

ICES WGAGFM Report 2008

ICES Mariculture Committee

ICES CM 2008/MCC:04

REF. ACOM

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

1–4 April 2008

Pitlochry, Scotland, UK



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. 2008. Report of the Working Group on the Application of Genetics in Fisheries and Mariculture (WGAGFM), 1–4 April 2008, Pitlochry, Scotland, UK. ICES CM 2008/MCC:04. 77 pp.

For permission to reproduce material from this publication, please apply to the General Secretary.

The document is a report of an Expert Group under the auspices of the International Council for the Exploration of the Sea and does not necessarily represent the views of the Council.

© 2008 International Council for the Exploration of the Sea

Contents

Contents	i
Executive summary	1
1 Introduction.....	3
1.1 Attendance.....	3
1.2 Venue.....	3
1.3 Meeting Format.....	3
2 ToR a): Review the potential for application of SNP's (single nucleotide polymorphisms) in fisheries genetics and aquaculture	4
2.1 Background	4
2.2 SNP discovery	6
2.3 SNP genotyping	7
2.4 Comparison with other markers.....	8
2.5 Range of Applications.....	9
2.6 Individual multi-locus genotype based analysis.....	10
2.7 Case studies – salmon	11
2.8 Case studies – cod.....	11
2.9 Case studies – other species	12
2.10 Recommendations	13
2.11 References	14
3 ToR b): Current and future prospects of QTL-based studies in fisheries and aquaculture.....	17
3.1 Introduction.....	18
3.2 Linkage Maps.....	18
3.3 Identifying QTL	19
3.3.1 Estimation of number of major genes influencing a trait.....	19
3.3.2 QTL mapping	19
3.3.3 Experimental design to map QTLs.....	19
3.3.4 Statistical methods.....	20
3.4 QTL mapping in aquatic species	20
3.5 Marker-Assisted Selection (MAS)	24
3.6 Potential application in wild populations.....	24
3.7 Recommendations	25
3.8 References	26
4 ToR c): Update progress on the establishment of a meta-database for genetic data on fish and shellfish genetics covered under the ICES remit	34
4.1 Context and relevance.....	34
4.2 Progress and further action	35

4.3	Development and implementation strategy	35
4.4	Recommendations	38
5	ToR d): Review progress for optimising the storing of otoliths and scales for DNA analysis.....	42
5.1	Introduction.....	42
5.2	Traditional Methods.....	42
5.3	Location of DNA.....	43
5.4	Previous Genetic Studies	43
5.5	Degradation of DNA.....	43
5.6	Optimal Methods.....	44
	5.6.1 Archival collections	44
	5.6.2 New Collections	44
5.7	Multiple Uses	45
5.8	Conclusions	45
5.9	Recommendations	45
5.10	References	46
6	ToR e): To evaluate prospects for the use of genetic monitoring for evaluating conservation status, intraspecific biodiversity, and stock “health” in fishes.....	49
6.1	Rationale	49
6.2	Breeding populations and fisheries.....	49
6.3	Structuring into breeding populations	50
6.4	Genetic assessment and monitoring	54
6.5	Conclusions	59
6.6	References	60
6.7	Recommendations	62
7	Group Business.....	63
7.1	Draft Resolutions for 2009	63
7.2	Meeting Places in 2009 and 2010.....	63
	Annex 1: List of participants.....	64
	Annex 2: Proposed Draft Resolutions for 2009	67
	Annex 3: Recommendations	70

Executive summary

The Working Group on the Application of Genetics in Fisheries and Mariculture (WGAGFM) met at Pitlochry, Scotland 1–4 April 2008. The meeting was very well attended; with a total of 25 representatives present from 11 countries (14 national delegates and 11 experts appointed by the Chair).

- Five Terms of Reference (ToRs) were on the agenda for 2008. The first issue addressed was to review the application of a new class of genetic markers, “SNP’s” (Single Nucleotide Polymorphisms), in fisheries genetics and aquaculture. The molecular genetic revolution has facilitated the investigation of a high number of genetic markers (SNP’s) spread throughout the genome of an organism. SNP’s have revolutionised many other fields of genetics, and is also expected to gain wider application and become “state of the art” for many purposes in fisheries genetics and aquaculture. For instance, the application of SNP analysis allows analysis of both demographic processes as well as natural and human induced selection. Likewise, SNP’s are very well suited for analysis of DNA from historical samples thereby facilitating genetic monitoring of populations. Still, different technical platforms for scoring SNP variability have to be assessed for reliability and cost effectiveness. Likewise, the development of new analytical tools to make full use of the large datasets generated should be encouraged. Finally, the group does not recommend to abandon previously employed genetic methods, but to critically evaluate the best suited method for different purposes.
- The second term of reference was to review current and future prospects of QTL (Quantitative Trait Loci) based studies in fisheries and aquaculture. There was a general agreement that QTL studies should be encouraged and supported as they are one of the most direct ways to understand the genetic basis of phenotypic variation, linking classic quantitative genetic and genomic studies. The application of QTL studies should, however, not only be restricted to Marker Assisted Selection (MAS). More effort should be diverted for identification of QTL in a wider variety of aquatic species. To aid this identification the current development of genomic resources – notably linkage and physical maps, genomic libraries and whole genome sequences – should be encouraged. Likewise, the development and maintenance of divergent lines, family data, or other biological material of interest for QTL mapping should be encouraged.
- The 3rd item on the agenda was to update progress on the establishment of a meta-data base for genetic data, which was a follow up from a ToR from 2007. The group has identified that a centrally administered, web accessible meta-data base could prove vital for integration among researchers within the ICES remit. The meta-database should serve as a portal cataloguing relevant primary information with respect to generated population genetic data, primary and secondary research reports, available biological samples, and point to the locations and contact points from which to get such data, samples and extended information. The WGAGFM should agree before September 2008 which types of meta-data to include and the database structure, as well as the functionalities of the interface providing access to the data and about other features to be made available.

Positive feedback from both ICES and the European Commission on the development and hosting of the data-base was achieved.

- The next ToR was to review progress for optimizing the storing of otoliths and scales. Historical collections represent an invaluable source for conducting retrospective genetic analysis and, accordingly, there is an urgent need to establish optimal storage conditions which will facilitate future use of such samples. It is recommended that ICES should request information on all scale and otolith collections held by fisheries laboratories, institutes and universities. This information should be accessible via the proposed ICES genetic database (see above). If disposal of archived material is envisaged, the scientific community should be notified via the ICES network. Archived material should be inspected on a regular basis and subsequently stored in controlled conditions. It should be borne in mind that most fungi prefer dark humid environments. Thus dry storage conditions are vital. Preferably a non-mineralised tissue should be sampled in the future, (e.g. a fin clip), solely for genetic analysis, due to the clear benefits of molecular approaches in fisheries biology. This is also important because of the competing uses for scale and otolith material (e.g. microchemistry).
- The final issue was to evaluate prospects for genetic monitoring for evaluating the conservation status, intraspecific biodiversity and population health in fishes. The group agreed that monitoring of fish stocks should be based on genetic as well as on demographic metrics, which are informative with regard to the status of their contributing breeding populations as this will help to provide insights needed for sustainable management. The insights should be generated by building working models of local, regional and global breeding population structure in commercial fish species of interest to ICES, using existing information on their breeding biology and genetics, to help realise management based on breeding populations and focus research directed at model improvement (a ToR to conduct this exercise for cod has been suggested for 2009). Genetic studies should be incorporated into existing fisheries research programmes to advance understanding of population structuring and to help assess the benefits of using genetic metrics to assess population status. Finally, the group strongly recommends that ICES promotes the establishment of direct links between individual species assessment working groups and the WGAGFM to facilitate the above.

1 Introduction

The Working Group on the Application of Genetics in Fisheries and Mariculture (WGAGFM) met at Pitlochry, Scotland from 1–4 April 2008. The ToRs were decided in Council Resolutions adopted at the ICES Statutory meeting held in Copenhagen, Denmark in 2007. Dr. E. E. Nielsen (Denmark) chaired the meeting, which opened at 0900 h on Tuesday, 1 April and closed at 12.00, Friday, 4 April 2008.

1.1 Attendance

Twenty-five persons from eleven countries (Canada, Denmark, Estonia, Latvia, France, Germany, Ireland, Italy, Norway, Poland, and United Kingdom) attended the meeting (Annex 2). Fourteen were official members (or substitutes) for their countries and eleven were appointed experts by the chair for 2008. The latter were registered with ICES prior to the meeting.

1.2 Venue

The meeting was held at the FRS Freshwater Laboratory, Pitlochry Scotland. The WG wishes to express their appreciation to the local hosts Dr. Eric Verspoor and Dr. John Gilbey and the rest of the staff at the institute for their kind hospitality. The meeting venue was ideal with accommodation available at the “castle” in Pitlochry and “taxi service” provided by our local hosts which took us to the Laboratory in the morning and to the hotel in the 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): Review the potential for application of SNP's (single nucleotide polymorphisms) in fisheries genetics and aquaculture

Geir Dahle, Torild Johansen, Anti Vasemägi, Gary Carvalho, Ann-Britt Florin, Anna Was, Paolo Prödohl. Anastasia M. Khrustaleva

2.1 Background

One of the more recent additions to the “molecular toolbox” is single nucleotide polymorphisms (SNPs), which comprise a usual alternative of only two possible nucleotides at a given position. Much of the continuing quest for new molecular markers stems from the realisation that no individual marker is ideal for all applications, and it is important to appreciate from the outset that SNPs are no exception to this. As new markers become available, it is necessary to assess their characteristics and potential utility, and to match them to appropriate questions, especially to identify new opportunities that become tractable.

SNP genotyping is rapidly becoming a powerful tool for assessing genetic variation in natural populations (reviewed in Brumfield *et al.*, 2003). Recent applications show that SNPs are extremely frequent and relatively easy to ascertain in many non-model organism genomes (Elfstrom *et al.*, 2006; Morin *et al.*, 2007a), and can be applied to a wide range of population studies, from individual identification to population structure and taxonomy (Glaubitz *et al.*, 2003; Seddon *et al.*, 2005; Elfstrom *et al.*, 2006). An important attribute of SNPs is that they are bi-allelic markers and not based on length polymorphism. This facilitates comparison of data from different sources and reduces the need for cross-laboratory calibration. Studies on conservation and management of natural resources will benefit from using markers that allow datasets to build over decades, and that are amenable to augmentation, both spatially and temporally. Microsatellites can be constrained in long-term studies that require addition of data from multiple laboratories because of the difficulties in calibrating size-based alleles, often of only a few differences in nucleotide size (Broquet and Petit, 2004; Hoffman *et al.*, 2006). A major attribute of SNPs is that SNP genotypes are based on detection of nucleotide sequence differences rather than PCR product size differences, so that genotype data are universally comparable and portable, without the need to include common controls, as in the case of size-based markers. Other benefits of using SNPs relative to other nuclear markers such as microsatellites include ease and efficiency of discovery and genotyping (e.g. Elfstrom *et al.*, 2006; Morin *et al.*, 2007a; Hinten *et al.*, 2007), ability to target variation in random genomic regions or known genes (Aitken *et al.*, 2004; Kohn *et al.*, 2006), and analytical tools and methods for assessing power and population parameters (Ryman 2006; Ryman and Palm 2006). The potential application of SNPs to degraded, historical, and ancient samples has been discussed (Asher and Hofreiter 2006; Poinar *et al.*, 2006; Römpler *et al.*, 2006), and there are some examples to date (Römpler *et al.*, 2006, Consuega *et al.*, 2002, Leonard *et al.*, 2000).

SNPs are a class of genetic markers that is well suited to a broad range of research and management applications, and has recently been used in various areas of ecology, evolution, conservation, and aquaculture (Brumfield *et al.*, 2003; Morin *et al.*, 2004). SNPs are the most frequent type of variation and thus provide the unique possibility to screen a larger part of the species' genome. SNPs have the advantages that they are found in coding and non-coding areas, whereas most microsatellites used in population genetics, for example, are typically in non-coding regions of the genome that is expected to be less influenced by selection. Based on direct sequencing

of 27 genes (not random ESTs) the frequency of SNPs in Atlantic salmon (*S. salar*) has been estimated to be 1/680 base pair (bp). The frequency varied between introns (1/405bp) and exons (1/1448 bp) (Ryynänen and Primmer 2006). The frequency of single nucleotide substitutions at the origin of SNPs has been estimated to between 1×10^{-9} and 5×10^{-9} in mammals, but 1–10 per 1000bp in humans (Lee *et al.*, 2005). The two most used main approaches is to develop SNPs either from known genes (target genes; Smith *et al.*, 2005a) or develop a large number of SNP markers from the alignment of Expressed Sequence Tag (EST) (Moen *et al.*, 2008). A limitation of the targeted-gene approach is that SNP discovery is restricted to the genes for which sequence data are available from sister taxa (Elfstrom *et al.*, 2005). EST based SNPs uncovers usually a larger number of loci, but not all loci might be identified in the model species (Moen *et al.*, 2008; Johansen *et al.*, submitted).

Table 1. Characteristics of different molecular markers and their applicability using various analysis methods (from Vignal *et al.*, 2002).

Marker name	Variation type			Information content		
	SNP ¹	Indel ²	VNTR ³	2 dominant alleles	2 co-dominant alleles	Multi allelic co-dominant
RFLP	+	(+) ⁴	(+)	–	+	(+) ⁵
PCR-RFLP	+	(+) ⁴	(+)	–	+	–
RAPD	+	(+) ⁴	(+)	+	–	–
AFLP	+	(+) ⁴	(+)	+	(+) ⁶	–
SSCP	+	(+) ⁴	(+)	–	+	(+) ⁵
Microsatellite	–	(+) ⁷	+	–	–	+
SNP	+	(+) ⁸	–	–	+	–

¹ Single nucleotide polymorphism: any kind of base substitution. The fact that SNPs appear both as a variation type and a marker name, is due to the fact that in reality, many genotyping techniques used for genotyping SNPs are grouped under this generic marker name.

² Insertions and deletions.

³ Variable number of tandem repeats.

