Review of Aquaculture Genetics and Genomics.

Brendan McAndrew,
Genetics and Reproduction Research Group,
Institute of Aquaculture,
University of Stirling,
Stirling FK9 4LA.
Scotland UK.

**General introduction**
Aquaculture production has continued to grow at an ever increasing rate from less than 1 million tonnes in the 1950s to nearly 60 million tonnes in 2004 (FAO 2006). Much of this increase occurring since the mid 1980s with the vast percentage of the production being from Asia and the Pacific rim, particularly China. The percentages by volume are China 69.7%, Rest of Asia 21.9% and Rest of the world 8.1% the respective values of the production are 51.2%, 29.3% and 19.5%. The Rest of the world percentage includes Western Europe which accounts for 7.7% of the world total aquaculture production by value. The European sector is dominated by Atlantic salmon production but, rainbow trout, mussels, oysters and marine species such as seabass and seabream also now make significant contributions. Despite the great importance of the present aquaculture production and the hope that its expansion in the future will help feed the ever growing world population very little of this production can be said to be coming from genetically managed or improved strains. Gjedrem (2005) estimated that less than 5% of world aquaculture production comes from scientifically managed breeding programmes and much of this from salmonids in northern Europe and north and South America. This leaves a large proportion of the world’s farmed fish and shellfish under traditional domestication processes. Many of the new farmed species are still at an early stage of development and still utilise wild caught fry or broodstock, particularly in Asia. Gjedrem (2005) postulates that despite the potential for improvements in line with the domestication of terrestrial animals the high fecundity of aquatic species resulted in the need for small numbers of broodstock and therefore rapid inbreeding and loss of viability. Farmers had to constantly replace stocks from the wild to maintain performance. The evidence is that traditional fish breeding and management continues to seriously degrade the genetic quality of stock cut off from continuous replacement by wild or managed fish (Pullin and Capilli, 1988; Eknath and Doyle 1990). The FAO State of World Aquaculture report (2006) says little about these problems and the potential role of genetics to help improve the sustainability and efficiency of world aquaculture.

**Aquaculture Genetics**

The early history of fish genetics has recently been reviewed in chapters in books by Dunham (2004) and Gjedrem (2005) (but also see Benzie, 1998; Gjedrem 2000; Hulata 2001). The science of aquaculture genetics has grown quickly over the last 30 years alongside the ever expanding aquaculture industry in Europe and North America. There has been an increase in the number of breeding programmes within species such as Atlantic salmon and rainbow trout but there is a growing recognition that genetics can result in immediate gains and should be incorporated at the very beginning of the domestication process to avoid genetic degradation of these stocks.
This is notable in the level of funding already advanced for genetic management and improvement and development of genomic tools for new species such as halibut and Atlantic cod in Canada and Norway.

Aquatic organisms offer many advantages to the geneticist over terrestrial animals.

**High fecundity/External fertilisation**
- Means that a large number of gametes can be collected and held and fertilised under controlled hatchery conditions. Typical fecundity for salmonids is $10^3$ eggs/kg, Cyprinids $10^5$ eggs/kg, Oysters $10^6$ eggs/kg per spawning.
- Enables relatively complex breeding designs to be constructed with large family sizes enabling more accurate estimates of genetic parameters and high selection pressures. Hybridisation (inter-specific) has been widely used.
- Short or no multiplication stages are needed between breeding nucleus and commercial production so genetic improvement can be seen by the industry immediately.
- Gametes relatively cheap so possible to use techniques that induce high mortality such as chromosome set manipulations or transgenesis but generate unique and highly valuable genotypes. E.g. Clones, YY males.

**Essentially wild organisms**
- High levels of genetic variation have been observed in most wild populations of aquatic organisms. Heterozygosity 6-10% in fish 10-20% in invertebrates.
- Genetic variation is a prerequisite for selective improvement programmes. Measured heritabilities tend to be medium to high across a wide range of traits.
- Most species sub-divided into isolated populations that may have genetically differentiated to adapt to differing environments. Comparisons possible to identify significantly better founder stocks for aquaculture.
- Few inbred strains available the most developed fish strains are usually ornamental.

**Plastic Phenotypes**
- The same genotype can have many possible phenotypes!
- Possible to change sex of many aquatic organisms by temperature or by administering hormones during sexually labile period of development. This enables single sex stock to be generated directly or through controlled breeding.
- It is possible to change the ploidy state and levels of genetic variation by chromosome set manipulation using temperature or pressure shocks. This enables sterile triploid fish or unique genotypes such as YY males or isogenic double haploid clonal lines to be produced.
- Possible to modify the morphology and meristics of scales, fin rays, muscle precursor cells or other structures by rearing at different temperatures.
- Possible to advance or delay sexual maturation or increase or decrease the age of onset of traits such as smoltification in salmonids by photoperiod manipulation.
Some disadvantages.

Despite all these advantages genetic progress has been hindered by the general lack of knowledge on the biology of many aquatic species. The different spawning behaviours and the small size of young aquatic organisms make it very difficult to individually physically mark them at an early stage in their development. In the past this meant that most improvement programmes used a mass selection approach to improve obvious phenotypic traits such as scale pattern, colour, growth rate, and body conformation. Mass selection can result in rapid inbreeding unless measures are taken to minimise the incidence of related mating. Inbreeding can be minimised through maintaining a large effective population number (Ne) and through the adoption of line crossing schemes to delay and reduce the relatedness of any future crosses (see Tave 1992). When well managed this approach has resulted in improvement (many of the Eastern European carp programmes came under this heading) but has also in many circumstances compromised the genetic resource of the species through the loss of genetic variation caused by genetic drift and inbreeding (the case today in many Asian carp hatcheries).

In order to undertake the more efficient but complex family selection programmes it was necessary to use family units so individual families could to be reared in separate tanks or cages up to a size for physical tagging. Tagging technology has also improved from the use of external tags or freeze branding that that disfigured fish or made them more susceptible to disease to the widespread adoption of Passive Integrated Transponder (PIT) tags that are injected intramuscularly or intra-peritoneal so avoiding loss. Family units and PIT tagging are now a common feature of salmon and trout improvement programmes worldwide.

