISSN 1205-7533.
- McGowan, C.R., Davidson, E.A., Woram, R.A., Danzmann, R.G., Ferguson, M.M. & Davidson, W.S. (2004). Ten Polymorphic Microsatellite Markers from Arctic Charr (*Salvelinus alpinus*): Linkage Analysis and Amplification in Other Salmonids. *Animal Genetics*, Vol. 35, No. 6, pp. 479-481, ISSN 1365-2052.
- Metzker, M.L. (2010). Sequencing Technologies - The Next Generation. *Nature Reviews. Genetics*, Vol. 11, No. 1, pp. 31-46, ISSN 1471-0056.
- Miller, K.M., Li, S., Kaukinen, K.H., Ginther, N., Hammill, E., Curtis, J.M., Patterson, D.A., Sierocinski, T., Donnison, L., Pavlidis, P., Hinch, S.G., Hruska, K.A., Cooke, S.J.,

- English, K.K. & Farrell, A.P. (2011) "Genomic Signatures Predict Migration and Spawning Failure in Wild Canadian Salmon. *Science*, Vol. 331, No. 6014, pp. 214-217, ISSN 0036-8075.
- Moen, T., Baranski, M., Sonesson, A.K. & Kjøglum, S. (2009a), "Confirmation and Fine-mapping of a Major QTL for Resistance to Infectious Pancreatic Necrosis in Atlantic Salmon (*Salmo salar*): Population-level Associations Between Markers and Trait", *BMC Genomics*, Vol.10, No.1, ISSN 1471-2164.
- Moen, T., Delghandi, M., Wesmajervi, M., Westgaard, J. & Fjalestad, K. (2009b), "A SNP/Microsatellite Genetic Linkage Map of the Atlantic Cod (*Gadus morhua*)", *Animal Genetics*, Vol. 40, No.6, pp. 993-996, ISSN 1365-2052.
- Moen, T., Fjalestad, K.T., Munck, H. & Gomez-Raya, L. (2004). A Multistage Testing Strategy for Detection of Quantitative Trait Loci Affecting Disease Resistance in Atlantic Salmon. *Genetics*, Vol. 167, No. 2, pp. 851-858, ISSN 0016-6731.
- Moen, T., Hayes, B., Baranski, M., Berg, P.R., Kjøglum, S., Koop, B.F., Davidson, W.S., Omholt, S.W. & Lien, S. (2008a). A Linkage Map of the Atlantic Salmon (*Salmo salar*) Based on EST-derived SNP Markers. *BMC Genomics*, Vol. 9, No.1, ISSN 1471-2164.
- Moen, T., Hayes, B., Nilsen, F., Delghandi, M., Fjalestad, K.T., Fevolden, S.E., Berg, P.R. & Lien, S. (2008b). Identification and Characterisation of Novel SNP Markers in Atlantic Cod: Evidence for Directional Selection. *BMC Genetics*, Vol. 9, No.1, ISSN 1471-2156.
- Moen, T., Hoyheim, B., Munck, H. & Gomez-Raya, L. (2004). A Linkage Map of Atlantic Salmon (*Salmo salar*) Reveals an Uncommonly Large Difference in Recombination Rate Between the Sexes. *Animal Genetics*, Vol. 35, No. 2, pp. 81-92, ISSN 1365-2052.
- Moen, T., Sonesson, A.K., Hayes, B., Lien, S., Munck, H. & Meuwissen, T.H. (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*, Vol. 8, No.1, ISSN 1471-2156.
- 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*, Vol. 277, No. 6, pp. 647-661, ISSN 1617-4623.
- Muir, W.M., Wong, G.K., Zhang, Y., Wang, J., Groenen, M.A., Crooijmans, R.P., Megens, H.J., Zhang, H., Okimoto, R., Vereijken, A., Jungerius, A., Albers, G.A., Lawley, C.T., Delany, M.E., MacEachern, S. & Cheng, H.H. (2008). Genome-wide Assessment of Worldwide Chicken SNP Genetic Diversity Indicates Significant Absence of Rare Alleles in Commercial Breeds. *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 105, No. 45, pp. 17312-17317, ISSN 0027-8424.
- Nelson, J.S. (2006). *Fishes of the World*, John Wiley and Sons Inc., ISBN 978-0-471-25031-9, New Jersey, USA.
- Ng, S.H., Artieri, C.G., Bosdet, I.E., Chiu, R., Danzmann, R.G., Davidson, W.S., Ferguson, M.M., Fjell, C.D., Hoyheim, B., Jones, S.J., de Jong, P.J., Koop, B.F., Krzywinski, M.I., Lubieniecki, K., Marra, M.A., Mitchell, L.A., Mathewson, C., Osoegawa, K., Parisotto, S.E., Phillips, R.B., Rise, M.L., von Schalburg, K.R., Schein, J.E., Shin, H.,