⁴ Although the RAPD, AFLP, RFLP, PCR-RFLP and SSCP techniques will detect base substitutions in the vast majority of cases, the two other types of DNA variation can also be analysed.

⁵ In some instances, more than two alleles can be analysed.

⁶ With an automatic sequencer, some markers can be scored as co-dominant.

⁷ Variations in PCR product length can be due to a deletion in the sequence flanking the microsatellite.

⁸ Many SNP detection techniques can also be used for scoring small insertions or deletions (indels).

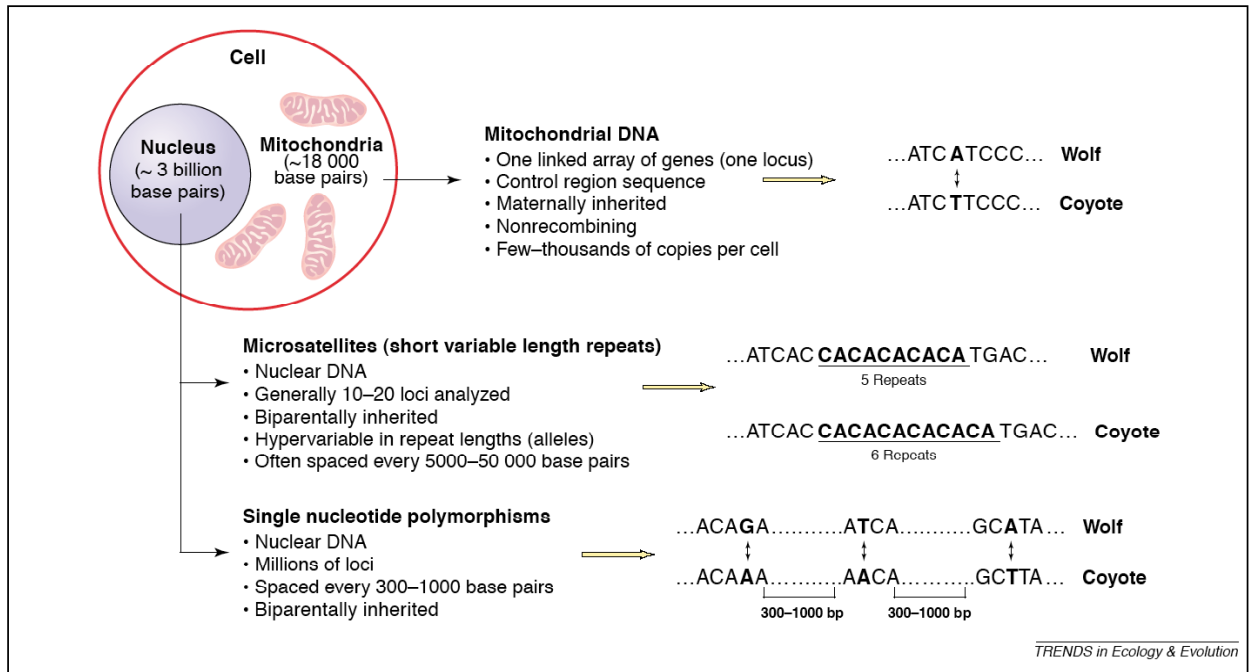


Figure 1. Comparison of the characterization of mtDNA, microsatellites and single nucleotide polymorphism (SNPs) as genetic markers. The figure show example of possible DNA sequence differences between wolf and coyote; these types of differences can be found within or between taxa). (from Morin *et al.*, 2004).

Most population genetic statistic theory is based on neutral markers. Thus, markers that are suspected to be under the influence of selection are often discarded for subsequent estimation of population parameters including evolutionary relationships (Luikart *et al.*, 2003). While markers displaying strong population differentiation might be useful in the delimitation of units of management, full evaluation of their statistical properties should be undertaken.

2.2 SNP discovery

There are many available approaches for SNP discovery. The simplest possibility is comparison of PCR products from different individuals using direct sequencing of target fragment. A very convenient method is comparison of sequence data from EST production projects, which is potentially effective for discovering large numbers of SNPs. Alternatively we can capitalize on the use of sequences which are readily available in genomic databases. There are also many other non-sequencing methods to identify new SNPs such as Denaturing High Performance Liquid Chromatography (DHPLC), Denaturing Gradient Gel Electrophoresis (DGGE), Temperature Gradient Gel Electrophoresis (TDGE), Single-Stranded Conformation Polymorphism analysis, mismatch-targeting using endonucleases such as CEL I (also known as EcoTILLING) or transposons (Orsini *et al.*, 2007).

Ascertainment bias in SNP discovery:

An appropriate sampling design in population genetic studies is of particular relevance for the identification of SNPs (Renwick *et al.*, 2003, Clark *et al.*, 2007). In most instances, SNPs are identified by sequencing regions of nuclear and/or mtDNA genome for a limited number of individuals. Considering that the probability that a

SNP is identified in a particular sample is a function of its frequency, the use of inadequate sample sizes and/or geographical coverage during SNP discovery, will result in misrepresentation of true levels of genetic variation (i.e. a number of potentially relevant variations will not be detected). A primary consequence of this ascertainment bias is that any population genetic statistics that rely on frequency data including nucleotide diversity and index of population substructuring (e.g. F_{st}) will be potentially compromised. For instance, because several SNPs might be missing, the average heterozygosity estimates for polymorphic sites will be overestimated. Similarly, the overall average heterozygosity estimates (across all sites encompassing the assayed SNPs) will be underestimated. Insofar as population structuring is concerned, ascertainment bias will result in an underestimation of among-population genetic heterogeneity because the most common SNPs are more likely to be shared among populations. Clarke *et al.* 2007 have extensively discussed this particular problem during an empirical study in human populations using different sets of SNPs. The authors suggested that careful consideration is required in the identification and use of SNPs in population genetic studies. The most appropriate strategy in many cases will therefore be to collect individuals from a representative geographic or ecological range of a species, thereby maximising the inclusion of any population-specific polymorphisms.

2.3 SNP genotyping

Since SNP genotyping technologies vary widely, they allow a choice of systems to match the cost, throughput and equipment capacity of each laboratory. Techniques can vary from simple and standard (e.g. electrophoresis-based systems) to highly multi-plexed systems for rapid screening (e.g. microarrays). One key feature of most SNP genotyping techniques, apart from those based on direct hybridisation, is the two step separation: 1. generation of allele-specific molecular reaction products; 2. separation and detection of the allele specific products of their identification.

The enzyme-based methods utilize a broad range of enzymes including DNA ligase, DNA polymerase and nucleases (e.g. flap endonuclease that catalyzes structure-specific cleavage). They include Tetra-primer ARMS-PCR, primer extension, in which two oligonucleotides are used, each with a 30 nucleotide complementary to one of the SNP alleles. Since only perfectly matched oligonucleotides will prime DNA polymerase extension with dNTPs (allele specific PCR), 'mini-sequencing' reactions, in which DNA polymerase extends the hybridized primer by adding a base that is complementary to the SNP nucleotide, Taqman assay in which Taq DNA polymerase's 5'-nuclease activity is used for SNP genotyping, oligonucleotide ligation assay (OLA), invasive cleavage of oligonucleotide probes (invader assay), next-generation sequencing technologies such as pyrosequencing etc.

The hybridization-based methods include dynamic allele-specific hybridization (DASH) that is based on the differences in the melting temperature in DNA that results from the instability of mismatched base pairs from hybridization signals. Most hybridization techniques are derived from the Dot Blot, in which DNA to be tested, either genomic, cDNA or a PCR reaction, is fixed on a membrane and hybridized with a probe, usually an oligonucleotide. In the Reverse Dot Blot technique, it is the oligonucleotide probes that are immobilized. When using allele specific oligonucleotides (ASOs), genotypes can be inferred. Different kinds of fluorescent labelled probes, such as molecular beacons can be used. Dozens or even thousands of the probes can be placed on a small chip, allowing for many SNPs to be genotyped simultaneously.

Depending on a specific requirements of the analyses it is possible to choose most appropriate SNP genotyping platform that either maximises the sample throughput allowing genotyping of thousands of individuals or the number of SNPs genotyped allowing to screen thousands of SNPs in relatively limited number of samples. For example, in snow crab Smith and co-workers 2005 were able to screen XX SNPs in 384 samples in five minutes. On the other side of the scale, several commercial array-based genotyping technologies enable analysis of over 100 000 SNPs in model organisms but such high-throughput genotyping is usually applicable to a limited number of samples.

2.4 Comparison with other markers

The fact that SNPs are bi-allelic markers could be considered a step backwards when compared to the highly informative multi-allelic microsatellites, but there are several major drawbacks with microsatellites and other markers presently used in genetic studies. Alternative marker types like allozymes, and microsatellites suffer from attributes that negatively impact either through-put rates or inter-laboratory standardization, and would lead to development of redundant data bases. The problem of defining, often numerous, microsatellite alleles often with small length-based differences at a locus, poses a major disadvantage in the use of these markers. One common error is the misinterpretation of microsatellite genotypes. Especially problematic is distinguishing between homozygous and adjacent-allele heterozygous genotypes, and large allele drop-out. The fact that it is difficult to compare microsatellites data produced by different laboratories, due to inconsistencies in allele size calling, is usually not a problem for family based studies, but it can be a serious issue in between population comparisons. It is believed that these inconsistencies are mainly due to variety in sequencing machine/platform, fluorescent dye and allele calling software.

When all population genetic and analytical considerations are considered, SNPs seems superior to microsatellites for elucidating historical demography. Mutation rates at microsatellite loci are difficult to estimate, and vary across loci and across alleles within the same locus. Perhaps most importantly for comparative studies of historical demography, the difference in evolution and variation of the same microsatellite locus in even closely related species make them ill suited for interspecific comparisons of genomic variability. Many more tests for deviations from neutrality, for population size changes and for recombination exist for SNP data than for microsatellites, and the fit of models to data is probably better for SNPs. The recent focus on microsatellites might have caused an unconscious ascertainment bias in the estimation of genomic variability for many species. In addition, measures of population differentiation, such as F_{st} , can be very sensitive to the level of within-population variation, resulting in suspiciously low values in many microsatellite studies (Brumfield *et al.*, 2003). However, because of their low variability, bi-allelic SNP markers are expected to be less effective than microsatellites in detecting recent demographic events such as very recent genetic bottlenecks. In addition, some statistical test for detecting genetic bottlenecks is only suitable for markers that evolve in stepwise manner (Garza and Williamson 2001) and thus, are not suitable for SNP data.

Compared to the microsatellites, SNP data should be easier to standardize across chemistries, hardware platform, and laboratories. They are conducive to high throughput genotyping platforms, and the assays and data are readily transferred among laboratories via unambiguous allele reporting. (i.e. base identities rather than

relative electrophoretic mobilities). The allele nomenclature problem is much simpler in the case of SNPs, for which the results can just be coded as a YES/NO problem, in which each of the two alleles can be simply considered as being present or absent. This simplification in the scoring of alleles will enable the data analysis step of genotyping to be automated to a higher degree than with microsatellites, which still require a great investment of time for reading the data, even with the use of analytical software or automated allele analysis methods. It is also likely that SNPs can provide better resolution to discriminate populations compared to microsatellites because of the effects of size homoplasy in microsatellite markers can be substantial (O'Reilly *et al.*, 2004). The bi-allelic nature of SNPs necessitates the examination of variation at many more loci than is typically the case for microsatellites. A recent study in Atlantic salmon, for example, demonstrated that even a relatively small set of SNP markers can be sufficient for obtaining concordant results compared to microsatellites in various population genetic analyses, although estimates of genetic distance are generally more concordant than estimates of genetic diversity (Ryynänen *et al.*, 2007).

The other advantage of using SNPs as population-level markers is the ability to efficiently target coding and non-coding regions of the genome simultaneously and even to predict the functional importance of the SNP depending on the position of the polymorphism (i.e. amino acid changing, silent, regulatory mutation). Although the presumed neutrality of most microsatellite markers is of use to infer demographic processes (ref), the potential for detecting genes under selection is likely higher when targeting coding regions (Zayed and Whitfield 2007; Barreiro *et al.*, 2008).

2.5 Range of Applications

Most SNP studies in fisheries and aquaculture to date have focused on gene mapping, QTL analysis and functional genomics (refs). Here, we focus our attention on population genetic and individual-multilocus genotype based applications, and review briefly examples that illustrate their performance relative to other marker systems (Table 1).

Using neutral SNPs, population structure can be estimated from genetic distance measures such as F_{st} , R_{st} and Nei's D . The use of such population parameters for SNPs under selection has yet to be fully evaluated. Microsatellites and allozymes have historically been the most commonly used molecular markers in most aquatic species. However, there is an increased interest in the development and use of SNPs in evolutionary and ecological studies of natural populations. Such markers should prove useful not only for population genetic studies, but also for monitoring genetic interactions between resident/anadromous fish, the reconstruction of pedigrees and the application of gene-based tagging techniques (Anderson and Garza, 2006).

There is an increasing requirement for traceability of fish and fish products, both for consumer protection and for regulatory enforcement, in particular with respect to illegal, unreported and unregulated (IUU) fishing. For example, in the UK, the Marine Stewardship Council encourages consumers to eat particular landings of cod that are taken from 'stocks maintained within safe limits'. A traceability system based on regional stocks is necessary to preclude fraudulent allocations. IUU fishing is a serious global problem and one of the main impediments to the achievement of sustainable world fisheries. Traceability is defined by the EU (Anon, 2002) as "the ability to trace and follow a food, feed, food-producing animal or substance intended to be, or expected to be, incorporated into a food or feed, through all stages of

production, processing and distribution". There is an urgent need to identify traceability markers that can be used throughout the food supply chain, from on-board samples, to processed product ("fish to fork"), and which exhibit minimal variance. Furthermore, it is essential that such tools are validated to internationally recognise forensic standards. Only under such stringent conditions can traceability data be used for legal enforcement and as evidence in a court of law.