Not all hatchery producers can afford a family unit and not all species can be individually stripped and their young be reared separately for reasons of biology or logistics. It was with the recent advances in DNA marker development that it became possible to individually identify offspring from known genotyped parents making it possible to reconstruct pedigrees retrospectively after a period of rearing, even in large communal populations. With this technology broodstock management, replacement and improvement becomes available to a much greater range of species and farm based environments. This approach also opens up the possibilities for the traditional breeding programmes to work with greater numbers of fish in commercial environments in which tags would be unacceptable, and take advantage of serendipitous events, e.g. disease outbreaks in untagged populations, regenerate pedigrees from untagged backup populations in case of losses in breeding nuclei, compare results from family unit and communally reared fish to assess magnitude of environmentally induced differences.

It is now possible to overcome the main disadvantages of working with aquatic organisms and undertake broodstock management, replacement and genetic improvement using a variety of approaches that can be matched to the biology as well as the available resources for most farmed aquatic species. The speed of technological development particularly in genomics has been rapid and its uptake in the aquatic sciences has become widespread. It is still unclear how much of this will eventually benefit the aquaculture industry. We have been in a phase of developing the tools needed to effectively use this technology to inform farmers, breeders, feed companies,
and vaccine manufacturers on how they should integrated the knowledge gained into their existing procedures. Some of the issues outlined will be expanded in the following sections.

**Development of Genetic Tools**

**Classical breeding programmes**

These have been applied to all of the major animal and plant agricultural species and have proved highly efficient in improving the yield and quality of the food we eat today. Man has selected the animals and plants through a domestication process that has been going on for thousands of years without ever understanding the biological or genetic processes involved. In many cases it is impossible to assess the degree of change since many of these ancestral populations no longer exist. Scientifically driven genetic improvement of terrestrial animals and plants began in the 1930s as the theoretical basis of quantitative genetics was developed (Lush 1949). Quantitative genetics was widely adopted supported and funding made available for national and international institutions aimed at the improvement of terrestrial farm animals. More recently these activities have been taken over by multinational breeding companies. This work has resulted in significant improvements in yield particularly over the last 40 years in poultry (meat yield and FCR >300%) Havenstein et al. (2003), dairy cattle milk yield in 20 years (>100%) (Shook, 2006) and pig meat yield (>200%) in animals that where nominally already domesticated and genetically improved. This was seen as being a cost effective way of improving production with cost benefits ranging between 1:5 to 1:50 depending on the species.

Common carp.

In fish classical selective improvement programmes were established after the second world war for European common carp in many Central and Eastern European countries resulting in strains being developed to improve yield from semi-intensive and intensive pond culture systems (Kirpichnikov 1981). These strains have subsequently been widely distributed around the world and contribute to the vast bulk of freshwater farm production in Europe, China and Asia. The history of the development of some of these carp strains is not well documented. Carp breeding and testing stations such as that at Szarvas in Hungary maintain landrace strains collected from around the country as well as improved strains developed from these using mass selection and crossbreeding (Bakos et al. (2002) also Network of Aquaculture Centers in Central and Eastern Europe (NACEE). There have been programmes to further develop these carp in the countries to which they have been imported to better adapt them to new conditions. Such programmes have been well documented from Israel, Vietnam and India (Moav and Wohlfarth 1968: Penman et al. 2005 also Network of Aquaculture Centres in Asia Pacific (NACA)). It is now recognised that many of these strains no longer respond to selective improvement for growth performance and many of these programmes utilise some form of multi-strain cross breeding programme. This is probably because of high selection intensities and the lack of parentage assignment resulted in a rapid reduction in Ne and loss of variation in many of these strains. This has to be contrasted with the continued gains in chickens and pigs from properly managed breeding flocks and herds that have been closed for over 30 years. The radically different tropical environments into which these carp have been introduced are causing new production problems not seen in carp within their normal temperate range such as early or precocious maturation.
Rainbow trout
Rainbow trout has also been the subjected to long-term selective improvement in the USA initially by Donaldson and Olson (1955) and later more scientifically by Kincaid et al. (1977), Gall and Huang (1988 a, b) and Gunnes and Gjedrem (1981) and Gjedrem (1992) for sea grown trout. The largest supplier of eggs to the rainbow trout industry is Trout lodge and they have only been involved in a family based breeding programme for the past 5 years. There are breeding programmes in Finland (Kause et al. 2002, 2003) and Denmark (Henryon et al. 2002). Rainbow trout is another species that has had a long history of domestication despite this there are few scientifically managed breeding programmes and defined improved strains. It is difficult to assess how much of the vast literature on the biology, physiology, endocrinology and genetics of rainbow trout that has been produced has actually contributed towards the improvement of rainbow trout as a farmed species. Much of this work has been done in research institutes and little has been adopted by the industry. Short term improvements such as all female production or all female triploids have been widely adopted but the majority of eggs still come from mass selected domesticated strains.

Atlantic salmon
Probably the best documented of the modern scientifically based fish breeding programmes is that of the Norwegian salmon established in 1971, not long after the first salmon were being ongrown in pens. The breeding programme was well funded and benefited from the input of established animal breeders. The breeding programme assessed the performance of over 40 individual families from 40+ different river strains of Atlantic salmon under a range of culture environments along the Norwegian coast (Gjedrem et al. 1991). This work showed that there was little genotype x environment interaction, the best families grew well under all conditions, and that most of the genetic variation resided within rather than between populations. Therefore the breeding programme was setup using a synthetic strain that combined the best families from a number of rivers. The history of the Norwegian salmon breeding programme has been described by (Gjedrem et al. 1991; Gjoen and Bentsen 1997; Gjedrem 2005). Today these breeding programmes are run by private companies Aquabreed and Salmobreed. These programmes have a number of similarities in that they utilise family units to ongrow individual families of salmon until the fish are large enough to be tagged (PIT) a proportion of these fish from each family are retained as potential broodstock and the remainder are ongrown under commercial condition and their individual and family performance are assessed for a wide range of commercial traits. (Growth, sexual maturation, body conformation) In recent years additional traits such as disease resistance (IPNV, Aeromonas, ISA) are also assessed in these families under controlled challenge conditions and this information is also included in a selection index. The estimated gains in the Norwegian programme are between 8-10% per generation (Gjoen and Bentsen 1997).