- Siddiqui, A., Thorsen, J., Wye, N., Yang, G. & Zhu, B. (2005). A Physical Map of the Genome of Atlantic Salmon, *Salmo salar*. *Genomics*, Vol. 86, No. 4, pp. 396-404, ISSN 0888-7543.
- 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*). *Animal Genetics*, Vol. 34, No. 2, pp. 102-115, ISSN 0268-9146.
- Nichols, K., Felip-Edo, A., Wheeler, P. & Thorgaard, G. (2008). The Genetic Basis of Smoltification-related Traits in *Oncorhynchus mykiss*. *Genetics*, Vol. 179, No. 3, pp. 1559-1575, ISSN 0016-6731.
- Norris, A.T., Bradley, D.G. & Cunningham, E.P. (2000). Parentage and Relatedness Determination in Farmed Atlantic Salmon (*Salmo salar*) Using Microsatellite Markers. *Aquaculture*, Vol. 182, No. 1-2, pp. 73, ISSN 0044-8486.
- Ødegård, J., Olesen, I., Gjerde, B. & Klemetsdal, G. (2007). Positive Genetic Correlation Between Resistance to Bacterial (Furunculosis) and Viral (Infectious Salmon Anaemia) Diseases in Farmed Atlantic Salmon (*Salmo salar*). *Aquaculture*, Vol. 271, No. 1-4, pp. 173, ISSN 0044-8486.
- Ohno, S. (1970). *Evolution by Gene Duplication*, Springer-Verlag, ISBN 0-04-575015-7, New York, USA.
- O'Malley, K.G., McClelland, E.K. & Naish, K.A. (2010). Clock Genes Localize to Quantitative Trait Loci for Stage-specific Growth in Juvenile Coho Salmon, *Oncorhynchus kisutch*. *The Journal of Heredity*, Vol. 101, No. 5, pp. 628-632, ISSN 0022-1503.
- O'Malley, K.G., Sakamoto, T., Danzmann, R.G. & Ferguson, M.M. (2003). Quantitative Trait Loci for Spawning Date and Body Weight in Rainbow Trout: Testing for Conserved Effects Across Ancestrally Duplicated Chromosomes. *Journal of Heredity*, Vol. 94, No. 4, pp. 273-284, ISSN 1465-7333.
- O'Reilly, P.T., Herbinger, C. & Wright, J.M. (1998). Analysis of Parentage Determination in Atlantic Salmon (*Salmo salar*) Using Microsatellites. *Animal Genetics*, Vol. 29, No. 5, pp. 363-370, ISSN 1365-2052.
- Palti, Y., Genet, C., Luo, M.C., Charlet, A., Gao, G., Hu, Y., Castano-Sanchez, C., Tabet-Canale, K., Krieg, F., Yao, J., Vallejo, R.L. & Rexroad, C.E.,3rd (2011). A First Generation Integrated Map of the Rainbow Trout Genome. *BMC Genomics*, Vol. 12, No. 1, ISSN 1471-2164.
- Palti, Y., Luo, M.C., Hu, Y., Genet, C., You, F.M., Vallejo, R.L., Thorgaard, G.H., Wheeler, P.A. & Rexroad, C.E.,3rd (2009). A First Generation BAC-based Physical Map of the Rainbow Trout Genome. *BMC Genomics*, Vol. 10, No. 1, ISSN 1471-2164.
- Perry, G.M., Danzmann, R.G., Ferguson, M.M. & Gibson, J.P. (2001). Quantitative Trait Loci for Upper Thermal Tolerance in Outbred Strains of Rainbow Trout (*Oncorhynchus mykiss*). *Heredity*, Vol. 86, No. 3, pp. 333-341, ISSN 0018-067X.
- Phillips, R.B., Keatley, K.A., Morasch, M.R., Ventura, A.B., Lubieniecki, K.P., Koop, B.F., Danzmann, R.G. & Davidson, W.S. (2009). Assignment of Atlantic salmon (*Salmo salar*) Linkage Groups to Specific Chromosomes: Conservation of Large Syntenic