One option to secure high statistical power is to use a full battery of all available traceability methods. However, to provide laboratories within several European countries with stand-by certified procedures for a wide range of traceability methods would prove difficult and costly. Instead, we think focus should be on those methods with the highest discriminatory power, greatest reproducibility, simplest validation and most flexibility with respect to type of tissue and degree of processing. While most traceability methods can be applied to a whole fresh fish, only DNA analysis can be readily applied to all tissue types throughout the food supply chain. If the same method(s) can be used for monitoring potential changes in the population structure and intrapopulation diversity, this should be considered a highly desirable by-product. On the basis of the above considerations, we conclude that applying forensically validated genetic methods (SNPs), supplemented with the most robust high power non-genetic alternatives, such as otolith morphometrics and microchemistry methods, and is the optimum strategy to facilitate technology uptake and integration across laboratories.

SNPs have recently started to replace or complement microsatellites, especially where the application of individual assignment (IA) or mixed stock analysis (MSA) is most advanced (e.g. Pacific salmonids; Smith, CT *et al.*, 2005 *N Am J Fish Manage* 25, 944–953; Smith, CT *et al.*, 2005 *T Am Fish Soc* 134, 207–217). Although SNP markers have been limited by sampling difficulties, number of available loci and levels of polymorphism, the higher reproducibility and transferability among laboratories for SNPs outweigh the advantages of microsatellites. For traceability and associated forensic applications, it is essential that a tool is employed that exhibits high *reproducibility, comparability and stability*, thereby facilitating precise signatures of identity and collation across databases. SNPs offer the most promising category of molecular genetic markers for such application, both in terms of their ability to discriminate populations and in their relative ease of validation as enforcement tools.

2.6 Individual multi-locus genotype based analysis

Today the molecular markers of choice for parentage inference are highly polymorphic, repetitive loci such as the microsatellites (Queller *et al.*, 1993). In contrast, SNPs have not been widely employed for parentage inference and other forms of relationship estimation, because, possessing only two alleles, each SNP have a lower resolving power per locus than most microsatellites (Glaubitz *et al.*, 2003). In a simulation study Glaubitz *et al.* 2003 explored the potential for relationship estimation with SNP, and their conclusion was that 16–20 "typical" microsatellites would be expected to provide information equivalent to that given by 100 independent SNPs, each with a minor allele frequency of 0.2. Seddon *et al.* 2005 also concluded that the number of SNPs required for population studies will vary considerably based on the species objectives of the study, distinctiveness of the taxonomic units of interest, information content of the species markers available, and the desired probability level.

By selecting SNPs that are highly informative it is possible to increase their statistical power significantly, and SNPs have recently been employed for individual

identification and paternity inference in large herds of cattle (Heaton *et al.* 2002; Werner *et al.*, 2004). Selecting SNPs that had a minor allele frequency of at least 0.1 and that were not closely linked, 32 SNPs gave a probability of identity (PI) of less than 10^{-13} and a resolving power of 99.9% for multi-breed U.S. beef cattle (Heaton *et al.*, 2002). Using the same selection criteria 58 SNPs gave a PI of less than 10^{-13} and a resolving power of 99.99% for European dairy breeds (Werner *et al.*, 2004). SNPs have also been developed for human forensic purposes (Lee *et al.*, 2005), and based on frequency data from 30 unrelated Koreans, the estimated PI and probability of paternity exclusion (PE) with 22 highly informative (allele frequencies close to 0.5) autosomal SNP loci, were 1.905×10^{-10} and 98.9%, respectively, which is comparable to that of nine microsatellite loci for the same population (PI = 2.31×10^{-12} , PE = 99.9%) (Lee *et al.*, 2005).

In an experimental study at the Institute of Marine Research, 20 out of 96 individual wild cod from one fjord sample in Norway were grouped randomly into ten families. The FAP-program (Taggart, 2007) predicted a resolving power for 40 SNP loci to be 0.9847, while 2 microsatellite loci data from the same 20 cod and an additional 20 (altogether 20 families) resulted in resolving power of 1.000 (Dahle, data not shown).

2.7 Case studies – salmon

Genetic markers like allozymes and microsatellites have historically been the most commonly used molecular markers in population studies of salmonids, but there has been an increasing interest in the development and use of SNPs in evolutionary and ecological studies of natural populations of Pacific salmon; chum salmon (Smith *et al.*, 2005b; Elfstrom *et al.*, 2007) coastal steelhead (*O. mykiss*) (Aguilar and Garza, 2007), sockeye salmon (*O. nerka*) (Elfstrom *et al.*, 2006), and coho salmon (*O. kisutch*) (Smith *et al.*, 2006). Indeed, microsatellite-based Mixed Stock Analysis (MSA), though the dominant approach, is beginning to replace MSA and IA in routine studies of population-based management in Pacific salmon, primarily to assure reproducibility and transferability of data.

In a survey of 19 populations of chinook salmon (*O. tshawytscha*) a comparison of microsatellite stock identification power with an existing nine-locus SNP baseline indicated that nine SNPs or one single microsatellite locus can provide accurate and reasonably precise estimates of stock composition to country of origin. In comparisons of population-specific estimation, the nine-SNP baseline was approximately equivalent to a single microsatellite locus with 17–22 alleles (Beacham *et al.*, 2008). While nine SNPs were sufficient for addressing international treaty objectives for *O. tshawytscha* in the Yukon River (Smith *et al.*, 2005c), a larger number of loci will be necessary for analyses of salmon collected on the high seas and for projects in other species. Assuming it is possible to generalize the number necessary in other species, approximately 500 loci would be needed to be sequenced in a species in order to discover an adequate SNP panel (Smith *et al.*, 2005c). However, recent breakthroughs in massively parallel sequencing technologies have considerably increased the efficiency and speed of SNP discovery making it possible to identify thousands of SNPs in very short time (Hudson 2008).

2.8 Case studies – cod

Whereas many markers display significant structuring of cod in the North Atlantic, the question has been raised to which degree the genetic variation reflects breeding structure or the effect of varying selectional constraints. In marine species, the number of alleles at microsatellite loci, which are at present the most used marker in

population genetic studies of cod, is in general high. However, the relatively low levels of F_{st} values observed between populations of cod at microsatellite loci (Knutzen *et al.*, 2003; O'Leary *et al.*, 2007; Sarvas and Fevolden, 2005), might indicate that microsatellites are less suitable for assigning individuals to local populations. Moreover, some microsatellite markers are possibly affected by selection (Nielsen *et al.*, 2006; Westgaard and Fevolden, 2007). To confirm the validity of the suggested genetic structure of cod, it is important to increase the number of variable loci. There is also a need for new markers to be applied in family identification related to the growing cod aquaculture industry.

For cod, as in many marine fishes, no large scale nuclear SNP-based MiSA project has been conducted to our knowledge, though single SNPs have been included in population genetic analysis of marine fishes. The Pan I locus (Pogson, GH, 2001 *Genetics* 157, 317–330) has been widely used for population studies of gadoids, in particular Atlantic cod. In some cases this locus alone can provide almost unambiguous assignment to population (Case, RAJ *et al.*, 2006 *T Am Fish Soc* 135, 241–250). The reason for the high divergence among populations is most likely divergent selection on the two different allele classes identified. The finding highlights the potential of SNPs to detect selection in marine fish with low levels of neutral divergence, either through direct effects on target loci or through hitch-hiking selection. Thus, the identification of SNPs subject to differential selection in marine fish species ⁴³ would greatly improve the power and resolution in relation to traceability

From aligned Expressed Sequence Tags (EST) provided by the Institute of Marine Research in Bergen, 318 loci have been described by Moen *et al.* 2008. Of these loci, 270 were screened for 96 randomly collected wild cod from nine locations along the Norwegian coast from the Skagerrak up to the White Sea, and although the study was based on a limited sample size, the samples reflect the geographic distribution of the Atlantic cod in the northeast Atlantic and should therefore give a good indication on the performance and potential of the selected SNP markers. (Johansen *et al.*, submitted).

2.9 Case studies – other species

Commercial aggregations of the weathervane scallop (*P. caurinus*) are found in Alaskan waters, and growth rates and age at maturation appears to differ between different areas in these waters. In order to improve the knowledge about the genetic stock structure of weathervane scallop in Alaskan waters, SNP markers were developed (Elfstrom *et al.*, 2005).

There is no clear picture of the world wide stock structure of the sperm whale (*P. macrocephalus*), although it is still on the US Endangered Species List and currently protected from commercial hunting under the International Whaling Commission treaties. In order to complement existing mtDNA and microsatellite analysis for the sperm whale a targeted approach was used to detect SNPs in a panel of sperm whale samples resulting in a genotyping assay for 18 independent loci (Morin *et al.*, 2007)

Very few DNA markers have been described for gilthead seabream (*S. aurata*) compared to other species like salmon and sea bass, but Cendelli *et al.* 2007 have introduced SNP markers for use in studies of natural populations in aquaculture and selective breeding programmes of seabream.

2.10 Recommendations

- 1) Prior to inter-laboratory use of SNPs the performance of different platforms available should be evaluated to minimize potential bias in presently unknown error-rates.
- 2) SNP discovery should be based on a panel of individuals that optimises the “ascertainment width” (geographic coverage) and “ascertainment depth” (number of chromosomes and individuals per location) to avoid ascertainment bias.
- 3) Use SNPs in addition to other markers, not as a replacement.
- 4) SNPs should be the preferred markers for establishing long-term genetic datasets and subsequent genetic monitoring.
- 5) As with other markers the extent to which the employed SNPs are influenced by selection (and other evolutionary forces) should be determined.
- 6) Applying a combination of selected and neutral SNPs is possible and is likely to yield novel information on population structure.
- 7) The development and application of new analytical tools should be a research priority for examining large datasets such as generated by SNP studies.

2.11 References

- Aitken, N., Smith, S., Schwarz, C., Morin, P.A. 2004. Single nucleotide polymorphism (SNP) discovery in mammals: a targeted-gene approach. *Molecular Ecology*, 13: 1423–1431.
- Anon. 2002. REGULATION (EC) No 178/2002 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL. http://eur-lex.europa.eu/pri/en/oj/dat/2002/l_031/l_03120020201en00010024.pdf
- Aguilar, A., Garza, J.C. 2008. Primer Note: Isolation of 15 single nucleotide polymorphisms from coastal steelhead, *Oncorhynchus mykiss* (Salmonidae). *Molecular Ecology Resources* 8, 659–662
- Anderson, E.C., Garza, J.C. 2006. The Power of Single-Nucleotide Polymorphisms for Large-Scale Parentage Inference. *Genetics* 172, 2567–2582.
- Asher, R.J., Hofreiter, M. 2006. Tenrec phylogeny and the noninvasive extraction of nuclear DNA. *Systematic Biology*, 55: 181–194.
- Beacham, T.D., Wetklo, M., Wallace, C., Olsen, J.B., Flannery, B.G., Wenburg, J.K., Templin, W.D., Antonovich, A., Seeb, L.W. 2008. The Application of Microsatellites for Stock Identification of Yukon River Chinook Salmon. *North American Journal of Fisheries Management* 28 (1), 283–295.
- Broquet, T., Petit, E. 2004. Quantifying genotyping errors in non-invasive population genetics. *Molecular Ecology*, 13: 3601–3608.
- Brumfield, R.T., Beerli, P., Nickerson, D.A., Edwards, S.V. 2003. The utility of single nucleotide polymorphisms in inferences of population history. *Trends in Ecology and Evolution* 18, 249–256.
- Cenadelli, S., Maran, V., Bongioni, G., Fusetti, L., Parma, P., Aleandri, R. 2007. Identification of nuclear SNPs in gilthead seabream. *Journal of Fish Biology* 70, 399–405.
- Cossins, A.R., Crawford, D.L. 2005. Fish as models for environmental genomics. *Nature Reviews Genetics* 6, 324–333.
- Elfstrom, C.M., Gaffney, P.M., Smith, C.T., Seeb, J.E. 2005. Characterization of 12 single nucleotide polymorphisms in weathervane scallop. *Molecular Ecology Notes* 5, 406–409.
- Elfstrom, C.M., Smith, C.T., Seeb, J.E. 2006. Thirty-two single nucleotide polymorphism markers for high-throughput genotyping of sockeye salmon. *Molecular Ecology Notes* 6, 1255–1259.
- Elfstrom, C.M., Smith, C.T., Seeb, L.W. 2007. Primer Note: Thirty-eight single nucleotide polymorphism markers for high-throughput genotyping of chum salmon. *Molecular Ecology Notes* 7, 1211–1215.
- Ferguson, A. 1994. Molecular genetics in fisheries: current and future perspectives. *Reviews in Fish Biology and Fisheries* 4, 379–383.
- Glaubitz, J.C., Rhodes, O.E., Dewoody, J.A. 2003. Prospects for inferring pairwise relationships with single nucleotide polymorphisms. *Molecular Ecology* 12, 1039–1047.
- Heaton, M.P., Harhay, G.P., Bennett, G.L., Stone, R.T., Grosse, W.M., Casas, E., Keele, J.W., Smith, T.P.L., Chitko-McKown, C.G., Laegreid, W.W. 2002. Selection and use of SNP markers for animal identification and paternity analysis in U.S. beef cattle. *Mammalian Genome* 13, 272–281.
- Hinten, G.N., Hale, M.C., Gratten, J., Mossman, J.A., Lowder, B.V., Mann, M.K., and Slate, J. 2007. SNP SCALE: SNP scoring by colour and length exclusion. *Molecular Ecology Notes* 7, 377–388.
- Hudson, M.E. 2008. Sequencing breakthroughs for genomic ecology and evolutionary biology. *Molecular Ecology Resources* 8, 3–17.

