ICES WGAGFM Report 2007

ICES Mariculture Committee
ICES CM 2007/MCC:03
REF. DFC, ACME AND ACE (WGECO)

Report of the Working Group on the Application of Genetics in Fisheries and Mariculture (WGAGFM)

19–23 March 2007

Ispra, Italy

ICES International Council for the Exploration of the Sea
CIEM Conseil International pour l'Exploration de la Mer

International Council for the Exploration of the Sea
Conseil International pour l'Exploration de la Mer

H. C. Andersens Boulevard 44–46
DK-1553 Copenhagen V
Denmark
Telephone (+45) 33 38 67 00
Telefax (+45) 33 93 42 15
www.ices.dk
info@ices.dk

Recommended format for purposes of citation:

ICES. 2007. Report of the Working Group on the Application of Genetics in Fisheries and Mariculture (WGAGFM), 19-23 March 2007, Ispra, Italy. ICES CM 2007/MCC:03. 70 pp.

https://doi.org/10.17895/ices.pub.8911

Contents

Executive Summary

There is an urgent need for ICES member states to secure and provide proper storage conditions for historical tissue collections such as scales and otoliths, since they contain invaluable DNA evidence of natural historical demographic processes in fish and shellfish populations and allow for evaluating genetic effects climate change and harvesting.

The Working Group on the Application of Genetics in Fisheries and Mariculture (WGAGFM) met at the European Joint Research Centre (JRC), Ispra, Italy 19–23 March 2007. The meeting was well attended; with in total 20 representatives present from 9 countries (10 national delegates and 10 experts appointed by the chairman). In addition we were very pleased to be hosted by the European Science Centre (JRC), which was also represented at the meeting.

Six terms of reference (ToR) were on the agenda for 2007. The first issue addressed was to update the general knowledge on the population genetics of eel. A number of scientific papers have been produced since our last evaluation, generally suggesting some degree of genetic structuring among eels collected at different European sites. Eel fisheries management should be cognisant of the putative genetic structure suggested. Accordingly, long distance transplantations of eel should be avoided. More emphasis should be put on elucidating the biology of the species, in particular in the marine phase, and genetic sampling should be conducted as close to the suspected spawning grounds as possible (adults or larvae).

The second issue was on the application of large scale genome wide DNA methods, so-called "genomics". The technological revolution in high throughput DNA and RNA based methods offer many new opportunities in fisheries and aquaculture. In relation to fisheries, the application of many markers will lead to higher power for detection of population structure and more precise genetic assignment of individuals to population and mixed stock analysis. Also, by studying functional genes, direct evidence of adaptations to local environmental conditions can be obtained. In aquaculture, genomics will, enable easy construction of DNA pedigrees and "marker assisted selection", i.e. breeding on a trait of interest by using the information on the association between the trait and numerous genome wide markers facilitating a faster selection response. ICES should promote the implementation of genomics in fisheries and aquaculture by supporting international collaborative networks and open access web-based resources.

The next ToR emphasized the importance of correct tissue storage for both historical and contemporary samples for DNA analyses, and of choosing the most suitable DNA extraction method for 'valuable' (irreplaceable) samples, thus allowing for maximum 'information yield' from stored samples, that may be jeopardized by suboptimal storage. The group recommend storing tissue (when available) as opposed to DNA. Fresh tissue should generally be stored in ethanol. Otoliths and scales should, for now, be stored in paper envelopes under dry conditions. ICES should support research in relation to optimal storage of historical samples as well as optimal methods for DNA extraction and storage for unique samples. Also, a meta-data base including an inventory of historical collections on a pan-European level should be supported to avoid the loss of invaluable historical DNA.

Novel statistical methods are now available to combine molecular population genetic results with geographical features data(often referred to as landscape genetics). These methods show great potential for delineating population boundaries without making assumptions on population membership by non-objective pooling of samples. It is recommended to fisheries managers that these methods be used in conjunction with geographical information systems to define the spatial and temporal 'footprint' of breeding populations in order to allow population focused management. We recommend further investigations of the relationship between

geographical information and population genetics and the optimal design of such studies, so that maximal use can be made of the synergy between these two fast developing fields.

With the increasing number of DNA based studies of marine fish species, it is vital that integration be encouraged and supported. This can be achieved by the development centrally administered; web accessible meta-data bases on existing primary data sets, DNA and tissue archives, and of those actively engaged in research on the various species. The group recommends that ICES should host such a meta-data base and that the European commission represented by JRC should facilitate the development. The members of WGAGFM should be responsible for input and updating of the data base.

The final ToR was based on a request from WGNAS regarding recommendations on the application of state-of-the-art Genetic Stock Identification (GSI) methods, with particular emphasis on Atlantic salmon (AS). GSI has advanced and diversified in recent years resulting in two types of approach, Mixed Stock Analysis (MSA) and Individual Assignment (IA), in which the goal is either to estimate the proportions of contributing stocks in the catch mixture, or to solve the origin of an individual fish. The group is convinced that in most circumstances IA can give valuable information for Atlantic salmon management and specifically identify the population of origin of individual Atlantic salmon with relatively high probabilities. Presently, according to performed comparison tests, the Bayesian approach of Pella & Masuda (2001) appears to provide the most accurate results with regard to individual assignment, and we therefore recommend its use. Current methods utilise microsatellite technology, however the development of novel markers such as SNPs (Single Nucleotide polymorphisms) should be investigated with regard to their diagnostic usefulness as river and regional specific identifiers.

1 Introduction

The Working Group on the Application of Genetics in Fisheries and Mariculture (WGAGFM) met at Ispra, Italy 19–23 March 2007. The ToRs were decided in Council Resolutions adopted at the ICES Statutory meeting held in Copenhagen, Denmark in 2006. Dr. E. E. Nielsen (Denmark) chaired the meeting, which opened at 0900 h on Tuesday, and closed at 12.00, Friday, April 27.

1.1 Attendance

Nineteen persons from ten countries (Canada, Denmark, Finland, France, Germany, Ireland, Italy, Norway, Poland and United Kingdom) attended the meeting (Annex 2). Nine were official members (or substitutes) for their countries and nine were appointed experts by the chairman for 2007. The latter were registered with ICES prior to the meeting.

1.2 Venue

The meeting was held at the EU joint research centre (JRC), Ispra, Italy. The WG wishes to express our appreciation to our local host Dr. Jann Martinsohn and the rest of the staff at the institute for their kind hospitality. The meeting venue was ideal with accommodation available in Ispra and "shuttle bus service" provided by our local hosts which took us to the Institute in the morning and evenings.

1.3 Meeting Format

WGAGFM has an established framework for completing its ToRs. Prior to the meeting, small ad hoc working groups, under the leadership of one person, are established to prepare position papers related to specific issues in the Terms of Reference. The leader of the ToR is responsible for presenting the position paper in plenary at the meeting and chairing the discussion. Thereafter, volunteers undertake the task of editing and updating position papers according to points raised in the plenary discussions. The ToR leader is responsible for preparing the final report text from their sessions. Prior to the meeting an agenda is circulated to all members.

2 ToR (a): Update and review the available information on the genetics of the European Eel (*Anguilla anguilla*) including importance for recovery plans

Jochen Trautner & Phil McGinnity

2.1 Introduction

In the latest report of the Joint EIFAC/ICES Working Group on Eels is a review (ICES CM 2006/ACFM:16) of the available information on the status of the stock and fisheries of the European eel, which supports the view that the stock as a whole has declined in most of its distribution area, that the stock is outside safe biological limits and that current fisheries are not sustainable. Furthermore, recruitment is at a historical minimum, the level of eel productivity observed since 1990 being below 20% of the level observed not more than three generations ago, that the most recent observations do not indicate recovery and that the opportunities for the protection and restoration of spawner escapement are fading. Fishing pressure, parasites, viral and bacterial infections, oceanic and climatic changes and the impact of other human activities such as the generation of hydroelectric power are implicated in the decline.

In response to the decline in eel stocks the Commission of the European Communities is to develop a Community Action Plan for the management of European Eel. A number of restorative eel management responses are envisaged including; 1) the translocation of glass eel within the natural range of the species using glass eels from sources where there is still a demonstrable surplus and 2) the stocking of eels sourced from aquaculture production (justified on the basis that these are developed entirely on the basis of wild seed). Knowledge of population structuring will have some bearing on the appropriateness of trans-locating eels between river basins and between regions such as between the Mediterranean and the Atlantic and the North Sea and the Baltic. To transfer eels between genetically different populations maybe counter productive to the long-term health of the resource (McGinnity *et al*. 2003, 2004).

In a previous report (ICES Working Group on the Application of Genetics in Fisheries and Mariculture, 2004), no firm conclusions were arrived at with regard genetic structuring of the European eel. However the group did recommend that the precautionary principle be adopted to protect as of yet unresolved genetic variability, and as a consequence the transfer of glass eels between basins should be avoided. Since our last review three years ago there have been a number of new genetic studies. A brief summary of this work and recommendations are provided below.

2.2 Genetic structure within the European Eel

Maes and Volckaert (2006) and Dannewitz *et al*. (2005) have provided comprehensive reviews of the population genetics of the European eel and should be consulted for a more detailed synthesis of the most recent research. In the review by Maes and Volkeart (2006) the suggestion that the eel be classified as a marine fish is a significant insight on how the eel should be viewed in terms of its likely population organisation, at least from the genetic perspective. The eel therefore, because of its assumed reproductive biology i.e. a prolonged spawning period, variance in age-at-maturity, high variability in parental contribution and reproductive success, might be expected to exhibit a high level of genetic variability, high exchange between populations (gene flow) resulting in low genetic differentiation (low genetic signal/noise ratio) and a high genetic population size, all of which are characteristics observed in other typically marine species such as cod, *Gadus morhua* (Nielsen *et al*., 2006) and herring, *Clupea harengus* (Bekkevold *et al*. 2005). Also, as has been observed by Rousset (1997), widely distributed species are rarely fully panmictic (mating randomly), but are

commonly divided into subgroups in a pattern that can be described by one of the classical population models, such as the island model, stepping-stone model or Isolation-by-Distance (IBD) model. In populations composed of a mixture of individuals reproducing at different times within a reproductive season, temporal differentiation can supplement possible geographical partitioning. Under these conditions, gene flow is expected to be limited between early and late reproducers, possibly creating a pattern of Isolation-by-Time (IBT) (Hendry & Day, 2005; Maes *et al*., 2006). Additionally, temporal heterogeneity in the genetic composition of recruits is likely to result from a large variance in parental reproductive success driven by the unpredictability of the marine environment (Waples, 1998). Under the hypothesis of "sweepstakes reproductive success" (Hedgecock 1994), chance events determine which adults are successful in each spawning event, attributing the variation in reproductive success of adults to spatio-temporal variation in oceanographic conditions, occurring within and among seasons. Many marine species split their reproductive effort among several events during a protractive spawning season, to maximize their reproductive success (Hutchings & Myers, 1993; Maes *et al*., 2006).

It is not surprising therefore that Daemen *et al*. (2001); Wirth and Bernatchez (2001); Maes and Volckaert (2002) should independently detect genetic structure indicative of isolation by distance. Ocean currents, resulting in a differential distribution of eel larvae, have recently been suggested to explain this observed genetic structure (Kettle and Haines, 2006). Again not surprising that more recently Maes *et al*. (2006) should detect a significant correlation between genetic distance and temporal distance among recruitment waves indicative of isolation by time. Yet, despite these glimpses of putative structuring, Dannewitz *et al*. (2005) still concluded from their detailed investigations that European eels sampled along the coasts of Europe and Africa most probably belong to a single spatially homogeneous population. However the existence of discrete and stable spawning aggregations is not completely fanciful. In explaining the high incidence of American and European eel (*Anguilla rostrata* and *Anguilla anguilla*) hybrids in Icelandic rivers, Albert *et al*. (2006) suggest that intermediate larval development times for the hybrids are plausible with the effect that ocean currents will deliver the hybrids to rivers positioned in the middle of the eels distribution. Larval development times would have to be adaptive (transporting American eels into American rivers and European eels into European and African rivers) and therefore would have to have at least some heritable basis. That American and European eels are described as two distinct species, in itself, suggests that possibility of structuring and maintenance of structuring over time, as it has been suggested that the spawning grounds of both species overlap in space and time (McCleave, 1987). It is also plausible that larvae and glass eel imprint during ocean transport and that this allows homing of adult eel to natal spawning areas.

Identifying and sampling discrete reproductive aggregations in the spawning areas will most effectively achieve the resolution of the genetic structure of the European eel. This is difficult because the European eel spawns in an area that is not well delineated or sufficiently accessible to fishing. Since Schmidt (1923) identified concentrations of eel leptocephali around the Sargasso Sea in the 1920s there has been little progress in locating eel spawning areas. However it is likely that recent advances in physical oceanography (Kettle and Haines, 2006) offer a reasonable opportunity of overcoming this deficit in the near future. In addition, tagging and tracking of fish has progressed such that monitoring from feeding to spawning ground is feasible. An international project is currently (Spring 2007) underway to recover geo-locational pop up tags in the Sargasso Sea from adult eels previously tagged leaving European rivers and is the main target of a Danish research cruise (http://www.Galathea3.dk).

This information should allow eels to be captured on the spawning grounds for the first time.

There is now sufficient evidence available to suggest that small but significant levels of genetic structuring exist in European eel and that this diversity should be protected.

Within a precautionary principle framework, eel fisheries management should be cognisant of the putative genetic structure suggested by recent studies and that management strategies designed for recovering stocks should incorporate this possibility.

2.3 Recommendations

- Most of the original recommendations from 2004 are still relevant i.e. efforts should be made to ensure sufficient adult recruitment to maintain a large and spatially representative silver eel population; that support be given to current EIFAC/ICES initiatives to collect better spawning stock size data; that a genetic baseline (using otoliths) from pre-decline historical collections be established for critical long-term monitoring of genetic composition; that information on the sex composition of migrating silver eels should be routinely collected in ongoing and future monitoring programmes.
- Knowledge of the biology of the eel in the marine environment is almost non-existent. This lack of knowledge has limited our ability to determine the propensity of the organism to form discrete populations. It is recommended that investigations directed at determining the biology of the eel in the sea should be actively pursued.
- It is recommended that surveys be undertaken to sample and to delineate eel spawning grounds and that new technology based opportunities be taken to locate spawning individuals.
- In light of emerging information suggesting putative stock structure of European eel it is recommended from the genetic viewpoint that glass eels, elvers and other life history stages should not be trans-located between river basins for restocking purposes. However, if it is decided that this should be done in order to avoid imminent collapse of specific river stocks, where possible the translocation should be done within geographically proximate areas e.g. within the Mediterranean basin, within the North Sea, within the Baltic Sea.

2.4 References

Albert, V., Jónsson, B., Bernatchez, L. (2006). Natural hybrids in the Atlantic eels (*Anguilla anguilla*, *A. rostrata*): evidence for successful reproduction and fluctuating abundance in space and time. Molecular Ecology 15, 1903-1916.

Bekkevold, D., André, C., Dahlgreen, T.G., Clausen, L.A.W., Torstensen, E., Mosegaard, H., Carvalho,G.R., Christensen, T.B., Norlinder, E. & Ruzzante, D.E. 2005. Environmental correlates of population differentiation in Atlantic herring. Evolution, 59, 2656-2668 .

Daemen E, Cross T, Ollevier F, Volckaert FAM (2001) Analysis of the genetic structure of European eel (*Anguilla anguilla*) using microsatellite DNA and mtDNA markers. Mar. Biol. 139:755–764

Dannewitz, J., Maes, G.E., Johansson, L., Wickström, H., Volchaert, F.A.M., and Järvi, T. (2005). Panmixia in the European eel: a matter of time? Proc. Royal Society of London, 272: 1129-1137.

Hedgecock, D. 1994: Temporal and Spatial Genetic Structure of Marine Animal Populations in the California Current. California Cooperative Oceanic Fisheries Investigations Reports, 35:73-81.

Hendry, A.P. and Day, T. 2005: Population Structure Attributable to Reproductive Time: Isolation by Time and Adaptation by Time. Molecular Ecology, 14:901-916.

Hutchings, J. A. & Myers, R. A. 1993. Effect of age on the seasonality of maturation and spawning of Atlantic cod, *Gadus morhua*, in the northwest Atlantic. Can. J. Fish. Aquat. Sci. 50, 2468–2474.

ICES CM 2006/ACFM:16. Report of the 2006 session of the Joint EIFAC/ICES Working Group on Eels.

ICES WGAGFM 2004. Report of the Working Group on the Application of Genetics in Fisheries and Mariculture 2004:

Kettle AJ, Haines K (2006) How does the European eel (*Anguilla anguilla*) retain its population structure during its larval migration across the North Atlantic Ocean? Canadian Journal of Fisheries and Aquatic Sciences, 63, 90–106.

Maes, G.E., and Volckaert, F.A.M. 2002. Clinal genetic variation and isolation by distance in the European eel *Anguilla Anguilla* (L.). Biol. J. Linn. Soc. 77: 509–521.

Maes, E.G. Pujolar, J.M., Raeymaekers, C., Joost, D. J. & Volckaert, F. (2006). Microsatellite conservation and Bayesian individual assignment in four Anguilla species. Marine Ecology Progress Series 319: 251-261.

Maes, E. G., Pujolar, J. M., Hellemans, B., Volckaert, F. A. M. (2006). Evidence for isolation by time in the European eel (*Anguilla anguilla* L.). Molecular ecology 15. 2095-2107.

McCleave, J.D. (1987). Migration of Anguilla in the ocean: signposts for adults! Signposts for leptochephali? In: Signposts in the sea. Proceedings on a Multidisciplinary Workshop on Marine Animal Orientation and Migration (eds Herrnkind, W.F., Thistle, A.B.) pp. 102-117. Florida State University, Tallahasse, Florida.

McGinnity P, Prodohl P, Ferguson A, Hynes R, Maoileidigh O, Baker N, Cotter D, O Hea B, Cooke D, Rogan G, Taggart J & Cross T (2003). Fitness reduction and potential extinction of wild populations of Atlantic salmon, *Salmo salar*, as a result of interactions with escaped farm salmon. Proceedings of the Royal Society of London Series B 270: 2443-2450.

McGinnity P, Prodohl P, Maoileidigh O, Hynes R, Cotter DN, Baker N, O'Hea B & Ferguson A (2004). Differential lifetime success and performance of native and non-native Atlantic salmon examined under communal natural conditions. Journal of Fish Biology 62: 1-15.

Nielsen, E.E., Hansen, M.M. & Meldrup D. (2006) Evidence of microsatellite hitch-hiking selection in Atlantic cod (*Gadus morhua* L.): Implications for inferring population structure in non-model organisms. Molecular Ecology, 15, 3219-3229.

Pujolar, M., Maes, E.G. & Volckaert, F (2006). Genetic patchiness among recruits in the European eel *Anguilla anguilla*. Marine Ecology Progress Series 307: 209-217.

Rousset, F. 1997: Genetic Differentiation and Estimation of Gene Flow From F-Statistics Under Isolation by Distance. Genetics, 145:1219-1228.

Schmidt, J. 1923: The breeding places of the eel. Phil. Trans. R. Soc. B. 211: 179–208.

Waples, R.S. 1998: Separating the Wheat From the Chaff: Patterns of Genetic Differentiation in High Gene Flow Species. Journal of Heredity, 89:438-450.

Wirth T, Bernatchez L (2001) Genetic evidence against panmixia in the European eel. Nature 409:1037–1040

3 ToR (b): Critically review the potential application of genomics in fisheries management and aquaculture

Roman Wenne[1], Pierre Boudry[2], Jakob Hemmer-Hansen[3], Krzysztof P. Lubieniecki[4], Anna Was[5], Antti Kause[6]

1- Department of Genetics and Marine Biotechnology, Institute of Oceanology, Polish Academy of Sciences, Sw. Wojciecha 5, 81-347 Gdynia, Poland, E-mail: RWenne@iopan.gda.pl or RWenne@cbmpan.gdynia.pl

2-Laboratoire IFREMER de Genetique et Pathologie, 17390 La Tremblade, France, (E-mail: pboudry@ifremer.fr)

3- Danish Institute for Fisheries Research, Department of Inland Fisheries, Vejlsøvej 39, 8600 Silkeborg, Denmark (E-mail: Jhh@difres.dk)

4- Department of Molecular Biology and Biochemistry, Simon Fraser University, 8888 University Dr., Burnaby, BC, V5A 1S6 Canada, (E-mail: Klubien@sfu.ca)

5- Mayo Insitute of Technology, Galway, Ireland (E-mail: Anna.Was@gmit.ie)

6- MTT Agrifood Research Finland, Biotechnology and Food Research, Biometrical Genetics, 31600 Jokioinen, Finland (E-mail: Antti.Kause@mtt.fi)

3.1 Abstract

At present a very dynamic development and application of genomics has been facilitated in a number of fields by the availability of new methodologies and tools, such as high throughput DNA sequencing and cDNA microarrays. Genomic tools are already used in research on commercially important fish and shellfish species. Thousands of ESTs are now available for some of these species and sequencing of complete genomes of cod, salmonids, flatfishes, sea bass and Pacific oyster has already been initiated. Microarray technology through the expression studies of thousands of genes at a time allow for identification of candidate genes involved in the function of multiple physiological, morphological and behavioural traits of interests in organisms and populations from different environments, which can be subject to selective pressure from e.g. fishery and aquaculture. This ToR will review the current development of genomic technologies and pinpoint their potentially beneficial applications and implications for fisheries management and aquaculture.

3.2 Introduction

It is well recognised that fisheries catches have reached a plateau in recent years. Due to the high demand for fish and shellfish on the global market, aquaculture production contributes an increasing amount to the food supply. Management of exploited wild stocks is undergoing systematic improvement. Results of population genetic investigations have recently been incorporated as a useful tool in stock identification in addition to morphological, biological and physiological traits. Aquaculture industry has expanded especially in South America and Asia, from which aquaculture products are also imported world-wide. To increase further the competitiveness of the fisheries and aquaculture industry, major development work should be conducted. Simultaneously, industry practices should be sustainable and marine biodiversity should be maintained. Genomics tools combined with the already well-established aquaculture and fisheries management practices can serve as a novel framework in such development work. Genomics is a field of science that deals with the structure, function and evolution of genomes. Many current DNA and RNA-based studies fall into this field, even if they are often not strictly part of it. Genomics often simply implies the use of high throughput DNA- or RNA-based methods. It comprises comparative, functional and environmental genomics.

Comparative genomics examines whole genomes, their gene content, gene order, structure, evolution and taxonomy. Functional genomics investigates the biochemical and physiological role of gene products and their interactions on a large or small scale. Environmental genomics encompasses studies of molecular variation in natural or artificial populations of different taxa and their response to environmental conditions such as temperature or pollutants. One of the main efforts in genomics has been to obtain high numbers of large pieces of sequences of genomes and to assemble them into full sequences of chromosomes. Another goal has been to study the expression of thousands of genes using techniques such as microarrays or other high throughput expression RNA profiling (i.e. transcriptomics). The analysis of the immense amount of data generated by such approaches often requires the use of specific computerised methods, or "bioinformatics". The knowledge of genomics opens new perspectives for the biotechnology of marine organisms, with implications for fisheries and aquaculture.