- Blocks Corresponding to Whole Chromosome Arms in Rainbow Trout (*Oncorhynchus mykiss*). BMC Genetics, Vol. 10, No. 1, ISSN 1471-2156.
- Phillips, R.B., Nichols, K.M., DeKoning, J.J., Morasch, M.R., Keatley, K.A., Rexroad, C., 3rd, Gahr, S.A., Danzmann, R.G., Drew, R.E. & Thorgaard, G.H. (2006). Assignment of Rainbow Trout Linkage Groups to Specific Chromosomes. Genetics, Vol. 174, No. 3, pp. 1661-1670, ISSN 0016-6731.
- Presá, P. & Guyomard, R. (1996). Conservation of Microsatellites in Three Species of Salmonids. Journal of Fish Biology, Vol. 49, No. 6, pp. 1326-1329, ISSN 1095-8649.
- Quinn, N.L., Boroevich, K.A., Lubieniecki, K.P., Chow, W., Davidson, E.A., Phillips, R.B., Koop, B.F. & Davidson, W.S. (2010). Genomic Organization and Evolution of the Atlantic Salmon Hemoglobin Repertoire. BMC Genomics, Vol. 11, No.1, ISSN 1471-2164.
- Quinn, N.L., Levenkova, N., Chow, W., Bouffard, P., Boroevich, K.A., Knight, J.R., Jarvie, T.P., Lubieniecki, K.P., Desany, B.A., Koop, B.F., Harkins, T.T. & Davidson, W.S. (2008). Assessing the Feasibility of GS FLX Pyrosequencing for Sequencing the Atlantic Salmon Genome. BMC Genomics, Vol. 9, No. 1, ISSN 1471-2164.
- Quinn, N.L., McGowan, C.R., Cooper, G.A., Koop, B.F. & Davidson, W.S. (2011a). Identification of Genes Associated with Heat Tolerance in Arctic Charr Exposed to Acute Thermal Stress. Physiological Genomics, Vol. 43, No. 11, pp 685-96, ISSN 1094-8341.
- Quinn, N.L., McGowan, C.R., Cooper, G.A., Koop, B.F. & Davidson, W.S. (2011b). Ribosomal Genes and Heat Shock Proteins as Putative Markers for Chronic, Sub-lethal Heat Stress in Arctic Charr: Applications for Aquaculture and Wild Fish. Physiological Genomics, Jul 12 2011 [epub ahead of print].
- Ramos, A.M., Crooijmans, R.P., Affara, N.A., Amaral, A.J., Archibald, A.L., Beever, J.E., Bendixen, C., Churcher, C., Clark, R., Dehais, P., Hansen, M.S., Hedegaard, J., Hu, Z.L., Kerstens, H.H., Law, A.S., Megens, H.J., Milan, D., Nonneman, D.J., Rohrer, G.A., Rothschild, M.F., Smith, T.P., Schnabel, R.D., Van Tassell, C.P., Taylor, J.F., Wiedmann, R.T., Schook, L.B. & Groenen, M.A. (2009). Design of a High Density SNP Genotyping Assay in the Pig Using SNPs Identified and Characterized by Next Generation Sequencing Technology. PloS One, Vol. 4, No. 8, pp. e6524, ISSN 1932-6203.
- 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, Vol. 94, No. 2, pp. 166-172, ISSN 0018-067X.
- Rexroad, C.E., 3rd, 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, Vol. 6, No.1, ISSN 1471-2164.
- Rise, M.L., von Schalburg, K.R., Brown, G.D., Mawer, M.A., Devlin, R.H., Kuipers, N., Busby, M., Beetz-Sargent, M., Alberto, R., Gibbs, A.R., Hunt, P., Shukin, R., Zeznik, J.A., Nelson, C., Jones, S.R., Smailus, D.E., Jones, S.J., Schein, J.E., Marra, M.A., Butterfield, Y.S., Stott, J.M., Ng, S.H., Davidson, W.S. & Koop, B.F. (2004). Development and Application of a Salmonid EST Database and cDNA Microarray:

- Data Mining and Interspecific Hybridization Characteristics. *Genome Research*, Vol. 14, No. 3, pp. 478-490, ISSN 1088-9051.
- 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*, Vol. 7, No.1, ISSN 1471-2164.
- 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*, Vol. 155, No. 3, pp. 1331-1345, ISSN 0016-6731.
- Sanchez, C.C., Smith, T.P., Wiedmann, R.T., Vallejo, R.L., Salem, M., Yao, J. & Rexroad, C.E.3rd (2009). Single Nucleotide Polymorphism Discovery in Rainbow Trout by Deep Sequencing of a Reduced Representation Library. *BMC Genomics*, Vol. 10, No.1, ISSN 1471-2164.
- Sanchez, J.A., Clabby, C., Ramos, D., Blanco, G., Flavin, F., Vazquez, E. & Powell, R. (1996). Protein and Microsatellite Single Locus Variability in *Salmo Salar* L. (Atlantic salmon). *Heredity*, Vol. 77, No. 4, pp. 423-432, ISSN 0018-067X.
- Sarropoulou, E., Franch, R., Louro, B., Power, D.M., Bargelloni, L., Magoulas, A., Senger, F., Tsalavouta, M., Patarnello, T., Galibert, F., Kotoulas, G. & Geisler, R. (2007). A Gene-based Radiation Hybrid Map of the Gilthead Sea Bream *Sparus aurata* Refines and Exploits Conserved Synteny with *Tetraodon nigroviridis*. *BMC Genomics*, Vol. 8, No.1, ISSN 1471-2164.
- Scribner, K.T., Gust, J.R. & Fields, R.L. (1996). Isolation and Characterization of Novel Salmon Microsatellite Loci: Cross-species Amplification and Population Genetic Applications. *Canadian Journal of Fisheries and Aquatic Sciences*, Vol. 53, No. 4, pp. 833-841, ISSN 1205-7533.
- Slettan, A., Olsaker, I. & Lie, O. (1995). A Polymorphic Dinucleotide Repeat Microsatellite in Atlantic Salmon, *Salmo salar* (SSOSL436). *Animal Genetics*, Vol. 26, No. 5, pp. 368, ISSN 1365-2052.
- Smith, C.T., Elfstrom, C.M., Seeb, L.W. & Seeb, J.E. (2005). Use of Sequence Data From Rainbow Trout and Atlantic Salmon for SNP Detection in Pacific Salmon. *Molecular Ecology*, Vol. 14, No. 13, pp. 4193-4203, ISSN 0962-1083.
- Somorjai, I.M., Danzmann, R.G. & Ferguson, M.M. (2003). Distribution of Temperature Tolerance Quantitative Trait Loci in Arctic Charr (*Salvelinus alpinus*) and Inferred Homologies in Rainbow Trout (*Oncorhynchus mykiss*). *Genetics*, Vol. 165, No. 3, pp. 1443-1456, ISSN 0016-6731.
- Sun, X. & Liang, L. (2004). A Genetic Linkage Map of Common Carp (*Cyprinus carpio* L.) and Mapping of a Locus Associated with Cold Tolerance. *Aquaculture*, Vol. 238, No. 1-4, pp. 165, ISSN 0044-8486.
- Thorsen, J., Zhu, B., 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*, Vol. 6, No. 1, ISSN 1471-2164.