- Johansen, T., Moen, T., Westgaard, J.-I., Berg, P., Fevolden, S.-E., Dahle, G., Nilsen, F., Jørstad, K. 2008. 80 SNP markers for application in genetic and ecological studies of North Atlantic cod (*Gadus morhua*) Submitted
- Knutsen, H., Jorde, P.E., Andre, C., Stenseth, N.C. 2003. Fine-scaled geographical population structuring in a highly mobile marine species: the Atlantic cod. *Molecular Ecology* 12, 385–394
- Kohn, M.H., Murphy, W.J., Ostrander, E.A., Wayne, R.K. 2006. Genomics and conservation genetics. *Trends in Ecology and Evolution*, 21: 629–637.
- Lee, H.Y., Park, M.J., Yoo, J.-E., Chung, U., Han, G.-R., Shin, K.-J. 2005. Selection of twenty-four highly informative SNP markers for human identification and paternity analysis in Koreans. *Forensic Science International* 148, 107–112.
- O'Leary, D.B., Coughlan, J., Dillane, E., McCarthy, T.V., Cross, T.F. 2007. Microsatellite variation in cod *Gadus morhua* throughout its geographic range. *Journal of Fish Biology* 70, 310–335.
- Moen, T., Hayes, B., Nilsen, F., Delghandi, M., Fjalestad, K.T., Fevolden, S.-E., Berg, P.R. Lien, S. 2008. Identification and characterisation of novel SNP markers in Atlantic cod: evidence of directional selection. In press.
- Morin, P.A., Luikart, G., Wayne, R.K. 2004. SNP-Workshop Group 2004 SNPs in ecology, evolution and conservation. *Trends in Ecology and Evolution*, 19: 208–216.
- Morin, P.A., LeDuc R.G., Robertson, K.M., Hedrick, N.M., Perrin, W.F., Etnier, M., Wade, P., Taylor, B.L. 2006. Genetic analysis of killer whale (*Orcinus orca*) historical bone and tooth samples to identify western U.S. ecotypes. *Marine Mammal Science*, 22: 897–909.
- Morin, P.A., Aitken, N.C., Rubio-Cisneros, N., Dizon, A.E., Mesnick, S.L. 2007. Primer Note: Characterization of 18 SNP markers for sperm whale (*Physeter macrocephalus*). *Molecular Ecology Notes* 7, 626–630.
- Morin, P.A., Hedrick, N.M., Robertson, K.M., LeDuc, C.A. 2007a. Comparative mitochondrial and nuclear quantitative PCR of historical marine mammal tissue, bone, baleen and tooth samples. *Molecular Ecology Notes* 7, 404–411.
- 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 nonmodel organisms. *Molecular Ecology* 15, 3219–3229.
- Poinar, H.N., Schwarz, C., Qi, J., Shapiro, B., MacPhee, R.D.E., Buigues, B., Tikhonov, A., Huson, D.H., Tomsho, L.P., Auch, A., Rampp, M., Miller, W., Schuster, S.C. 2006. Metagenomics to paleogenomics: large-scale sequencing of mammoth DNA. *Science* 311, 392–394.
- Queller, D.C., Strassmann, J.E., Hughes, C.R. 1993. Microsatellites and kinship. *Trends in Ecology and Evolution* 8, 285–288.
- Römler, H., Rohland, N., Lalueza-Fox, C., Willerslev, E., Kuznetsova, T., Rabeder, G., Bertranpetit, J., Schöneberg, T., Hofreiter, M. 2006. Nuclear gene indicates coat-color polymorphism in mammoths. *Science*, 313, 62.
- Ryman, N. 2006. CHIFISH: a computer program testing for genetic heterogeneity at multiple loci using chi-square and Fisher's exact test. *Molecular Ecology Notes* 6, 285–287.
- Ryman, N., Palm, S. 2006. POWSIM: a computer program for assessing statistical power when testing for genetic differentiation. *Molecular Ecology Notes* 6, 600–602.
- Ryynänen, H.J., Primmer, C.R. 2006. Single nucleotide polymorphism (SNP) discovery in duplicated genomes: intron-primed exon-crossing (IPEC) as a strategy for avoiding amplification of duplicated loci in Atlantic salmon (*Salmo salar*) and other salmonid fishes. *BMC Genomics* 7, 192.

- Rynnänen, H.J., Tonteri, A., Vasemägi, A., Primmer, C.R. 2007. A comparison of biallelic markers and microsatellites for the estimation of population and conservation genetic parameters in Atlantic salmon (*Salmo salar*). *Journal of Heredity* 98, 692–704.
- Sato, S., Moriya, S., Azumaya, T., Suzuki, O., Urawa, S., Abe, S., Urano, A. 2004. Genetic Stock Identification of Chum Salmon in the Central Bering Sea and Adjacent North Pacific Ocean by DNA Microarray During the Early Falls of 2002 and 2003. North Pacific Anadromous Fisheries Commission Document 793. available at www.npafc.org/.
- Sarvas, T.H., Fevolden, S.E. 2005. Pantophysin (Pan I) locus divergence between inshore v. offshore and northern v. southern populations of Atlantic cod in the north-east Atlantic. *Journal of Fish Biology* 67, 444–469.
- Seddon, J.M., Parker, H.G., Ostrander, E.A., Ellegren, H. 2005. SNPs in ecological and conservation studies: a test in the Scandinavian wolf population. *Molecular Ecology* 14, 503–511.
- Smith, C.T., Elfstrom, C.M., Seeb, L., Seeb, J.E. 2005a. Use of sequence data from rainbow trout and Atlantic salmon for SNP detection on Pacific salmon. *Molecular Ecology* 14, 4193–5203
- Smith, C.T., Baker, J., Park, L., Seeb, L.W., Elfstrom, C., Abe, S., Seeb, J.E. 2005b. Primer Note: Characterization of 13 single nucleotide polymorphism markers for chum salmon. *Molecular Ecology Notes* 5, 259–262.
- Smith, C.T., Templin, W.D., Seeb, J.E., Seeb, L. 2005c. 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, 944–953.
- Smith, C.T., Park, L., Vandoornik, D., Seeb, L.W., Seeb, J.E. 2006. Primer Note: Characterization of 19 single nucleotide polymorphism markers for coho salmon. *Molecular Ecology Notes* 6, 715–720.
- Taggart, J.B. 2007. FAP: an exclusion-based parental assignment program with enhanced predictive functions. *Molecular Ecology Notes* 7, 412–415.
- Vignal, A., Milan, D., SanCristobal, M., Eggen, A. 2002. A review on SNP and other types of molecular markers and their use in animal genetics. *Genetic Selection Evolution* 34, 275–305.
- Werner, F.A.O., Durstewitz, G., Habermann, F.A., Thaller, G., Krämer, W., Kollers, S., Buitkamp, J., Georges, M., Brem, G., Mosner, J., Fries, R. 2004. Detection and characterization of SNPs useful for identity control and parentage testing in major European dairy breeds. *Animal Genetics* 35, 44–49.
- Westgaard, J.I., Fevolden, S.E. 2007. Atlantic cod (*Gadus morhua* L.) in inner and outer coastal zones of northern Norway display divergent genetic signature at non-neutral loci. *Fisheries Research* 85, 306–315.

3 ToR b): Current and future prospects of QTL-based studies in fisheries and aquaculture.

Pierre Boudry¹, John Gilbey², Anti Vasemägi³, Delphine Lallias⁴ and Elisabeth Gosling⁵

¹ UMR M100 Physiologie et Ecophysiologie des Mollusques Marins, Ifremer – Technopole de Brest-Iroise, BP 70, 29280 Plouzane, France. ² FRS Freshwater Laboratory, Faskally, Pitlochry Perthshire, Scotland, PH16 5LB, UK. ³ Division of Genetics and Physiology, Department of Biology (Vesilinnantie 5) 20014 University of Turku, Finland. ⁴ School of Ocean Sciences, Bangor University, Menai Bridge, Anglesey, LL59 5EY, United Kingdom. ⁵ Molecular Ecology Research Group School of Science, Galway-Mayo Institute of Technology, Galway, Ireland.

Abstract. An increasing number of studies aim to identify quantitative trait loci (QTLs) in species of interest for fisheries and/or aquaculture. Such studies commonly imply the availability of polymorphic markers, linkage maps, informative biological material and high through-put genotyping facilities. Until now, most of the QTL studies ultimately 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 review the present status and prospects of QTL mapping in fisheries and aquaculture in the light of the current research in plant and animal genetics and breeding.

3.1 Introduction

Many phenotypic traits in plants and animals are quantitative in nature, i.e. influenced by a number of genes. The individual loci controlling a quantitative trait are called polygenes or quantitative trait loci (QTLs). On the other hand, a few major genes control some qualitative phenotypic variation e.g. albinism in rainbow trout (*Onchorynchus mykiss*), which is determined by a recessive mutation at a single locus (Thorgaard *et al.*, 1995). A QTL is a chromosomal region containing DNA polymorphism that has a significant effect on a phenotypic trait. A QTL is associated with a given percentage of the variance of a trait. Identification and mapping of QTLs requires appropriate biological resources i.e. informative segregating families, information on phenotype and together with genomic resources i.e. polymorphic markers. The combination of individual phenotypes and genotypes allows QTL detection. Determining the number of QTLs that explain variation in the phenotypic trait tells us about the genetic architecture (the way phenotypes map to genotypes) of a trait, whether a particular trait is controlled by many genes with small effect, or by few “major” genes with a large effect. Genotypic information can be used to establish linkage maps which provide a more elaborate framework for identification and localisation of QTL. Determination of the number of QTLs that contribute significantly to the variation of traits both within and between populations, and identification of markers linked to these QTL contributes to the improvement of breeding schemes (in mariculture and wild situations) using marker assisted selection (MAS). Identification of QTL may also be of use in the management of wild populations. QTL mapping has the potential to identify genetic markers that have greater discriminatory power when used to in population genetic analysis. QTL mapping can also aid in the understanding of the genetic basis of within population phenotypic variation, which in turn may be of use in managing the exploitation of different within population phenotypes.

3.2 Linkage Maps

A linkage map is an ordered collection of genes and genetic markers occurring along the chromosomes of a species, with distances between them estimated on the basis of the number of recombination events observed in the data, rather than as specific physical distance along each chromosome. Linkage maps provide a framework for identification and localization of QTLs. A variety of markers (allozymes, RAPDs, AFLPs, microsatellites, SNPs) have been used for the construction of linkage maps for several major aquaculture species (See Table 1 in Wenne *et al.*, 2007). The most frequently used markers are microsatellites (SSRs, single sequence repeats) chiefly because they are co-dominant and highly informative, are evenly distributed through the genome and are locus specific. More recently, expressed sequence tags (ESTs) have been explored as a new source of SSR markers. In species where large numbers of ESTs are available, mining them for SSRs is a cost-effective alternative (Vasemägi *et al.*, 2005). In addition, EST-SSRs may allow increased cross-species transferability and comparative mapping. Single nucleotide polymorphisms (SNPs) are regarded as the marker of choice for the future because of the need for high density of genetic markers, and the recent progress in polymorphism detection and genotyping techniques. Increasing numbers of SNPs have been identified for a number of species, including Atlantic salmon (Hayes *et al.*, 2007), cod (Delghandi *et al.*, 2007) and Pacific oyster (Sauvage *et al.*, 2007). For most linkage maps, the average distance between markers is 2–10 cM. However, generation of much denser linkage maps (<1cM) are in progress for several economically important aquatic species.

3.3 Identifying QTL

Two approaches have been developed to achieve this goal: detection of major genes relying on multimodal distribution and complex segregation analysis, and QTL mapping.

3.3.1 Estimation of number of major genes influencing a trait

Most of the variation of quantitative traits can be explained by the segregation of a few major genes, and there are several methods used to detect these. The simplest rely on features of the phenotypic distribution: multimodal distribution or departure from normality can be taken as an indicator of the presence of a major gene. Another approach is based on offspring/parent resemblance. Under polygenic inheritance (no major gene), the mean of offspring will resemble more closely the mid-parent value than single parents. But if a major gene is segregating, the mean of offspring will resemble more closely one of the parents. The most powerful approach, complex segregation analysis, relies on quantitative genetics and consists of finding parameter estimates for a series of increasingly complicated hypotheses: a pure environmental model, a single gene model, a polygene model and the full mixed model (Lynch and Walsh 1998).

3.3.2 QTL mapping

QTL mapping relies on linkage between marker alleles and QTL alleles, which result in marker-trait associations, i.e. different marker genotypes having different expected values for characters influenced by QTLs linked to these markers. Therefore, in QTL mapping one needs to analyse segregation of markers showing disequilibrium, using either crosses between inbred lines (fixed for different alleles) or analysis of relatives. Mapping QTLs consists of genotyping individuals at the marker locus and scoring their phenotype for the quantitative trait. The presence of a QTL linked to the marker is inferred if there is a difference in phenotypic value among marker genotype classes.

3.3.3 Experimental design to map QTLs

Both natural pedigrees and various types of artificial crosses have been used for QTL identification. The most powerful approaches are those derived from isogenic lines, inbred lines, nearly inbred lines and double haploids (Müller-Belecke 2005). Most statistical methods developed are suitable for this kind of experimental design. However, the most widely used experimental designs in aquatic organisms are backcrosses or F₂ intercrosses of hybrids between two strains, or populations or species with strong phenotypic divergence, in order to maximise the likelihood of identifying QTLs. Alternatively, F₁ crosses have also been used (e.g. Wang *et al.*, 2006).

Other critical factors affecting the power of QTL detection include: number of individuals per family, number of families, number and type of markers, genome coverage of markers, accuracy of phenotyping and genotyping, and heritability of the trait. Commonly, several hundreds of individuals are genotyped for more than one hundred markers in QTL studies. However, more cost effective approaches have been developed such as selective genotyping, which consists of scoring individuals for the trait and then genotyping only individuals selected from the tails of the character distribution. This is particularly useful in QTL studies of aquatic organisms because of high progeny numbers. For example, Darvasi and Soller 1992 showed that

selective genotyping can reduce the number of individuals genotyped by seven fold for a given power of detecting QTL.

3.3.4 Statistical methods

A variety of statistical methods and software have been developed to identify the most likely QTL position and effects, and to test their significance (Manly and Olsen 1999, see also Table 1 in Slate 2005).

The most simple and commonly used QTL identification method is single marker trait association because it does not require linkage map information (Broman 2001). Other more elaborate and powerful approaches include: Interval mapping, multiple regression analysis, composite interval mapping and Bayesian methods (Lynch and Walsh 1998). Among these, interval mapping has been the most widely used QTL mapping method, which explores the interval between pairs of markers for the presence of QTL (Lander and Botstein, 1989). Trait information from each adjacent pair of marker loci is examined and the information is used to infer the likelihood of a QTL being at any given position between them. The test of significance is based on a log likelihood ratio, or LOD score ($\log_{10}L/L_0$ L being the observed likelihood assuming one QTL in the interval, and L_0 the likelihood under the null hypothesis of no QTL segregating). A LOD score greater than 3.0 has been traditionally considered evidence for linkage; a score of 3.0 means the likelihood of observing the given pedigree if the two loci are *not* linked is less than 1 in 1000. However, several statistically more justified approaches have been developed to determine the appropriate statistical thresholds when testing for evidence of QTL.