3.3 Sequencing and analyses of genomes

Sequencing of genomes facilitates the development of variety of DNA-based genetic markers to be used for the management of the wild populations and aquaculture. Expressed Sequence Tags (ESTs) are obtained by sequencing cDNA libraries. Such libraries can be obtained from tissue specific libraries or generated by Suppressive Subtractive Hybridization (SSH). EST databases for important marine species have been established (e.g. in oyster: http://www.ifremer.fr/GigasBase/) and most sequences are submitted to databases (http://www.ncbi.nlm.nih.gov/Genbank/; http://compbio.dfci.harvard.edu/tgi/). ESTs are the first step towards full length cDNA and gene sequence. Sequencing of whole fish and shellfish genomes contributes not only to the understanding of vertebrate and invertebrate evolution but also to environmental genomics and aquaculture (Crollius and Weissenbach 2005, Cossins and Crawford 2005). Full genome sequences are now available for a few model fish species such as zebrafish, fugu, puffer fish, medaka and stickleback. To date, only one commercially important fish – Tilapia (Cichlid Genome Consortium, http://hcgs.unh.edu/cichlid/) has obtained founding for sequencing; however, knowledge of the genome sequences of other commercially important species is critical for an efficient identification of economically important genes and polymorphisms. Recently, international collaborative initiatives have been undertaken with the aim of obtaining full or partial genomic sequences of some commercially important fish and shellfish species as salmon and rainbow trout (cGRASP, www.cgrasp.org), cod (www.codgene.ca), sea bass *Dicentrarchus labrax* (Chini *et al*. 2006) and the Pacific oyster *Crassostrea gigas* (Hedgecock *et al*. 2005). Full genome sequences can be based on the sequencing and assemblage of bacterial artificial chromosomes (BACs) containing DNA fragments of the whole target genome, or shotgun approaches. BAC libraries are available for Atlantic salmon (Thorsen *et al*. 2005), rainbow trout (Palti *et al*. 2004), sea bass (Whitaker *et al*. 2006), channel catfish (Quiniou *et al*. 2003) and oysters (Cunningham *et al*. 2006), and can also be used for physical mapping (i.e. BAC fingerprinting).

3.4 Fisheries

3.4.1 Discrimination of wild populations

Population genetic research has contributed substantially to our understanding of how fish and shellfish species are genetically structured into reproductively isolated populations across their distributions. Such knowledge is of major importance for fisheries management because local populations are often considered worth conserving due to their unique contribution to the genetic diversity of the species, which may allow them to sustain productivity in changing environments (e.g. Hilborn *et al*. 2003). In addition, local populations have often adapted to local environmental conditions and are therefore characterised by unique morphological, physiological and life history traits that have a genetic basis and are therefore of conservation interest. Moreover, the vitality of such populations are often of great economic interest.

However, the actual genetic basis of quantitative traits remains mostly unknown, because their analysis has, until recently, been logistically difficult and time consuming to conduct in most species. The identification of local adaptations in natural populations has recently been highlighted as worth special attention in the years to come (e.g. Moritz 2002; van Tienderen *et al*. 2002; ICES 2006).

Many fish and shellfish populations have been over exploited or reduced by changes in local environments. These populations are endangered and some indigenous populations are already extinguished (e.g. Dulvy *et al*. 2003; Reynolds *et al*. 2005) and hence there is an urgent need for knowledge of the basic population structure of many species. Various genetic markers have been used in order to identify and characterise populations. Studies of such markers as allozymes, mtDNA polymorphism, microsatellites, RFLPs and recently, AmpFLP/RAPD have successfully demonstrated significant genetic differences between populations of many species. However, these markers represent only a small fraction of the total genomic polymorphism. Furthermore, they are mostly believed to be selectively neutral, and have thus mainly been used to draw inferences about the interplay of gene flow, genetic drift and historical processes, thereby limiting inferences about local adaptations in the species under study. Single nucleotide polymorphism (SNP) analysis is a new and powerful method that will be very helpful in managing natural and captive populations in the future. SNPs can be identified in inter-individual comparison of genomic DNA sequences or sequences derived from ESTs. Thus both coding and noncoding DNA sequences can be used to identify SNPs.

Newly developed techniques enable screening for polymorphisms throughout the whole genome. Screening for many more loci will open new possibilities in population genetic research, moving to population genomics (Luikart *et al*. 2003). Global gene expression can be examined through the use of microarray techniques, enabling the simultaneous analysis of thousands of genes. Different expression levels can be observed in specimens originating from different localities and differing in functional traits (Rise at al. 2004a). To verify these correlations, many SNPs with known locations on a linkage map can be used to identify QTLs, which can subsequently be genotyped in natural populations. Microarrays have been developed for a number of fish and shellfish species (such as salmon (Rise *et al*. 2004a; von Schalburg *et al*. 2005a), killifish (Oleksiak *et al*. 2002), carp (Gracey *et al*. 2004), zebrafish (Ton *et al*. 2002), catfish (Li and Waldbieser 2006), medaka (Kimura *et al*. 2004), European flounder (Williams *et al*. 2003), Japanese flounder (Kurobe *et al*. 2005), mussel (Venier *et al*. 2006) and oysters (Lang *et al*. 2006). However, even if arrays have not been developed for a species, reliable estimates of gene expression may still be achieved through alternative measures such as cross species hybridization (Renn *et al*. 2004) or the use of non-array based measures of gene expression, such as cDNA-AFLP or differential display (Breyne *et al*. 2003; Venkatesh *et al*. 2005). It should be noted that gene expression analyses essentially measure expression phenotypes. The degree of heritability of gene expression traits has rarely been assessed, but is often assumed (Gibson and Weir 2005). Still, for these measures to be used to illustrate population genetic differences, the environment needs to be controlled to rule out environmental effects on gene expression. Therefore, analyses of global gene expression require common garden approaches. Still, if properly designed such studies have great potentials to disclose the genetic basis of adaptations in local populations of fish and shellfish (e.g. Whitehead and Crawford 2006)

Genome scans are another important group of genomic tools applying the screening of a high number of markers to cover the entire genome of a species under study (e.g. Luikart *et al*. 2003; Storz 2005). Genome scans allow for the identification of outlier loci that are potentially under selection or linked to a locus under selection, i.e. hitch-hiking selection (Maynard Smith and Haigh 1974), thereby facilitating detection of the genetic basis of local adaptation in natural populations. Outliers can be detected using model based (e.g. Beaumont and Nichols 1996) or model free (e.g. Schlötterer 2002; Kauer *et al*. 2003) methods of tackling.

Conclusions with respect to outlier status of particular loci will often be considerably strengthened, if signals of selection are supported by several different analytical approaches as well as different pairwise population comparisons (e.g. Vasemägi *et al*. 2005; Bonin *et al*. 2006). Both allozymes, microsatellites and AFLPs have been used in genomes scans in non-model species (e.g. Storz and Dubach 2004; Vasemägi *et al*. 2005; Bonin *et al* 2006), but SNPs could also be very useful for these approaches. A major advantage of genome scans is that they can be applied to natural populations, thereby increasing the number of species for which such approaches are possible.

3.4.2 Conservation issues

Genomics offer new possibilities for conservation genetics in two ways. The first is that it can increase the number of neutral genetic markers available. This will likely improve estimates of the effects of demographic processes, such as population declines and bottlenecks, effective population sizes, identification of wild and farmed individuals etc. It may also result in increased statistical power to detect minute levels of population structuring and assigning individuals of unknown origin to known baseline populations. However, many of these aspects are already relatively well covered with existing population genetic approaches that use genetic markers, such as microsatellites or AFLPs at moderate scales (Kohn *et al*. 2006). The second, a more prosperous application of genomics in relation to conservation is the detection of the genetic basis of local adaptation. Given that we know very little about this in most species of fish and shellfish, such knowledge will greatly improve our ability to manage genetic diversity in natural populations.

3.4.3 Case study using Atlantic salmon

Atlantic salmon is one of the species of relevance to fisheries management and aquaculture, where genomic resources are building up rapidly. Hence this species could serve as an important case study to demonstrate the resources that may become available in other species in the near future, as well as the potential applications of these resources. Genomic approaches in Atlantic salmon have targeted both RNA and DNA levels of variation. For instance, a salmonid microarray containing cDNAs representing 16,006 genes have been developed. The genes spotted on the array have been carefully selected from Atlantic salmon and rainbow trout expressed sequence tag (EST) databases. This array will serve as an important resource for genetic, physiological and ecological studies as well as many other fields of salmonid research (von Schalburg at al. 2005a). Gene expression patterns determined either for target genes or using microarrays have already been used to investigate the salmonid immune response, several disease processes and disease resistance (Lindenstrom *et al* 2003; Rise *et al*. 2004b; Sigh *et al* 2004; Bridle *et al* 2006a; b; Fast *et al*. 2006; Lindenstrom *et al*. 2006; Martin *et al*. 2006; Purcell *et al*. 2006). Moreover, they have been used to survey the genes involved in the maturation and development of the rainbow trout ovarian and testicular tissues (von Schalburg *et al*. 2005b; 2006; Bonnet *et al*. 2007), to examine brain gene expression profiles in male salmon with different life history strategies (Aubin-Horth *et al*. 2005a; b), to carry out toxicogenomic profiling of hepatic tumour promoters in rainbow trout (Tilton *et al*. 2006), to investigate the response of the rainbow trout transcriptome to model chemical contaminants (Koskinen *et al*. 2004) and to study gene expression in atrophying muscle (Salem *et al*. 2006). Microarrays have also been applied to discriminate between farmed and wild Atlantic salmon using genome wide transcription profiles, showing that similar transcription profiles characterised farmed strains from Norway and Canada, suggesting adaptation via gene expression to common captive environments (Roberge *et al*. 2006).

The DNA level has been targeted through markers such as microsatellites and SNPs. For example, Vasemagi *et al*. (2005) used EST linked microsatellites in a genome scan of Atlantic salmon populations and identified a number of outliers potentially under selection and it has been shown that these markers can be used in other salmonids (Ng at al. 2005). SNPs have

been identified through different approaches. For example, five populations of chinook salmon *Oncorhynchus tshawytscha* from the Pacific North America were surveyed for SNPs at 19 loci by sequencing (Campbell and Narum 2007). Of these 13 were chosen for Taqman assays (5' exonuclease assays) out of 58 SNPs. Similarly, 1195 SNPs have been identified from ESTs and 121 of these have been characterised by pedigree analysis. As a result of the genome duplication that took place in common ancestor of extant salmonids 25 – 120 MYA, it has been estimated that up to 15-20% of salmonid loci have functional duplicates. Therefore, when identifying putative SNPs in the Atlantic salmon EST database, it is important to be able to distinguish between true SNPs (i.e., those corresponding to alleles at a single locus) and paralogous sequence variants (PSVs; i.e., sequence differences between duplicate loci), (Wright at al. 2007).

SNPs with known locations on a linkage map can be used to identify QTLs, which can subsequently be genotyped in natural populations. Boulding *et al*. (2007) present an example of such research using Atlantic salmon in which 4 full sib families (backcrossed F1 males to a female from European and American parental populations) of salmon were used to identify SNPs for known QTLs. The identified SNPs in traits under selection will be genotyped in endangered wild Atlantic salmon to demonstrate genetic differences in functional traits among these endangered salmon populations and may help in their conservation (Boulding *et al*. 2007).

3.5 Aquaculture

3.5.1 Using genomic information in aquaculture breeding

3.5.1.1 Constructing DNA pedigrees

In breeding programmes, information on family relations of individuals is used when estimating genetic parameters (heritabilities and genetic correlations) and breeding values for traits, and when optimising selection and mating in order to avoid inbreeding. Similarly, pedigrees are useful in management of conservation programmes and wild populations (Wilson and Ferguson 2002), e.g. when controlling inbreeding. In addition to physical individual tagging, pedigrees can be determined using DNA markers. This procedure is quite straightforward using microsatellites. Typically 10-20 variable genetic markers are needed to assign >95% of individuals to single pairs of parents (e.g., Vandeputte *et al*. 2006). To do this, one needs tissues samples from both potential parents and their offspring, and several freely available softwares exists for parental assignment (reviewed by Jones and Ardren 2003). Physical individual markers are useful when large facilities exist where family groups can be held separately until fish are large enough to be individually tagged. For instance, in salmonid breeding programmes fish are typically held in hundreds of family tanks until individually tagged at weight around 50g (Kause *et al*. 2005). Using DNA pedigree is useful in many aquaculture breeding and conservation programmes when individual tagging is difficult or when facilities for family tanks do not exist. Three examples of using DNA pedigrees are detained below: walk-back selection, estimation of genetic parameters, and conservation programmes.

Walk-back selection refers to a selection programme where a group of superior individuals are first selected, and then only the selected animals are genotyped for family relations. Using the established pedigree, only those superior animals that are not too closely related are used in matings (Doyle and Herbinger 1994, Sonesson 2005). This is an improved mass selection scheme to obtain genetic improvement while simultaneously controlling for inbreeding. This is cost-effective because only some hundreds of individuals among the potentially (tens of) thousands of individuals reared need to be genotyped. This is especially useful for species for which no extensive resources are available or for new species those reproduction cannot be fully controlled. Furthermore, there are studies showing that microsatellite markers are useful

for determining the effective number of parents and their individual reproductive success (e.g. Boudry *et al*. 2002), and level of inbreeding in mass spawning condition mass selection scheme in the Nile tilapia (Fassehaye *et al*. 2007) as well as the flat oyster *Ostrea edulis* (Launey *et al*. 2001). A full population can be genotyped for parental analysis, which allows one to estimate heritabilities and genetic correlations to traits of interest. Such an approach has been used e.g. Vandeputte *et al*. (2005) on common carp.

Likewise, DNA pedigrees are useful in conservation programmes of wild fish and shellfish, especially when aiming at controlling inbreeding. Microsatellite genotyping in the induced mass spawns of lion-paw scallop demonstrated that some parents contribute much higher percent to the progeny than expected (Petersen *et al*. 2007). Similarly, the impact of hatchery practices on the genetic variability of progenies can be monitored (Taris *et al*. 2006). That is, when the effective population size is reduced and inbreeding can decrease hatchery stocks and impact natural population in the case of introgression. Mutliplexing (i.e. simultaneous PCR amplification) of microsatellites (e.g. Taris *et al*. 2005) and SNP-based parentage assignment (Rengmark *et al*. 2006; Anderson and Garza 2006) are now greatly facilitating this type of studies.

3.5.1.2 Marker assisted selection

Marker assisted selection (MAS) refers to a selection procedure which is improved using information from genetic markers. Allelic variation in genetic markers can be linked to the variation in traits of economic interest, and thus the marker provides DNA level information on the inheritance of the traits. Marker assisted selection is especially useful for traits that are difficult to breeder using traditional means. Such traits can be costly or difficult to record (feed efficiency, disease resistance, omega-3-acids), they may require slaughtering of individuals (fillet quality, body composition), can be recorded from only one sex (caviar production), or cannot be directly recorded from breeding candidates (e.g., sea performance when breeding candidates are held at fresh water breeding station). Moreover, MAS can be used early in life to breed for traits that are expressed later in life (e.g., caviar production, maturity age), allowing one to cull the population to save feed and management costs (e.g. Martinez *et al*. 2005).

The practical use of markers in selection can be roughly divided into three classes: 1) removing genetic disorders, 2) marker breeding value-selection, and 3) genomic selection. These three methods differ in the complexity of computational selection tools needed and requirements of the genomic data. Recessive genetic disorders determined by a simple Mendelian one-locus way can be effectively removed from a population using a gene test made from a small tissues sample. Individuals carrying a deleterious allele are culled, and no computationally demanding selection tools are needed. Such tests are in practical use in terrestrial farm animals (e.g., Sironen *et al*. 2006). Marker breeding values of individuals can be estimated by combining information on phenotypes and a single or several QTLs (quantitative trait loci) segregating within a pedigreed population (Fernando and Grossman 1989). When estimating breeding values, genetic variation can be explained by the QTL effect(s) and the remaining polygenetic parts. For a QTL to be useful here, a genetic marker needs to be located very close to the actual gene, i.e., within less than 1 cMs. If this is not the case, then it is unlikely that the QTL will be applicable across the whole population, and the linkage between the marker and the gene will be broken down by recombination during the next few generations. Thus, QTL fine mapping is needed for the QTL to be practically usefully. Marker breeding values are used in dairy cow selection, e.g., in France and Germany (Hayes *et al*. 2006).

Genomic selection refers to selection directed on allelic variation identified across the whole genome. Allelic variation in thousands of loci as well as their affects on economic traits can be estimated, and genomic breeding values can be thus estimated (Meuvissen *et al*. 2001).

After the effects of the alleles have been established, no phenotypic information on animals are needed in selection. Single nucleotide polymorphism (SNP) is the most promising method for such whole-genome analysis. Using current technology, variation in tens of thousands of SNPs can be simultaneously estimated. For this method to be effective, however, Hayes *et al*. (2006) suggested that 10-20 QTLs need to be found for each trait and up to 30 000 SNPs may be needed to obtain enough dense marker map. Methods to perform such analyses and genomic selection tools are currently under development. Thousands of SNPs are already available for Atlantic salmon (Hayes *et al*. 2007).

3.5.1.3 Linkage maps and QTLs in aquatic species

Linkage maps are needed for mapping of major chromosome regions influencing phenotypic traits (quantitative trait loci, QTLs). Examples of published linkage maps for several major aquaculture and model species are given in Table 1. For most of the marker maps, the average distance between markers is 2-10 cM. Already average marker distance of 20 cM would be suitable for location a QTL to a correct chromosome arm. A useful feature is that the male maps are often shorter than the female maps. Thus, the initial QTL mapping can be more easily performed using male offspring. For fine mapping, marker distance of 1 cM or less is needed.

A variety of markers has been used for identification of populations and strains in the wild and aquaculture with the aim of improving management. These markers can be used for construction of high-resolution genetic linkage maps and search for quantitative trait loci (QTL), and finally to marker assisted selection (MAS) (Liu and Cordes 2004; Sarropoulou *et al*. 2005a; Senger *et al*. 2006; Silverstein *et al*. 2006; Montano-Perez *et al*. 2006). Table 2 lists QTL studies performed in several aquaculture species. Two observations can be done from these studies. First, most of the studies are on growth-related traits, followed by disease resistance traits. Only a few or none studies exits for quality or feed utilisation traits. MAS will be especially useful for disease and quality traits. Second, most of the studies on aquatic species have not proceeded to fine-mapping. Consequently, a lot of effort must be put on this area, in order for genomic studies to be useful in practical breeding programmes.

Studies of large scale gene expression using microarrays containing clones from cDNA library are helpful in discovery of candidate genes for particular/multifactorial traits (Sarropoulou *et al*. 2005b). However, the challenge is to find between-individual variation in gene expression that could be exploited in selective breeding. By simply knowing that a certain gene is expressed or not is not enough, the gene must also display alternative gene variants that can be selected. Hedgecock *et al*. (2007) recently reported transcriptomic analysis of growth heterosis in larvae using megacloning and massively parallel signature sequencing (MPSS) in the Pacific oyster. Microarrays have been produced in USA (Lang *et al*., 2006, Jenny *et al*., in press) and are currently under development in Europe to study summer mortality following SSH approaches (Huvet *et al*. 2004; Saavedra and Bachere 2006).

3.5.1.4 Identification of sex

The understanding of sex determination systems is one of the most sought after aspects of genomics in finfish aquaculture. Aquaculture farmers often prefer to farm only one of sexes, because of its superior characteristics (Kause *et al*. 2003). Moreover, production of sterile animals (e.g. using triploidy) will enable further reduction of risks related to escape effects of farmed animals on natural marine populations. Males are the heterogametic sex in salmon and Arctic charr. Several microsatellite markers are linked to the sex-determining factor (SEX) in the linkage analysis (Woram *et al*. 2003, Artieri *et al*. 2006; Fujiki *et al*. 2007; Kwitkowski *et al*. 2007). BACs or fosmid clones positive for these microsatellites were isolated from libraries. Fluorescence in situ hybridisation (FISH) was used to identify their positions on chromosomes. Fosmids, BACs and BAC-ends sequences were used for identification of the

SEX candidate genes. Despite the discovery of a finding sequence characterised amplified region (SCAR), the search for the SEX factor has to be continued.

3.5.1.5 Cost-benefit analyses

The development and extensive use of genomic tools in selection are resource demanding. Consequently, a cost-benefit analysis would be useful for determining the advantage of using genomic tools. Break-even cost of genotyping depends on the efficiency of MAS selection (relative to traditional selection), duration until the loci selected is fixed, the size of the producer level and costs of genotyping, as shown by the analysis of pig enterprise by Hayes *et al*. (2004). Likewise, breeders should in advance determine the way marker assisted selection can be most effectively used. For instance, should all individuals be genotyped, should individuals be genotyped early in life or at maturity, should only pre-selected breeding candidates be genotyped, could within-family selection be effective, and what are the economical benefits and practical constraints of the alternative selection strategies. There are no such studies performed in aquatic species.

Table 1. A list of linkage maps on aquaculture species. If two values are given for a parameter, the first refers to the male and the second to the female map.

SPECIES	NRO OF MARKERS	MAP LENGTH (CM)	LINKAGE GROUPS	AVERAGE DISTANCE BETWEEN MARKERS (CM)	REFERENCE
Sea bream	204	1242	26	6.1	Franch *et al*. (2006)
Sea bass	162	567 / 906	25	3.5 / 5.6	Chistiakov *et al*. (2005)
Atlantic salmon	251 / 230	103 / 901	31 / 33	0.41 / 3.9	Moen *et al*. (2004a)
Atlantic salmon	64	na	15	na	Gilbey *et al*. (2004)
Rainbow trout	476	2628	31	5.5	Young *et al*. (1998)
Rainbow trout	1359	4359	30	3.2	Nichols *et al*. (2003)
Rainbow trout	903	2750	31	3.0	Guyomard *et al*. (2006)
Rainbow trout	208	na	29	na	Sakamoto *et al*. (2000)
Brown trout	288	346 / 912	37	1.2 / 3.2	Gharbi *et al*. (2006)
Pink salmon	22	na	8	na	Matsuoka *et al*. (2004)
Arctic charr	327	390 / 992	46	1.2 / 3.0	Woram *et al*. (2004)
Tilapia	162	704	30	4.3	Kocher *et al*. (1998)
Tilapia	546	1311	24	2.4	Lee *et al*. (2005)
Tilapia	214	1632	24	7.6	Agresti *et al*. (2000)
Common carp	268	4111	50	15.3	Sun & Liang (2004)
Channel catfish	418	1593	44	3.8	Liu *et al*. (2003)
Walking catfish	134	2037	31	17.1	Poompuang & Na-Nakorn (2004)
Pacific oyster	119	1031	11	9.5	Li & Guo (2004)
Pacific oyster	102	616 / 770	11 / 22	6.0 / 7.6	Hubert & Hedgecock (2004)
Blacklip abalone	102 / 98	621 / 766	17 / 20	7.3 / 9.8	Baranski *et al*. (2006)
Pacific abalone	94 / 119	1366 / 1774	19 / 22	18.2 / 18.3	Liu *et al*. (2006)
Blue mussel	116 / 121	825 / 863	14	8.0 / 8.1	Lallias *et al*. (submitted)
European flat oyster	137 / 149	471 / 450	9 / 10	4.9 / 4.2	Lallias *et al*. (submitted)

SPECIES	NRO OF MARKERS	MAP LENGTH (CM)	LINKAGE GROUPS	AVERAGE DISTANCE BETWEEN MARKERS (CM)	REFERENCE
Medaka	633	1354	24	2.1	Naruse *et al*. (2000)
Pufferfish	200	697 / 1213	22 / 22	3.5 / 6.1	Kai *et al*. (2005)
Swordtail	290	2178	24	7.5	Walter *et al*. (2004)
Ayu	195	1660	36	8.5	Watanabe *et al*. (2004)

n.a. - Not available.