Similar programmes have also been developed in other salmon producing countries (e.g. see web sites for Landcatch Natural Selection in Scotland and Chile, Stofnífiskur in Iceland, Aquachile in Chile). Today most farmed Atlantic salmon eggs come from scientifically managed breeding programmes that have significantly improved the performance of this species.
Need for Genetic Management of Broodstock fish.
The vast majority of farmed fish worldwide are still not under managed programmes but under a variety of broodstock replacement strategies. The domestication of aquatic organisms usually follows a set pattern with either the initial collection of young fish or mature adults from local populations of the wild species. Once the farmers have managed to control the growth and reproduction of the species then seed supply from hatchery produced fish is usually preferred as this reduces the risk of introducing diseases from the wild and reduces variability in the fry as they are likely to be the same age and size. The decision to close a hatchery population, stop using wild fish as broodstock, needs to be managed as it will be critical for the long term viability of that particular population. The early domestication process and husbandry methods can result in a dramatic reduction in the actual number of wild broodstock that contribute offspring to the replacement broodstock pool (Effective Breeding number Ne). The reproductive success of the original wild fish and the survival of their offspring can be highly variable and without some form of tagging it may be impossible to identify the Ne of the hatchery strain. Ne is particularly difficult to estimate in marine species that mass spawn. Recent studies using genetic markers in mass spawning species such as seabream (Brown et al., 2006) and cod (Herlin et al. 2008) have shown that the offspring of a few individual fish can dominate in potential broodstock replacements. Even in species that can be stripped but still have high hatchery mortality, so fry batches are frequently pooled, the Ne of the hatchery population can still be dramatically reduced in species such as halibut (Jackson et al. 2003).

Message
Selective improvement works in aquatic organisms and improvement has been measured for a wide range of different traits in many commercially important species.

The potential improvement in a trait will depend on the initial levels of additive genetic variation present in the strain(s) and how well they are managed in the long-term.

It is never too early to start the genetic management of a new farm species. Identification of potential source material and comparative testing of strains can result in major improvements in performance and ensure genetic variation is retained for future improvement.

The use of tags, genetic markers or some form of line management is required, no matter what type of selection approach is used, to minimise inbreeding and associated loss of genetic variation.

Chromosome set manipulation.
The plastic phenotype displayed by many aquatic species and the lack of genetic imprinting in many invertebrate and lower vertebrates enables a range of genetic and environmental manipulations to be undertaken in these species that can have profound
effects on their subsequent performance. These manipulations are not seen as genetic modification as many of the outcomes have been observed naturally in wild fish populations.

External fertilisation in fish presents a unique opportunity for geneticists to manipulate the gametes to generate structured breeding groups or novel genotypes in ways not possible in other farmed animals. One of the most powerful technologies is the ability to induce parthenogenesis and derive offspring from wholly maternal or paternal origins (see Komen and Thorgaard (2007) and Dunham (2004) for reviews of these technologies). These technologies are achieved by destroying the nuclear DNA in either the egg or sperm, using ionising radiation (Gamma or UV); the treated gamete is then fused with an untreated sperm or egg, respectively, to produce a haploid embryo. The haploid embryo could continue to develop but would normally die before hatching. Haploid embryos make a good resource for gene-mapping as they are effectively large single gametes that contain enough DNA for gene-mapping purposes (Kocher et al. 1998). However, it is possible to make the embryo diploid by inhibition of the second meiotic division or first mitotic division; such individuals retain a duplicated set of chromosomes from the untreated gamete. Eggs fertilised by UV treated milt can become meiotic gynogenetic offspring by shocks that interfere with the second meiotic division, causing the retention of the second polar body. If the shock is delayed to the first mitotic division then two haploid copies of the maternal chromosomes are retained to produce a mitotic gynogenetic offspring. Meiotic gynogenetic offspring are on average homozygous at 50% of their loci because they retain a pair of chromosomes, sister chromatids, which have just undergone recombination. Mitotic gynogenetic offspring are 100% homozygous because they arise from the duplication of a single chromosome set and are described as double haploid (DH) individuals. Eggs treated with UV or gamma radiation, to destroy their nuclear DNA, and then fused with a normal sperm produce a haploid embryo that can be made diploid by a disruption of the first mitotic division to produce an androgenetic DH that is 100% homozygous.

Double haploid individuals are homozygous at all loci but different DH individuals in the same family will be fixed for different alleles at any given locus depending on the recombinant event that generated that gamete. Double haploid individuals from either a gynogenetic or androgenetic background can be used to generate clonal or isogenic lines as all gametes produced by an individual will be identical, even after recombination. When these gametes are used in a second round of parthenogenesis all of the offspring will be identical and clonal. It is therefore possible to generate clonal lines in as little as two generations for any particular species or strain of fish. Apart from the models species such as zebrafish (Streisinger et al 1981) clonal lines have been produced in tilapia (Müller-Belecke and Hörstegen Schwark 1995: Hussain et al. 1998 and Sarder et al. 1999) rainbow trout ( Scheerer et al. 1986 and Thorgaard et al. 1990. Quillet et al. 2007) and common carp (Komen et al. 1991; Bongers et al. 1997)

If eggs are fertilised with normal sperm and similar shock treatment are applied then an early shock at the 2nd meiotic division will generate an embryo that contains 3 chromosome sets and is triploid (3n). A late shock at the first mitotic division will retain two diploid sets and produce an embryo with 4 sets of chromosomes and is tetraploid (4n). These types of animals with extra sets of chromosomes are known as
polyploids. Polyploidy is lethal in higher vertebrates and birds but has been widely applied in plants and aquatic organisms (Chourrout et al 1986).

The disruption of the meiotic and mitotic divisions can be achieved using a variety of shocks, most commonly, hot and cold temperature shocks and pressure shocks although anaesthetics and chemical treatments can also be used. Optimisation experiments are needed to determine the timing of the onset of the shock, be it the 2nd meiotic or 1st mitotic division, at a given ambient water temperature and the strength and duration of the shock to maximize the effect for a given species. Dunham (2004) has produced a list of over 30 different fish and shellfish species and the published optimised shock treatments and outcomes.