3.4 QTL mapping in aquatic species

To date, a number of QTL studies have been carried out, particularly in salmonids (Table 1). QTLs have been detected for traits of economic and evolutionary importance in fisheries: life history, meristic and morphological, behavioural, disease and parasite resistance, and thermal tolerance traits. The individual QTLs detected explained a different amount of the phenotypic variance of the trait, ranging from 1 up to 64%. Commonly, several QTLs were found for a single trait, confirming the polygenic inheritance of the traits. Below, we outline recent studies that illustrate the application of QTL mapping in fisheries management and aquaculture.

Case study 1 – *Gyrodactylus salaris* resistance in Atlantic salmon: linking QTL mapping, population genetics and fishery management

Gyrodactylus salaris Malmberg, 1957 is a freshwater monogenean ectoparasite of salmonids endemic in the Baltic area. It was first recorded in Norway in 1975, and has been responsible for extensive epizootics of wild Atlantic salmon (*Salmo salar* L.) leading to an annual loss of 250 to 500 t of salmon (NOU 1999) and a total cost exceeding 500 million dollars (Bakke *et al.*, 2004). Identification of resistant individuals and of the genes involved in the host response to infection can be incorporated into selective breeding programs to produce both wild and aquaculture stocks which are less susceptible to parasite attack (Jones *et al.*, 2002).

In order to examine the genetic basis of resistance to *G. salaris*, Gilbey *et al.* 2006 carried out a classical QTL mapping study. They produced crosses using Baltic salmon from the river Neva, which have been shown to be relatively resistant to *G. salaris* infection (Bakke *et al.*, 1990, Cable *et al.*, 2000), and Scottish salmon from the river Conon, which have been shown to be extremely susceptible (Bakke and MacKenzie 1993, Dalgaard *et al.*, 2003). They challenged the fish with *G. salaris* and

counted parasite loads on individuals over a 30-day period. Each fish was then screened at 39 microsatellite loci chosen from the partial linkage map of Gilbey *et al.* 2004. Marker/trait associations and the amount of trait variance associated with each significant association were determined using GLM. Ten genomic regions were identified associated with heterogeneity in both the innate and acquired immune resistance, explaining up to 27.3% of the total variation in parasite loads. They found that both innate and acquired parasite resistance in Atlantic salmon are under polygenic control, and that salmon would be well suited to a management program designed to increase resistance to *G. salaris* in wild or farmed stocks (Gilbey *et al.*, 2006).

C. Thompson (2006, unpublished) went on to examine the markers found to be associated with *G. salaris* resistance in the QTL mapping study in relation to their ability to determine wild Atlantic salmon population structuring. She took samples from three rivers in Scotland (Conon, North Esk and Spaddagh) and three from the Baltic (Torne, Simojoki and Numendalslagen) and compared F_{ST} estimates both within and between the two areas using 5 of the loci linked to resistance and 10 randomly chosen microsatellite markers. F_{ST} values calculated using microsatellite markers linked to resistance in the QTL mapping study were found to be higher than those at randomly chosen markers both within regions and between regions, with the greatest difference found in comparisons between regions (Figure. 1).

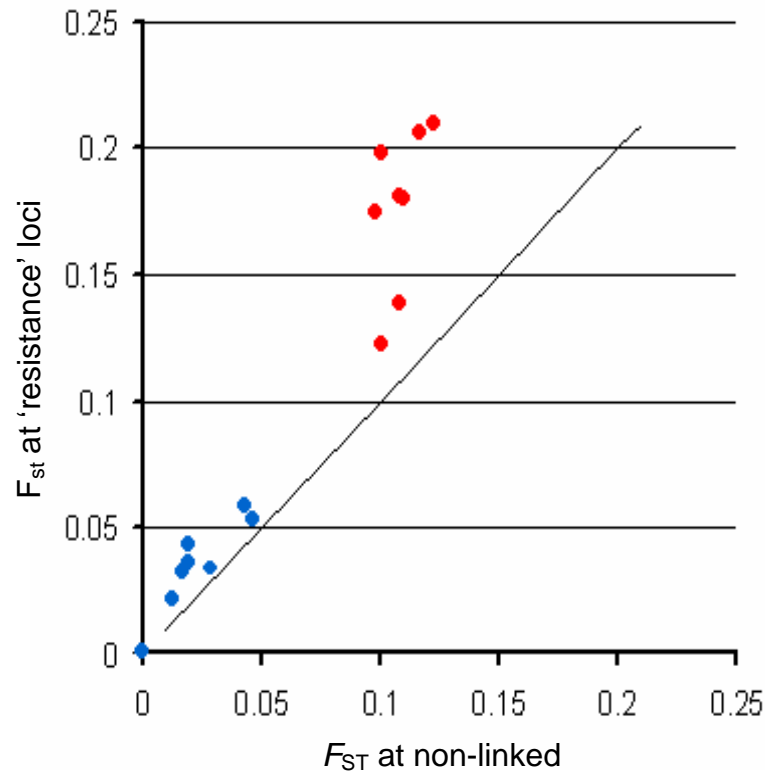


Fig. 1. Comparisons of F_{ST} values between pairs of populations within Scottish and Baltic regions (blue) and between Baltic and Scottish populations (red) at microsatellite markers found to be linked to *G. salaris* resistance in the QTL mapping study compared to randomly chosen microsatellite markers.

QTL mapping of markers linked to *G. salaris* resistance has thus achieved three outcomes: 1) it has helped in the understanding of the genetic architecture of parasite resistance; 2) it has been able to identify genetic markers which give greater discriminatory power when used in population genetic analysis of wild stocks; 3) it has provided markers which could aid in the establishment of more resistant stocks in both an aquaculture and wild fisheries context and so aid in the management of salmon stocks.

Case study 2 – QTL for disease resistance in oysters

Oyster farming is strongly impacted by diseases in many production areas (Renault, 1996). The most dramatic cases are probably the Eastern oyster *Crassostrea virginica*, infected by Dermo (*Perkinsus marinus*) and MSX (*Haplosporidium nelsoni*) and the European flat oyster *Ostrea edulis* infected by *Martelia refringens* and *Bonamia ostreae*. Additionally, summer mortality affects the Pacific oyster *Crassostrea gigas* in many countries, but its causal factors are complex, being related due to different environmental factors and pathogens. Taking into account the specificities of bivalves' biology (notably their lack of adaptive immune system) and aquaculture (*i.e.* extensive rearing in tidal or shallow coastal areas), breeding of resistant strains has been proposed as one of the most efficient way to counteract negative effects of pathogens. Selective breeding programs to improve resistance to *Bonamia ostreae* (Naciri-Graven *et al.*, 1998; Culloty *et al.*, 2004), Dermo, or MSX (Ford and Haskin,

1987; Ford *et al.*, 1990, Guo *et al.*, 2003) have been initiated, giving encouraging results. However, in most cases, no proper estimate of heritability of resistance is available do to the technical difficulty to set up such experiments.

Yu and Guo 2006 identified QTLs for resistance to *Perkinsus marinus* in the American oyster *C. virginica* by comparing genotype frequency at \approx 110 segregating AFLP loci before and after mortality in two F1 families (the first resulting from a cross, between two oysters from a disease-resistant strain, the second family being an hybrid cross between a wild female and a selected male from the selected strain). In this study, the genome of the eastern oyster was scanned with AFLP) markers before and after \approx 60 % mortalities. Significant shifts in genotype frequency were detected at a large number of loci and linkage analysis revealed that most these markers were closely linked on the genetic map, identifying twelve QTLs. This method does not however evaluate the proportion of phenotypic variance associated with these identified QTLs. In continuation to this work, the research group led by Pr. Guo (Haskin Shellfish Laboratory, Rutgers University, NJ, USA) is currently mapping SSRs and SNPs in a backcross family.

Lallias 2007 used an F2 family derived from a bi-parental cross between a wild oyster and an individual from one of the flat oyster families selected for improved resistance to *Bonamia ostreae* to identify QTL for that trait. F2 individuals from the family cohabited with wild oysters that were injected with the parasite. Selective genotyping was used and the F2 progeny was scored for 309 segregating AFLP loci and 17 polymorphic microsatellites. Firstly, using the multistage testing strategy for detection of quantitative trait loci affecting disease resistance in Atlantic salmon developed by Moen *et al.* 2004, fifteen AFLPs (out of 304) showed a significant difference of allele frequencies between the two phenotypically extreme groups. Secondly, QTL interval mapping approach identified 3 QTLs in the first parental map and 2 in the second. Some of the 15 markers identified in the survival analysis were mapped to the same locations as the QTLs. In continuation to this work, the research group led by Dr. Sylvie Lapègue (Ifremer, La Tremblade, France) and Andy Beaumont (Bangor University, Wales) are currently developing additional SSR markers. The comparison between MAS and non-MAS is proposed to validate the utility of QTLs to improve resistance to *Bonamia ostreae*.

Sauvage 2008 used Pacific oyster lines selected of high or low survival to summer mortality (Samain and McCombie, 2008) to generate F2 families, among which 3 presented significant levels of mortality over their first summer. Interestingly, individual viral load was also estimated by qPCR (Pepin *et al.*, 2008) for ostreid herpes virus-1 (OsHV-1), known to be associated with oyster mortality (Renaud and Novoa, 2004). QTL detection was successful for the two studied traits. Over the three F2s, five QTL explained 49 and 33% of mortality and viral load variances respectively. Most of the QTL co-localized for the two traits. These results support the high heritability of resistance to summer mortality (Dégremont *et al.*, 2007) and open new perspectives in terms of causal factors of summer mortality. Current research notably aims to map polymorphic genes (Sauvage *et al.*, 2007) that are differentially expressed in resistant and sensitive lines (Huvet *et al.*, 2004; Fleury *et al.*, 2008) to test for their eventual co-localisation with these QTL.

These QTL mapping studies for disease resistance in oysters have provided interesting new information concerning the genetic basis of disease resistance in oysters. They contributed to better understand the studied disease by setting up

novel experimental designs, establishing new links between oyster genetics, pathology and genomics.

3.5 Marker-Assisted Selection (MAS)

Once QTLs have been identified one major application is incorporation of QTL-linked marker information into selective breeding programs. This is known as marker-assisted selection (MAS). Allelic variation in genetic markers can be linked to the variation in traits of interest, and thus the marker provides DNA level information on the inheritance of the traits and the breeding value of individuals. This in turn can increase the efficiency of breeding programs by a significant factor over non-MAS programs (e.g. > 50% in simulations of MAS in aquaculture breeding programs, Sonesson 2007). The efficiency of MAS depends on three factors: heritability of the trait, proportion of genetic variance associated with marker (s), and the selection scheme used. MAS is of particular value for traits of low heritability and as a way of improving accuracy of selection to reduce generation interval by early selection before maturity e.g. carcass trait (Bertrand *et al.*, 2001). To our knowledge MAS has not been applied to any aquatic species. However, MAS has been successfully used in many breeding programs in livestock and crop plants (Mohan *et al.*, 1997; Dekkers and Chakraborty 2004; Tsilo *et al.*, 2008). If the gene affecting a trait is known, then gene-assisted selection (GAS) can be applied which is expected to be even more effective than MAS. For example, this selection method is being applied to increase scrapie resistance in sheep (Brochard *et al.*, 2006).

The identification and location of QTL for several traits and in several species highlights the potential for MAS in breeding programs in mariculture and fisheries management. There are several challenges ahead: in order to test the feasibility of MAS it is important to assess the role of genomic background and epistasis (when action of one gene is modified by one or several other genes) on the effects of these QTL. Most of the studies report the location of QTLs in a quite large area, spanning 10 to 60 cM. Therefore, it is not clear if a QTL might be made up of several genes with more effects, some acting in opposing directions. Before applying MAS it is therefore important restrict the region of interest to a more narrow area to increase the likelihood of successful MAS. A further step will be to go from the QTL down to the specific genes and ultimately to identification of the causative polymorphism (sometimes referred as Quantitative Trait Nucleotide, QTN). Various genomic resources can contribute to this task including physical maps, large EST databases and BAC (bacterial artificial chromosome) libraries. For example, BAC libraries have been constructed in several aquaculture species, e.g. Japanese flounder (Katagiri *et al.*, 2000), rainbow trout, Atlantic salmon, carp, tilapia (Katagiri *et al.* 2001) and the Pacific oyster (Cunningham *et al.*, 2006).

3.6 Potential application in wild populations

QTL mapping allows the identification of markers linked to genes influencing phenotypic traits and, using fine-scale QTL mapping and other techniques, the potential for eventual detection of the actual genes. Identification of such genetic markers provides information that is of use in helping to understand the basic genetic architecture of quantitative traits, and this information can also then be of use in the management of wild populations.

An important first step in the effective management of such populations is to define the actual breeding units of interest. This can be achieved using classical population genetic analysis techniques (e.g. Carvalho and Hauser 1994; Ward, 2000; Hallerman,

2003). However, the ability to identify population structures in a species using these techniques is dependant on the discriminatory power of the genetic markers used. QTL mapping allows markers to be identified that are linked to functional genes potentially under differential selection in the different breeding units. These markers might then be better able to define population structuring, as was seen in the Atlantic salmon case study outlined above (see also Le Corre and Kremer 2003; Rogers and Bernatchez, 2005; Raeymaekers *et al.*, 2007; Leinonen *et al.*, 2008).

QTL mapping may also be of use when managing populations, which contain individuals exhibiting different life-history phenotypes, and which may be under differential exploitation rates. For example, Atlantic salmon populations contain fish that may return to the rivers after their ocean migrations as 1 sea-winter 'grilse' or 2 sea-winter 'salmon'. The larger salmon tend to suffer greater fisheries exploitation than the smaller grilse, and proportions of these fish have been falling in recent years. QTL mapping could help in the identification of genetic markers for the different life-history types, which in turn would be of great use in helping to understand within population dynamics and allow basic fishery management questions to be addressed (e.g. ocean mortality rates of the different types).