Table 2. A list of quantitative trait loci (QTL) studies on aquaculture species

SPECIES		
Traits studied	Reference	
Rainbow trout	Hatching time, embryonic length, weight	Martinez *et al*. (2005)
Rainbow trout	Embryonic development rate	Robison *et al*. (2001)
Rainbow trout	Development rate	Sundin *et al*. (2005)
Rainbow trout	Body length, thermotolerance	Perry *et al*. (2005)
Rainbow trout	Growth, condition factor, maturity age	Martyniuk *et al*. (2003)
Rainbow trout	Spawning time	Sakamoto *et al*. (1999)
Rainbow trout	Length, pyloric caeca, nro of scales	Nichols *et al*. (2004)
Rainbow trout	Pyloric caeca	Zimmerman *et al*. (2005)
Rainbow trout	Thermotolerance	Jackson *et al*. (1998)
Rainbow trout	Thermotolerance	Danzmann *et al*. (1999)
Rainbow trout	Infectious pancreatic necrosis virus	Ozaki *et al*. (2001)
Rainbow trout	Infectious hematopoietic necrosis	Rodriquez *et al*. (2004)
Rainbow trout	*Ceratomyxa shasta* resistance	Nichols *et al*. (2003)
Rainbow trout	Killer-cell activity	Zimmerman *et al*. (2004)
Atlantic salmon/Rainbow trout/Arctic charr	Body weight, condition factor	Reid *et al*. (2005)
Atlantic salmon	Infectious salmon anaemia	Moen *et al*. (2004b)
Atlantic salmon	Infectious pancreatic necrosis virus, furunculosis, infectious salmon anaemia	Kjoglum *et al*. (2005)
Atlantic salmon	Furunculosis, infectious salmon anaemia	Grimholt *et al*. (2003)
Coho salmon	Fillet colour	Araneda *et al*. (2005)
Arctic charr	Growth rate	Tao & Boulding (2003)
Tilapia	Innate immunity, response to stress, growth	Cnaani *et al*. (2004)
Tilapia	Thermotolerance	Moen *et al*. (2004c)
Asian seabass	Growth traits	Wang *et al*. (2006)
Eastern oyster	Perkinsus marinus resistance	Yu and Guo (2006)
European Flat oyster	Bonamia ostreaea resistance	Lallias *et al*. (2007)

3.6 Cultured fish and shellfish health

Genomics can help to overcome problems related to infectious diseases by better understanding host defence systems and identifying QTLs or candidate genes. Aquaculture productivity is reduced by various pathogens. Examples of genomics based studies include oomycete *Saprolegnia parasitica* (Torto-Alalibo *et al*. 2005), a bacterial agent of cold-water disease *Flavobacterium psychrophilum* (Soule *et al*. 2005) and a parasitic protozoan ciliate *Ichthyophthirius multifiliis*, the agent of the white spot disease through virulent factors (Abernathy *et al*. 2007). The acute phase response following infection of catfish with

Edwarsiella ictaluri, causing enteric septicaemia, was studied by high density in situ oligonucleotide microarray (Peatman *et al*. 2007). Numerous of acute phase proteins were upregulated and many pathogen recognition receptors and chemokines were differentially expressed in the liver. These results were confirmed with real-time PCR. A candidate gene approach was employed to find markers associated with disease resistance in which 28 microsatellites located near and in the immune genes were developed (Karsi *et al*. 2007). Two microsatellites were associated with resistance and two with susceptible phenotypes. These markers have been incorporated to the catfish linkage map, which will facilitate finding resistance QTLs and will help in development of MAS programmes. Cytokines are important immune system regulators in fish and genomics and proteonomics can help to develop vaccines and immunostimulants for aquaculture (Savan and Sakai 2006). Bao *et al*. (2007) identified 26 chemotactic cytokine genes, sequenced them and studied their expression in catfish.

Whirling disease strongly affects western American hatcheries and natural populations of rainbow trout, is caused by *Myxobolus cerebralis*. A European hatchery strain Hofer exhibits almost complete resistance to this pathogen. Microarray technology was used to study differences in global gene expression between resistant and susceptible rainbow trout strains (Baerwald *et al*. 2007). Several candidate genes were discovered that indicates genetic mechanisms of resistance to whirling disease in fish. To pinpoint these genetic mechanisms of resistance rainbow trout full sib families under hatchery conditions were exposed to the bacterial pathogen *Yersinia ruckeri*, which causes enteric red mouth disease and *Flavobacterium psychrophilum* (Palti *et al*. 2007). Linkage disecquilibrium and the resistance to the pathogens was assessed by genotyping with microsatellites linked to the four major histocompatibility (MH) genomic regions, to toll-like receptor genes and to the two copies of tumour necrosis factor superfamily 13b. The MH sequences variation can be useful in selective breeding for resistance to the pathogens. Atlantic salmon T-cell receptor alpha/delta genes exhibit vast diversity for antigen recognition (Yazawa *et al*. 2007). Differences in susceptibility to infectious hematopoietic necrosis virus were studied with microarray technology between four salmonid species (*Salmo salar, Oncorhynchus nerka, O. keta* and *O. kisutch*). The observed differences were related to species-specific differences in viral ability to enter cells, and possibly to effectiveness in taking control over cellular mechanisms rather than from strength of the host immunological response (Miller *et al*. 2007).

Examples of invertebrate cultured species related studies are proteins AMPs and other elements of the immune system in paneid shrimps and Pacific oyster (*C. gigas*, Bachere *et al*. 2004). Another pathogen in shrimp is the white spot syndrome virus, which can be controlled by the anti-viral immunity of injected dsRNA molecules and siRNAs (Westenberg *et al*. 2005).

3.7 Alternative feed

Functional genomics can contribute to the production of new kinds of feed for cultured fishes. One of examples is finding a possibility of production of novel feed sources (such as plant based protein for carnivorous fish) or terrestrial transgenic plants (soybean or rapeseed) feed with marine – similar fatty acid composition (Opsahl – Ferstad *et al*. 2003). Present evidence indicates that in Salmonids, novel plant-based feed sources may induce only weak genotype-by-diet interactions that would need special attention from fish breeders (Quinton *et al*. 2007). However, novel feed ingredients may have deleterious effects on quality or biological efficiency of aquatic species, and it is a challenge to develop the feed and the animal material further. Moreover, genomic approaches could also contribute to a better understanding of lipid pathways and synthesis of long chain polyunsaturated fatty acids, facilitating the selection of genotypes that dictate a good performance when feed a low fish oil or protein diet.

3.8 Recommendations

- The implementation of genomic approaches should be encouraged in the fields of fisheries and aquaculture by supporting the development of genomic resources, such as BAC libraries, fine scale linkage maps, EST databases and expression profiling.
- International networks and large collaborative initiatives are essential so that projects such as full genome sequencing can be implemented and be exploited in various fields of fisheries and aquaculture.
- Open access web-based resources, joining available genomic data (ESTs, mapping data, BAC fingerprinting and annotation…) should be developed in order to favour integrated collaborations (see also ToR E).
- Studies of local adaptations in the wild and hatchery populations should incorporate genomic approaches to further understand the footprints of selection at a genome wide level.
- Potentials of molecular marker assisted selection and domestication process in aquaculture species should be further explored, benefiting from the development of new genomic resources and computational and analytical tools.

3.9 Literature cited

Abernathy J, Xu P, Li P, Xu D, Kucuktas H, Klesiun P, Arias C, Liu Z. 2007. A genomic resource for the analysis of developmental and virulent factors of the protozoan fish parasite *Ichthyophthirius multifiliis* using microarrays. Plant & Animal Genomes XV Conference, January 13-17, 2007, Town & Country Convention Center, San Diego, CA, P639, Abstract

Agresti JJ, Seki S, Cnaani A, Poompuang S, Hallerman EM, Umiel N, Hulata G, Gall GAE, May B. 2000. Breeding new strains of tilapia: development of an artificial center of origin and linkage map based on AFLP and microsatellite loci. Aquaculture 185: 43-56.

Anderson EC, Garza JC. 2006 Power of single-nucleotide polymorphisms for large-scale parentage inference. Genetics 172: 2567-2582

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

Artieri CG, Mitchell LA, Ng SHS, Parisotto SE, Danzmann RG, Hoyheim B, Phillips RB, Morasch M, Koop BF, Davidson WS. 2006. Identification of the sex-determining locus of Atlantic salmon (*Salmo salar*) on chromosome 2. Cytogenet. Genome Res. 112: 152-159.

Aubin-Horth N, Letcher BH, Hofmann HA. 2005a. Interaction of rearing environment and reproductive tactic on gene expression profiles in Atlantic salmon. Journal of Heredity 96: 261-278.

Aubin-Horth N, Landry CR, Letcher BH, Hofmann HA. 2005b. Alternative life histories shape brain gene expression profiles in males of the same population. Proceedings Biological Sciences, 272: 1655-1662.

Bachere E, Gueguen Y, Gonzalez M, de Lorgeril J, Garnier J, Romestand B. 2004. Insights into the anti-microbial defence of marine invertebrates: the penaeid shrimps and the oyster *Crassostrea gigas*. Immunological Reviews 198: 149-168.

Baerwald MR, Herdick RP, May B. 2007. Discovering Genes Associated With Whirling Disease Resistance Using Microarray Analysis. Plant & Animal Genomes XV Conference, January 13-17, 2007, Town & Country Convention Center, San Diego, CA, W30, Abstract

Bao B, Peatman E, Xu P, Baoprasertkul P, Wang G, Liu Z. 2007. Characterization of 23 CC chemokine genes and analysis of their expression in channel catfish (*Ictalurus punctatus*). Plant & Animal Genomes XV Conference, January 13-17, 2007, Town & Country Convention Center, San Diego, CA, P612, Abstract

Baranski M, Loughnan S, Austin CM, Robinson N. 2006. A microsatellite linkage map of the blacklip abalone, *Haliotis rubra*. Animal Genetics 37: 563-570.

Beaumont MA, Nichols RA. 1996 Evaluating loci for use in the genetic analysis of population structure. Proceedings of the Royal Society of London Series B-Biological Sciences 263: 1619-1626

Bonin A, Taberlet P, Miaud C, Pompanon F. 2006. Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (Rana temporaria) Molecular Biology and Evolution 23: 773-783

Bonnet E, Fostier A, Bobe J. 2007. Microarray-based analysis of fish egg quality after natural or controlled ovulation. BMC Genomics 8: 55.

Boudry P., Collet B., Cornette F, Hervouet V, Bonhomme F. 2002. High variance in reproductive success of the Pacific oyster (*Crassostrea gigas*, Thunberg) revealed by microsatellite-based parentage analysis of multifactorial crosses. Aquaculture 204: 283-296.

Boulding EG, Ang KP, Berg PL, Fleming IA, Glebe B, Lien S, MacDonald D, Moen T, Moron P, O'Reilly P, Schulte PM. 2007. Conservation genomics of Atlantic Salmon: towards understanding adaptive trait differences for the management of endangered populations. Plant & Animal Genomes XV Conference, January 13-17, 2007, Town & Country Convention Center, San Diego, CA, W 25, Abstract

Breyne P, Dreesen R, Cannoot B, Rombaut D, Vandepoele K, Rombauts S, Vanderhaeghen R, Inze D, Zabeau M. 2003. Quantitative cDNA-AFLP analysis for genome-wide expression studies Molecular Genetics and Genomics 269: 173-179.

Bridle A, Morrison R, Cupit PM, Cunningham BM, Nowak B. 2006a. Quantitation of immune response gene expression and cellular localisation of interleukin-1β mRNA in Atlantic salmon, *Salmo salar*, affected by amoebic gill disease (AGD). Veterinary Immunology and Immunopathology 114: 121-134.

Bridle AR, Morrison RN, Nowak BF. 2006b. The expression of immune-regulatory genes in rainbow trout, *Oncorhynchus mykiss*, during an amoebic gill disease (AGD) infection. Fish and Shellfish Immunology 20: 346-364.

Campbell NR, Narum SR. 2007. Evaluation Of Single Nucleotide Polymorphisms (SNPs) And Life History Variation In Chinook Salmon. Plant & Animal Genomes XV Conference, January 13-17, 2007, Town & Country Convention Center, San Diego, CA, P622, Abstract

Cnaani A, Zilberman N, Tinman S, Hulata G, Ron M. 2004. Genome-scan analysis for quantitative trait loci in an F-2 tilapia hybrid. Molecular Genetics and Genomics 272: 162-172.

Chini V, Rimoldi S, Terova G, Saroglia M, Rossi F, Bernardini G, Gornati R. 2006. EST-based identification of genes expressed in the liver of adult seabass (*Dicentrarchus labrax*, L.). Gene 376:102-106.

Chistiakov DA, Hellemans B, Haley CS, Law AS, Tsigenopoulos CS, Kotoulas G, Bertotto D, Libertini A, Volckaert FAM. 2005. A microsatellite linkage map of the European sea bass *Dicentrarchus labrax* L. Genetics 170: 1821-1826.

Cossins AR, Crawford DL. 2005. Opinion - Fish as models for environmental genomics Nature Reviews Genetics 6: 324-333.

Crollius HR, Weissenbach J. 2005. Fish genomics and biology. Genome Research 15: 1675-1682.

Cunningham C, Hikima JI, Jenny MJ, Chapman RW, Fang GC, Saski C, Lundqvist ML, Wing RA, Cupit PM, Gross PS, Warr GW, Tomkins JP. 2006. New resources for marine genomics: Bacterial artificial chromosome libraries for the eastern and pacific oysters (*Crassostrea virginica* and *C. gigas*). Marine Biotechnology 8: 521-533.

Danzmann RG, Jackson TR, Ferguson MM. 1999. Epistasis in allelic expression at upper temperature tolerance QTL in rainbow trout. Aquaculture 173: 45-58.

Doyle RW, Herbinger CM. 1994. The use of DNA fingerprinting for high-intensity within-family selection in fish breeding, Proceedings of the 5th World Congress on Genetics Applied to Livestock Production, 7-12 August 1994, Vol. 19, Department of Animal and Poultry Science, University of Guelph, Guelph, pp. 364-371.

Dulvy N.K., Sadovy Y., Reynolds J.D. 2003. Extinction vulnerability in marine populations. Fish and Fisheries 4: 25-64.

Fassehaye Y, Bovenhuis H, Crooijmans R, Komen H. 2007. Microsatellite DNA markers for optimal design of Nile tilapia (*Oreochromis niloticus*) breeding programs. Plant & Animal Genomes XV Conference, January 13-17, 2007, Town & Country Convention Center, San Diego, CA, P627, Abstract

Fast MD, Muise DM, Easy RE, Ross NW, Johnson SC. 2006. The effects of *Lepeophthirius salmonis* infection on the stress response and immunological status of Atlantic salmon (*Salmo salar*). Fish and Shellfish Immunology 21: 228-241.

Fernando RL, Grossman M 1989. Marker-assisted selection using best linear unbiased prediction. Genetics Selection Evolution 33: 209-229.

Franch R, Louro B, Tsalavouta M, Chatziplis D, Tsigenopoulos CS, Sarropoulou E, Antonello J, Magoulas A, Mylonas CC, Babbucci M, Patarnello T, Power DM, Kotoulas G, Bargelloni L. 2006. A genetic linkage map of the hermaphrodite teleost fish *Sparus aurata* L. Genetics 174: 851–861.

Fujiki K, Lubieniecki K, Koop B, Davidson WS. 2007. Construction of a BAC-contig covering three microsatellite loci in the sex-determining region of the Atlantic salmon genome. Plant & Animal Genomes XV Conference, January 13-17, 2007, Town & Country Convention Center, San Diego, CA, P620, Abstract

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

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

Gibson G, Weir B. 2005. The quantitative genetics of transcription. Trends in Genetics 21: 616-623.

Gracey AY., Fraser EJ., Li W, Fang Y, Taylor RR, Rogers J, Brass A, Cossins AR. 2004. Coping with cold: An integrative, multitissue analysis of the transcriptome of a poikilothermic vertebrate. Proceedings of the National Academy of Sciences of the USA 101: 16970–16975.

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

Guyomard R, Mauger S, Tabet-Canale K, Martineau S, Genet C, Krieg F, Quillet E. 2006. A type I and type II microsatellite linkage map of rainbow trout (*Oncorhynchus mykiss*) with presumptive coverage of all chromosome arms. BMC Genomics 7, art. No. 302.

Hayes BJ, Chamberlain AJ, Goddard ME. 2006. Use of markers in linkage disequilibrium with QTL in breeding programs. Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, August 13-18, 2006, Belo Horizonte, MG, Brasil.

Hayes B, Goddard ME. 2004. Break-even cost of genotyping genetic mutations affecting economic traits in Australian pig enterprises. Livestock Production Science 89: 235-242.

Hayes BJ *et al*. 2007. An extensive resource of single nucleotide polymorphism markers associated with Atlantic salmon (*Salmo salar*) expressed sequences. Aquaculture, in press.

Hedgecock D, Gaffney PM, Goulletquer P, Guo XM, Reece K, Warr GW. 2005. The case of sequencing the Pacific oyster genome. Journal of Shellfish Research 24: 429-441.

Hedgecock D, Lin J Z, DeCola S, Haudenschild CD, Meyer E, Manahan DT, Bowen B. 2007. Transcriptomic analysis of growth heterosis in larval Pacific oysters (*Crassostrea gigas*). Proceedings of the National Academy of Sciences of the United States of America 104: 2313-2318.

Hilborn R, Quinn TP, Schindler DE, Rogers DE. 2003. Biocomplexity and fisheries sustainability. Proceedings of the National Academy of Sciences of the USA 100: 6564–6568.

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

Huvet A, Herpin A, Degremont L, Labreuche Y, Samain JF, Cunningham C. 2004. The identification of genes from the oyster *Crassostrea gigas* that are differentially expressed in progeny exhibiting opposed susceptibility to summer mortality. Gene 343: 211-220.

ICES. 2006. Report of the Working Group on the Application of Genetics in Fisheries and Mariculture (WGAGFM), 24–27 March 2006, Newport, Ireland. ICES CM 2006/MCC:04. 59 pp, available from www.ices.dk.

Jackson TR, Ferguson MM, Danzmann RG, Fishback AG, Ihssen PE, O'Connell M, Ccrease TJ. 1998. Identification of two QTL influencing upper temperature tolerance in three rainbow trout (*Oncorhynchus mykiss*) half-sib families. Heredity 80: 143-151.

Jenny MJ, Chapman RW, Mancia A, Chen YA, McKillen DJ, Trent H, Lang P, Escoubas JM, Bachere E, Boulo V, Liu ZJ, Gross PS, Cunningham C, Cupit PM, Tanguy A, Guo X, Moraga D, Boutet I, Huvet A, De Guise S, Almeida JS, Warr GW. A cDNA Microarray for *Crassostrea virginica* and *C. gigas*. Marine Biotechnology (in press).

Jones AG, Ardren WR. 2003. Methods of parentage analysis in natural populations. Molecular Ecology 12: 2511-2523.

Kai W , Kikuchi K, Fujita M, Suetake H, Fujiwara A, Yoshiura Y, Ototake M, Venkatesh B, Miyaki K, Suzuki Y. 2005. A genetic linkage map for the tiger pufferfish, *Takifugu rubripes*. Genetics 171: 227-238.

Karsi A, Turkmut L, Waldbieser GC, Wolters WR. 2007. Determination of microsatellite markers sssociated with disease resistance in channel catfish. Plant & Animal Genomes XV Conference, January 13-17, 2007, Town & Country Convention Center, San Diego, CA, W26, Abstract

Kauer MO., Dieringer D, Schlotterer C. 2003. A microsatellite variability screen for positive selection associated with the "Out of Africa" habitat expansion of *Drosophila melanogaster*. Genetics 165: 1137-1148.

Kause A, Ritola O, Paananen T, Mäntysaari E, Eskelinen U. 2003. Selection against early maturity in farmed rainbow trout: the quantitative genetics of sexual dimorphism and genotype-by-environment interactions. Aquaculture 228: 53-68.

Kause A, Ritola O, Paananen T, Wahlroos H, Mäntysaari EA. 2005. Genetic trends in growth, sexual maturity and skeletal deformations, and rate of inbreeding in a breeding programme for rainbow trout. Aquaculture 247: 177-187.

Kimura T, Jindo T, Narita T, Naruse K, Kobayashi D, Shin-I T, Kitagawa T, Sakaguchi T, Mitani H, Shima A, Kohara Y, Takeda H. 2004. Large-scale isolation of ESTs from medaka embryos and its application to medaka developmental genetics. Mechanisms of Development 121: 915-932.

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

Kocher TD, Lee WJ, Sobolewska H, Penman D, McAndrew B. 1998. A genetic linkage map of a cichlid fish, the tilapia (*Oreochromis niloticus*). Genetics 148: 1225-1232.

Kohn MH, Murphy WJ, Ostrander EA, Wayne R.K. 2006. Genomics and conservation genetics. Trends in Ecology & Evolution 21: 629-637.

Koskinen H, Pehkonen P, Vehniainen E, Krasnov A, Rexroad CE, Afanasyev S, Molsa H, Oikari A. 2004. Response of rainbow trout transcriptome to model chemical contaminants. Biochemical and Biophysical Research Communications 320: 745-53.

Kurobe T, Yasuike M, Kimura T, Hirono I, Aoki T. 2005. Expression profiling of immune-related genes from Japanese flounder *Paralichthys olivaceus* kidney cells using cDNA microarrays. Developmental and Comparative Immunology 29: 515–523.

Kwitkowski CN, Howard LC, McGowan CR, Davidson EA, Mosher A, Ou G, McCarthy J, Phillips RB, Davidson WS. 2007. Identification of the sex-determining chromosomes in the tree river strain of Arctic char (*Salvelinus alpinus*). Plant & Animal Genomes XV Conference, January 13-17, 2007, Town & Country Convention Center, San Diego, CA, P611, Abstract.

Lallias D, Lapègue S, Haley C, Heurtebise S, Boudry P, Beaumont A. 2007. European flat oyster, *Ostrea edulis*: towards the identification of Bonamia resistance QTLs. 8th International Marine Biotechnology Conference. Eilat, Israel, 11-16 March 2007.

Lang P, Langdon CJ, Camara M. 2006. Transcriptome analysis of Pacific oyster *Crassostrea gigas* families exposed to heat stress. Journal of Shellfish Research 25: 746.

Launey S, Barre M, Gerard A, Naciri Graven Y. 2001. Population bottleneck and effective size in Bonamia ostreae-resistant populations of *Ostrea edulis* as inferred by microsatellite markers. Genetical Research 78: 259-270.

Lee BY , Lee WJ, Streelman JT, Carleton KL, Howe AE, Hulata G, Slettan A, Stern JE, Terai Y, Kocher TD. 2005. A second-generation genetic linkage map of tilapia (*Oreochromis* spp.). Genetics 170: 237-244.

Li L, Guo X. 2004. AFLP-based genetic linkage maps of the Pacific oyster *Crassostrea gigas* Thunberg. Marine Biotechnology 6: 26-36.

Li RW, Waldbieser GC. 2006. Production and utilization of a high-density oligonucleotide microarray in channel catfish, *Ictalurus punctatus*. BMC Genomics, 7: article 134.

Lindenstrom T, Sigh J, Dalgaard MB, Buchmann K. 2006. Skin expression of IL 1β in East Atlantic salmon, *Salmo salar* L., highly susceptible to Gyrodactylus salaris infections is enhanced compared to a low susceptibility Baltic stock. Journal of Fish Diseases 29: 123-128.

Lindenstrom T, Buchmann K, Secombes C. 2003. *Gyrodactylus derjavini* infection elicits interleukin 1 beta expression in rainbow trout skin. Fish and Shellfish Immunology 15: 107-115.

Liu ZJ, Cordes JF. 2004. DNA marker technologies and their applications in aquaculture genetics Aquaculture 242: 735-736

Liu X, Liu X, Guo X, Gao Q, Zhao H, Zhang G. 2006. A preliminary genetic linkage map of the pacific abalone *Haliotis discus hannai* Ino. Marine Biotechnology 8: 386-397.

Liu Z, Karsi A, Li P, Cao D, Dunham R. 2003. An AFLP-based genetic linkage map of channel catfish (*Ictalurus punctatus*) constructed by using an interspecific hybrid resource family. Genetics 165: 687-694.

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P. 2003. The power and promise of population genomics: From genotyping to genome typing. Nature Reviews Genetics 4: 981-994.