- Jackson, T.R, Ferguson, M.M, Danzmann, R.G, Fishback, A.G, Ihssen, P.E, O'Connell, M. & Crease, T.J.(1998). Identification of Two QTL Influencing Upper Temperature Tolerance in Three Rainbow Trout (*Oncorhynchus mykiss*) Half-sib Families. *Heredity*, Vol. 80, No. 2, pp. 143-151, ISSN 0018-067X.
- Tuberosa, R., Salvi, S., Sanguineti, M.C., Landi, P., Maccaferri, M. & Conti, S. (2002). Mapping QTLs Regulating Morpho-physiological Traits and Yield: Case Studies, Shortcomings and Perspectives in Drought-stressed Maize. *Annals of Botany*, Vol. 89 Spec No, pp. 941-963, ISSN 0305-7364.
- Van Tassel, C.P., Smith, T.P., Matukumalli, L.K., Taylor, J.F., Schnabel, R.D., Lawley, C.T., Haudenschild, C.D., Moore, S.S., Warren, W.C. & Sonstegard, T.S.(2008). SNP Discovery and Allele Frequency Estimation by Deep Sequencing of Reduced Representation Libraries. *Nature Methods*, Vol. 5, No. 3, pp. 247-252, ISSN 1548-7091.
- Veerkamp, R.F. & Beerda, B. (2007). Genetics and Genomics to Improve Fertility in High Producing Dairy Cows. *Theriogenology*, Vol. 68, Suppl. 1, pp. S266-273, ISSN 0093-691X.
- von 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*, Vol. 6, No. 1, ISSN 1471-2164.
- Waldbieser, G.C., Bosworth, B.G., Nonneman, D.J. & Wolters, W.R. (2001). A Microsatellite-Based Genetic Linkage Map for Channel Catfish, *Ictalurus punctatus*. *Genetics*, Vol. 158, No. 2, pp. 727-734, ISSN 0016-6731.
- Wang, S., Peatman, E., Abernathy, J., Waldbieser, G., Lindquist, E., Richardson, P., Lucas, S., Wang, M., Li, P., Thimmapuram, J., Liu, L., Vullaganti, D., Kucuktas, H., Murdock, C., Small, B.C., Wilson, M., Liu, H., Jiang, Y., Lee, Y., Chen, F., Lu, J., Wang, W., Xu, P., Somridhivej, B., Baoprasertkul, P., Quilang, J., Sha, Z., Bao, B., Wang, Y., Wang, Q., Takano, T., Nandi, S., Liu, S., Wong, L., Kaltenboeck, L., Quiniou, S., Bengten, E., Miller, N., Trant, J., Rokhsar, D., Liu, Z. & Catfish Genome Consortium (2010). Assembly of 500,000 Inter-specific Catfish Expressed Sequence Tags and Large Scale Gene-associated Marker Development for Whole Genome Association Studies. *Genome Biology*, Vol. 11, No. 1, pp. R8, ISSN 1465-6906.
- Woram, R.A., McGowan, C., Stout, J.A., Gharbi, K., Ferguson, M.M., Hoyheim, B., Davidson, E.A., Davidson, W.S., Rexroad, C. & Danzmann, R.G. (2004). A Genetic Linkage Map for Arctic Char (*Salvelinus alpinus*): Evidence for Higher Recombination Rates and Segregation Distortion in Hybrid Versus Pure Strain Mapping Parents. *Genome*, Vol. 47, No. 2, pp. 304-315, ISSN 0831-2796.
- Wringe, B., Devlin, R., Ferguson, M., Moghadam, H., Sakhrani, D. & Danzmann, R. (2010). Growth-related Quantitative Trait Loci in Domestic and Wild Rainbow Trout (*Oncorhynchus mykiss*). *BMC Genetics*, Vol. 11, No. 1, ISSN 1471-2156.
- Xia, J.H., Feng, F., Lin, G., Wang, C.M. & Yue, G.H. (2010). A First Generation BAC-based Physical Map of the Asian Seabass (*Lates calcarifer*). *PloS One*, Vol. 5, No. 8, pp. e11974, ISSN 1932-6203.

- Xu, P., Li, J., Li, Y., Cui, R., Wang, J., Wang, J., Zhang, Y., Zhao, Z. & Sun, X. (2011). Genomic Insight Into the Common Carp (*Cyprinus carpio*) Genome by Sequencing Analysis of BAC-end Sequences. *BMC Genomics*, Vol. 12, No. 1, ISSN 1471-2164.
- Xu, P., Wang, S., Liu, L., Thorsen, J., Kucuktas, H. & Liu, Z. (2007). A BAC-based Physical Map of the Channel Catfish Genome. *Genomics*, Vol. 90, No. 3, pp. 380, ISSN 0888-7543.



## **Aquaculture**

Edited by Dr. Zainal Muchlisin

ISBN 978-953-307-974-5

Hard cover, 390 pages

**Publisher** InTech

**Published online** 27, January, 2012

**Published in print edition** January, 2012

This book provides an understanding on a large variety of aquaculture related topics. The book is organized in four sections. The first section discusses fish nutrition second section is considers the application of genetic in aquaculture; section three takes a look at current techniques for controlling lipid oxidation and melanosis in Aquaculture products. The last section is focused on culture techniques and management, which is the larger part of the book. The book chapters are written by leading experts in their respective areas. Therefore, I am quite confident that this book will be equally useful for students and professionals in aquaculture and biotechnology.

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Nicole L. Quinn, Alejandro P. Gutierrez, Ben F. Koop and William S. Davidson (2012). Genomics and Genome Sequencing: Benefits for Finfish Aquaculture, Aquaculture, Dr. Zainal Muchlisin (Ed.), ISBN: 978-953-307-974-5, InTech, Available from: <http://www.intechopen.com/books/aquaculture/genomics-and-genome-sequencing-benefits-for-finish-aquaculture>

# **INTECH**

open science | open minds

### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.