Understanding the genetic architecture of phenotypic traits, and identifying markers linked to QTL, provides information that, as has been pointed out above, can be used in MAS programs. These programs are of obvious use in the mariculture industry, where they can be focused on traits of commercial interest, but might also be of use in the management of wild stocks, where, for instance, such stocks are threatened by a novel infective agent. MAS schemes could be envisaged that seek to increase resistance to the agent using broodstock collected from the wild and selected based on their QTL genotypes.

In summary, effective fishery management relies on an understanding of the structure of the populations of interest, and of the phenotypic composition of these populations. QTL mapping can aid in this understanding by helping to identify the genetic basis of the structure and phenotype compositions observed.

3.7 Recommendations

- 1) QTL studies should be supported in both in wild and farmed aquatic species as they are one of the most direct ways to understand the genetic basis of phenotypic variation, linking classic quantitative genetic and genomic studies.
- 2) QTL studies should not be restricted to MAS. The development of QTL studies should be supported as they can also contribute to a better understanding of the genetic architecture of adaptive traits of interest to fisheries and their management.
- 3) To aid identification of QTL in a wider variety of aquatic species, the current development of genomic resources – notably linkage and physical maps, EST and BAC libraries and whole genome sequences – should be encouraged.
- 4) The development of statistical methods and software adapted to aquatic species should be supported to facilitate the development of linkage maps and to identify QTLs.
- 5) The development and maintenance of divergent lines, segregating progenies, or other biological material of interest for QTL mapping should be encouraged.

3.8 References

- 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.
- Bakke, T.A., Harris, P.D., Hansen, H., Cable, J., Hansen, L.P. 2004. Susceptibility of Baltic and east Atlantic salmon *Salmo salar* stocks to *Gyrodactylus salaris* (Monogenea). *Diseases of Aquatic Organisms*, 58, 171–177.
- Bakke, T.A., Jansen, P.A., Hansen, LP. 1990. Differences in the host-resistance of Atlantic salmon, *Salmo salar*, to the monogenean *Gyrodactylus salaris* Malmberg 1957. *Journal of Fish Biology*, 37, 577–587.
- Bakke, T.A., Mackenzie, K. 1993. Comparative susceptibility of native and Norwegian stocks of Atlantic salmon, *Salmo salar*, to *Gyrodactylus salaris* Malmberg – Laboratory experiments. *Fisheries Research*, 17, 69–85
- Baranski, M., Robinson, N., Loughnan, S., Rourke, Hayes, B., Austin, C. 2007. Detection of QTL for growth rate in abalone using selective DNA pooling. *Aquaculture*, 272, S242–S243.
- Bertrand, J.K., Green, R.D., Herring, W.O., Moser, D.W. 2001. Genetic evaluation for beef carcass traits. *Journal of Animal Science*, 79, E190–E200.
- Brochard, M., Palhiere, I., Verrier, E., Moazami-Goudarzi, K., Neuts, E., Amigues, Y., Barillet, F., Bed'hom, B., Bibe, B., Bouix, J., Crihiu, E.-P., Francois, D., Leymarie, C., Pantano, T. 2006. Impact d'une sélection forte pour un gene majeur sur la variabilité, et les niveaux génétiques: cas de la sélection pour la résistance à la tremblante ovine. *Rencontres Autour des Recherches sur les Ruminants*, 13, 231–234.
- Broman, K.W. 2001. Review of statistical methods for QTL mapping in experimental crosses. *Lab Animal*, 30, 44–52.
- Cable, J., Harris, P.D., Bakke, T.A. 2000. Population growth of *Gyrodactylus salaris* (Monogenea) on Norwegian and Baltic Atlantic salmon (*Salmo salar*) stocks. *Parasitology*, 121, 621–629.
- Carvalho, G.R., and Hauser, L. 1994. Molecular genetics and the stock concept in fisheries. *Reviews in Fish Biology and Fisheries*, 4, 326–350.
- Cauwelier, E., 2007. The nature of intraspecific biodiversity in Atlantic salmon (*Salmo salar* L.). PhD Thesis, University of Aberdeen, Scotland, U.K.
- 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.
- Culloty, S.C., Cronin, M.A., Mulcahy, M.F. 2004. Potential resistance of a number of populations of the oyster *Ostrea edulis* to the parasite *Bonamia ostreae*. *Aquaculture*, 237, 41–58.
- Cunningham, C., Hikima, J.I., Jenny, M.J., Chapman, R.W., Fang, G.C., Saski, C., Lundqvist, M.L., Wing, R.A., Cupit, P.M., Gross, P.S., Warr, G.W., Tomkins, J.P. 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.
- Dalgaard, M.B., Nielsen, C.V., Buchmann, K. 2003. Comparative susceptibility of two races of *Salmo salar* (Baltic Lule river and Atlantic Conon river strains) to infection with *Gyrodactylus salaris*. *Diseases of Aquatic Organisms*, 53, 173–176.
- Danzmann, R.G., Jackson, T.R., Ferguson, M.M. 1999. Epistasis in allelic expression at upper temperature tolerance QTL in rainbow trout. *Aquaculture* 173, 45–58.
- Darvasi, A., Soller, M. 1992. Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus. *Theoretical and Applied Genetics*, 85, 353–359.

- Dégremont, L., Ernande, B., Bédier, E., Boudry, P. 2007. Summer mortality of hatchery-produced Pacific oyster spat (*Crassostrea gigas*). I. Estimation of genetic parameters for survival and growth. *Aquaculture*, 262, 41–53.
- Delghandi, M., Stenvik, J., Moen, T., Wesmajervi, M.S., Westgard, J.I., Fjalestad, K.T., Kettunen, A., Nilsen, F. 2007. Development and characterisation of microsatellite and single nucleotide polymorphism markers in Atlantic cod (*Gadus morhua* L.). *Aquaculture*, 272: S250–S250.
- Dekkers, J.C.M., Chakraborty, R. 2004. Optimizing purebred selection for crossbred performance using QTL with different degrees of dominance. *Genetics Selection Evolution*, 36, 297–324.
- Derayat, A., Houston, R.D., Guy, D.R., Hamilton, A., Ralph, J., Spreckley, N., Taggart, J.B., McAndrew, B.J., Haley, C.S. 2007. Mapping QTL affecting body lipid percentage in Atlantic salmon (*Salmo salar*). *Aquaculture*, 272, S250-S251.
- Drew, R.E., Schwabl, H., Wheeler, P.A., Thorgaard, G.H. 2007. Detection of QTL influencing cortisol levels in rainbow trout (*Oncorhynchus mykiss*). *Aquaculture*, 272, S183-S194.
- Fleury, E., Fabioux, C., Lelong, C., Favrel, P., Huvet, A. 2008. Characterization of a gonad-specific transforming growth factor- β superfamily member differentially expressed during the reproductive cycle of the oyster *Crassostrea gigas*. *Gene*, 410, 187–196.
- Ford, S.E., Haskin, H.H. 1987. Infection and mortality patterns in strains of oysters *Crassostrea virginica* selected for resistance to the parasite *Haplosporidium nelsoni* (Msx). *Journal of Parasitology*, 73, 368–376.
- Ford, S.E., Figueras, A.J., Haskin, H.H. 1990. Influence of selective breeding, geographic origin, and disease on gametogenesis and sex-ratios of oysters, *Crassostrea virginica*, exposed to the parasite *Haplosporidium nelsoni* (Msx). *Aquaculture*, 88, 285–301.
- Fotherby, H.A., Moghadam, H.K., Danzmann, R.G., Ferguson, M.M. 2007. Detection of quantitative trait loci for body weight, condition factor and age at sexual maturation in North American Atlantic salmon (*Salmo salar*) and comparative analysis with rainbow trout (*Oncorhynchus mykiss*) and Arctic charr (*Salvelinus alpinus*). *Aquaculture*, 272: S256–257.
- Fuji, K., Kobayashi, K., Hasegawa, O., Coimbra, M.R.M., Sakamoto, T., Okamoto, N. 2006. Identification of a single major genetic locus controlling the resistance to lymphocytes disease in Japanese flounder (*Paralichthys olivaceus*). *Aquaculture*, 254(1–4), 203–210.
- Fuji, K., Kobayashi, K., Hasegawa, O., Ozaki, A., Sakamoto, T., Okamoto, N. 2002 Quantitative trait locus (QTL) associated with resistance to Lymphocystis disease in Japanese flounder (*Paralichthys olivaceus*). Fourth International Symposium on Aquatic Animal Health, Abstracts. p 119.
- Gilbey, J. 2003. Genetic mapping of Quantitative Trait Loci influencing growth, development and morphology in Atlantic salmon (*salmo salar*, L.). PhD Thesis. University of Aberdeen, Scotland, U.K.
- Gilbey, J., Verspoor, E., McLay, A., Houlihan, D. 2004. A microsatellite linkage map for Atlantic salmon (*Salmo salar*). *Animal Genetics* 5, 98–105.
- Gilbey, J., Verspoor, E., Mo, T.A., Sterud, E., Olstad, K., Hytterod, S., Jones, C., Noble, L.R. 2006. Identification of genetic markers associated with *Gyrodactylus salaris* resistance in Atlantic salmon (*Salmo salar*). *Diseases of Aquatic Organisms*, 71, 119–129.
- Guo, X., Ford, S.E., DeBrosse, G., Smolowit, R. 2003. Breeding and evaluation of eastern oyster strains selected for MSX, dermo and JOD resistance. *Journal of Shellfish Research*, 22, 333–334.
- Grimholt, U., Larsen, S., Nordmo, R., Midtlyng, P., Kjoeglum, S., Storset, A., Saebo, S., Stet, R.J.M. 2003. MHC polymorphism and disease resistance in Atlantic salmon (*Salmo salar*);

- facing pathogens with single expressed major histocompatibility class I and class II loci. *Immunogenetics* 55, 210–219.
- Haidle, L., Janssen, J., Gharbi, K., Danzmann, R.G., Ferguson, M.M. 2007. Identifying regions of the genome responsible for early maturation and associated traits in rainbow trout (*Oncorhynchus mykiss*) *Aquaculture*, 272, S265–266.
- Hallerman, E. 2003. Population Genetics: Principles and Applications for Fisheries Scientists. (Eric Hallerman, ed.) American Fisheries Society, Bethesda, MD, USA.
- Hayes, B., Laerdahl, J.K., Lien, S., Moen, T., Berg, P., Hindar, K., Davidson, W.S., Koop, B.F., Adzhubei, A., Høyheim, B. 2007. An extensive resource of single nucleotide polymorphism markers associated with Atlantic salmon (*Salmo salar*) expressed sequences. *Aquaculture*, 265, 82–90.
- Hedgecock, D., Perry, G.M.L., Voigt, M.L. 2007. Mapping heterosis QTL in the Pacific oyster *Crassostrea gigas*. *Aquaculture*, 272, S267–S268.
- Houston, R., Guy, D.R., Hamilton, A., Ralph, J., Spreckley, N., Tabbart, J.B., McAndrew, B.J., Haley, C.S., Bishop, S.C. 2007. Mapping QTL affecting resistance to infectious pancreatic necrosis (IPN) in Atlantic salmon (*Salmo salar*). *Aquaculture*, 272: S269–S269.
- Houston, R.D., Haley, C.S., Hamilton, A., Guy, D.R., Tinch, A.E., Taggart, J.B., McAndrew, B.J., Bishop, S.C. 2008. Major quantitative trait loci affect resistance to infectious pancreatic necrosis in Atlantic salmon (*Salmo salar*). *Genetics*, 178, 1109–1115.
- Huvet, A., Herpin, A., Degremont, L., Labreuche, Y., Samain, J.F., and 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.
- Jackson, T.R., Ferguson, M.M., Danzmann, R.G., Fishback, A.G., Ihssen, P.E., O'Connell, M., Crease, T.J. 1998. Identification of two QTL influencing upper temperature tolerance in three rainbow trout (*Oncorhynchus mykiss*) half-sib families. *Heredity*, 80, 143–151.
- Jones, C.S., Lockyer, A.E., Verspoor, E., Secombes, C.J., Noble, L.R. 2002. Towards selective breeding of Atlantic salmon for sea louse resistance: approaches to identify trait markers. *Pest Management Science*, 58, 559–568.
- Katagiri, T., Asakawa, S., Hirono, I., Aoki, T., Shimizu, N. 2000. Genomic bacterial artificial chromosome library of the Japanese flounder *Paralichthys olivaceus*. *Marine Biotechnology*, 2, 571–576.
- Katagiri, T., Asakawa, S., Minagawa, S., Shimizu, N., Hirono, I., Aoki, T. 2001. Construction and characterization of BAC libraries for three fish species: rainbow trout, carp and tilapia. *Animal Genetics*, 32, 200–204.
- Khoo, S.K., Ozaki, A., Nakamura, F., Arakawa, T., Ishimoto, S., Nickolov, R., Sakamoto, T., Akutsu, T., Mochizuki, M., Denda, I., Okamoto, N. 2004. Identification of a novel chromosomal region associated with infectious hematopoietic necrosis (IHN) resistance in rainbow trout. *Fish Pathology*, 39, 95–102.
- 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.
- Lallias, D. 2007. Genetic linkage mapping in the blue mussel *Mytilus edulis* and the European flat oyster *Ostrea edulis*, and the search for Quantitative Trait Loci of resistance to a disease in *O. edulis*. PhD thesis, University of Wales, Bangor, College of Natural Sciences, School of Ocean Sciences.
- Lander, E.S., Botstein, D. 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, 121, 185–199.
- Le Corre, V., Kremer, A. 2003. Genetic variability at neutral markers, quantitative trait loci and trait in a subdivided population under selection. *Genetics*, 164, 1205–1219.