Martin SA, Blaney SC, Houlihan DF, Secombes CJ. 2006. Transcriptome response following administration of a live bacterial vaccine in Atlantic salmon (*Salmo salar*). Molecular Immunology 43: 2089-2106.

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

Martinez V, Thorgaard G, Robinson B, Sillanpää MJ. 2005. An application of Bayesian QTL mapping to early development in double haploid lines of rainbow trout including environmental effects. Genetical Research 86: 209-221.

Martyniuk CJ, Perry GML, Mogahadam HK, Ferguson MM, Danzmann RG. 2003. The genetic architecture of correlations among growth-related traits and male age at maturation in rainbow trout. Journal of Fish Biology 63: 746-764.

Matsuoka MP, Gharrett AJ, Wilmot RL, Smoker WW. 2004. Genetic linkage mapping of allozyme loci in even- and odd-year pink salmon (*Oncorhynchus gorbuscha*). Journal of Heredity 95: 421-429.

Maynard Smith J, Haigh J. 1974. Hitch-hiking effect of a favorable gene. Genet Res 23: 23-35

Meuwissen THE, Hayes BJ, Goddard ME. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157: 1819-1829.

Miller KM., Li S, Kaukinen KH, Ginther N, Richard J, Traxler G. 2007. Host transcriptional response to infectious hematopoetic necrosis virus (IHNV): A comparison among species with different susceptibilities. Plant & Animal Genomes XV Conference, January 13-17, 2007, Town & Country Convention Center, San Diego, CA, W29, Abstract

Moen T, Hoyheim B, Munck H, Gomez-Raya L 2004a. A linkage map of Atlantic salmon (*Salmo salar*) reveals an uncommonly large difference in recombination rate between the sexes. Animal Genetics 35: 81-92.

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

Moen T , Agresti JJ, Cnaani A, Moses H, Famula TR, Hulata G, Gall GAE, May B. 2004c. A genome scan of a four-way tilapia cross supports the existence of a quantitative trait locus for cold tolerance on linkage group 23. Aquaculture Research 35: 893-904.

Montano-Perez K, Villalpando - Canchola E, Vargas – Albores F. 2006. RFLP (amplified fragment length polymorphism) and its application in aquaculture. Interciencia 31: 563 - 569

Moritz C. 2002. Strategies to protect biological diversity and the evolutionary processes that sustain it. Systematic Biology 51: 238-254

Naruse K, Fukamachi S, Mitani H, Kondo M, Matsuoka T, Kondo S, Hanamura N, Morita Y, Hasegawa K, Nishigaki R, Shimada A, Wada H, Kusakabe T, Suzuki N, Kinoshita M, Kanamori A, Terado T, Kimura H, Nonaka M, Shima A. 2000. A detailed linkage map of Medaka, *Oryzias latipes*: Comparative genomics and genome evolution. Genetics 154: 1773-1784.

Ng SHS, Chang A, Brown GD, Koop BF, Davidson WS. 2005. Type I microsatellite markers from Atlantic salmon (*Salmo salar*) expressed sequence tags. Molecular Ecology Notes 5: 762-766.

Nichols KM, Bartholomew J, Thorgaard GH. 2003. Mapping multiple genetic loci associated with *Ceratomyxa shasta* resistance in *Oncorhynchus mykiss*. Diseases of Aquatic Organisms 56: 145-154.

Nichols KM, Young WP, Danzmann RG, Robison BD, Rexroad C, Noakes M, Phillips RB, Bentzen P, Spies I, Knudsen K, Allendorf FW, Cunningham BM, Brunelli J, Zhang H, Ristow S, Drew R, Brown KH, Wheeler PA, Thorgaard GH. 2003. A consolidated linkage map for rainbow trout (*Oncorhynchus mykiss*). Animal Genetics 34: 102-115.

Nichols KM, Wheeler PA, Thorgaard GH. 2004. Quantitative trait loci analyses for meristic traits in *Oncorhynchus mykiss*. Environmental Biology of Fishes 69: 317-331.

Oleksiak MF, Churchill GA, Crawford DL. 2002. Variation in gene expression within and among natural populations. Nature Genetics 32: 261-266.

Opsahl – Ferstad HG, Rudi H, Ruyter B, Refstie S. 2003. Biotechnological approaches to modify rapeseed oil composition for application in aquaculture. Plant Science 165: 349 – 357.

Ozaki A, Sakamoto T, Khoo S, Nakamura K, Coimbra MRM, Akutsu T, Okamoto N. 2001. Quantitative trait loci (QTLs) associated with resistance/susceptibility to infectious pancreatic necrosis virus (IPNV) in rainbow trout (*Oncorhynchus mykiss*). Molecular Genetics and Genomics 265: 23-31.

Palti Y, Gahr SA, Hansen JD, Rexroad IIICE. 2004. Characterization of a new BAC library for rainbow trout: evidence for multi-locus duplication. Animal Genetics 35: 130-133.

Palti Y, Johnson NA, Finchman MR, Rexroad IIICE, Hovatter KE, Welch TJ, Wiens GD, Silverstein JT, Vallejo RL. 2007. Associations of markers linked to major histocompatibility (MH) regions and to other immune response genes with disease resistance in rainbow trout. Plant & Animal Genomes XV Conference, January 13-17, 2007, Town & Country Convention Center, San Diego, CA, P630, Abstract

Peatman E, Baoprasertkul P, Terhune J, Xu P, Nandi S, Kucuktas H, Li P, Wang S, Somridhivej B, Lamkom T, Lee W, Dunham R, Liu Z. 2007. Global assessment of the acute phase response in catfish after infection with a gram negative bacterium. Plant & Animal Genomes XV Conference, January 13-17, 2007, Town & Country Convention Center, San Diego, CA, P613, Abstract

Perry GML, Danzmann RG, Ferguson MM, Gibson JP. 2001. Quantitative trait loci for upper thermal tolerance in outbred strains of rainbow trout (*Oncorhynchus mykiss*). Heredity 86: 333-341.

Perry GML, Ferguson M, Sakamoto T, Danzmann RG. 2005. Sex-linked quantitative trait loci for thermotolerance and length in the rainbow trout. Journal of Heredity 96: 97-107.

Petersen JL, Ibrarra AM, May BP. 2007. A mass spawn of lion-paw scallop (*Nodipecten subnodosus*): genetic contribution of spawners to the progeny and allelic inheritance. Plant & Animal Genomes XV Conference, January 13-17, 2007, Town & Country Convention Center, San Diego, CA, W21, Abstract

Poompuang S, Na-Nakorn U. 2004. A preliminary genetic map of walking catfish (*Clarias macrocephalus*). Aquaculture 232: 195-203.

Purcell MK, Nichols KM, Winton JR, Kurath G, Thorgaard GH, Wheeler P, Hansen JD, Herwig RP, Park LK. 2006. Comprehensive gene expression profiling following DNA vaccination of rainbow trout against infectious hematopoietic virus. Molecular Immunology 43: 2089-2106.

Quiniou SMA, Katagiri T, Miller NW, Wilson M, Wolters WR, Waldbieser GC. 2003. Construction and characterization of a BAC library from a gynogenetic channel catfish, *Ictalurus punctatus*. Genetics Selection Evolution 35:1–11.

Quinton C, Kause A, Koskela J. 2007. Breeding salmonids for feed efficiency in current fishmeal and future plant-based diet environments. Genetics Selection Evolution, in press.

Reid DP, Szanto Z, Glebe B, Danzmann RG, Ferguson MM. 2005. QTL for body weight and condition factor in Atlantic salmon (*Salmo salar*); comparative analysis with rainbow trout (*Oncorhynchus mykiss*) and Arctic charr (*Salvelinus aplinus*). Heredity 94: 166-172.

Rengmark AH, Slettan A, Skaala O, Lie O, Lingaas F. 2006. Genetic variability in wild and farmed Atlantic salmon (*Salmo salar*) strains estimated by SNP and microsatellites. Aquaculture 253: 229-237

Renn SCP, Aubin-Horth N, Hofmann HA. 2004. Biologically meaningful expression profiling across species using heterologous hybridization to a cDNA microarray. BMC Genomics 5: article no. 42.

Reynolds JD, Dulvy NK, Goodwin NB, Hutchings JA. 2005. Biology of extinction risk in marine fishes. Proceedings of the Royal Society of London B 272: 2337-2344.

Rise ML, von Schalburg KR, Brown GD, Devlin RH, Mawer MA, Kuipers N, Busby M., Beetz-Sargent M, Alberto R, Gibbs AR, Hunt P, Shukin R, Zeznik JA, Nelson C, Jones SMR, Smailus DE, Jones SJM, JE Schein JE, Marra MA, Butterfield YSN, Stott JM, Ng SH, Davidson WS, Koop BF. 2004a. Development and application of a salmonid EST database and cDNA microarray: data mining and interspecific hybridization characteristics. Genome Research 14: 478-490.

Rise ML, Jones SRM, Brown GD, von Schalburg KR, Davidson WS, Koop BF. 2004b. *Piscirickettsia salmonis* infection consistently regulates expression of a suite of genes in cultured Atlantic salmon macrophages and in hematopoietic kidney. Physiological Genomics 20:21-35.

Roberge C, Einum S, Guderley H, Bernatchez L. 2006. Rapid parallel evolutionary changes of gene transcription profiles in farmed Atlantic salmon. Molecular Ecology15: 9-20.

Robison BD, Wheeler PA, Sundin K, Sikka P, Thorgaard GH. 2001. Composite interval mapping reveals a major locus influencing embryonic development rate in rainbow trout (*Oncorhynchus mykiss*). Journal of Heredity 92: 16-22.

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

Saavedra C., Bachere E. 2006. Bivalve genomics. Aquaculture 256: 1 - 14

Sakamoto T, Danzmann RG, Gharbi K, Howard P, Ozaki A, Khoo SK, Woram RA, Okamoto N, Ferguson MM, Holm LE, Guyomard R, Hoyheim B. 2000. A microsatellite linkage map of rainbow trout (*Oncorhynchus mykiss*) characterized by large sex-specific differences in recombination rates. Genetics 155: 1331-1345.

Sakamoto T, Danzmann RG, Okamoto N, Ferguson MM, Ihssen PE. 1999. Linkage analysis of quantitative trait loci associated with spawning time in rainbow trout (*Oncorhynchus mykiss*). Aquaculture 173: 33-43.

Salem M, Kennedy PB, Rexroad CE, Yao JB. 2006. Microarray gene expression analysis in atrophying rainbow trout muscle: a unique non-mammalian muscle degradation model. Physiological Genomics 28: 33-45.

Sarropoulou E, Power DM, Magoulas A, Geisler R, Kotoulas G. 2005a. Comparative analysis and characterization of expressed sequence tags in gilthead sea bream (*Sparus aurata*) liver and embryos. Aquacultutre 243: 69 – 81.

Sarropoulou E, Kotoulas G, Power DM, Geisler R. 2005b. Gene expression profiling of gilthead sea bream during early development and detection of stress – related genes be application of cDNA microarray technology. Physiological Genomics 23: 182 - 191

Savan R, Sakai M. 2006. Genomics of fish cytokines. Comp. Biochem. Physiol. D – Genomics & Proteonomics, 1: 89 - 101

Schlötterer 2002. Towards a molecular characterization of adaptation in local populations. Current Opinion in Genetics & Develelopment 12: 683-687.

Senger F, Priat C, Hitte C, Sarropoulou E, Franch R, Geisler R, Bargelloni L, Power D, Galibert F. 2006. The first radiation hybrid map of a perch – like fish: The gilthead seabream (*Sparus aurata* L.). Genomics 87: 793 – 800.

Sigh J, Lindenstrom T, Buchmann K. 2004. Expression of pro-inflammatory cytokines in rainbow trout (*Oncorhynchus mykiss*) during an infection with *Ichthyophthirius multifiliis*. Fish and Shellfish Immunology 17:75-86

Silverstein JT, Weber GM, Rexroad CE, Vallejo RL. 2006. Genetics and genomics – Integration of molecular genetics into a breeding program for rainbow trout. Israeli J. of Aquaculture – Bamidgeh 58: 231 - 237

Sironen A, Thomsen B, Andersson M, Ahola V, Vilkki J. 2006. An intronic insertion in KPL2 results in aberrant splicing and causes the immotile short-tail sperm defect in the pig. Proceedings of the National Academy of Sciences of the United States of America 103: 5006-5011.

Somorjai IM, Danzmann RG, Ferguson MM. 2003. Distribution of temperature tolerance quantitative trait loci in Arctic charr (*Salvelinus alpinus*) and inferred homologies in rainbow trout (*Oncorhynchus mykiss*). Genetics 165: 1443-1456.

Sonesson AK. 2005. A combination of walk-back and optimum contribution selection in fish. Genetics Selection Evolution 37: 587-599.

Soule M, LaFrentz S, Cain K, LaPatra S, Call DR. 2005. Polymorphism in 16S rRNA genes from *Flavobacterium psychrophilum* correlate with elastin hydrolysis and tetracycline resistance. Diseases of Aquatic Organisms 65: 209 – 216.

Storz JF. 2005. Using genome scans of DNA poplymorphism to infer adaptive population divergence. Molecular Ecology 14: 671-688.

Storz JF, Dubach JM. 2004. Natural selection drives altitudinal divergence at the albumin locus in deer mice, *Peromyscus maniculatus*. Evolution 58: 1342–1352.

Sun X, Liang L. 2004. A genetic linkage map of common carp (*Cyprinus carpio* L.) and mapping of a locus associated with cold tolerance. Aquaculture 238: 165-172.

Sundin K, Brown KH, Drew RE, Nichols KM, Wheeler PA, Thorgaard GH. 2005. Genetic analysis of a development rate QTL in backcrosses of clonal rainbow trout, *Oncorhynchus mykiss*. Aquaculture 247: 75-83.

Tao WJ, Boulding EG. 2003. Associations between single nucleotide polymorphisms in candidate genes and growth rate in Arctic charr (*Salvelinus alpinus* L.). Heredity 91: 60-69.

Taris N, Baron S, Sharbel TF, Sauvage C, Boudry P. 2005. A combined microsatellite multiplexing and boiling DNA extraction method for high throughput parentage analyses in the Pacific Oyster (*Crassostrea gigas*). Aquaculture Research 36: 516-518.

Taris N, Ernande B, Boudry P. 2006. Phenotypic and genetic consequences of size selection at the larval stage in the Pacific oyster (*Crassostrea gigas*). Journal of Experimental Marine Biology and Ecology 333: 147-158.

Thorsen J, Zhu B, Frengen E, Osoegawa K, de Jong PJ, Koop BF, Davidson WS, Hoyheim B. 2005. A highly redundant BAC library of Atlantic salmon (*Salmo salar*): an important tool for salmon projects. BMC Genomics 6: 50

Tilton SC, Givan SA, Pereira CB, Bailey GS, Williams DE. 2006. Toxicogenomic profiling of the hepatic tumor promoters indole-3-carbinol, 17 β-estradiol and β-naphthoflavone in rainbow trout. Toxicological Sciences 90: 61-72.

Ton C. Stamatiou D, Dzau VJ, Liew CC. 2002. Construction of a zebrafish cDNA microarray: gene expression profiling of the zebrafish. Biochemical and Biophysical Research Communications 296: 1134-1142.

Torto-Alalibo T, Tian MY, Gajendran K, Waugh ME, van West P, Kamoun S. 2005. Expressed sequence tags from the oomycete fish pathogen *Saprolegnia parasitica* reveal putative virulence factors. BMC Microbiology, 5: Art. No. 46.

Vandeputte M, Mauger S, Dupont-Nivet M. 2006. An evaluation of allowing for mismatches as a way to manage genotyping errors in parentage assignment by exclusion. Molecular Ecology Notes 6: 265-267.

Vandeputte M , Kocour M, Mauger S, Dupont-Nivet M, De Guerry D, Gela D, Vallod D, Linhart O, Chevassus B. 2005. Heritability estimates for growth-related traits using microsatellite parentage assignment in juvenile common carp (*Cyprinus carpio* L.). Aquaculture 247: 31-32.

van Tienderen PH, de Haan AA, van der Linden CG, Vosman B. 2002. Biodiversity assessment using markers for ecologically important traits. Trends in Ecology & Evolution 17: 577-582.

Vasemägi A, Nilsson J, Primmer CR. 2005. Expressed sequence tag-linked microsatellites as a source of gene-associated polymorphisms for detecting signatures of divergent selection in Atlantic salmon (*Salmo salar* L.). Molecular Biology and Evolution 22: 1067-1076.

Venier P, De Pitta C, Pallavicini A, Marsano F, Varotto L, Romualdi C, Dondero F, Viarengo A, Lanfranchi G. 2006. Development of mussel mRNA profiling: Can gene expression trends reveal coastal water pollution? Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis 602: 121-134.

Venkatesh B, Hettwer U, Koopmann B, Karlovsky P. 2005. Conversion of cDNA differential display results (DDRT-PCR) into quantitative transcription profiles. BMC Genomics 6: Art. No. 51.

von Schalburg KR, Rise ML, Cooper GA, Brown GD, Gibbs RA, Nelson CC, Davidson WS, Koop BF. 2005a. Fish and Chips: Various methodologies demonstrate utility of a 16,006-gene salmonid microarray. BMC Genomics 6 Art. No. 126

von Schalburg KR, Rise ML, Brown GD, Davidson WS, Koop BF. 2005b. A comprehensive survey of the genes involved in maturation and development of the rainbow trout ovary. Biology of Reproduction 72: 687-699.

von Schalburg KR, McCarthy SP, Rise ML, Hutson JC, Davidson WS, Koop BF. 2006. Expression of morphogenic genes in mature ovarian and testicular tissues: potential stem-cell niche markers and patterning factors. Molecular Reproductive Development 73: 142-152.

Walter RB , Rains JD, Russell JE, Guerra TM, Daniels C, Johnston DA, Kumar J, Wheeler A, Kelnar K, Khanolkar VA, Williams EL, Hornecker JL, Hollek L, Mamerow MM, Pedroza A, Kazianis S. 2004. A microsatellite genetic linkage map for *Xiphophorus*. Genetics 168: 363-372.

Wang CM , Lo LC, Zhu ZY, Yue GH. 2006. A genome scan for quantitative trait loci affecting growth-related traits in an F1 family of Asian seabass (*Lates calcarifer*). BMC Genomics 7 Art. No. 274.

Watanabe T, Fujita H, Yamasaki K, Seki S, Taniguchi N. 2004. Preliminary study on linkage mapping based on microsatellite DNA and AFLP markers using homozygous clonal fish in Ayu (*Plecoglossus altivelis*). Marine Biotechnology 6: 327-334.

Westenberg M, Heinhuis B, Zuidema D, Vlak JA. 2005. SiRNA injection induces sequence-independent protection in *Penaeus monodon* against white spot syndrome virus. Virus Research 114: 133-139.

Whitaker HA, McAndrew BJ, Taggart JB. (2006) Construction and characterisation of a BAC library for the European sea bass *Dicentrarchus labrax*. Anim Genet 37:526

Whitehead A, Crawford DL. 2006. Neutral and adaptive variation in gene expression. Proceedings of the National Academy of Sciences of the USA 103: 5425-5430.

Williams TD, Gensberg K, Minchin SD, Chipman JK 2003. A DNA expression array to detect toxic stress response in European flounder (*Platichthys flesus*). Aquatic Toxicology 65: 141-157.

Wilson AJ, Ferguson MM. 2002. Molecular pedigree analysis in natural populations of fishes: approaches, applications, and practical considerations. Canadian Journal of Fisheries and Aquatic Sciences 59: 1696-1707.

Woram RA, Gharbi K, Sakamoto T, Hoyheim B, Holm LE, Naish K, McGowan C, Ferguson MM, Phillips RB, Stein J, Guymard R, Cairney M, Taggart JB, Powell R, Davidson WS, Danzmann RG. 2003. Comparative genome analysis of the primary sex-determining locus in salmonid fishes. Genome Research. 13: 272-280.

Woram RA , McGowan C, Stout JA, Gharbi K, Ferguson MM, Hoyheim B, Davidson EA, Davidson WS, Rexroad C, Danzmann RG. 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 47: 304-315.

Wright JJ, Lubieniecki KP, Ng SHS, Leong J, Koop BF, Davidson WS. 2007. Preliminary data of mapping single nucleotide polymorphisms sn Atlantic salmon (*Salmo salar*). Plant & Animal Genomes XV Conference, January 13-17, Town & Country Convention Center, San Diego, CA, P618, Abstract

Yazawa R, Cooper GA, Beetz-Sargent M, Robb A, Leong J, Davidson WS, Koop BF. 2007. High diversity for antigen recognition in the Atlantic salmon T-cell receptor alpha / delta locus. Plant & Animal Genomes XV Conference, January 13-17, 2007, Town & Country Convention Center, San Diego, CA, P621, Abstract

Young WP, Wheeler PA, Coryell VH, Keim P, Thorgaard GH. 1998. A detailed linkage map of rainbow trout produced using doubled haploids. Genetics 148: 839-850.

Yu Z, Guo X. 2006. Identification and mapping of disease-resistance QTLs in the eastern oyster, *Crassostrea virginica* Gmelin. Aquaculture 254: 160-170.

Zimmerman AM , Evenhuis JP, Thorgaard GH, Ristow SS. 2004. A single major chromosomal region controls natural killer cell-like activity in rainbow trout. Immunogenetics 55: 825-835.

Zimmerman AM, Wheeler PA, Ristow SS, Thorgaard GH. 2005. Composite interval mapping reveals three QTL associated with pyloric caeca number in rainbow trout, *Oncorhynchus mykiss*. Aquaculture 247: 85-95.

4 ToR (c): To identify and provide recommendations for the optimal extraction and storage of DNA from fish for molecular based studies

Martha O' Sullivan, Mary Cross, Geir Dahle and Elizabeth Gosling

4.1 Introduction

Insight into the impact of harvesting and climate change on a species biodiversity is vital for the sustainable conservation and management of commercially important animals. Understanding these impacts relies on analyzing some sort of time series data. Fisheries molecular biologists interested in studying temporal changes in genetic diversity in fish and shellfish stocks are extremely fortunate, as viable sources of time-series genetic data for many species can be found in fisheries institutes and universities throughout the world. In the majority of cases these 'sources of genetic information' take the form of scales or otoliths collected for aging and subsequently stored in cupboards or warehouses. In addition, naturally preserved tissues (e.g. ancient salmon, *Salmo salar* scales found in caves Consuegra *et al*., 2005) and formalin preserved samples (Chakraborty *et al*., 2006) found in natural history museums have in many cases been found to be viable sources of DNA.

The potential of such sources of genetic information has been acknowledged for many years, but it is only with relatively recent advancements in methods for DNA extraction, and subsequent analysis of DNA variation, that we have been able to exploit these sources. Some of these advancements have included relatively cheap and rapid DNA extraction methods that produce an impure DNA extract that can successfully be screened for variation using different types of molecular marker, thus cutting down on analysis time and expense. Though these quick extraction methods are obviously of huge benefit for large population studies they do have their drawbacks, such as reduced quality and quantity of the DNA extracted and shortened 'shelf life'. In general, under optimal storage conditions the purer the DNA extract, the longer it will last. Therefore it is vital in instances where there is only enough tissue for a single DNA extraction from a valuable source (as defined below), that a method be chosen that gives the highest possible yield and quality and in turn the longest possible life span. Therefore to prevent the inadvertent loss of information from these 'valuable' samples it is vital that the tissue is stored under optimal conditions, and that an extraction method be used that results in the highest possible quality end-product.

This ToR aims to emphasize the importance of correct tissue storage for both historical and contemporary samples, as well as to stress the importance of choosing the most suitable DNA extraction method for 'valuable' samples. This will allow for maximum 'information yield' from these samples. The Tor deals primarily with DNA as opposed to RNA or proteins. This is because when dealing with older samples, unless the tissue has been stored under very specific conditions it is not possible to analyze RNA or protein markers. However, we recommend that in future sampling programs where 'valuable' tissues are being collected that tissue is stored in such a way as to allow for analysis of RNA and proteins (see later).