Polyploid production in shellfish is different to that observed in fish because it is possible that both the 1st and 2nd polar body are still present after fertilisation (Tian et al. 1999 a,b). This means that a range of other polyploidy states may be present in a given experiment to produce triploids (aneuploid, triploid, tetraploid and pentaploid) in Pacific oysters depending on whether the treatment impacts on the 1st, 2nd or both polar bodies (Stephens and Downing 1988; Guo et al. 1989). Essentially the shock or the chemicals used interferes with the normal polymerisation of the microtubules that control the organisation and separation of the chromosomes within the nucleus during meiosis and mitosis.

The shock treatments used to produce 3n fish can have a long-term impact on the viability of the fish as many of these treatments are at the thermal limits of the species involved and often when carried out on a large scale under farm conditions the treatments can be sub-optimal. One possibility is to use gametes from 4n fish. Tetraploid fish having a balance number of chromosomes can still undergo meiosis and produce viable gametes, which are potentially already diploid. So a cross between a normal egg (1n) and a milt from a 4n male (2n) will generate 3n offspring directly. This has been successful in rainbow trout (Myers et al. 1986) once eggs capable of being fertilised by the larger diploid milt have been identified.

These techniques have been applied on a wide range of aquatic organisms (see Dunham 2004). Direct applications in aquaculture include the generation of triploid fish, usually from all-female strains, in a number of species in which sexual maturity can reduce the size and quality of the farmed fish. Triploid females tend to be effectively sterile as the unequal chromosome number interferes with the meiotic divisions needed to produce viable eggs resulting in the ovary of such fish being small and containing few if any developing ova. This technology has been applied to the production of larger rainbow trout (Bye and Lincoln 1986), channel catfish (Wolters et al. 1982), common carp (Basavaraju et al. 2002) and in other aquaculture species in which the onset of maturity in diploid strains slows growth. In shellfish the triploidy is now widely used to improve the growth rate and improve flesh quality in oysters (Allen and Downing 1988; Guo et al 2000). It can also be used as a biological control mechanism to stop spawning in exotic species, such as grass carp used for the control of aquatic weeds (Shelton and Jensen 1979: Wattendorf, 1986). To reduce the risk of introgression of genes from stocked or escaped farmed strains into native populations of the same species (e.g. Atlantic salmon in N.Europe and Eastern USA and Canada e.g. Brown trout in UK, rainbow trout and cutthroat trout in USA and Canada) (Kozfkay et al. 2006). Triploid sterility has also been put forward as a means to stop
the introgression of genes from transgenic fish strains that might escape into the wild (see review Mair et al. 2007).

Direct applications of gynogenetic or androgenetic techniques on production fish is rare and these techniques are usually undertaken in research laboratories. However, these techniques can be used to generate unique genotypes that can be used to generate new strains. The development of YY male tilapia used to produce all–male XY tilapia can be generated directly by androgenesis as any male offspring in *O. niloticus* will have the YY genotype (Myers et al 1986). Alternatively, eggs produced by neo-female (males sex-reversed to female phenotype) and crossed to normal males XY x XY will also generate XX females and XY and YY male offspring (Mair et al 1995). However, all the male fish in such crosses have to be progeny tested to identify their genotype. This approach has also been used in the development of all-female fish (Pongthana et al. 1995). In the silver barb (*Puntius goniontus*) females are the preferred sex as the ovaries are a delicacy. Meiotic gynogenetic offspring can be produced in large numbers and are all-female so can be sex-reversed to generate neo-males without the need for progeny testing. These neomales fish are supplied in large number to the hatcheries as broodstock for the commercial production of all female fry (Pongthana et al. 1999).

**Gene Transfer Technologies.**

The large number of gametes and the ease of collection, the control over the timing of fertilisation, the ease of manipulation because of the large size of the ova and *in vitro* incubation encouraged the application of gene transfer techniques in a number of fish species. Particularly so when you compare the relative ease of working with fish eggs to the complexity involved in mammals (Powers et al 1992). It was the much publicised work by Palmiter et al. (1982) and Palmiter and Brinster (1986) that showed huge growth improvements in mice, induced by the integration of growth hormone (GH) constructs, that stimulated much of the early work on fish. Transgenic fish containing the human GH gene (hGHg) were quickly produced in several commercial species including goldfish (Zhu et al. 1985), rainbow trout (Chourrout et al. 1986; Maclean et al. 1992; Penman et al. 1990; Guyomard et al. 1989a,b), channel catfish (Dunham et al. 1987) and Nile tilapia (Brem et al. 1988). These early studies utilised the genes that happened to be available and spliced together constructs from a wide range of genes and promoter sequences from different sources and injected these into the cytoplasm near the nucleus in the fish egg.

The results from these studies suggested that less than 5% of the injected fish ended up with a copy of the construct integrated into the host genome. The problems included cytoplasmic amplification and expression of the construct that often resulted in an overestimate of the success of the procedure. Integration of the constructs appears to occur after the first cell division so that only a proportion of the cells within the host will contain the construct resulting in mosaicism. Breeding from these mosaic fish means that only those that have the construct stably integrated into germ tissue will pass the construct onto their offspring. Even with the advantages offered by fish the process of microinjection is time consuming and tedious and was resulting in relatively low levels of success. Other techniques have been tried that would enable mass transfer of constructs into the large numbers of egg available from aquatic species. These included electroporation, retroviral integration, liposomal reverse phase evaporation, sperm mediated transfer and microprojectiles (see Chen and
Powers et al. (1992) believe that electroporation offers the best possibility for mass production of first generation transgenic fish.

The earlier work concentrated on the enhancement of growth through the use of GH constructs. With the improvements in the technique and the move to better constructs growth enhancement was observed in a number of commercial species including rainbow trout (Guyomard et al. 1989; Penman et al. 1991) common carp (Chen et al. 1989; Zhang et al. 1990) channel catfish (Dunham et al. 1992) tilapia (Martinez et al. 1996; Rahman et al. 1998, Rahman and Maclean 1999). These studies resulted in a wide range of gains in growth performance, between 10% -500%. However, the most dramatic improvements were observed in the work undertaken in Devlin’s laboratory in Vancouver. This laboratory developed total piscine gene constructs that utilised either an ocean pout antifreeze promoter (ocAFP) used to control a chinook salmon GH cDNA or a sock-eye salmon metallothionein (MT) promoter controlling a full length sock-eye GH gene. When successfully integrated these constructs resulted in significant elevation of circulating GH levels (40X), particularly in younger fish (Devlin et al. 1994; Devlin et al. 1997). Growth enhancements of transgenic individual over controls ranged from 5-30 fold increase in growth up to 1 year of age (Du et al. 1992; Devlin 1994, 1995) and successfully passed this performance onto their offspring. The same construct had similar effects when used in other salmonid species, coho salmon 10-30X improvement, rainbow trout 3.2-10X improvement (Devlin et al. 1995a ; 1997b), cutthroat trout 10X improvement, a 6.2X improvement in chinook salmon (Devlin et al. 1995a) and 3-6X improvement in Atlantic salmon (Du et al. 1992; Devlin et al. 1997; Cook et al. 2000).