- Leinonen, T., O'Hara, R.B., Cano, J.M., Merila, J. 2008. Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. *Journal of Evolutionary Biology*, 21, 1–17.
- Li, Y.T., Dierens, L., Byrne, K., Miggiano, E., Lehnert, S., Preston, N., Lyons, R. 2006. QTL detection of production traits for the Kuruma prawn *Penaeus japonicus* (Bate) using AFLP markers. *Aquaculture*, 258, 198–210.
- Lynch, M., Walsh, B. 1998. *Genetics and analysis of quantitative traits*. Sinauer Associates, Sunderland, 980 pp.
- Manly, K.F., Olson, J.M. 1999. Overview of QTL mapping software and introduction to map manager QT. *Mammalian Genome*, 10, 327–334.
- Martinez, V., Thorgaard, G., Robison, B., Sillanpaa, M.J. 2005. Posterior evidence of multiple QTL influencing early development in double haploid lines of rainbow trout, *Oncorhynchus mykiss*. *Aquaculture*, 247, 25–25.
- Martyniuk, C.J., Perry, G.M.L., Mogahadam, H.K., Ferguson, M.M., Danzmann, R.G. 2003. The genetic architecture of correlations among growth-related traits and male age at maturation in rainbow trout. *Journal of Fish Biology*, 63, 746–764.
- Moen, T., Agresti, J.J., Cnaani, A., Moses, H., Famula, T.R., Hulata, G., Gall, G.A.E., May, B. 2004. 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.
- Moen, T., Sonesson, A.K., Hayes, B., Lien, S., Munck, H., Meuwissen, T.H.E. 2007. Mapping of a quantitative trait locus for resistance against infectious salmon anaemia in Atlantic salmon (*Salmo Salar*): comparing survival analysis with analysis on affected/resistant data. *BMC Genetics*, 8: 53.
- Moen, T., Fjalestad, K.T., Munck, H., Gomez-Raya, L. 2004. A multistage testing strategy for detection of quantitative trait Loci affecting disease resistance in Atlantic salmon. *Genetics* 167, 851–858.
- Moghadam, H.K., Poissant, J., Fotherby, H., Haidle, L., Ferguson, M.M., Danzmann, R.G. 2007. Quantitative trait loci for body weight, condition factor and age at sexual maturation in Arctic charr (*Salvelinus alpinus*): comparative analysis with rainbow trout (*Oncorhynchus mykiss*) and Atlantic salmon (*Salmo salar*). *Molecular Genetics and Genomics*, 277, 647–661.
- Mohan, M., Nair, S., Bhagwat, A., Krishna, T.G., Yano, M., Bhatia, C.R., Sasaki, T. 1997. Genome mapping, molecular markers and marker-assisted selection in crop plants. *Molecular Breeding*, 3, 87–103.
- Muller-Belecke, A. 2005. Relations between egg size, reproductive success and growth performance of progeny in isogenic *Oreochromis niloticus* lines. *Aquaculture*, 247, 127–134.
- Naciri-Graven, Y., Martin, A.G., Baud, J.P., Renault, T., Gerard, A. 1998. Selecting the flat oyster *Ostrea edulis* (L) for survival when infected with the parasite *Bonamia ostreae*. *Journal of Experimental Marine Biology and Ecology*, 224, 91–107.
- Nichols, K.M., Bartholomew, J., Thorgaard, G.H. 2003. Mapping multiple genetic loci associated with *Ceratomyxa shasta* resistance in *Oncorhynchus mykiss*. *Diseases of Aquatic Organisms*, 56, 145–154.
- Nichols, K.M., Wheeler, P.A., Thorgaard, G.H. 2004. Quantitative trait loci analyses for meristic traits in *Oncorhynchus mykiss*. *Environmental Biology of Fishes*, 69, 317–331.
- Nichols, K.M., Broman, K.W., Sundin, K., Young, J.M., Wheeler, P.A., Thorgaard, G.H. 2007. Quantitative Trait Loci x Maternal Cytoplasmic Environment Interaction for Development Rate in *Oncorhynchus mykiss*. *Genetics*, 175, 335–347.
- Norges offentlige utredninger 1999:9. Til laks åt alle kan ingen gjera. Oslo.
- O'Malley, K.G., Sakamoto, T., Danzmann, R.G., Ferguson, M.M. 2003. Quantitative trait loci for spawning date and body weight in rainbow trout: testing for conserved effects across ancestrally duplicated chromosomes. *Journal of Heredity*, 94, 273–284.

- Ozaki, A., Sakamoto, T., Khoo, S., Nakamura, K., Coimbra, M.R., 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., Parsons, J.E., Thorgaard, G.H. 1999. Identification of candidate DNA markers associated with IHN virus resistance in backcrosses of rainbow (*Oncorhynchus mykiss*) and cutthroat trout (*O. clarki*). *Aquaculture*, 173, 81–94.
- Palti, Y., Shirak, A., Cnaani, A., Hulata, G., Avtalion, R.R., Ron, M. 2002. Detection of genes with deleterious alleles in an inbred line of tilapia (*Oreochromis aureus*). *Aquaculture*, 206, 151–164.
- Pepin, J.F., Riou, A., Renault, T. 2008. Rapid and sensitive detection of ostreid herpesvirus 1 in oyster samples by real-time PCR. *Journal of Virological Methods*, 149, 269–276.
- Perry, G.M.L., Ferguson, M.M., Sakamoto, T., Danzmann, R.G. 2005. Sex-linked quantitative trait loci for thermotolerance and length in the rainbow trout. *Journal of Heredity*, 96, 97–107.
- Qin, Y., Liu, X., Zhang, H., Zhang, G., Guo, X. 2007. Genetic mapping of size-related quantitative trait loci (QTL) in the bay scallop (*Argopecten irradians*) using AFLP and microsatellite markers. *Aquaculture*, 272, 281–290.
- Raeymaekers, J.A.M., van Houdt, J.K.J., Larmuseau, M.H.D., Geldof, S., and Volckaert, F.A.M. 2007. Divergent selection as revealed by P_{ST} and QTL-based F_{ST} in three-spined stickleback (*Gasterosteus aculeatus*) populations along a coastal-inland gradient. *Molecular Ecology*, 16, 891–905.
- Reid, D.P., Szanto, Z., Glebe, B., Danzmann, R.G., Ferguson, M.M. 2005. QTL for body weight and condition factor in Atlantic salmon (*Salmo salar*); comparative analysis with rainbow trout (*Oncorhynchus mykiss*) and Arctic charr (*Salvelinus alpinus*). *Heredity*, 94, 166–172.
- Renault, T., Novoa, B. 2004. Viruses infecting bivalve molluscs. *Aquatic Living Resources* 17, 397–409.
- Renault, T. 1996. Appearance and spread of diseases among bivalve molluscs in the northern hemisphere in relation to international trade. *Revue Scientifique et Technique de l'Office International des Epizooties*, 15, 551–561.
- Robison, B.D., Wheeler, P.A., Sundin, K., Sikka, P., Thorgaard, G.H. 2001. Composite interval mapping reveals a major locus influencing embryonic development rate in rainbow trout (*Oncorhynchus mykiss*). *Journal of Heredity*, 92, 16–22.
- Rodriguez, M.F., LaPatra, S., Williams, S., Famula, T., May, B. 2004. Genetic markers associated with resistance to infectious hematopoietic necrosis in rainbow and steelhead trout (*Oncorhynchus mykiss*) backcrosses. *Aquaculture*, 241, 93–115.
- Rogers, S.M., Bernatchez, L. 2005. Integrating QTL mapping and genome scans towards the characterization of candidate loci under parallel selection in the lake whitefish (*Coregonus clupeaformis*). *Molecular Ecology*, 14, 351–361.
- Sakamoto, T., Danzmann, R.G., Okamoto, N., Ferguson, M.M., Ihssen, P.E. 1999. Linkage analysis of quantitative trait loci associated with spawning time in rainbow trout (*Oncorhynchus mykiss*). *Aquaculture*, 173, 33–43.
- Samain, J.F., McCombie, H. 2008. Summer Mortality in Pacific Oyster *Crassostrea gigas*. The Mores Project, Versailles, Editions Quae.
- Sauvage, C. 2008. Développement de marqueurs moléculaires liés à la résistance à la mortalité estivale chez l'huître creuse *Crassostrea gigas* – approche QTL. Thèse de l'Université de La Rochelle. 271 pp.

- Sauvage, C., Bierne, N., Lapègue, S., Boudry, P. 2007. Single Nucleotide polymorphisms and their relationship to cordon usage bias in the Pacific oyster *Crassostrea gigas*. *Gene*, 406, 13–22.
- Slate, J. 2005. Quantitative trait locus mapping in natural populations: progress, caveats and future directions. *Molecular Ecology*, 14, 363–379.
- Somorjai, I.M.L., Danzmann, R.G., Ferguson, M.M., 2003, Distribution of temperature tolerance quantitative trait loci in Arctic charr (*Salvelinus alpinus*) and inferred homologies in rainbow trout (*Oncorhynchus mykiss*), *Genetics*, 165, 1443–1456.
- Sonesson, A.K. 2007. Marker-assisted selection in aquaculture breeding programs. *Aquaculture* 272: S311–S311.
- Sun, X.W., Liang, L.Q. 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, K.H., Drew, R.E., Nichols, K.M., Wheeler, P.A., Thorgaard, G.H. 2005. Genetic analysis of a development rate QTL in backcrosses of clonal rainbow trout, *Oncorhynchus mykiss*. *Aquaculture*, 247, 75–83.
- Tao, W.J., Boulding, E.G. 2003. Associations between single nucleotide polymorphisms in candidate genes and growth rate in Arctic charr (*Salvelinus alpinus* L.). *Heredity*, 91, 60–69.
- Thorgaard, G.H., Spruell, P., Wheeler, P.A., Scheerer, P.D., Peek, A.S., Valentine, J.J., Hilton, B. 1995. Incidence of albinos as a monitor for induced triploidy in rainbow trout. *Aquaculture* 137, 121–130.
- Tsilo, T.J., Jin, Y., Anderson, J.A. 2008. Diagnostic microsatellite markers for the detection of stem rust resistance gene Sr36 in diverse genetic backgrounds of wheat. *Crop Science*, 48, 253–261.
- Vasemagi, A., Nilsson, J., Primmer, C.R. 2005. Expressed sequence tag-linked microsatellites as a source of gene-associated polymorphisms for detecting signatures of divergent selection in Atlantic salmon (*Salmo salar* L.). *Molecular Biology and Evolution*, 22, 1067–1076.
- Wang, C.M., Lo, L.C., Zhu, Z.Y., Yue, G.H. 2006. A genome scan for quantitative trait loci affecting growth-related traits in an F1 family of Asian seabass (*Lates calcarifer*). *BMC Genomics*, 7, Art. No. 274.
- Ward, R.D. 2000. Genetics in fisheries management. *Hydrobiologia*, 420, 191–201.
- Wenne, R., Boudry, P., Hemmer-Hansen, J., Lubieniecki, K.P., Was, A., Kause, A. 2007. What role for genomics in fisheries management and aquaculture? *Aquatic Living Resources*, 20, 241–255.
- Yu, Z.N., Guo, X.M. 2006. Identification and mapping of disease-resistance QTLs in the eastern oyster, *Crassostrea virginica* Gmelin. *Aquaculture* 254, 160–170.
- Zimmerman, A.M., Evenhuis, J.P., Thorgaard, G.H., Ristow, S.S. 2004. A single major chromosomal region controls natural killer cell-like activity in rainbow trout. *Immunogenetics*, 55, 825–835.
- Zimmerman, A.M., Wheeler, P.A., Ristow, S.S., Thorgaard, G.H. 2005. Composite interval mapping reveals three QTL associated with pyloric caeca number in rainbow trout, *Oncorhynchus mykiss*. *Aquaculture*, 247, 85–95.

Table 1. A list of quantitative trait loci (QTL) studies in aquaculture species.

SPECIES	TRAITS STUDIED	REFERENCE
Rainbow trout	Hatching time, embryonic length, weight	Martinez <i>et al.</i> 2005
	Embryonic development rate	Robison <i>et al.</i> 2001
	Development rate	Sundin <i>et al.</i> 2005
	Body length, thermotolerance	Perry <i>et al.</i> 2005
	Weight, age at sexual maturity	Fotherby <i>et al.</i> 2007
	Growth, condition factor, maturity age	Martyniuk <i>et al.</i> 2003
	Spawning time	Sakamoto <i>et al.</i> 1999
	Spawning time	O'Malley <i>et al.</i> 2003
	Length, pyloric caeca, no of scales	Nichols <i>et al.</i> 2004
	Embryonic growth rate	Nichols <i>et al.</i> 2007
	Pyloric caeca	Zimmerman <i>et al.</i> 2005
	Thermotolerance	Jackson <i>et al.</i> 1998
	Thermotolerance	Danzmann <i>et al.</i> 1999
	Infectious pancreatic necrosis virus	Ozaki <i>et al.</i> 2001
	Infectious hematopoietic necrosis	Rodriguez <i>et al.</i> 2004
	Ceratomyxa shasta resistance	Nichols <i>et al.</i> 2003
	INHV resistance	Palti <i>et al.</i> 1999
	IHNV resistance	Khoo <i>et al.</i> 2004
	Killer-cell activity	Zimmerman <i>et al.</i> 2004
	Cortisol levels	Drew <i>et al.</i> 2007
Early maturation	Haidle <i>et al.</i> , 2007	
Atlantic salmon/Rainbow trout/Arctic charr	Body weight, condition factor	Reid <i>et al.</i> 2005
Atlantic salmon	Growth, development and morphology	Gilbey 2003
	Infectious salmon anaemia	Moen <i>et al.</i> 2004
	IPN, furunculosis, ISA	Kjoglum <i>et al.</i> 2005
	Furunculosis, infectious salmon anaemia	Grimholt <i>et al.</i> 2003
	Resistance to Gyrodactylus salaris	Gilbey <i>et al.</i> 2006
	PKD mortality	Cauwelier 2007
	Resistance to ISA	Moen <i>et al.</i> 2007
	Condition, percentage lipid composition	Derayat 2007
	Resistance to IPN	Houston <i>et al.</i> (2007, 2008)
Coho salmon	Fillet colour	Araneda <i>et al.</i> 2005
Arctic charr	Growth rate	Tao and Boulding 2003
	Thermotolerance	Somorjai <i>et al.</i> 2003
	Sexual maturity, condition, weight	Mogahadam 2007
Common carp	Cold tolerance	Sun and Liang 2004
Tilapia	Innate immunity, response to stress, growth	Cnaani <i>et al.</i> 2004
	Thermotolerance	Moen <i>et al.</i> 2004c
	Growth under salt challenge, sex ratio, susceptibility to inbreeding	Palti <i>et al.</i> 2002
Asian seabass	Growth traits	Wang <i>et al.</i> 2006