4.2 Tissue storage

Once the mechanism for DNA repair in a living organism is terminated at death, DNA quickly deteriorates. However, this process can be slowed down, or even stopped, if a suitable preservation method is used. Therefore, to undertake any molecular genetic study, the source material must be preserved to facilitate the analysis of targeted molecules, e.g. DNA, RNA or proteins.

The storage of tissue for DNA analysis in 96-100% alcohol is common practice in many molecular genetic laboratories today. However, storage vials and ratio of tissue to alcohol is

still not standardised in many sampling programs. It is important that tissue is stored in well sealed vials with a tissue to alcohol ratio of approximately 1 to 10. Storage in 2ml screw cap eppendorf tubes with an O-ring (seal) is regarded as the most successful storage method, and should ideally be used in all cases of wet tissue sampled for DNA extraction. These samples should then be stored at a constant temperature between 4 and 20oC to ensure maximum yields of high quality DNA (G. Dahle *pers com.* 2007).

Samples of tissue stored at –800C can still be found in many laboratories. Although this tissue can yield high quality DNA, this method of preservation does have its disadvantages. Not least are the requirements for space and the cost of running the freezers. There is also the danger of freezer malfunction with consequent loss of samples. Therefore, a better strategy is to transfer tissue to vials containing alcohol, as described above.

Shorter strands of DNA can be successfully extracted from dried scales and otoliths that are stored in paper envelopes. Such sources of genetic material are widespread with many institutions holding years of scale and otolith samples from decades of sampling in storage. However, optimal storage conditions are not well worked out. At present these collections tend to be generally kept in laboratory cupboards or warehouses. The degeneration rate of the DNA in this material is unknown, and before this source of genetic information is lost forever it is important that optimal conditions for storage e.g., temperature and humidity, are determined.

It should be noted that tissue stored in ethanol cannot be used for RNA or protein analysis with storage in liquid nitrogen being regarded as the best method of tissue preservation if these molecules are to be analysed. As this is not always possible an alternative method for RNA work is to add an RNA stabilizing solution to the tissue, while in the case of proteins storage at -800C is generally regarded as suitable.

Smith & Burgoyne (2004) recommend a method of storage where tissue is stored on filter papers that contain chemicals that lyse cells, denature proteins and protect nucleic acids from nucleases, oxidation and UV damage (eg Watman FTA cards) They successfully extracted and amplified DNA from the filter papers following storage for four years at a room temperature which fluctuated between 18oC and 42oC. Livia *et al*. (2006) used filter papers to store buccal and body mucus from brown trout (*Salmo trutta*) and northern pike (*Esox lucius*), from which they successfully extracted DNA seven months later and carried out both mtDNA RFLP and nDNA microsatellite analysis. However, it is still unclear how successful this method of tissue storage is over a longer time scale, and for now should be avoided in the case of valuable samples.

4.3 Extraction methods

Several methods exist for the successful extraction of DNA from fish. The earlier successful methods of extraction followed a phenol-chloroform based method (Taggart *et al*., 1992) with various modifications, depending on the source tissue (Nielsen *et al*., 1999). This results in a relatively high quality, easily quantified, DNA extract. When these extracts are stored at -20^{o}C or -80^{o}C many laboratories have been able to go back and reuse the extract for analysis many years later (T. Cross *pers com* 2007).

As well as the standard phenol–chloroform extraction method, and modifications of it, there are now several kits on the market: Qiagen DNeasy™ (Qiagen), Illustra ™Tissue DNA extraction kit (GE Healthcare) Picopure™ (Arcturus), for the extraction of high quality DNA. These tend to be relatively simple, non-toxic methods which produce high yields of intact high quality DNA.

With the introduction of the Polymerase Chain Reaction (PCR) the importance of having a highly purified DNA extract for some types of DNA analyses decreased. Fragments of

targeted DNA can easily be amplified from a relatively crude extract, and as a result it has become common practice to use such quick extraction methods in many studies involving large samples sizes. The advantages are clear in that these methods tend to be much cheaper, less laborious and less time consuming. However, it is unknown as to how long these crude extracts can be stored. Some laboratories have found it necessary to go back and re-extract DNA from tissue within a few months of storing the original extract at -4oC (M. Cross *pers com* 2007), while others have been able to use the same extract over 10 to 12 years later (E. Verspoor *pers com* 2007). Because of the uncertainty of how long these extracts will last it is recommended that these protocols are avoided when dealing with valuable samples.

4.4 Valuable samples

Recent advances in molecular techniques offer potential for the extraction of DNA from a diverse range of material, previously considered intractable (see DeSalle & Bonwich 1996 for a review) with biochemical methods moving ahead to such an extent that DNA sequences can now be retrieved from ancient sources. This means that DNA can be amplified and sequenced from individuals representing extinct populations, or from individuals representing entire species that are now extinct. In many cases these tissue sources may be considered unique. However, such ancient tissue sources are not the only DNA sources that should be treated with care. The term 'unique' may also be applied to tissue samples collected several decades ago from populations that are still being sampled today. The term may also be applied to samples taken from 'hard to sample areas'. In many cases only a very small piece of tissue (e.g. a single scale sample) is available, and it is important that we make the best possible use of the genetic information that the sample could potentially provide.

In cases where there is only enough tissue for one extraction from these valuable samples extraction protocols should be chosen which give a high yield high quality DNA extract. This should result in a DNA extract that lasts over time if stored at -80^{o}C.

4.5 Conclusion

Although we can be fairly certain that tissue stored in alcohol, as described above, and the use of extraction methods that yield a purified DNA extract, will give us an extended time frame over which we can analyse samples, there are still a number of questions that remain to be answered if we are to make maximum use of samples available to us. Top priority for the WGAGFM is to determine the optimal storage conditions for worldwide repositories of scales and otoliths, in order to minimise DNA degradation. At present the majority of these samples are kept in laboratory cupboards in uncontrolled conditions with large variations in factors such as temperature and humidity. Another question concerns the longevity of DNA extracts. Experience has taught us that the more exacting extraction methods result in a higher quality DNA extract, but how long this product will last when stored at -80oC is unknown. Whether the DNA obtained by different pure extraction methods varies in terms of extended shelf life is also unknown.

There is also the question as to which method, storage of the original tissue, or storage of the actual DNA extract, gives the longest lasting source of genetic information. From the point of storage cost and space, keeping tissue in alcohol is more manageable but we cannot be certain as to just how long such tissue will yield viable DNA. Similarly, we are uncertain as to how long we can store pure DNA extracts such that the DNA is still informative.

Finally, a large amount of time and expense goes into the collecting and extracting of DNA for molecular studies. Members of the WGAGFM feel that it is extremely important that the potential availability of these samples be made known to other laboratories when a study finishes and particularly in cases where the material is being disposed of. This is vital where valuable samples are concerned.

4.6 Recommendations

1) A working group be established to investigate the best methods for DNA extraction from tissue stored by various methods, together with optimum storage methods of the extracted product to ensure its longevity.

2) Otoliths and scales should for now continue to be stored in paper envelopes in dry conditions, with mollusc shells being stored in dry conditions. A working group should be set up as soon as possible to determine best conditions for storage. This is a matter of urgency if we are to preserve the genetic information contained in the thousands of samples now in storage.

3) Tissue should be stored in preference to DNA when enough tissue is available

4) Fresh tissues for DNA analysis should be stored in ethanol as outlined in text. Freezer -stored specimens should be removed and stored in ethanol.*

5) When sampling rare, or difficult to access, samples for DNA work, tissue should be also stored for RNA and protein analysis.

6) Extraction methods for unique samples should be chosen to give a product with highest possible yield, quality and longevity. Until such a method is agreed upon (following Recommendation 1) a technique should be chosen that is known to give a high quality, purified extract, in order to minimise its degradation over time.

7) Prior to disposal of tissue or DNA samples, all relevant institutions should be made aware of the availability of samples. This would be made easier if a meta data base was operational. Consideration should be given to setting up a working group to look into the feasibility of setting up institutionally-based repositories.

*This method is not suitable for storage of tissue for RNA or protein analysis.

4.7 References

Chakraborty, A., Sakai, M., and Iwatsuki, Y. 2006. Museum fish specimens and molecular taxonomy: A comparative study on DNA extraction protocols and preservation techniques. Journal of Applied Ichthyology 22 (2): 160-166

Consuegra, S., Verspoor, E., Knox, D., and de Leaniz, C.G. 2005. Asymmetric gene flow and the evolutionary maintenance of genetic diversity in small, peripheral Atlantic salmon populations . Conservation Genetics 6 (5): 823-842

DeSalle, R., and Bonwich, E., 1996. DNA isolation, manipulation and characterization from old tissues. Genetic Engineering (ed. Setlow J), pp.13-22. Vol. 18. Plenum, New York.

Livia, L. Antonella, P., Hovirag, L.,Mauro, N., and Panara, F. 2006. A nondestructive, rapid, reliable and inexpensive method to sample, store and extract high-quality DNA from fish body mucus and buccal cells. Molecular Ecology Notes 6 (1): 257-260

Nielsen, E.E., Hansen, M.M., and Loeschcke, V. 1999. Analysis of DNA from old scale samples: technical aspects, applications and perspectives for conservation. Hereditas 130 (3): 265-276 1999

Smith, L.M., and Burgoyne, L. A. 2004. Collecting, archiving and processing DNA from wildlife samples using FTA® databasing paper. BMC Ecology, 4, 1-11.

Taggart, J.B., Hynes, R.A., Prodhol, P.A. and Ferguson, A. 1992. A simplified protocol for routine total DNA isolation from salmonid fishes. Journal of Fish Biology 40: 963-965

5 ToR (d): To assess, through a case study with anadromous salmonids, the potential of genetic and spatial data analysis methods for resolving spatial boundaries of finfish and shellfish populations, and for gaining insight into the geographic and ecological factors controlling the development of population boundaries.

Tom Cross, Elizabeth Gosling, Ellen Kenchington, Eric Verspoor, Dorte Bekkevold, Geir Dahle, Eileen Dillane, Torild Johansen, Philip McGinnity

Novel methods are now available to combine molecular population genetic results with geographical features (often referred to as landscape genetics). Under this TOR these methods are described and illustrated with examples from finfish and shellfish species. A genetic population can be defined as a group of freely interbreeding individuals within a species, which is genetically differentiated from other such groups. For example, in anadromous Atlantic salmon (*Salmo salar*) populations are highly differentiated during the freshwater phase, often occurring in different areas within large river systems (Primer *et al*., 2006). On the other hand, fully marine species such as cod have a much less defined population structure, with the same population apparently often occurring over large geographic areas.

Within any fish or shellfish population, spatial extent and position ("population footprint") varies according to the life cycle stage. As for fisheries and aquaculture population genetics in general, a particular stage and time (Bekkevold *et al*. 2005) in the life cycle, where populations are most likely to be discrete (based on the biology of the species), will be selected for analysis of landscape genetics. This is usually the spawning stage in finfish species (but see Appendix 3 with reference to cod spawning off the Lofoten Islands). With most mollusks and crustacean species, this pattern may not apply. For example, in bivalve mollusks the sedentary adult stage will be chosen, but sexual maturity is not relevant in the context of sampling. What is important is not to choose the motile larval stage. Therefore, the biology of each individual species needs to be considered separately.

Taking account of all available fisheries, geographic and hydrographic data (Grant & Waples, 2000), areas for sampling are selected. Appropriate samples are then collected, depending on the type of genetic marker being used. Samples are screened for these genetic markers (generally 10 or more microsatellite loci, at present) and the results analysed using appropriate population genetic software. To consider the genetic results in a geographic context, several landscape genetic programmes are available.

5.1 Landscape genetic analyses

Geographical latitude and longitude co-ordinates or other x-y plane positioning data (e.g. position in a grid) can be used along with genetic data to visualize the genetic landscape. This relatively new field has been recently reviewed by Storfer *et al*. (2006), who provide an overview of the literature, as well as a summary of analytical tools and associated sampling strategies. Here we focus on applications used in the recent marine literature.

BARRIER (F_{ST} BASED, Appendix 2; Manni *et al*., 2004) and AIS (IA based, Appendix 1; Miller, 2005) are just two examples of software packages which use Monmonier's Maximum Difference algorithm (Monmonier, 1973) to identify putative genetic barriers across a connectivity network between sampling locations, based on Delaunay triangulation (Watson, 1992). The triangulated network is the simplest and most common tessellation technique for the creation of surfaces. Contiguous, non-overlapping triangles are created by linear interpolation of the variable.

Monmonier's algorithm identifies neighbouring sample-pairs associated with the highest rate of change in a given distance measure, which in this case is a measure of genetic distance. BARRIER does not calculate the genetic distance matrix, which must be separately generated and input into the programme, along with the corresponding geographic distance matrix. Usually the genetic distance matrix is an F_{ST} distance matrix between populations. AIS software generates a connectivity network internally using Nei's genetic distance (Nei *et al*., 1983) applied to pairs of individuals (Manel *et al*., 2003), instead of pairs of populations, to produce the genetic matrix. This subtle difference is important in that AIS does not require *a priori* assumptions of populations, as is needed for F_{ST} calculations (see Appendix 1). If precise information on individuals is not available, e.g. when numerous individuals are collected in a single tow, AIS allows for multiple observations from the same spatial co-ordinate and average between locality inter-individual genetic distances are used to infer putative barriers.

Monmonier's algorithm identifies, as the initial barrier segment, the greatest genetic distance between any two localities joined in the connectivity network. Additional segments are added until an external edge or a previously defined internal barrier segment is encountered (see Miller 2005 for further details). When genetic distances are equal, both directions are followed, resulting in bi- or multi-furcating barriers. When the initial barrier is halted in one direction, the opposite direction is followed along the same steps. BARRIER allows for testing the significance of the barriers using a bootstrap matrices analysis, which is not implemented in AIS. For the latter the putative barriers can be tested *a posteriori* for statistical differences between them, using any set of genetic or other variables. Additionally, a hierarchical multi-locus AMOVA can be used to partition the total variance into variance components distributed among regions defined by the putative barriers, among populations within regions and within populations, using programs such as ARLEQUIN (Schneider *et al*., 2000).

Graphical representations of genetic distance patterns can be visualized using AIS. Three-dimensional surface plots (see Appendix 1) examine the whole range of genetic variation, rather than focusing only on areas of low gene flow. The method uses the same connectivity network and triangulation procedures discussed above (X-Y plane), and assigns a calculated inter-individual or inter-sample genetic distance in the Z-axis. A simple inverse-distance-weighting interpolation procedure (Watson, 1992), commonly used in stock assessments (e.g., Simard *et al*., 1992) is used to infer genetic distances between samples. The programme generates a graphical representation of the genetic landscape, which can be rotated to view from different angles.

Genetic landscape surfaces can then be input into any number of geographic information systems (GIS), which allow different data layers to be overlaid on each other to visually evaluate the spatial distribution of features. In GIS software packages (e.g. ArcGIS; http://www.esri.com), different types of geographic features (e.g. bathymetry, substrate, temperature contours and historical features, such as position of ice during glaciations) can be stored as different map layers. Organizing data as spatially referenced layers provides flexibility to select and combine layers in various ways, in order to evaluate which features tend to co-localize or to explain the presence of another feature, as a consequence of its spatial relationship. Then, further analyses can be performed to test the explanatory power of the selected variables.

Once the populations have been differentiated, the space occupied by each at the spawning stage can be investigated using survey and commercial fisheries data, and using geographical information such as spawning habitat heterogeneity. In addition, for fully marine species, information on the hydrography of the region in question must be considered.

The anadromous Atlantic salmon, *Salmo salar*, is a good model to demonstrate the effectiveness of the landscape genetics approach for two reasons. Firstly, during the

freshwater phase of the lifecycle populations are highly differentiated even within large river systems, probably because of accurate natal homing and small size of spawning population, with the latter accentuating genetic drift effects. Secondly, a great deal of geographical data has been accumulated on most rivers occupied by salmon, and detailed GIS plots are often available. Salmon require suitable spawning habitat, which is often not homogeneously distributed within a river. Further, since salmon parr tend to remain adjacent to spawning redds at least during their first year of life, then sampling can be informed by geographical knowledge.

An example of such an approach is an investigation of population structure of salmon in the river Moy and adjacent smaller rivers in north western Ireland (Dillane, McGinnity and Cross- in preparation). The Moy is one of the bigger salmon rivers in western Europe and the native populations are largely unaffected by interactions with reared salmon (as either farm escapes or as the result of stocking). Spawning areas (Figure 1) were identified by direct observation during the spawning period and 0+ and 1+ (n=50 of each) parr were subsequently sampled from each area by electrofishing. Samples were screened for 12 microsatellite loci and standard population genetic analyses, both group based (e.g. F_{ST}) and individual based, were undertaken using the resulting genotypic data. No temporal variability was observed within most sampling areas, so pooled year-classes were used in subsequent analyses. Significant pairwise differences were observed between all spatially distinct samples but these populations fell into several more distinct groupings (shown in different colours in the PCA plot in Figure 2).

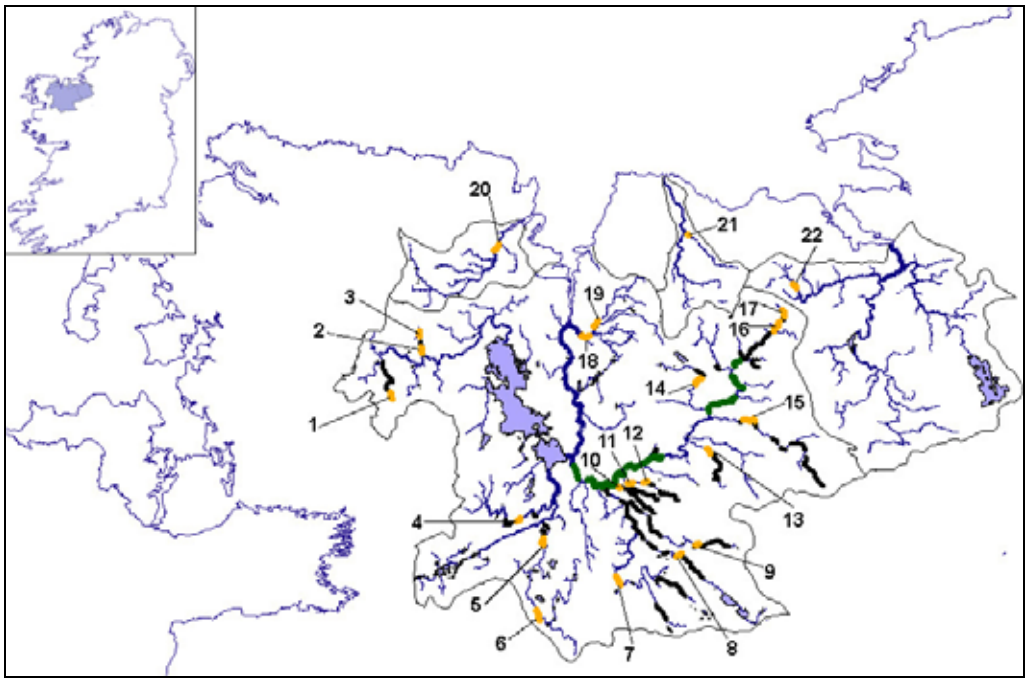


Figure 1. The Moy, Bunrea, Palmerstown, Easkey and Ballysadare catchments. Sampling areas are shown in yellow. Adjacent spawning zones are shown in black. The green area represents parts of the main Moy channel which consisted of an extensive spawning area prior to drainage work carried out in the 1960s. (1 Glendavolagh; 2 Lower Shanvolahan; 3 Upper Shanvolahan ; 4 Clydagh; 5 Lower Manulla; 6 Upper Manulla; 7 Pollagh; 8 Glore; 9 Trimoge; 10 Killeen; 11 Upper Spaddagh; 12 Lower Spaddagh; 13 Sonnagh; 14 Eighnagh; 15 Owengarve; 16 Lower Cloonacool; 17 Upper Cloonacool; 18 Lower Bunrea; 19 Upper Bunrea; 20 Palmerstown; 21 Easkey; 22 Ballysadare)

BARRIER analysis was then undertaken on the spatial samples, resulting in the distinctions between groupings as superimposed on a map of the system in Figure 2. Work is currently underway to further investigate these barriers using various geographical techniques, such as GIS. In applying similar analyses to fully marine species there are greater uncertainties, since

population structure is usually far less defined and less is known about the geographical parameters of marine areas. None-the-less, the approach can be very useful as demonstrated with scallops and herring (Appendix 1 and 2 respectively).

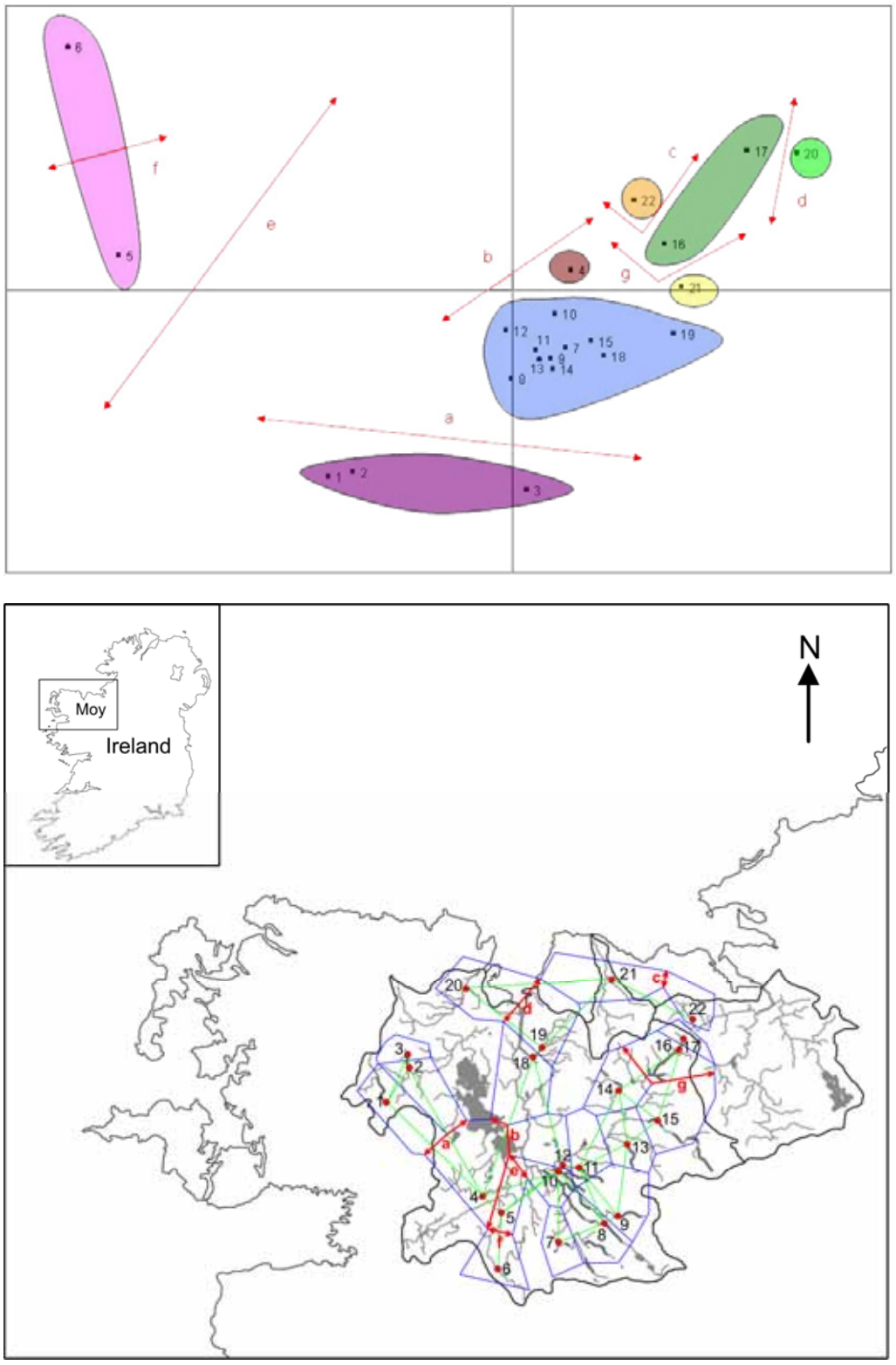


Figure 2. Above: a principal components plot, showing how populations cluster in two dimensions, with barriers inserted. Below: BARRIER output; red arrowed lines show the main barriers separating populations in the Moy and surrounding catchments (a, b, c etc. refer to the first, second and subsequent barriers, in order of importance).