It is now clear that fish species respond very differently to the effects of GH, salmonids appear to be very sensitive and this is probably related to their physiology. Growth rate in salmonids is generally slow particularly when water temperature is low and appears to be partially controlled by the levels of circulating GH. An increased level of GH in transgenic salmon effectively uncouples this regulatory control enabling the fish to grow rapidly in the colder winter months in comparison to control. This early advantage is further magnified as these fish smolt earlier and the associated growth enhancement of this process gives these fish a further advantage that can be subsequently magnified (Dunham and Devlin 1998).

Warm water species such as tilapia and carp can grow rapidly throughout the year and are probably not as reliant on GH regulation. Transgenic fish produced from selected lines do not show the dramatic improvements seen when wild fish are used. Devlin et al. (2001) produced transgenic individuals from a wild rainbow trout and a domesticated rainbow trout strain. The wild transgenics showed a 17X improvement over non transgenics, but they did not grow better than the fast growing non transgenic domestic trout. Introducing the gene into the domesticated strain only improved overall performance by 4.4%.

It must also be remembered that it can take 4-5 generations of breeding to develop a stable transgenic fish line to obtain a single step improvement in growth performance. It appears at this stage that selective breeding can, not only achieve the growth enhancements but also improvement across a number of other commercially important traits that will enhance the overall value of strains in a cumulative manner. However, the large amount of functional genomic work being undertaken is already
identifying possible candidate transgenic genes especially in the area of improved disease resistance (Dunham 2008). Increased resistance has been observed in channel catfish and medaka transgenic for the lytic peptide cecropin B. Pleiotropic effects on disease resistance have been observed in fish transgenic for non disease constructs (Dunham 2008).

To date there is no authenticated release of transgenic fish for aquaculture production because of the difficulties in assessing the environmental risk of any escapes. Any assessment of the phenotype of a given transgenic strain under laboratory conditions and extrapolating this to an ever changing aquatic environment is almost impossible (Devlin et al. 2006). This is likely to result in transgenic organisms remaining in contained facilities until we better understand their characteristics and we have reliable means of biological containment Mair et al. (2007).

**Molecular techniques**

**Genetic Markers**

There has been continuous development of genetic markers since the 1960’s as our ability to separate and visualise protein and more recently DNA molecules has improved. There have been a number of reviews on the development of molecular markers and there application in aquatic organisms (see Liu and Cordes (2004) also Liu 2007). The development of marker technology has been critical to our better understanding of the population structure and evolution of aquatic organisms and will underpin the latest developments in aquaculture genetics and genomics.

There are essentially 2 classes of markers Type 1 or actual genes of known function (coding sequence) and Type 2 or anonymous DNA segments. In terms of direct application, to date, Type 2 markers and in particular microsatellite markers have had a dramatic impact on our ability to manage farmed fish in recent years. Genetic markers based on proteins (allozymes) have been widely used from the 1960’s and has proved particularly useful in developing our understanding of the population genetics of aquatic organisms and the subsequent management and conservation of fisheries. However, the relatively low number of loci that could be studied and the low levels of observed polymorphism meant that their use in mapping and parentage assignment was very limited. The development of DNA based markers initially in mitochondrial DNA and various other restriction fragment analysis techniques such as RAPDs and AFLP require no prior knowledge of DNA sequence and have been applied to identify high levels of polymorphism but these techniques are not easily transferable across laboratories and the dominant inheritance pattern of the sequences being studied reduces their utility (Liu. 2007).

It was the identification of a new class of microsatellite DNA (Tautz, 1989) made up of small tandem repeat sequences (CA, CAT, GATA) e.g. (CA)10 would be a simple microsatellite (CA repeated 10 times). Identification of these microsatellite loci initially required the development and sequencing of microsatellite enriched libraries constructed from the DNA of the species being studied. However, today with the widespread sequencing efforts being undertaken in many species many new microsatellite loci can be identified directly from databases. In the better studied
species (such as salmonids, tilapia seabass, sebream and carp) large numbers of microsatellite loci have already been generated for genetic mapping studies.

**Why are these markers so useful?**

The fact that they are analysed using a PCR approach means that we can use a very small sample of tissue (scale, fin clip, drop of blood) to generate enough DNA of the markers we are interested in without damaging the fish. We normally have a large number of microsatellite loci available for a given species and would select loci that are the most informative, those having a large number of different alleles, in the population under study. Many microsatellite loci can have between 10-50 alleles in a given population. Although there are a large number of different alleles in a population any individual fish can only have two at any given locus. These can either be the same making the individual homozygous or different making the individual heterozygous at that particular locus, simple Mendelian genetics. Therefore, if you have a set of parents that you can genotype accurately at a number of loci (say 10 loci with 10 alleles each) all of these parents should have a unique subset of 20 alleles from the 100 in the population. If you then breed from these parents it should be possible to predict the genotypes of the offspring arising from any given pair.

Several studies have shown that it is possible to assign a parentage to an untagged fish grown under commercial conditions with varying degrees of success (Ferguson and Danzmann 1998; Jackson et al. 2003; Norris et al. 2000). The success is higher in controlled breeding, closed system, where the exact parentage of each cross is known and in the quality of the microsatellite loci and PCR reactions have been optimised to reduce genotyping errors. Success is usually reduced in uncontrolled mass spawning, open system, in which any individual male could potentially spawn with any individual female giving rise to a large number of potential families and the use of too few poorly optimised loci. Simulations based on the genotypes of the parents often suggest that high levels of assignment can be achieved with relatively few loci (4-6) but in reality errors associated with PCR, the fragment analysis technology and mutations in the locus or PCR primer site requires more loci to be screened to achieve high levels of assignment (Liu and Cordes, 2004; Duchesne and Bernatchez, 2007).