5.2 Populations at other life stages

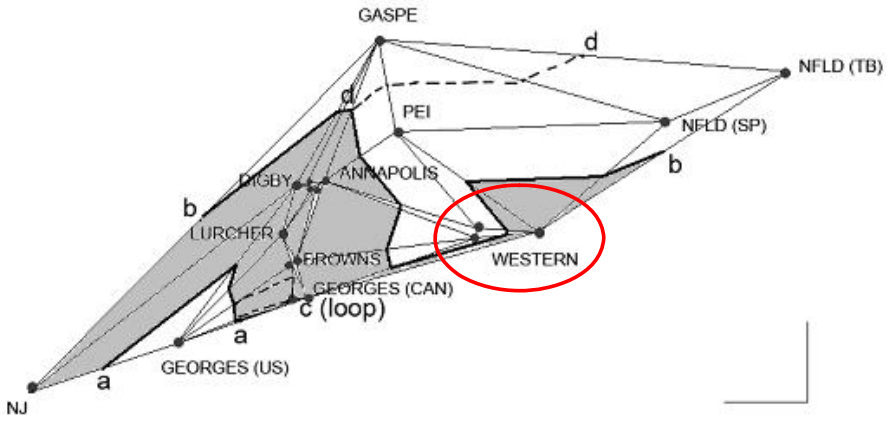
At other (usually later) stages of the life cycle individuals from a population may occupy much larger areas than the spawning area and mixing may occur between different populations, so mixed stock analysis (MSA) and individual assignment (IA) is required where the learner sample consists of the constituent populations as defined above (see TOR f, also Appendixes 2 and 3). A similar situation may occur in the planktonic stage of mollusk and crustacean species, unless mechanisms that retain larvae are operating.

5.3 Appendix 1 *Placopecten magellanicus* (Sea Scallop)

Contributed by E. Kenchington

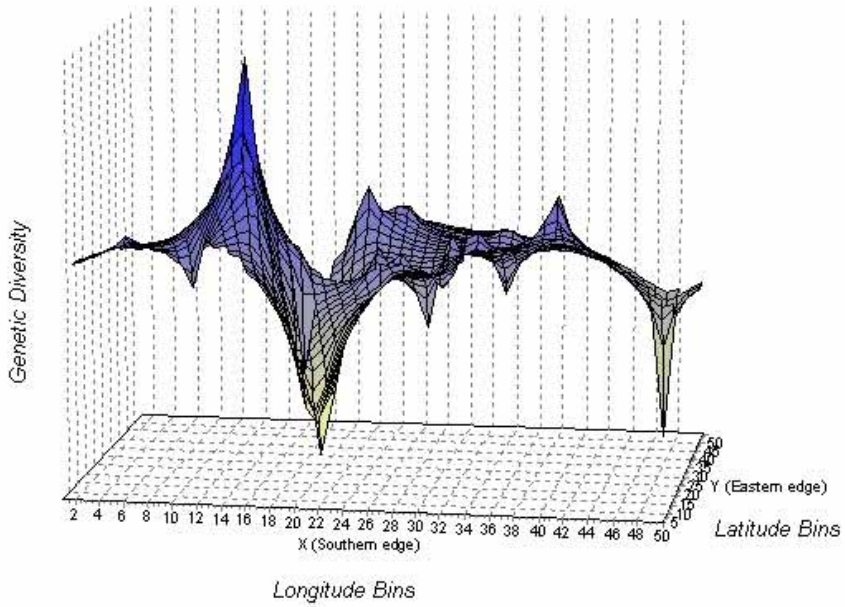
The genetic structure of the sea scallop has recently been characterized with microsatellite DNA markers throughout its geographic range from Newfoundland, Canada to New Jersey, USA by Kenchington *et al*. (2006). This species is a broadcast spawner with a larval dispersal stage of approximately 30 days. It is a relatively long-lived (< 20 years) species and is commercially fished in both Canada and the United States. Traditional, genetic analyses rejected a model of panmixia, and there was some evidence for isolation by distance. However, further insight into the genetic structure of this species was gained using a landscape genetic approach.

Alleles in Space (AIS) software (Miller 2005) was used to detect areas of low gene flow, using a joint analysis of spatial and genetic information. The two major regions of low gene flow (barriers) inferred by Monmonier's algorithm (delineated by Barrier b in App 1 Fig.1) were then used to define regions for an analysis of molecular variance (AMOVA). That analysis showed that a significant but low percentage (1.2%) of the variation is partitioned among the two regions. It also showed that two of the three samples from Western Bank (Circled in red; Western in App 1 Fig. 1) were in a different genetic region from the third.



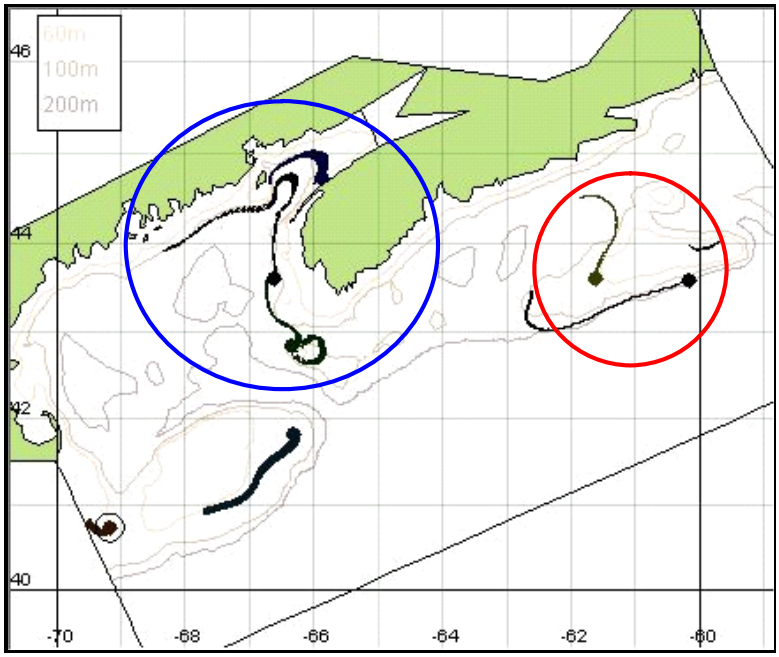
Appendix 1 Figure 1. Areas of low gene flow identified between scallop samples are indicated by lines labelled a-d. The label order indicates the successive strength of the barriers with lesser barriers further indicated by dashed lines. Figure redrawn from Kenchington *et al.* (2006), see publication for details.

The spatial pattern of genetic diversity is visualized below in geo-referenced space (App 1 Fig. 2). This figure, also produced with AIS software, shows that the greatest genetic diversity is found in the southern portion of the range, coincident with the first barrier (a) in Appendix 1 Figure 1, with two areas of below average genetic diversity, one at the northern edge of the range and another in the central area of the distribution coincident with the light grey coloured region in Figure 1.



Appendix 1 Figure 2. A spatial map of genetic diversity in the sea scallop. Areas of high genetic diversity are indicated in blue and are expressed as peaks. Areas of low genetic diversity are indicated in yellow and are expressed as troughs.

By visualizing the genetic structure in a geo-referenced space, hypotheses regarding environmental explanatory variables were created. It was concluded that prominent currents were concordant with the demarcation of the regions, while particle tracking software (Hannah *et al*. 2001) mimicking larval dispersal was used to interpret within-region genetic patterns. This modeling showed that larvae from the most easterly sample on Western Bank would be entrained in a current running along the shelf-break contour – a result consistent with the grouping of this sample in Region 2 of the BARRIER analyses. It also demonstrated connectivity between the samples in Region 2 with larvae having the potential to travel from Browns Bank to Lurcher Shoal or from those beds to Digby within a single generation (App 1 Fig. 3). Further, the general surface currents indicated connectivity from the Bay of Fundy along the New England coast and south around Cape Cod. The authors conclude that contemporary oceanographic circulation patterns play a large role in shaping the genetic structure of this species.

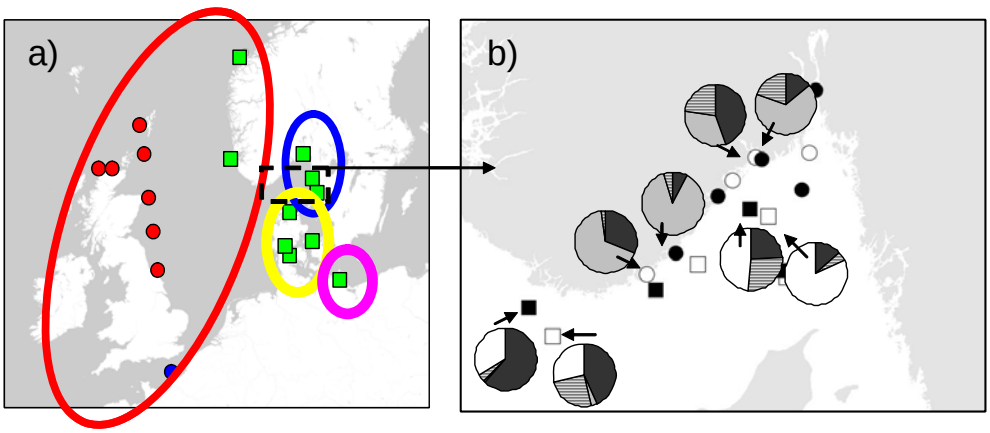


Appendix 1 Figure 3. Velocity streamlines (solid lines) for mean fall flow fields for near surface depth layer based on particles released from 1 September to 1 October 2004. This figure illustrates the method used for forecasting larval dispersal and hence gene flow. The red circle highlights the different flow patterns of two of the Western Bank samples, which contribute to their separation into different regions. The blue circle highlights larval trajectories leading to high gene flow between the enclosed samples resulting in the characterization of these samples in a single region.

5.4 Appendix 2 Atlantic herring *Clupea harengus*

Contributed by D. Bekkevold

The Atlantic herring *Clupea harengus* is a highly migratory species that occurs throughout the North Sea, Skagerrak and Baltic Sea, where spawning, feeding and transition areas commonly overlap among spawning components. Herring moreover seem to have a capacity for changing migratory behaviour and spawning sites in response to shifts in population composition and environmental conditions (Corten 2001). Herring fisheries commonly exploit mixed stocks that may comprise individuals originating from a multitude of local spawning areas. Morphological analyses indicate that herring feeding and wintering in the Skagerrak comprise 1) juveniles which have immigrated from spawning sites spanning the western North Sea from North to South, 2) adults from the very large spawning component at Rügen in the western Baltic and 3) individuals from other, relatively less productive spawning components in the Skagerrak, Kattegat and inner Danish waters. Although demographically differentiated local stocks long have been recognised based on differences in, for example, spawning time, morphology and growth rate, the presumably highly plastic migratory behaviour of the species has led to the suggestion that there is little scope for reproductive isolation among local spawning populations and that recruitment only is weakly affected by local processes. Recently, however, a series of genetic studies applying microsatellite analysis for large samples of herring collected on spawning sites distributed throughout the North Sea, Skagerrak and the Baltic Sea, demonstrated genetic population structure, indicating natal homing and that recruitment occurs on a local scale (Bekkevold *et al*. 2005; Jørgensen *et al*. 2005; Mariani *et al*. 2005, see Appendix 2 Fig. 1a). The identification of genetic signatures of local spawning population subsequently formed the basis for a genetic determination of migratory behaviour outside the spawning season (genetic mixed-stock analysis) (Ruzzante *et al*. 2006; Bekkevold *et al*. 2007; Bekkevold *et al*., unpublished data), demonstrating hitherto unrecognised geographical and temporal separation among populations in their use of common feeding and wintering areas in the Skagerrak (Appendix 2 Fig. 1b).



Appendix 2 Fig. 1 (a) Schematic illustration of herring population structure in the North Sea and western Baltic, based on microsatellite DNA analysis (Bekkevold *et al.* 2005; Mariani *et al.* 2005). Landscape genetic analyses indicated that samples of autumn/winter spawning populations (circles) and spring spawning populations (squares) show clear genetic separation following an isolation-by-distance model, but also that major breaks in gene flow occur among geographic areas, indicated by circles grouping population samples. (b) Genetic analyses of herring on mixed feeding and wintering grounds in the Skagerrak demonstrate that different population components are distributed differently in time and space. Pie diagrams give examples of proportions of herring originating from respectively the North Sea (black), Skagerrak (grey), inner Danish waters (hatched) and Rügen (white) in two consecutive years (open/closed symbols = 2002/2003) in winter (circles) and summer (squares).

5.5 Appendix 3 Atlantic cod (*Gadus morhua*)

Contributed by G. Dahle

Atlantic cod (*Gadus morhua*) encompasses many different populations or stocks, which in part are managed separately. In the northeast Atlantic, cod is divided into two main management units; northeast Arctic cod and Norwegian coastal cod stocks (NEAC and NCC). There is some overlap in spawning areas of these two population groupings, since major spawning for both occur around the Lofoten Islands. Normally, the NEAC stock is predominant on the offshore banks west of the islands, while the NCC stock spawning is found mainly near the coastline, and inside the Vestfjord and other northern Norwegian fjords (Hylen, 1964; Møller, 1968; Dahle & Jørstad, 1993; Nordeide, 1998). Depending on environmental conditions, especially temperature, some NEAC will spawn in the inner parts of Vestfjorden, on the same spawning sites as the NCC. The detailed mechanisms that limit or prevent interbreeding between the two groups are unknown, but recent studies on depth distribution (Nordeide 1998), spawning cod sound (Nordeide & Kjellsby, 1999), modeling of egg and larvae dispersal (Vikebø *et al*. 2005) and population specific egg buoyancy (Stenevik & Sundby, 2005), have suggested several factors that possibly are involved. Thus, in such intermingling areas, the harvest is a typical mixed stock fishery. Thus Mixed Stock Analysis (MSA) and the related individual assignment tests were applied.

For the assignment and mixed-stock analysis of the samples collected in 2002 and 2003, base populations for presumed NCC and NEAC (as determined by otolith structure) were constructed by pooling all individuals from selected geographically-discrete stations as either NCC or NEAC. The programme GENECLASS2 (Piry *et al*., 2004) was used for individual assignment analyses, and mixed stock analysis (MSA) was conducted in the program BAYES (Pella and Masuda, 2001).

The two baseline data sets constructed using baselines defined from geographical distribution of samples (from outer and inner stations) in 2003 were highly differentiated, with an F_{ST} value of 0.11 across 13 loci. The largest differentiation was observed at the *Pan* I locus, which had a very large F_{ST} value of 0.68. A self-assignment test conducted in GENECLASS2

resulted in correct assignment of 402 of the 456 individuals (88.2%) into the correct baseline group. In the actual IA and MSA tests, percentage contributions of NEAC varied from approximately 30% to more than 90% in seven sampling stations investigated, where both forms occurred.

5.6 References

Bekkevold D, Andre C, Dahlgren TG, Clausen LAW, Torstensen E, Mosegaard H, Carvalho GR, Christensen TB, Norlinder E, Ruzzante DE (2005) Environmental correlates of population differentiation in Atlantic herring. Evolution, 59, 2656-2668.

Grant WS, Waples, RS (2000) Spatial and temporal scales of genetic variability in marine and anadromous species: implications for fisheries oceanography. In Harrison, PJ & Parsons, TR (eds) Fisheries Oceanography: an integrative approach to fisheries ecology and management. Blackwell Science. 61-93.

Hannah CG, Shore J, Loder JW, Naimie CE (2001) Seasonal circulation on the western and central Scotian Shelf. Journal of Physical Oceanography, 31, 591-615

Kenchington EL, Patwary MU, Bird CJ, Zouros E (2006) Genetic differentiation in relation to marine landscape in a broadcast-spawning bivalve mollusc (*Placopecten magellanicus*). Molecular Ecology, 15, 1781-1796.

Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. Trends in Ecology and Evolution, 18, 189-197.

Manni F, Guérard E, Heyer E (2004) Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by "Monmonier's algorithm". Human Biology, 76, 173-190.

Miller MP (2005) Alleles in Space (AIS): Computer software for the joint analysis of interindividual spatial and genetic information. Journal of Heredity, 96, 722-724.

Monmonier MS (1973) Maximum-difference barriers: an alternative numerical regionalization method. Geographical Analysis, 5, 245-261.

Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. Journal of Molecular Evolution, 19, 153-170.

Primmer CR, Veselov AJ, Zubchenko A, Poututkin A, Bakhmet I, Koskinen MT (2006) Isolation by distance within a river system: genetic population structuring of Atlantic salmon, *Salmo salar*, in tributaries of the Varzuga River in northwest Russia. Molecular Ecology, 15, 653-666.

Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN, Version 2.000: A Software for Population Genetic Data Analysis. Genetics and Biometry Laboratory. University of Geneva, Geneva, Switzerland.

Simard Y, Legendre P, Lavoie G, Marcotte D (1992) Mapping, estimating biomass, and optimizing sampling programs for spatially autocorrelated data: Case study of the northern shrimp (Pandalus borealis). Canadian Journal of Fisheries and Aquatic Sciences, 49(1), 32-45.

Storfer A, Murphy MA, Evans JS, Goldberg CS, Robinson S, Spear SF, Dezzani R, Delmelle E, Vierling L, Waits LP (2006) Putting the 'landscape' in landscape genetics. Heredity 98, 128–142.

Watson DF, Philips GM (1985) A refinement of inverse distance weighted interpolation. Geo-Processing, 2, 315–327.

5.7 Recommendations

1) Before starting a sampling programme for a particular species we recommend that all available information on biological and physical parameters, including geographical features, hydrographical data and geological information, be taken into consideration. In terms of biological parameters we need information such as: migration pattern, spawning areas, extent of philopatry, spawning time, feeding grounds, growth rate, natural and fishing mortality.

2) In order to compare genetic and geographic information it is necessary to identify stage in the life cycle where populations are most discrete. This will generally be the spawning stage. We recommend that an optimal sampling strategy be devised, depending on the species,

3) We recommend using the most appropriate molecular methods in a comprehensive spatial survey, incorporating as far as possible a temporal component.

4) It is recommended that genetic and geographic information be combined using the most appropriate landscape genetics approaches, eg, currently BARRIER and AIS.

5) We recommend attempting to explain results from landscape genetics software in terms of available physical and biological information in order to improve predictive capacity and make best use of the results of analysis.

6) To delineate the spatial extent of each population using survey, we recommend using both physical and genetic data.

7) Having defined populations that at other stages of the life cycle where population mixing may occur we recommend that approaches based on MSA/IA be used to estimate proportions in the mixture and population identity of individual animals.

8) As an overarching recommendation, given that methods are now available for many species of identifying structuring into breeding populations, it is recommended to fisheries managers that these methods be used in conjunction with geographical information systems to define the spatial and temporal ‘footprint’ of these breeding populations in order to allow population focused management.

9) We recommend that future work involve further investigations of the relationship between geographical information and population genetics, so that maximal use can be made of the synergy between these two fast developing fields.

6 ToR (e): To identify the structural and institutional requirements for developing meta-data bases for genetics of fish species covered under the ICES remit.

Eric Verspoor, Jann Thor Martinsohn, Torild Johansen and Tom Cross

6.1 Justification/Purpose

A large number of studies have been carried out into the nature and extent of genetic variation within stocks of wild finfish and shellfish species. These provide valuable insight into the structuring of stocks into genetic populations and the proportional contributions of different populations to mixed fisheries, as well as the impacts on fisheries of exploitation and global climate change. The value of this work could, however, be considerably enhanced by making results, findings and basic data accessible ensuring the research community as a whole, in a comprehensive and timely fashion, and by putting in place mechanisms which help to facilitate the integration of data across studies, both temporally and spatially. There are numerous examples where data integration has resulted in valuable novel insights, both within and outside fisheries science. The need for greater integration of existing research, data and materials related to population genetic studies in the North Atlantic and adjacent EU waters is recognised by the ICES WGAGFM.

The European Union, as well as national governments and research councils, have financed most of these studies, including both the collection of material and their genetic analysis. Unfortunately, when most projects are complete summary data on samples analysed are published and, frequently, researchers move on, or leave science. As published accounts in most cases include only summary statistics such as sample allele frequencies, raw genotype data on individuals in analysed samples are difficult to access and risk being lost. The same is true of samples. Research, universities and research institutions have over the years accumulated large collections of potentially valuable historical samples. In most cases these samples are not archived, and at risk of being forgotten and lost, seriously compromising the possibility of using them for research in the future. Where material and data are lost the scope for integrating data sets and gaining broader insights can be seriously compromised. In particular, their loss severely constrains our capacity to conduct studies of genetic changes overtime, something which is essential to meaningful studies of the genetic impacts, for example, related to climate change, fisheries induced evolution, and introgressive changes to wild populations from stocking of non-natives or interbreeding with escaped cultured fish.

The solution to this problem is to catalogue and assemble existing primary data, information and biological materials related to the genetics of finfish and shellfish species. However, this represents a logistically complex and financially costly exercise, and the resources for assembling and managing such a comprehensive database of information and materials are unlikely to be available. Yet what is entirely feasible with existing resources, and be a significant step in the right direction, would be the building of a metadatabase, to catalogue past and on-going research projects, and existing genetic data and biological materials and where it can be accessed. This would facilitate both increased awareness and use of existing data and materials, as well as the eventual integration of the primary data sets themselves. This database would serve as a portal, allowing stakeholders and fellow researchers to gather all relevant information related to projects of interest.

Over the last decade, genetic research has provided an increasing level and value of insight into the biology of marine finfish and shellfish species in the North Atlantic and adjacent EU waters, in support of marine conservation and fisheries management. To support of this work in the future and increase the benefits which will be gained from it, a regional

METADATABASE on important species should be set up which catalogues and cross references:

- existing raw genetic data sets, including those on allozyme, mtDNA, microsatellite, SNP, EST and coding gene variation, and links to where they can be located,
- existing collections of historical biological material, which can be exploited for DNA studies in the future, where they are archived, and contacts for queries on their use.
- past and on-going projects on population genetics of marine organisms, linked to individuals or locations where more detailed information can be found, to facilitate identifying and accessing both published and unpublished studies completed or underway;
- profiles of currently and recently active researchers working on the genetics of marine species in the North Atlantic and adjacent EU waters and links to contact them.

By cataloguing and cross referencing this information, the metadatabase will be able to provide an invaluable overview of:

- researchers working on a species or in an area, and who to contact for expert opinion on a particular species or geographical region
- species and regional stocks which have been genetically characterised by different population genetic markers
- what data exists for which parts of the species range in the North Atlantic and adjacent EU waters to help define data needs
- available historical and contemporary materials for genetic analysis and where these are located

This overview will facilitate and promote:

- better communication between scientists and other stakeholders increasing synergistic effects and stimulate new projects
- a reduction in unproductive overlapping of studies in the future and encouragement of research collaborations and joint projects
- more effective exploitation of existing data sets and materials
- finding partners to develop collaborative national and trans-national projects
- more effective, productive designs of new research projects

The potential benefits from linking studies are growing with an increasing number of genetic studies of marine species being carried out, made possible by the development of rapid, cost-effective DNA based analyses of variation at the molecular level. These benefits are already likely to be substantial. The existence of a metadatabase on existing work, data and biological materials, should immediately lead to greater integration and increase the value of currently on-going and planned work genetic work. This immediate impact will arise as a result of the increased communications among researchers which would occur as part of the cataloguing exercise associated with the development of the metadatabase.

Accepting the value of creating a metadatabase of genetic information on as described, the ICES WGAGFM have identified the structural and institutional requirements for the establishment and maintenance of a centrally administered, web-accessible meta-data base and make recommendations for its implementation.

6.2 System Architecture

A Database Management System (DBMS) is proposed which is optimally designed for the uploading, storage and accessibility of the data contained in the above described catalogue. To be maximally useful, the metadatabase constructed must incorporate the following features:

- An input/output design suited to user/client needs
- User/client accessibility
- User/client friendliness
- Stability of access and design
- Maintenance and regular updating

To ensure access by the intended user group, which encompasses a widely dispersed set of researchers, access to the metadatabase will need to be provided through a web application (Interface). Its design will need to ensure optimal support for data upload (submission) by users, as well as for data download (retrieval) by clients. Furthermore, its usefulness for both users and clients would be significantly enhanced if it provided a platform for a discussion forum for both sets of stakeholders.

Uploading of data will need to be carried out via an electronic submission form. Before being transferred to the server, the data will need to be routed to a control gate, by automatically sending the application to assigned database consortium members. The controllers will review the application by checking it against agreed quality criteria. It is important to note that this control gate will not be a profound peer review. Its purpose is to avoid major flaws in general data structure and quality rather than to carry out a detailed data analysis (see System Maintenance). Where data do not meet basic quality levels, an electronic dialogue between the controller and submitter will be initiated and the submission brought up to standard. When data meets basic quality standards, it will be uploaded onto the metadatabase server.

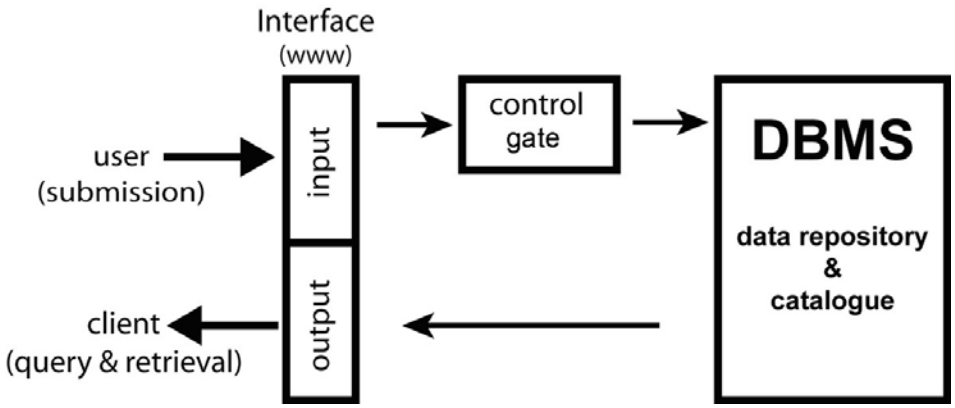


Figure 1. Basic architecture of proposed metadatabase on genetic variation in marine finfish and shellfish species stocks.

6.3 Data Structure

To provide the required catalogue of information, the metadatabase will need to encompass a number of primary data fields. These will need to include:

- Research Institution
- Individual researcher
- Type of data
- Species
- Country
- Primary Biogeographical Region

- Secondary Biogeographical Region

The first two would be provided by the login through the registration information and would be the contact for gaining more detail about the metadata entry. The final five would be specified using set lists in pop-down menus and would define the nature of the information catalogued. Of these "Primary Biogeographical Region" would be restricted by the choice of "Country" and "Secondary Biogeographical Region" by the choice of "Primary Biogeographical Region".

Three basic types of data would be catalogued:

- Genetic data sets
- Biological materials
- Completed and On-going Projects

Depending on the type of data being catalogued, a further linked secondary data field would need to be created. For "Genetic" and "Biological" data types, the entries would have to be sample based.

For "Genetic" data, the secondary data field would need to encompass:

- Type of genetic data
- Location from which data derive
- Year from which data derive
- Sample size
- Number of loci
- Names of loci
- Details/Comments on Data

For "Biological" data, the secondary data field would need to encompass:

- Type of biological material
- Method of storage
- Location from which material derive
- Year from which material derive
- Sample size
- Details/Comments on Material

For "Project" data, the secondary data field would need to encompass:

- Title of Project
- Start of Project
- End of Project
- Type of Study
- Objective of Study
- Country of Study
- Region of Study
- Summary of Study
- Genetic data collected

6.4 Data Input

The input of data into the metadatabase would be by data "owners". "Owners" would be individuals whom are the persons whom have generated or are in possession of the

information catalogued. Data entry would be via a "Home Page" which would encompass the following:

- A statement of the purpose/mission of the metadatabase
- A statement organisational support, ownership, and management
- A link to a page with information on contacting database managers
- A link to a page for submission of data
- A link to a page for retrieval of data
- A link to a page for information exchange among users and clients

The submission page for data would have links to two pages:

- One, to register as a user
- A second, once registered, to login and input new information

Following the login, a user would need to be directed to a data submission page. On the submission page, data could be input via a largely menu driven or choice selection form. This approach would provide for the rapid entry of data and ensure a uniform data structure and format to provide an efficient framework for cataloguing and accessing information. An example of a prototype data entry form for primary data fields is given in Figure 2. Secondary data fields would follow a similar format.

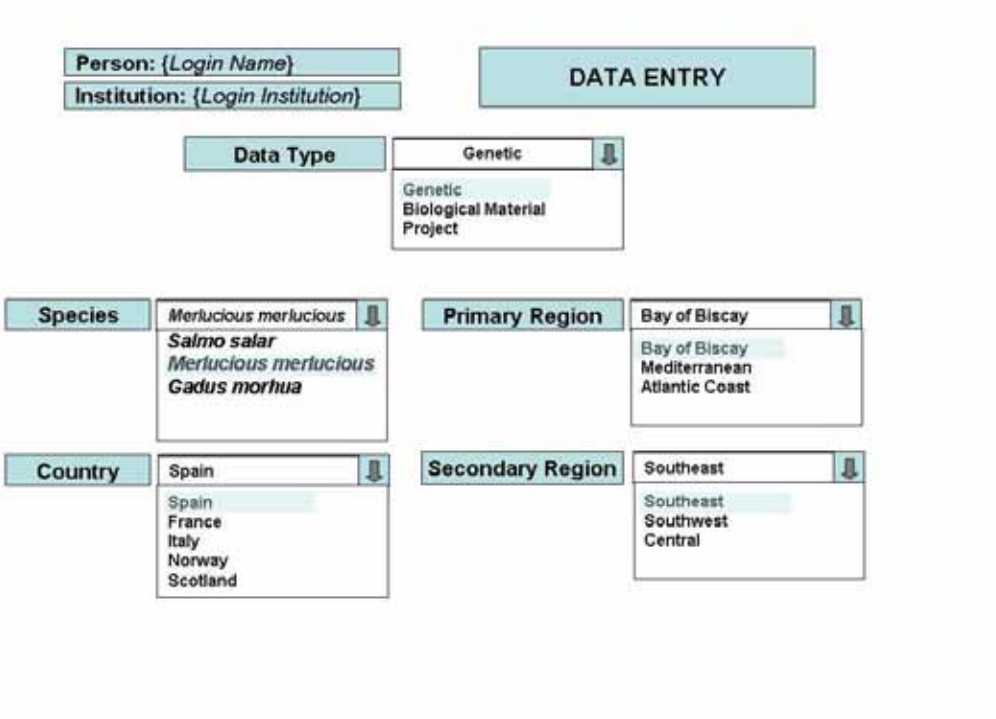


Figure 2 Example of a prototype screen for website data entry

6.5 Data Access

Access to the catalogue of information would be through the "Home Page" of the website by clicking on a link for to a page with a data retrieval form. The form would be used by the client to specify a search of available information using one or more data fields and focus on providing summaries of available data on a species basis in relation to data type, and within data type, according country, region, type of genetic data, etc. The data entries matching the query conditions would then be listed and detailed, including giving the person and institution

to be contracted for further information. These would be output to the screen in a form suitable for printing or downloading by the client.

6.6 System Maintenance

The project should be fundamentally democratic with the users responsible for the quality of the provided data. Stringent control of content will be assured by clients who would be encouraged to report observations on data quality and content to the database consortium. However, as explained above (System Architecture), an intermediate control gate to the data base will assure that qualitatively poor or irrelevant data will not be uploaded onto the server.

Implementation assessment would be carried out by monitoring the progress of uploaded data (user) as well as by counting the number of visits to the database website (client). The response to the database on the client side will be monitored in further detail by counting the visits per level (i.e. starting from the highest level, the homepage, towards deeper, more specific sections such as "species" or "region").

6.7 Organisational Context

The creation of the proposed metadatabase would involve a development and a production phase. The development phase needs the expertise of IT specialists with experience in the design and development of DBMS. In contrast, the production phase needs to be handled by an institution should be responsible which has experience with hosting databases, and can provide an appropriate institutional infrastructure. The latter is crucial to ensure professional database management on the required time scale which is far longer than the average life time of an individual research project.

For both the development as well as the production phase, knowledge and expertise in fisheries and mariculture genetics will need to be involved to guarantee the optimal adaptation of the database to the needs of stakeholders working in these fields.

Within the European/ North Atlantic context, a potentially workable framework for implementation would be that the development of the database is entrusted to the DG Joint Research Centre (JRC) of the European Commission, subject to the agreement of JRC, while the database could be hosted and managed by the ICES secretariat, subject to the agreement of ICES. JRC is well positioned to play this role as it has extensive expertise in IT technologies, including DBMS development, and also works in fisheries related areas in the context of the Common Fisheries Policy. At the same time, ICES is already host to numerous fisheries related databases. Hosting of the proposed metadatabase by ICES would ensure that the management is carried out according to the new ICES data strategy (see "ICES Data Strategy 2006-2010"), ensuring optimal integration and compatibility within their database system.

6.8 Recommendations

1) The WGAGFM, for marine species of interest to ICES and the EU, establish a web-based meta-database of experts, projects, existing genetic data sets, and archival collections of material suitable for DNA analysis, in support of studies of genetic population structuring and of fish product traceability at the intraspecific level.

2) The meta-database should be web-accessible for input and output by working group memebers and interested stakeholders, and used to disseminate information on population genetics and encourage the development, by institutions and scientists, of integrated primary genetic databases and.

3) The WGAGFM Chairman seek support and approval for ICES to host the meta-database and oversee administration of web site and metadata repository

4) Following agreement by ICES, the Chairman approach JRC to develop the meta-database

5) Following agreement on 3) and 4), the WGAGFM work with JRC to define the content, structure, and functional requirements for the meta-database

6) The WGAGFM manage input into and updating of the meta-database

7 ToR (f): to review and provide recommendations on the application state of the art Genetic Stock Identification methods, with particular emphasis on evaluating the precision for identifying the population of origin of individual Atlantic salmon

Dorte Bekkevold, Tom Cross, Eileen Dillane, Riho Gross, Marja-Liisa Koljonen, and Phil McGinnity

7.1 Introduction

Genetic stock identification (GSI) techniques have been successfully used to resolve a number of, what were presumed until relatively recently intractable, salmon fisheries management questions. For example, in determining the relative proportions of contributing populations in mixed stock fisheries: at the macro scale, separating stocks of North American and European origin at West Greenland (King *et al*. 2001); at the meso scale, apportioning catches from the Baltic Sea (Koljonen, 2006) and the Irish coastal fisheries (unpublished data) to individual river stocks; and even within river systems (micro scale) such as the Moy and Foyle catchments in Ireland (unpublished data), and the Teno river in Finland, allocating catches to individual river tributaries (e.g. Vähä *et al*. 2007). GSI, in the context of parental assignment, was also critical in the successful determination of the relative fitness of the progeny of farm escape, wild and hybrid salmon spawning in the wild (McGinnity *et al*. 2003) and in spatially and temporally determining levels of farm and hatchery introgression in wild populations (Clifford *et al*., 1998, Nielsen *et al*. 1999, 2001).

Salmon stocks have declined in both Europe and North America and all evidence points to there having been environmental changes in the ocean phase. One of the key issues in increasing the knowledge of the marine ecology of the species is the understanding of the differences which may occur between the sea distribution of different regional stock groups and river/tributary populations. This requires an ability, not only to determine the proportions of contributing populations in a given sample of fish captured in the marine environment, but also to identify the individual fish in that sample to their river or region of origin. Information on individual fish can then be used to map the distribution and migration patterns of different genetic stock groups. This in turn will provide the basis for the development of more informative eco-genetic models linking growth performance, environmental conditions, and the distribution of food organisms.

Genetic individual assignment has many advantages over physical tagging methods such as coded wire tags. The information derived from genetic identification can be done from the fish in the wild, thereby overcoming the experimental error introduced by either handling wild fish when physically tagging them, or using hatchery fish as a surrogate for wild salmon. All individuals captured in the experimental fisheries are of equal value and can be used in subsequent analysis, providing a significant cost advantage over conventional tagging where only those individuals that have tags can be used. With genetic methods there is no loss of tags and no bias due to viability or catchability effects as with external tags. In addition, the time and place of sampling can be chosen more freely and precisely, as preceding tag and release programmes are not required. Moreover, genetic identification is not dependent on fishermen in returning tags or on the detection of internal tags. Furthermore, all samples previously collected from marine surveys, i.e. historic archives of scales/otoliths, etc. can be used, and these data are of value in elucidating temporal trends in migration and distribution patterns.

Genetic methods also have some limitations. The extent of inter-population differentiation will affect the resolution power that can be achieved. Statistically significant differences in allele frequencies often occur, but quantitatively they may be too small to meet the assignment accuracy and precision requirements of the managers. Genetic assignment estimates give

probabilistic information about the origin of individuals or populations rather than absolute information, a limitation common to many techniques in fisheries biology, including stock assessment, although satisfactory levels of statistical confidence within regions and individual rivers can generally be achieved for fisheries management requirements. Genetic identification of populations and individuals of Atlantic salmon has been recently reviewed by Koljonen *et al*. (2007).

7.2 Brief Review of the methods

The genetic analysis of the composition of population mixtures has advanced and diversified in recent years resulting in two types of approach, Mixed Stock Analysis (MSA) and Individual Assignment (IA), in which the goal is either to estimate the proportions of contributing stocks in the catch mixture, or to solve the origin of an individual fish. A commonly used Statistical Program for Analyzing Mixtures (SPAM) is available for the Windows environment (Debevec *et al*., 2000). SPAM searches for maximum likelihood estimates of population proportions using three numerical algorithms: conjugate gradient (CG), iteratively reweighed least squares (IRLS) and expectation-maximation (EM). However, Bayesian modelling has been shown to provide the most reliable estimates of the relative contributions of different populations in mixed stock fisheries when compared with other methods (Beacham *et al*. 2006) and also for individual assignment (IA) when compared to GENECLASS (Cornuet *et al*. 1999, Luikart & England 1999)(see Koljonen *et al*. 2005). A number of statistical packages are available, which provide a range of methods for assigning individuals to population of origin, and these have been evaluated in the literature (e.g. Manel *et al*. 2005; Hauser *et al* 2006). The IA option incorporated into the software BAYES of Pella and Masuda (2001) seems to offer the best levels of correct assignment.

For example, up to 95% of 700 salmon caught in a recreational fishery in the tidal part of the Moy fishery in Ireland assigned to the Moy river catchment (with high levels of confidence). In this test fishing nearly 100% fish could be assumed to originate from the River Moy. Table 1 gives a comparison of this result with that from some other packages and methods.

Table 1

	cBAYES	cBAYES	SPAM	GENECLASS
method	*IA*	*MSA*	*MSA*	*IA*
Moy	95.3	89.8	77.5	43.6
Proximate catchments	2.7	5.6	10.8	11.6
Regional catchments	1.4	2.5	4.6	8.2
Outside region	0.6	2.1	7.1	36.6

In the Moy example, as well as providing the highest levels of correct assignment within the regions expected, the cBAYES assignment method gave high degrees of confidence for these assignments with 61% of the fish assigned with greater that 95% probability, and 86% with greater than 75% probability and the remainder assigned with 50% probability. These results were achieved using just 10 microsatellite loci (we recommend 15-20 for such studies which will substantially improve levels of confidence in the assignment). The probabilities may, however, be lower, when more stocks are contributing into the mixture.

When the identification of the stock of origin of the individual fish has been unsuccessful, in the sense that the probabilities for each of the stock of origin are low or too even (for example 0.3, 0.3, 0.3), individuals can be assigned to originate from groups of genetically similar stocks or regional grouping with higher levels of confidence (higher probabilities).

7.3 Sampling

Currently the most reliable individual assignments are achieved in combination with mixed-stock-analysis, in which information obtained from the genetic composition of the mixture can be utilized in addition to the multilocus genotype information of the particular individual, to determine the river of origin (Koljonen *et al*. 2005). Variation in the estimates may thus be derived from the mixture sample, the baseline sample, or both. However, the bias of the proportion estimates is mainly due to the baseline data, and is at its greatest when genetically similar stocks differ markedly in abundance. The variance resulting from mixture sampling depends on the size and stock composition of the mixture sample. To achieve high levels of precision, the number of fish per stock sampled within the mixture is important. If the number of contributing stocks is high, a large mixture sample is needed for reliable estimates. In the mixed-stock analysis of Atlantic salmon catches in the Baltic Sea, the 95% probability interval (confidence interval) for the stock group estimates was about $\pm$10%, when mixture sample size of 300 fish, 8 microsatellite loci, and a baseline with 32 river stocks were used (Koljonen 2006).

7.4 Issues related to data quality

A number of issues in relation to assuring data quality must be addressed when planning and carrying out MSA and/or IA for Atlantic salmon. These include issues related to

1) assumptions for baseline samples,
2) genotyping errors and
3) choice of genetic marker system to be applied.

7.4.1 Assumptions for baseline samples

Several factors in baseline sampling may affect the amount of variation and bias in the estimates:

Hardy-Weinberg equilibrium

All currently available statistical procedures for MSA and IA assume that the genotypes used for individual baseline population information conform to Hardy-Weinberg equilibrium proportions. Including information for samples not exhibiting Hardy-Weinberg proportions may therefore pose a problem to the outcome of the statistical analyses. Reasons for failure to conform include the presence of cryptic population sub-structure, sib-group sampling and genotyping errors (for the latter, see section 2 below). The effects of such non-representative sampling on IA performance are likely to vary among specific analysis aims, and precautions can be taken by performing detailed examinations of baseline data prior to performing IA analyses, and by not causing H-W deviations by unjustified pooling of baseline data (e.g. by statistically testing for presence of sub-structure and/or sib-groups).

Temporal stability of allele frequencies

If allele frequencies vary within baseline populations over time, such changes will affect the performance of MSA and IA procedures, reducing the statistical power for correct assignment of individuals to specific rivers, and should be taken into account as a potential source of error in GSI. In theory, any effects of changes caused by genetic drift can be compensated for to a marked extent by collecting baseline data over several years (Waples, 1990). The importance of repeated sampling depends on the life history of the species concerned and on the degree of overlapping in the year-classes. Atlantic salmon have overlapping generations, and partly for that reason the temporal variation of allele frequencies in large natural stocks may be of little significance to IA performance, whereas in small natural populations or in hatchery stocks, genetic drift can cause pronounced changes. Regular validation and updating of population

samples used to define baselines need to be planned as part of the estimation routine to encompass such variation.

Baseline sample size

The precision and accuracy of the estimates can be improved by increasing the baseline sample size for each baseline stock from the commonly used about 50 to about 100 fish, and ensuring that it is representative of at least two cohorts (50 from each).

Gaps in the baseline

Standard individual assignment procedures (like BAYES or GENECLASS) does not allow for the assignment of individuals to unknown source populations (i.e. populations not present in the baseline). This shortcoming may be overcome by the program HWLER (Pella and Masuda 2006), which allows unknown individuals to be assigned to a hypothetical baseline sample or samples. However, the statistical properties and levels of attainable resolution of such approaches are likely to differ from, for example, the standard method proposed by Pella and Masuda (2001).

Introgression

In relation to allele frequency stability, an important issue for Atlantic salmon is concerned with genetic effects of farmed salmon escapes and deliberate stocking from hatcheries and subsequent introgression in wild populations. In such cases MSA and IA performance may be negatively affected, as baselines generated using information for non-introgressed populations may at some point no longer adequately reflect the genetic composition of contemporary catches. Correspondingly, baseline information may have been collected at a point of time where one or more of the baseline populations were affected by genetic input from reared salmon that was lost over subsequent generations (e.g. due to selection against reared salmon genotypes under natural conditions). The impact of such introgression dynamics on MSA and IA can be assessed through simulation studies and needs to be routinely monitored for running MSA and IA programmes, by frequent re-sampling of populations/spawning rivers potentially affected by escapes or stocking in order to update baseline allele frequency information. Moreover, baseline samples need to include information for farmed salmon strains/populations. Especially for small farm brood-stocks, allele frequencies may change rapidly over time, and such variation also needs to be incorporated into sampling strategies. Simulation analyses can be used to assess the effects of introgression on MSA and IA performance.

7.4.2 Genotyping errors

Microsatellite genotyping errors

Microsatellite genotyping has been shown to be error-prone (reviewed by DeWoody *et al*. 2006) and even modest error rates can bias estimates of population allele and genotype frequencies and thus, cause artefact deviations from Hardy-Weinberg equilibrium, which is a fundamental assumption for baseline samples of many MSA and IA methods. Compounded over multiple loci, even a small per-locus genotyping error rate can result in relatively large probabilities of a multilocus genotype containing at least one error (Creel *et al*. 2003; Bonin *et al*. 2004; Hoffman & Amos 2005), although error rates are rarely equal across loci, and dropping a single locus may provide a disproportionate decrease in error rate. Checks can be built into the system to address these (Taberlet *et al*. 2004). See Appendix 1for Quality Control measures.

7.4.3 Marker selection

Number of loci

The required number of loci depends on the level of differentiation among stocks and is case-specific. However, there obviously is a level, where there is no diagnostic advantage in having additional markers. At the moment 15 microsatellites should be adequate for most MSA and IA analyses. In specific situations, however, the optimum number of loci has to be determined.

New markers

To date, most MSA and IA analyses in Atlantic salmon have been carried out employing genetic information from microsatellite DNA markers, which has proven to perform well in terms of statistical properties for assignment and technical reproducibility. Nonetheless, other approaches also exist, such as analysis of Major Histocompatibility gene Complex variation and single nucleotide polymorphisms (SNPs), which are being routinely employed for GSI approaches in other species, including salmonids (e.g. Beacham *et al*. 2004; Smith *et al*. 2005). In comparison with microsatellite markers, SNP screening is expected to be less affected by DNA quality and inter-laboratory variation. However, as SNPs are commonly bi-allelic, GSI analyses normally require screening of a larger number of loci (commonly > than 2-3-fold) compared with microsatellites. It is envisaged that in the future SNPs are going to be the population markers of choice across fish species and taxa, but whether Atlantic salmon GSI approaches would benefit from including SNP marker application remains to be examined.

Further research is needed to identify DNA markers associated with protein variation and other genetic variation defining regional groupings of populations, which can be used to achieve regional assignment in a practical cost effective way in support of marine ecological studies. There are now available a number of classes of DNA markers which could be applied for this purpose. For example, existing work shows point mutations in mtDNA with highly restricted regional distributions that could be informative for some regional groups but further work is needed to confirm their diagnostic potential and to identify a suite of markers to comprehensively cover the European range of salmon. Additionally work to date shows regionally restricted distributions of microsatellite alleles and varying levels of regional differentiations among different microsatellite loci. Approximately 1700 microsatellite loci have been identified in Atlantic salmon and those optimal for use in regional discrimination remain to be identified. Furthermore, as can be inferred from genetic protein studies of loci such as MEP-2*, there is considerable potential for identifying single nucleotide polymorphisms (SNPs) (McMeel *et al*. 2001) with the capacity to contribute regional assignment (Rengmark *et al*., 2006).