**Genetic mapping**

The large number of type II genetic markers available in many species has to a great extent been driven by groups interested in the construction of linkage maps. (See websites listed). To date most of the maps are low to medium density with an average map distance of between 2-15 centiMorgan (cM) between markers (a cM is equivalent to a recombination rate of 1% between adjacent markers and is equivalent approximately to 1 million bases). The process of developing a linkage map is very well described by Danzmann and Gharbi (2007) in this review they also give a list of all mapping efforts to date in a range of farmed aquatic organisms. The availability of a linkage map in a given species enables the potential identification of Quantitative Trait Loci (QTL) or sections of the genome that contain a gene or genes influencing important phenotypic traits. To date, a large number, over 14, QTLs have been identified in the rainbow trout by the Danzmann group and collaborators (see Korol et al. 2007 review). There is a steady increase in the rate of publication of QTL in other species such as Atlantic salmon, tilapia and common carp. In principle the marker needs to be within 1cM of the trait to have widespread application in populations
other than within the one used to discover the QTL. Most maps in commercial fish species still do not have this level of coverage. A strategy that has proved useful in Atlantic salmon is to take advantage of the large difference in recombination rate, and therefore map size between the sexes, this is in the order of 1:5 in male : female recombination rate (Hayes et al. 2006). Therefore by using a limited number of markers spread through each of the linkage groups discovered to date it is possible to localise a QTL to a single chromosome arm, within approximately 20cM using male parents. Once the QTL has been localised it is possible to undertake the fine mapping using more of the existing markers or generating new markers within the linkage groups of interest using the female parents. This strategy has been used successfully to identify a QTL for resistance to Infectious Pancreatic Necrosis (IPN) in a commercial Scottish strain of Atlantic salmon (Houston et al. 2008).

The above approach shows the great benefit, at this early stage in the development of these technologies in fish for collaboration between a commercial breeding programme using traditional selective breed approaches and molecular geneticists. In that the breeding programme had identified a medium to high heritability for IPN resistance in the strain and could identify high and low responding families with a high degree of accuracy, enabling a highly focused approach to localise the QTL (Guy et al. 2007). The majority of the QTL that have been published are for easily measured performance traits but as the technology and tools develop combining a QTL approach for traits difficult to measure on the breeding candidate such as flesh quality (e.g. colour) and other metabolic traits (e.g. lipid metabolism and adiposity). The combined approach should result in improved overall selection response.

It is important to continue the development of ever more dense linkage maps (see Danzmann and Gharbi, 2007) to enable the move from using a marker close to the gene to identifying the gene or genes and the allelic variation that gives rise to the enhanced performance. This is known as Marker Assisted Selection (MAS) and has potential to help in the improvement of traits with low heritability, identify possible candidates early, before maturity, and in traits only observed in one of the sexes (Lande and Thompson 1990). Despite the potential, the number of examples of its successful application in livestock is still low and the attitude of the established livestock breeders to MAS was described as one of “cautious optimism” by Dekkers (2004) in his review. The application of MAS technology will increase in aquaculture. De Santis and Jerry (2007) have listed a number of potential candidate genes based on what we know from livestock. They have concentrated on genes associated with growth such as GH, GHR, ILGF, GNRH, leptin, myogenin and showed evidence that these might play a role in fish. It is clear that this list can only increase as function genomic studies become more common.

**Type 1 Markers**

It is the rapid development in the identification of Expressed Sequence Tags (EST) that is now paving the way for the inclusion of Type 1 markers or actual genes being mapped to a species. The inclusion of Type 1 markers in a map means that we can compare the relative position of these genes between the genomes of different species. In aquaculture this is particularly important as it enables us to compare our map poor species with the map rich model fish or mammalian maps. Sarrpoulou et al. (2008) have shown the benefit of this comparative approach by using recently sequenced model fish species, fugu, medaka and stickleback, and comparing them with three
commercial species gilthead seabream, seabass and tilapia. This approach showed that the stickleback was the best intermediary, with 58% synteny with the gilthead seabream, and to use this information to compare seabream with the other two species.

Expressed Sequence Tags are generated by single pass 3’ sequencing of complementary DNA (cDNA) derived from messenger RNA (mRNA) from tissue libraries. The mRNA is present because a particular protein is being expressed in that tissue. It can be seen that by extracting mRNA from a wide range of different tissues and from different tissues in animals that are being exposed to a range of novel environmental conditions or at different developmental stages it will be possible to see which genes are expressed and how a given gene or subset of genes is affected. Most large fish genomics project have included a component to generate large numbers of ESTs (see the web sites listed, many list the libraries used and the EST generated in the respective species). In any given tissue many 1000s of genes could be expressed and the common housekeeping genes possibly in quite large amounts. In order to identify the gene being expressed it is necessary to sequence a strand of DNA that is complementary to the mRNA and compare this to other sequences already collected and identified and held by the various databases worldwide.

Sequencing of cDNA libraries can be random but this approach can result in multiple sequences from common transcripts. Improvements in the development of normalised libraries, multiple copy cDNA is selectively removed, so sequencing is more efficient. Subtractive Substitutional Hybridisation (SSH) is another technique used to remove cDNA common to a control sample and that from a challenged sample, so only differentially expressed genes will be sequenced. The sequence is then BLAST searched against all existing genetic data bases, hopefully identifying homology of the sequence to a known gene in another species (see Liu, 2007 for detailed review of these technologies).

Direct linkage mapping of ESTs is extremely difficult as we usually have no information on the allelic variation present for a given EST in the population under study. EST sequences can be checked for the presence of microsatellites, recent mining of EST data base sequences has shown around 10% of ESTs contain a microsatellite. It is necessary to have sequence from several individuals in order to identify whether a given gene is polymorphic. EST sequence can be also be aligned to identify Single Nucleotide Polymorphisms (SNP). These are point mutations that give rise to alternative bases at a given nucleotide position within the locus. SNP discovery is still very challenging in aquatic organisms as many EST libraries have not been constructed to look for intraspecific variation at a locus within a population, which will hinder progress in SNP discovery. SNPs are however very common within most genomes but it has only been recently possible to visualise and analyse these single base differences routinely with new technological developments. The background to the various techniques of SNP discovery and genotyping technologies are described in detail by Liu (2007). It is clear that the ability to identify and map large numbers of SNPs and automation of subsequent genotyping efforts has lead to rapid development and application of SNP technology in human and animal sciences. Sequencing efforts in Atlantic salmon rainbow trout and Tilapia will mean that it should be possible to identify large numbers of SNPs within these species over the next 5 years.
In species with better developed genetic tools EST can be labelled and used as probes to identify clones in an arrayed (macroarray filter) large insert libraries (e.g. Bacterial Artificial Chromosome libraries BAC) that contain the EST of interest. If the BAC library has been end sequenced and the various BAC clones aligned the EST will be part of a contig (several adjacent BAC clones) and its position will be known on the Physical and Genetic map of the species (He et al.2007)

The genetic tools for aquatic organisms are not as well advanced at this stage and BAC libraries are either not available or have yet to be fully sequenced and contiged. BAC libraries are available in a few farmed species Atlantic salmon (Davidson 2007 also see cGRASP web site), tilapia (Katagari et al. 2001), seabass (Whitaker et al. 2006) and the Pacific Crassostrea gigas (Cunningham et al. 2006). The availability of a BAC library opens up the possibility of positional or map-based cloning of genes or QTLs. The identification of markers flanking a gene of interest or QTL means that the appropriate BAC clones containing the markers can be identified and contiged and the clone containing the gene identified and sequenced a process call chromosome walking. BAC libraries are also a useful as a physical mapping resource as individual clones can be labelled with fluorescent dyes and then hybridised to chromosome spreads of the species. A number of BAC based physical mapping projects are underway in Atlantic salmon (Ng et al 2005), tilapia (Katagiri et al. 2005) and Channel catfish (Xu et al. 2007). BAC libraries have also been used as a resource in the sequence of several different animal and plant species. The BAC clones are first physically mapped to the genome and the overlapping clones that give the minimum distance can then be individually sequenced and a full sequence constructed (see Liu 2007 for detailed review).

Direct mapping of ESTs is possible in species in which a Radiation Hybrid mapping panel (RH panel) is available. Radiation hybrid lines have proved difficult to produce in fish, one of the first successes was in the zebrafish (Geisler et al. 1999) and the same group helped in the construction of the first panel in a commercial species, the gilthead seabream Sparus auratus (Senger et al. 2006). RH panel mapping does not require a polymorphism only knowledge of the sequence so the relevant primers can be used for PCR to confirm the presence or absence of the given sequence in each of the hybrid lines. Radiation Hybrid panels enable large numbers of markers to be mapped quickly without the restrictions of the recombination based approaches. Work is in progress to develop new RH panels for other species such as seabass (AQUAFIRST). The development of more RH mapping panels in other aquaculture species has been on hold as to some extent with the rapid development in the other tools and the expectation of more genome sequencing efforts in aquaculture species making the cost and time commitment in developing such panels unnecessary. However, despite international species community efforts to get their species sequenced few have been adopted making the likelihood of more RH initiatives likely (Rexroad, 2007). The addition of Type 1 markers to maps allows comparative genomic approaches to be adopted to help in the identification of possible candidate loci. Candidate loci or QTL identified in model species and the loci associated with them may identify regions of synteny in the less well studied commercial species.

Functional Genomics
The other driver for the discovery of ever more ESTs is the construction of microarray chips that contain cDNA clones or more specific oligonucleotide sequences derived
from critical parts of the gene in question. These microarray chips can contain 000’s of ESTs derived from the whole animal or more focussed subsets of genes specific to a tissue or a biological function, such as response to disease. The work in salmonids has progressed the furthest and a description of the development of the genetic resources including the various microarrays and the protocols used in these types of study are reviewed by Rise et al. (2007).

Microarrays can be used to study a wide range of potential traits of interest to the aquaculture sector. To date there has been a great deal of interest in disease related responses (e.g. *Piscirickettsia salmonis* (Rise et al. 2004), *Aeromonas salmonicida* (Ewart et al. 2005), Ameobic gill disease (Morrison et al. 2006) immune response to LPS challenge (MacKenzie et al. 2006), live bacterial vaccines (Martin et al. 2006) and DNA vaccination (Purcell et al. 2006). *Gyrodactylus* species, Lindenstrom et al. 2003, 2006: Fast et al 2006)). Other traits of interest include growth responses in transgenic salmon (Rise et al. 2006) and stress responses associated with handling (Krasnov et al. 2005) and temperature (Vornanen et al. 2005).

With large numbers of genes being monitored for their expression under a range of different conditions it is clear that this type of study may identify potential candidate genes that have large effects on the phenotype and may be of commercial importance. The QTL approach outlined earlier cannot presently give the levels of resolution needed to identify potential candidate genes because of the low density of the markers available in many species. Projects that integrate the QTL mapping approach alongside global gene expression studies may well identify possible candidate genes expression patterns that correlate with the traits being studied (Haley and de Koning 2006). Advance in our understanding of the disease resistance in aquatic organisms will identify pathways that are responding to infection than can be enhanced such as the immune system or supported so that the fish can recover much quicker from such challenges. Other traits of major importance in aquaculture for many species include the control and delay of sexual maturation. This can be to improve the reproductive biology so that fish will spawn and produce high quality gametes when required essential for a sustainable industry. The converse is also true in species in which maturation needs to be delayed or avoided.

Web based resources

Many new genetic tools and technologies have had to be developed over the last 20 years driven initially by the needs of Human Genome Project to sequence and order 3000 million bases. This story has been told elsewhere (Ventner et al 2001). These genetic tools and technologies have now been applied to sequence a wide range of other important model species, including fish. The following are the webites available for a range of model fish species.