7.5 Recommendations

1) As an overriding recommendation we are convinced that in most circumstances IA can give valuable information for Atlantic salmon management and specifically identify the population of origin of individual Atlantic salmon with relatively high probabilities.

2) We recommend that genetic stock identification methods be applied to addressing salmon biology and fisheries management questions, e.g. contribution of individual rivers to fisheries, identification of migration and distribution patterns, introgressions between hatchery and wild fish, temporal changes in population structure. This should be reported within the appropriate probabilistic framework.

3) Presently, according to performed comparison tests, the Bayesian approach of Pella & Masuda (2001) appears to provide the most accurate results with regard to individual assignment, and we therefore recommend its use. In specific cases other methods can be useful.

4) Ideally all contributing stocks should be included in the baseline, and discrete entities within rivers should be recognised, as well as those among rivers (i.e. the sampled units should be in Hardy-Weinberg equilibrium). Efforts must be made to make baseline samples representative of each population, which may subsequently be sampled at sea. Sampling programs should be prioritised by targeting the most productive rivers, rivers where conservation limits are not being reached, and should also include hatchery stocks (if not identifiable by physical markers).

5) Baseline sample sizes of 100, representative of at least two cohorts (50 from each), from each population are recommended. Sample size should be consistent across baselines.

6) It is recommended that baseline populations be re-sampled every 5-10 years. However, where introgressions from cultured fish are suspected, or where population sizes are quite small, it is recommended that this should be done on a more frequent basis.

7) The sufficient size of a mixture sample is dependent on the required precision level and the number of stocks occurring in the mixture, and can be determined by simulations prior to sampling plan.

8) If microsatellites are the marker of choice, recent reviews suggest, that a minimum of 15 polymorphic loci should be used for individual assignments. These should be investigated to ensure that null alleles, linkage, allelic dropouts and stutter bands do not cause problems e.g. microsatellite loci having a tetranucleotide repeat motif are preferred to loci having a dinucleotide repeat motif as they are easier to score and typically display no stuttering patterns.

9) To guarantee consistency of genotyping results, calibration of allele sizes among participating labs is warranted by exchange of reference samples, and baseline samples should be preferably genotyped in a single lab.

10) Accuracy and precision (depending on level required) of the assessment estimates must be examined by simulation studies, and using true mixtures of fish of known origin, not included in the baseline data set, before applying to true catch data. Simulations should be used for defining the needed catch mixture sample sizes, in order to achieve the required confidence. This information can be incorporated into the fisheries sampling protocols.

11) The highest resolution to be aimed for should be assignment to river of origin (corresponding H-W unit). In cases where the probability levels of the assignments for individual rivers remain low, pooling of probabilities over individual rivers can be employed to achieve a regional assignment for individuals. In designing programmes and models for handling assignment data, managers should be aware of the probabilistic nature of the data and the required level of probabilities and statistical confidence needed for utilising the data for combined studies, (e.g. for some ecological or fisheries studies, it may be sufficient to combine information from high and low assignment values to make a composite sample).

12) Current methods utilise microsatellite technology, however the development and application of novel microsatellite markers, as well as nuclear and mtDNA SNPs etc. should be investigated with regard to their diagnostic usefulness as river and regional specific identifiers. Also, some consideration should be given to the integration of genetic and biological data for assignment.

13) Development of statistical methods for utilising probabilistic individual assignment results is recommended.

7.6 References

Beacham, T.D., M. Lapointe, J.R. Candy, B. McIntosh, C. MacConnachie, A. Tabata, K. Kaukinen, L. Deng, K.M. Miller, and R.E. Withler. 2004. Stock identification of Fraser River sockeye salmon (*Oncorhynchus nerka*) using microsatellites and major histocompatibility complex variation. Trans. Am. Fish. Soc. 133: 1106-1126.

Beacham, T.D., Candy, J.R., Jonsen, K.L., Supernault, J., Wetklo, M., Deng, L., Miller, K.M. & Withler, R.E. (2006) Estimation of stock composition and individual identification of Chinook salmon across the Pacific rim by use of microsatellite variation. Transacations of the American Fisheries Society 135, 861-888.

Bonin A, Bellemain E, Bronken Eidesen P, Pompanon F, Brochmann C, Taberlet P (2004) How to track and assess genotyping errors in population genetic studies. Molecular Ecology, 13, 3261–3273.

Clifford, S.L., McGinnity, P. & Ferguson, A. (1998) Genetic changes in Atlantic salmon (*Salmo salar*) populations of northwest Irish rivers resulting from escapes of adult farm salmon. Canadian Journal of Fisheries and Aquatic Sciences 55, 358-363.

Cornuet, J.-M., Piry, S., Luikart, G., Estoup, A., and Solignac, M. 1999. New methods employing multilocus genotypes to select or exclude populations as origins of individuals. Genetics 153: 1989-2000.

Creel S, Spong G, Sands J. L. *et al*. (2003) Population size estimation in yellowstone wolves with error-prone noninvasive microsatellite genotypes. Molecular Ecology, 12, 2003–2009

Debevec, E.M., Gates, R.B., Masuda, M., Pella, J., Reynolds, J. & Seeb L.W. (2000) SPAM (Version 3.2): Statistics program for analyzing mixtures. Journal of Heredity 91, 509-510.

Dewoody J, Nason JD, Hipkins VD (2006) Mitigating scoring errors in microsatellite data from wild populations. Molecular Ecology Notes, 6 , 951–957.

Hauser, L., Seamons, T.R., Dauer, M., Naish, K.A. & Quinn, T.P. (2006) An empirical verification of population assignment methods by marking and parentage data: hatchery and wild steelhead (*Oncorhynchus mykiss*) in Forks Creek, Washington, USA. Molecular Ecology 15, 3157-3173.

Hoffman JI, Amos W (2005) Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. Molecular Ecology, 14, 599–612.

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, 10, 807-821.

Koljonen, M-L. (2006). Annual changes in the proportions of wild and hatchery Atlantic salmon (*Salmo salar*) caught in the Baltic Sea. ICES Journal of Marine Science 63: 1274-1285.

Koljonen, M.-L., Pella, J. J., Masuda, M. (2005). Classical individual assignments versus mixture modeling to estimate stock proportions in Atlantic salmon (*Salmo salar*) catches from DNA microsatellite data. Canadian Journal of Fisheries and Aquatic Sciences 62, 2143-2158.

Koljonen, M-L. King, T. and Nielsen, E. (2007). Genetic Identification of populations and individuals. In: The Atlantic salmon: Genetics, Management and Conservation (eds. E.Verspoor, J. Nielsen and L. Stradmeyer). Blackwell Publishing. In press.

Luikart, G., and England, P.R. 1999. Statistical analysis of microsatellite DNA data. TREE 14: 253-256.

Manel, S., Gaggiotti, O.E. & Waples, R.S. (2005). Assignment methods: matching biological questions with appropriate techniques. Trends in Ecology and Evolution 20, 136-142.

McMeel, O.M., Hoey, E.M., Ferguson, A. (2001) Partial nucleotide sequences, and routine typing of polymerase chain reaction-restriction fragment length polymorphism, of the brown trout (*Salmo trutta*) lactate dehydrogenase, LDH-C1*90 and *100 alleles. Molecular Ecology, 10, 29-34.

McGinnity, P., Prodöhl, P., Ferguson, A., Hynes, R., Ó Maoiléidigh, N., Baker, N., Cotter, D., O'Hea, B., Cooke, D., Rogan, G., Taggart, J. & Cross, T. (2003) Fitness reduction and

potential extinction of wild populations of Atlantic salmon *Salmo salar* as a result of interactions with escaped farm salmon. Proceedings of the Royal Society of London Series B 270, 2443-2450. DOI 10.1098/rspb.2003.2520.

Nielsen, E. E., Hansen, M. M. & Loeschke, V. (1999). Genetic variation in time and space: microsatellite analyses of extinct and extant populations of Atlantic salmon. Evolution 53, 261-268.

Nielsen, E. E., Hansen, M. M. & Bach, L. A. (2001). Looking for a needle in a haystack: Discovery of indigenous Atlantic salmon (*Salmo salar* L.) in stocked populations. Conservation Genetics 2, 219-232.

Pella, J. & Masuda, M. (2001) Bayesian methods for analysis of stock mixtures from genetic characters. Fisheries Bulletin 99, 151-167.

Pella, J. & Masuda, M. (2006) The Gibbs and split-merge sampler for population mixture analysis from genetic data with incomplete baselines. Canadian Journal of Fisheries and Aquatic Sciences 63, 576-596.

Rengmark, A.H., Slettan, A., Skaala, O., Lie, O., Frode, L. (2006) Genetic variability in wild and farmed Atlantic salmon (*Salmo salar*) strains estimated by SNP and microsatellites. Aquaculture, 253, 229-237.

Smith, C.T., Templin, W.D., Seeb, J.E., Seeb, U.W. (2005). Single nucleotide polymorphisms provide rapid and accurate estimates of the proportions of US and Canadian Chinook salmon caught in Yukon River fisheries. North American Journal of Fisheries Management 25 (3): 944-953.

Taberlet P, Griffin S, Goossens B *et al*. (1996) Reliable genotyping of samples with very low DNA quantities using PCR. Nucleic Acids Research, 24, 3189–1394.

Vähä, J-P, Erkinaro, J., Niemela, E., Primmer, C.R. (2007) Life-history and habitiat features influence within-river genetic structure of Atlantic salmon. Molecular Ecology, in press.

Waples, R.S. (1990) Temporal changes of allele frequency in Pacific salmon: implications for mixed-stock fishery analysis. Canadian Journal of Fisheries and Aquatic Sciences 47, 968-976.

7.7 Appendix 1 Quality Control Measures

Sources of errors

The numerous sources of scoring errors in microsatellite data have been thoroughly discussed by Taberlet *et al*. (1996), Bonin *et al*. (2004) and Hoffman & Amos (2005). These errors may arise in a number of ways. When the template DNA is of low quantity and/or quality, as is typical of studies employing noninvasive tissue-sampling, PCR amplification can become unreliable. A common problem is the stochastic failure of one allele to amplify, leading to heterozygotes appearing to carry only one allele (**'allelic dropout'**). Another source of artefact is the incorrect calling of alleles due to the presence of **'stutter bands'**, generated by slippage of *Taq* polymerase during PCR. The magnitude and shape of stuttering patterns varies across loci, with some markers displaying very little stuttering, and others consistently producing two or more stutter peaks. Interpreting patterns at stutter-prone loci becomes particularly difficult in the case of adjacent-allele heterozygotes at loci having a dinucleotide repeat motif. Such stutter can cause these heterozygotes to be scored as homozygotes for the larger allele. Consistent mistyping of this form will bias allele frequencies towards larger alleles and decrease observed heterozygosity. It is also possible to mistype a true homozygote as an adjacent-allele heterozygote, but these errors are less likely if the marker has been screened sufficiently, and the shape of a single allele is well known. However, even when large quantities of high-quality DNA are available, serious genotyping errors may still occur. These include e.g. large-allele dropout, 'null alleles' (allele non-amplification), errors due to electrophoresis artefacts, mis-scoring of allele banding patterns etc.

Large-allele dropout

Large-allele dropout results from the preferential amplification of the smaller allele in a heterozygous genotype and differs from allelic dropout in low-quality samples as it is a function of allele size (increases with allele size and is more prevalent in loci with large differences in allele sizes) and not a stochastic sampling error of template. If undetected, large-allele dropout will cause allele frequencies of shorter alleles to be overestimated and may result in rare large alleles being omitted from the data s*et al*together. Also, as with mistyping due to stuttering, large-allele dropout will decrease observed heterozygosity and increase the apparent level of inbreeding at affected loci.

Null alleles

Null-alleles result typically from primer binding site mutation and are particularly difficult to detect because they fail to produce a visible product. When present in a data set, samples that are heterozygous for a null allele will be mistyped as homozygous for the alternate, visible allele, whereas samples that are homozygous for the null allele will appear to have a failed reaction. Similar to the scoring errors described above, null alleles will bias allele frequencies (visible alleles will be overestimated), decrease observed heterozygosity and increase the apparent level of inbreeding.

Preventing and mitigating scoring errors

Many of the procedures used in preventing scoring errors are expensive and time-consuming (Taberlet *et al*. 1996; Ewen *et al*. 2000; Bonin *et al*. 2004), and others are not appropriate for studies of species with poorly described genomes or when pedigree information is not available. Dewoody *et al* (2006) recommended a protocol that involved implementing quality assurance procedures at different stages of a study:

1) careful screening of microsatellite loci prior to data collection allows suspect loci (those producing inconsistent or difficult to interpret patterns) to be further tested or omitted; if scoring errors due to null alleles, stuttering, or large-allele dropout are detected, primers should be redesigned and/or PCR conditions optimized to reduce the overall error rate;

2) reanalysing a subset of samples provides the opportunity to identify and quantify scoring errors;

3) when scoring the data set, combining automated allele calling with visual inspection of each sample is recommended to avoid additional scoring errors, e.g. inconsistent binning of alleles;

4) testing for scoring errors by available statistical methods or software, e.g. ML-NullFreq (Kalinowski & Taper 2006) or MICRO-CHECKER (van Oosterhout *et al*. 2004);

5) mitigating errors in downstream analyses by reanalysing or dropping one or a few problematic loci.

Other measures relevant to GSI and individual assignments:

- microsatellite loci having a tetranucleotide repeat motif are preferred to loci having a dinucleotide repeat motif as they are easier to score and typically display no stuttering patterns;
- to guarantee consistency of genotyping results, calibration of allele sizes among participating labs is warranted by exchange of reference samples;
- baseline samples should be preferably genotyped in a single lab

7.8 References

Bonin A, Bellemain E, Bronken Eidesen P, Pompanon F, Brochmann C, Taberlet P (2004) How to track and assess genotyping errors in population genetic studies. Molecular Ecology, 13, 3261–3273.

Creel S, Spong G, Sands JL *et al*. (2003) Population size estimation in yellowstone wolves with error-prone noninvasive microsatellite genotypes. Molecular Ecology, 12, 2003–2009.

Dewoody J, Nason JD, Hipkins VD (2006) Mitigating scoring errors in microsatellite data from wild populations. Molecular Ecology Notes, 6 , 951–957.

Ewen KR, Bahlo M, Treloar SA *et al*. (2000) Identification and analysis of error types in high-throughput genotyping. American Journal of Human Genetics, 67, 727–736.

Hoffman JI, Amos W (2005) Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. Molecular Ecology, 14, 599–612.

Kalinowski ST, Taper ML (2006) Maximum likelihood estimation of the frequency of null alleles at microsatellite loci. Conservation Genetics, 7, 991-995.

van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) micro-checker: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes, 4, 535–538.

Taberlet P, Griffin S, Goossens B *et al*. (1996) Reliable genotyping of samples with very low DNA quantities using PCR. Nucleic Acids Research, 24, 3189–1394.

Annex 1: Proposed Draft Resolutions for 2008

The **Working Group on The Application of Genetics in Fisheries and Mariculture** [WGAGFM] (Chair: E. Eg Nielsen, Denmark) will meet in Pitlochry, Scotland in 2008 to:

a. Review the potential for application of SNP´s (single nucleotide polymorphisms) in fisheries genetics and aquaculture;

b. Review current and future prospects of QTL based studies in fisheries and aquaculture;

c. Update progress on the establishment of a meta-data base for genetic data

d. Review progress for optimizing the storing of otoliths and scales;

e. To evaluate prospects for genetic monitoring for evaluating the conservation status, intraspecific biodiversity and population health in fishes.

Supporting Information

Priority:	The current activities of this Group will lead ICES into issues related to the ecosystem affects of fisheries and mariculture, especially with regard to the application of the Precautionary Approach. Consequently these activities are considered to have a very high priority.
Scientific Justification and relation to Action Plan:	Action Plan references: a)-2.5, -2.6 b) -2.5, - 1.10, c)-1.10 , d)-1.10,-3.7 e) 1.10 Term of Reference a) Single nucleotide polymorphisms (SNPs) seem to have become the marker of choice for most genetic studies, and the marker have had an increasing popularity in population genetic studies. These markers can be analysed in large numbers and access variability around the chromosomes. Most present population genetic studies normally include a number of SNPs. The Gene Conservation Laboratory in Alaska routinely performs SNP analyses during the fishing season for stock identification studies. The laboratory provides inseason estimates of the composition of the Cook Inlet sockeye salmon commercial harvest and of the Kenai River to aid in the management of the drift and set net fisheries. WGAGFM acknowledge the need to review the present state of these markers, including the possible applications and limitations, in addition to possibilities for fisheries management. (lead: J. G. Dahle, T. Johansen) Term of Reference b) An increasing number of studies aim to identify quantitative trait loci (QTLs) in species of interest for fisheries and/or aquaculture. Such studies imply the availability of medium to high density linkage maps, informative biological material and high through-put genotyping facilities. Until now, most of these studies relate to marker assisted selection (MAS) in species of major aquaculture interest. QTL mapping is also of more general interest to better understand the genetic architecture of quantitative traits. The identity and number of loci controlling quantitative trait variation are indeed central to the understanding of their evolutionary potential and patterns of population differentiation. However, the usage of QTL-based approaches in the assessment of genetic variability in adaptive traits and for prediction of trait values from known QTLs in natural populations of fish and shellfish remains largely untested. We will review the present status and prospects of QTL mapping in fisheries in aquaculture in the light of the current research in plant and animal genetics and breeding. (lead: E Gosling, P. Boudry)

	Term of Reference c) An initiative was put in place as part of the response to the 2006-2007 ToR e) to set up a metadata base on genetic information related to Atlantic salmon stocks. This ToR will provide an update on this initiative. (Lead E. Verspoor, J. Martinsohn). Term of Reference d) Understanding the impact of fishing pressure and climate change on a fish population relies on identifying biological changes incurred by that population over time. Such a task is potentially possible for fisheries molecular geneticists due to the availability of unique sources of time series data in the form of scales and otoliths collected over the decades and subsequently stored in cupboards. Technical advances mean DNA can now be extracted and analysed from these hard tissues, and in some instances the collections provide very comprehensive time series of molecular data. The WGAGFM accepts that such sources of data are invaluable and that the optimal conditions for storage of these samples be considered as a priority. At present samples are generally stored in uncontrolled environmental conditions. This ToR aims to identify optimal conditions for the storage of these tissues to ensure minimum DNA degradation until such time as the samples are used. (lead: M. O´Sullivan, E. Cross, E. Gosling). Term of Reference e) Advances in molecular biology over the last decade provide fisheries geneticists with cost-effective tools for resolving unprecedented levels of genetic diversity within the genomes of marine fish and shellfish species. Analysis of the amount and distribution of this diversity can be highly informative, not only as regards the structuring of a species into breeding populations but also, potentially, about the reproductive status of the populations themselves. Departures of the distribution association of alleles within and among loci from random reflect historical levels of mutation, natural selection, gene flow, patterns of mating, numbers of breeders and variance in the contribution of breeders to the next generation family survival to breeding (genetic drift). Theory indicates that these factors, and how they have acted historically, have characteristic effects on the pattern of distribution of allelic variation within and among individuals and loci, in a population. Using recently developed Bayesian statistical methods it is now possible, in principal, to assess the pattern of distribution to determine parameters such the effective numbers of breeders, effective population sizes, and levels of gene flow, in addition to classic indicators such as allelic diversity and heterozygosity. Estimation of these parameters using cost-effective molecular markers offers a potentially valuable approach to monitoring the reproductive and evolutionary “health” of a breeding population, and therefore has the potential to be a valuable management tool. The extent to which this is, and is likely to be, the case in practise is considered in this ToR. (lead: E. Verspoor, P. McGinnity)

Annex 2: List of participants

The Working Group on the Application of Genetics in Fisheries and Mariculture (WGAGFM)

Ispra, Italy 19–23 March 2007

NAME	ADDRESS	PHONE	FAX	EMAIL
Dorte Bekkevold	Danish Institute for Fishery Research Department of Inland Fisheries Vejlsøvej 39 DK-8600 Silkeborg Denmark			db@difres.dk
Pierre Boudry	IFREMER La Tremblade Station Ronce-les-Bains P.O. Box 133 F-17390 La Tremblade France			pierre.boudry@ifremer.fr
Mary Cross	University College Cork Department of Zoology, Ecology and Plant Science Distillery Fields North Mall Ireland			m.cross@ucc.ie
Tom F. Cross	University College Cork Department of Zoology, Ecology and Plant Science Distillery Fields North Mall Ireland	+353 21 4904191	+353 21 4270562	t.cross@ucc.ie
Geir Dahle	Institute of Marine Research Nordnesgt 33 P.O. Box 1870 N-5817 Bergen Norway	+47 55 238500	+47 55 238531	Geir.Dahle@imr.no
Eileen Dillane	University College Cork Department of Zoology, Ecology and Plant Science Distillery Fields North Mall Ireland	+353 21 4904343/4904368	+353 21 4270562	e.dillane@ucc.ie
Elizabeth Gosling	Galway-Mayo Institute of Technology Dublin Road Galway Ireland			
Riho Gross	Estonian University of Life Sciences Institute of Animal Science Kreutzwaldi 1 EE-51014 Tartu Estonia	-313124	-313064	rgross@eau.ee

NAME	ADDRESS	PHONE	FAX	EMAIL
Jakob Hemmer-Hansen	Danish Institute for Fishery Research Department of Inland Fisheries Vejlsøvej 39 DK-8600 Silkeborg Denmark	33963129		jhh@difres.dk
Torild Johansen	University of Bergen Department of Biology P.O. Box 7800 N-5020 Bergen Norway			torild.johansen@ifm.uib.no
Ellen L. Kenchington	Dept. of Fisheries & Oceans Bedford Institute of Oceanography 1 Challenger Rd P.O. Box 1006 B2Y 4A2 Dartmouth NS Canada			kenchingtone@mar.dfo-mpo.gc.ca
Marja-Liisa Koljonen	Finnish Game and Fisheries Research Institute Viikinkaari 4 P.O.Box 2 FI-00791 Helsinki Finland			marja-liisa.koljonen@rktl.fi
Jann Thorsten Martinsohn	DG Joint Research Centre JRC.G.3 - AGRIFISH Unit P.O. Box 051 IT-21020 Ispra (Va) Italy			jann.martinsohn@jrc.it
Einar E. Nielsen	Danish Institute for Fishery Research Department of Inland Fisheries Vejlsøvej 39 DK-8600 Silkeborg Denmark			een@dfu.min.dk
Martha O' Sullivan	Fisheries Research Services FRS Marine Laboratory 375 Victoria Road P.O. Box 101 AB11 9DB Aberdeen United Kingdom	+44 (0)1224 295326	+44 (0)1224 295511	osullivanm@marlab.ac.uk
Jochen Trautner	Federal Research Centre for Fisheries Institute for Fishery Ecology Palmaille 9 D-22767 Hamburg Germany			jochen.trautner@ifo.bfa-fisch.de
Eric Verspoor	Fisheries Research Services FRS Marine Laboratory 375 Victoria Road P.O. Box 101 AB11 9DB Aberdeen United Kingdom			verspoor@marlab.ac.uk

NAME	ADDRESS	PHONE	FAX	EMAIL
Anna Was	Galway-Mayo Institute of Technology Molecular Ecology Research Group Galway Ireland			anna.was@gmit.ie
Roman Wenne	Institute of Oceanology Powstancow Warszawy 55 PO Box 68 PL-81-712 Sopot Poland	+48 58 551 72 83		rwenne@cbmpan.gdynia.pl