Zebra danio (*Danio rerio*) ([http://www.sanger.ac.uk/Projects/D_rerio/](http://www.sanger.ac.uk/Projects/D_rerio/)) :
Puffer fish *Tetraodon nigroviridis* ([http://www.genome.gov/11008305](http://www.genome.gov/11008305)):
Medaka (*Oryzias latipes*) ([http://dolphin.lab.nig.ac.jp/medaka/index.php](http://dolphin.lab.nig.ac.jp/medaka/index.php)):
Stickleback (*Gasterosteus aculeatus*) ([http://www.genome.gov/12512292](http://www.genome.gov/12512292))

This review is essentially aimed at the application of genetics and genomics to aquaculture and will be restricted to existing or potential farmed species. There have
been a number of projects that have worked with important aquaculture species in recent years. These projects have in many cases stated that one of their main goals was to generate knowledge that would benefit the aquaculture sector.

**Atlantic salmon (Salmo salar)**
Collaborative EU funded project called SALMAP that aimed to develop new markers and better maps for salmonids (Atlantic salmon, rainbow trout and brown trout). EU funded project called SALGENE that aimed to generate large numbers of EST for functional genomic applications. Resources from these projects are now part of a large international collaboration GRASP and more recently cGRASP that has continued to develop libraries, mapping resolution and functional genomic resources for salmonids. ([http://www.cgrasp.org/](http://www.cgrasp.org/)). The Norwegian contribution to the SALMAP and SALGENE projects and continued collaboration and national efforts in salmon genomics can be viewed at ([http://www.salmongenome.no/cgi-bin/sgp.cgi](http://www.salmongenome.no/cgi-bin/sgp.cgi)). Recent projects based on these international initiative have utilised microarrays based on the ESTs collected to study gene expression under a variety of disease challenge. The GRASP chip has been used to study Immune and disease responses to *Gyrodactylus* species, Lindenstrom et al. 2003, 2006; *Piscirickettsia* Rise et al. 2004; *Ichthyophthirius* Sigh et al.2004: Amoebic gill disease Brindle et al 2006 a,b: sealice and vaccination, Martin et al 2006, Purcell et al. 2006).

Salmon Breeding companies.

AquaChile ([http://www.aquachile.com/spanish/inicio.html](http://www.aquachile.com/spanish/inicio.html))

Aquagen ([http://www.aquagen.no/eng/](http://www.aquagen.no/eng/))

Landcatch ([http://www.landcatch.co.uk/](http://www.landcatch.co.uk/))

Salmobreed ([http://www.salmobreed.no/SideMal01Eng.asp?ID=64](http://www.salmobreed.no/SideMal01Eng.asp?ID=64))

Stofnfiskur ([http://www.stofnfiskur.is/?modID=1&id=67](http://www.stofnfiskur.is/?modID=1&id=67))

**Rainbow trout. (Oncorhynchus mykiss)** The EU STRESSGENES project overall aim was to identify in fish candidate genes associated with resistance to stress conditions and thus provide the physiological and genetic basis for new marker-assisted selection strategies. The rainbow trout has been selected as the focus of this project for several reasons: it is an important aquaculture species throughout Europe; there is a large body of physiological and genetic data available for this species; strains with divergent responsiveness to stress are available; all the partners have worked extensively on this species. ([http://www.irisa.fr/stressgenes/](http://www.irisa.fr/stressgenes/)) This web site is closed to non members of the consortium.

The recent development of genomic tools, particularly microarray technology, allows systematic gene expression analysis of biological material and provides an integrated overview of the global response at the level of gene expression. Such information is of major importance for identifying genes responsible for genetic variation in response to stress and for further development of a sufficient number of molecular markers, a critical requirement for marker-assisted selection schemes. The present
study proposes, for the first time, to apply this new functional genomic approach to examine a complex physiological problem: analysis of the pattern of gene expression at the tissue level in response to exposure to a stressor. It is proposed to examine gene expression profiles in rainbow trout during exposure to a range of stressors typical of those encountered in the aquaculture environment. This should lead to **characterisation of stress-responsive genes as potential candidate gene markers.**

Tilapia. The genomic resources available for *Oreochromis niloticus* and other cichlid species can be viewed [here](http://hcgs.unh.edu/cichlid/#maps)

European Seabass. (*Dicentrarchus labrax*) The recent work undertaken by a multinational collaboration (BASSMAP) to improve the genetic tools for this species and can be viewed at [BASSMAP](http://www.bassmap.org/) This project developed and mapped over ?? type II markers and developed ??EST and mapped ?? type I markers. It also generated a 65K BAC library with 8X genome coverage. The second phase of this project has been included in a new project that will continue to improve the tools (RH panel end sequence BAC library) can be viewed at [AQUAFIRST](http://aquafirst.vitamib.com/)

Seabream (Sparus auratus) Another EU funded multinational project (BRIDGEMAP) to develop tools for the Gilthead seabream The resources generated can be viewed at [BRIDGEMAP](http://www.bridgemap.tuc.gr). This project developed the first RH hybrid panel in a fish and used this to map ?? type I and ?? type II markers (Franch et al 2006). The second phase of this project has been included in the EU AQUAFIRST project. ([AQUAFIRST](http://aquafirst.vitamib.com/))

AQUAFIRST is an EU project that seeks to identify genes associated with stress and disease resistance in fish and molluscs to provide a physiological and genetic basis for marker-assisted selective breeding. It builds on earlier EU funded projects: BRIDGEMAP, BASSMAP and STRESSGENES. To counter poor growth performance, impaired reproduction and increased susceptibility to disease, growers have increasingly used antibiotics and drugs. The development of selective breeding programmes using genetic markers offers a means to reduce, if not eliminate, the use of antibiotics. ([AQUAFIRST](http://aquafirst.vitamib.com/))

Halibut *Hippoglossus hippoglossus* and Senegal sole (*Solea senegalensis*) (PLEUROGENE) ([PLEUROGENE](http://pleurogene.ca/index.php)) This is a Canadian and Spanish initiative to develop genetic tools for these potentially important farmed marine flatfish. Other flatfish sequence and genetic resources can be obtained via the GENIPOL website.

Cod *Gadus morhua* Atlantic cod genomics and broodstock development in Canada ([CODGENE](http://codgene.ca/index.php))

Common Carp. (*Cyprinus carpio*) EUROCARP project Genetic Libraries NACEE ([NACEE](http://www.agrowebee.net/subnetwork/nacee/))
References


De-Santis, C. and Jerry, D. R. 2007. Candidate growth genes in finfish — where should we be looking? Aquaculture, 22-38.