SPECIES	TRAITS STUDIED	REFERENCE
Japanese flounder	Lymphocystis disease resistance	Fuji <i>et al.</i> (2002; 2006)
Kuruma prawn	Growth traits	Li <i>et al.</i> 2006
Blacklip abalone	Growth rate	Baranski 2007
Pacific oyster	Heterosis	Hedgecock 2007
	Summer mortality resistance	Sauvage 2008
Eastern oyster	Perkinsus marinus resistance	Yu and Guo 2006
Bay scallop	Size-related traits	Qin <i>et al.</i> 2007
European flat oyster	Bonamia ostreaea resistance	Lallias <i>et al.</i> 2007

4 ToR c): Update progress on the establishment of a meta-database for genetic data on fish and shellfish genetics covered under the ICES remit

Jann Th. Martinsohn and Eric Verspoor

4.1 Context and relevance

A wide range of studies have been carried out worldwide into the structuring of fish and shellfish stocks into genetic populations, into the effect of selective pressures such as fisheries exploitation and environmental change on levels of genetic variation, and into the genomic architecture of heritable performance traits important to local adaptation and aquaculture economics. The results of these studies represent a tremendously valuable source of biological understanding which is highly valuable not only to fisheries management and conservation, and to aquaculture, but also to fisheries control and enforcement issues related to fish and fish product traceability. The output from such work is continually expanding, driven by rapid technological advances in the field of molecular genetics. Yet, in contrast, the capacity to access and integrate the increasing body of information has advanced relatively slowly. In the main, primary population data sets generated are highly dispersed and at significant risk of being lost after the conclusion of research projects; few if any are published in detail and there is no formal mechanism for keeping them in the public domain as there is for genomic information using publicly accessible electronic data bases such as GENBANK.

This increasing discrepancy between data generation and data storage and management means that maximal benefits are unlikely to be gained by the research community and fisheries managers from the work carried out. To address this issue, the WGAGFM proposed in 2007 to establish a meta-database cataloguing existing data in the field of fish and shellfish genetics. Its existence would significantly help to promote coherence across research programmes, enhance research progress, and facilitate the translation of results from fundamental research to support the above described applications in the fisheries and aquaculture sector¹.

The complete solution to the loss of primary population genetic data would lie in its assembly into an all-embracing databank. However this concept, as recognised by the WGAGFM, is unrealistic and unlikely to be cost-effective due to the extremely large effort required and relatively large inherent costs needed for database implementation and maintenance. Instead, the WGAGFM recommended building a public online portal (metadata base) cataloguing existing genetic data sets and biological materials, as well as their location where they can be accessed, and which could catalogue historical and contemporary research projects for the species of interest. This would be expected to be a much more cost-effective option and advance science and its application in two ways. Firstly, it would allow researchers to more easily gain a comprehensive overview of existing population genetic information for a given fish or shellfish species of commercial or conservation interest, enhancing the capacity to carry out meaningful reviews to underpin advice and for rapidly developing new optimally targeted research programmes. Secondly, it could provide stakeholders, such as policy makers or fisheries managers, with a one-stop location

¹ Report of the Working Group on the Application of Genetics in Fisheries and Mariculture (WGAGFM); 19–23 March 2007; Ispra, Italy.

for rapidly identifying where information can be found which they can use to assess the state of the art for a given species, or assess work carried out on the population genetics on species generally, as well as the extent of progress made with respect to applications of the research to fisheries and mariculture.

4.2 Progress and further action

The ICES response to the 2007 recommendation of establishing a meta-database on fish genetics was positive. The ICES data centre is in principle willing to host the database, subject to clarification of some critical issues (see below)².

The general support for the development of such a database within the WGAGFM, as regards the need for such a data base, and for its development, was assessed by a formal internal Group survey carried out during the 2008 WGAGFM meeting. This survey confirmed that Group members saw development of such a database as adding a very valuable fisheries research tool to their science programmes. Positives mentioned repeatedly included a boosting effect on scientific cooperation, the prevention of duplicated efforts, ease of acquisition of knowledge about available data and biological material and where to obtain it. Furthermore it was stressed that such a database would not be an alternative to “classical” information resources such as literature databases (Pubmed etc.) or genetic databanks (e.g. EBI and Genbank). Rather, appropriately constructed, it could fill significant database gaps and complement existing resources, significantly facilitating day to day research activities. However, some useful refinements to the initial proposal were suggested and discussed.

4.3 Development and implementation strategy

The general consensus view was that the development of the proposed meta database would require considerable effort and has little if any potential for attracting specific funding to support its creation. This observation, combined with recognition of the need for a coordinated input into its development by all WGAGFM members, lead to the conclusion that making an operational on-line version the Fish Genetic meta-database online would not be something which could be achieved quickly and in the short-term. Rather, its creation would require working creatively with existing organisational infrastructure and research programmes, and taking a staged, progressive approach over the medium term i.e. the next 12–18 months. However, to do this, it was recognised the first step will be to produce a detailed database description and implementation scheme which is agreed upon by all involved stakeholders. The objective of producing such a plan by September 2008 was agreed. The most crucial issue, both in this first step and throughout the database development progress to ensure the successful implementation of the fish genetic meta-database will be to establish in this first stage a close dialogue between the WGAGFM (as the entity providing the input and a major user), the JRC (as the proposed developing entity), and ICES (as the hosting entity – see figure 1).

A further crucial issue is to ensure that the database developed is capable of dealing with real data and real data access needs. To help in this regard, as a template for initial data base development, it was agreed to focus on using Atlantic salmon metadata which is being compiled as part of the NASCO Salmon at Sea programme

² Neil Holdsworth; Head of ICES Data Centre (<http://www.ices.dk/datacentre/>) – personal communication

(SALSEA; <http://www.nasco.int/sas/salsea.htm>) under the EU funded FP7 SALSEA-MERGE project. This will help to identify key information types to be included, to assess access and data presentation features needed, and allow the underlying IT requirements to be specified in detail and realistically tested.

It was pointed out by ICES, in response to the meta-database recommendation 2007, that pre-existing reference code lists (species, country etc.) derived from the ICES online database RECO should be used for compatibility reasons and as to facilitate implementation. This was agreed to be an important issue which should be incorporated into the database design.

From discussions, it emerged that a number of specific issues from both the IT and scientific angle require to be clarified during this first step before embarking on subsequent steps in the development of a functioning database prototype.

IT-related issues:

Access and login control: This aspect is crucial for ensuring database quality. For the login control the WGAGFM group, who would run the database, has first to decide who will be endowed with what level of accessibility and administrative rights. From initial discussions, it was proposed that there will be a main administrator (level 1 – access/administrative rights for all data; see figure 2) nominated and that additionally volunteers, appointed for a distinct period of time and responsible for a defined set of data (e.g. species related) will get administrative rights for all data relevant to their respective fields (level 2). Users of the database, wishing to upload data will establish a user account by applying for a password and access to the database. The application will solely be admitted provided all relevant information such as institution affiliation and user contact details are submitted and after verification by level 2 database administrators. Upon admission, the user has the possibility to upload her/his data onto the database (level 3). However, to assure data quality the uploaded data sets will be scrutinized by second level administrators (see also 2b). Given this arrangement, it will be crucial to ensure that the system used for login control can be accommodated by the host. This must be clarified in the first step of the development process.

Database size: The size of the database is not yet predictable, in the absence of knowledge of the final database design and the extent of the data to be accommodated for all the species covered. Additionally, this is likely to increase over time as new information is produced annually. Presently it is planned to start out with a database “fed” with data supplied by WAGFM members. However, once established, it is expected that the database will be advertised and data entry by fellow researchers encouraged. In light of this, maximal flexibility should be assured with respect to the size of the database as ideally a constant addition of data should eventually set in.

Underlying technology: Given different partners will be involved in the development and hosting of the database, the database and interface technology used has to be compatible between the developer and host, which may be problematic if they use different technologies. If so, the technological issues and solutions will have to be clarified before starting the development, and ways to circumvent incompatibilities, where they exist, have to be found as part of the dialogue established between developer and host.

Upgrading: To remain fully functional, the database will need to be upgraded and versioned at regular intervals to ensure that the database itself and the interface are

supported by the latest software versions. This should be done by the database host (if needed with support from the developer) while the actual database content should be the responsibility of the WGAGFM members who are in the best position to vet database content (see “User friendliness and data quality control” below).

Content relationship: The content of the database has to be decided upon with regards to scientific relevance and the scope of the meta-database. However for the development of the database it is important to identify the content items at early stages and to define their respective relationships. This will help to design an entity relationship scheme which in turn underlies the database development (see figure 3).

Scientific issues:

From the research point of view to achieve the database goals as outline, information and data accessibility have not only to be ensured but also greatly facilitated. Additionally the quality and validity of information and data should be closely monitored, of which the most relevant issues are:

Content: The basic content of the meta-database was discussed during the WGAGFM meeting 2007³ and, at this point, it is worth stressing that, as currently conceptualised, the proposed fish genetics meta-database is not seen as containing primary genetic data. To encompass primary genetic data would represent a much greater development challenge and, most of all, require considerably more effort to manage and maintain. Not only would this increase the risk of the data based being problematic to develop and less user-friendly, it is also likely to be have a much higher cost to benefit ratio, bringing into serious question its sustainability with the available limited resources. For these reasons, it is seen to focus only on building a portal which simply catalogues what information exists, lists past and ongoing projects, collections of biological material which are available for genetic analysis in different host institutions, and available raw genetic databases with contact details for those wishing to obtain more information. However, there are a number of different types of metadata which it would potentially be of interest to include. Therefore, for database development, it will be crucial to define in detail the nature of the meta data to be covered and uploaded. This may include lists of researchers, research projects, project summaries, data bases and descriptions of data base contents, lists of archival material, and lists of known primary and secondary publications on the population genetics of species of interest. Optionally, the database may also wish to include pdf files which provide brief potted reviews or summaries of what is known about the population genetics of a given species. As well it may have a bulletin board to facilitate information exchange among researchers, as well as between researchers and stakeholders, where people seeking information or highlighting its availability can advertise to the fish population genetics community as a whole.

User friendliness and data quality control: Keeping both the upload and download interfaces easy to use and intuitively comprehensible will be critical to database success. The upload of data and information should be highly automated and take advantage of already routinely used formats e.g. by providing pre-formatted spreadsheets (e.g. MS Excel) which can be submitted for upload. In this regard, in reaction to the WGAGFM report 2007, the ICES DC unit indicated that their

³ Report of the Working Group on the Application of Genetics in Fisheries and Mariculture (WGAGFM); 19–23 March 2007; Ispra, Italy

EcoSystem Data team possesses very considerable experience with the development of data entry web pages and would be happy to assist the development of the fish genetics Meta database. This help would be both by providing inside into the design and construction of their existing portals and by helping to develop a simple, functional logic structure for the fish population genetic meta database. To help with this, it will be important for WGAGFM to clearly specify the relatedness of different types of meta data to be included to the database developer as this will be a crucial determinant of the portal architecture and design. Finally, in a largely autonomous system which relies on the input of its end-users, data quality assurance is a highly critical issue. It is envisioned that data validation and control can most effectively be carried out by assigned "second-level" administrators as outlined in 1a). To place this within an organisational context where it can be appropriately managed and monitored, it is proposed to include data base review as part of annual WGAGFM meetings as a reporting ToR.

Response and user monitoring: Database use should be monitored to provide feedback on functionality of data base design and the usefulness of different database components. This can in large measure be achieved by documenting levels of use (in the simplest case a "contact" link), and the number of visitors using different interface levels (*e.g.* Google™ analytics – <http://www.google.com/analytics/>). The results of this feedback, in combination with the recorded user frequency, should be evaluated at regular intervals by the WGAGFM to ensure a constant adaptation of the meta-database to user needs and to establish the merit of its continued existence.

Links to other marine meta databases: The underlying motivation for the development of a fish genetic meta-database is to address the problem of ensuring general awareness of the increasing amounts and types of data available to workers in the field of fish population genetics which is relevant to fisheries research, and to ensure avoid its loss over time. This can potentially be achieved in a variety of ways and it will be of value to examine other similar efforts, assess to what extent these efforts are relevant to addressing the objectives of the proposed fish genetic meta-database, and establish to what extent it would be worth developing linkages to promote these objectives. For example, the European Commission is currently investigating the possibilities of setting up a European Marine Observation and Data Network (EMODNET)⁴. EMODNET intends to provide a sustainable focus for improving systematic observation (*in situ* and from space), interoperability, and increasing access to data related to maritime affairs and fisheries, based on robust, open and generic ICT solutions. It should be explored to what extent the WGAGFM fish genetic Meta-Database might fit into this large database project to maximize its benefits to the research and resource management communities.

4.4 Recommendations

WGAGFM recommends:

- 1) The development and implementation of a web-based fish population genetic meta-database, under the responsibility of WGAGFM, within the remit of ICES and in collaboration with the European Commission, as proposed in the WGAGFM report 2007, is progressed;

⁴ GREEN PAPER (2007) Towards a future Maritime Policy for the Union: A European vision for the oceans and seas (presented by the Commission; SEC(2006) 689).

- 2) The meta-database serve as a portal cataloguing relevant primary information with respect to generated population genetic data, primary and secondary research reports, available biological samples, and point to the locations and contact points from which to get such data, samples and extended information;
- 3) The WGAGFM agree before September 2008 which types of meta data to include and the database structure, and the functionalities of the interface providing access to the data as well as about other features to be made available;
- 4) The WGAGFM agree access regulation and control by database end-users as well as strategies to ensure data quality and validity before September 2008;
- 5) Decisions regarding 3) and 4) the WGAGFM report to the database developer and host before September 2008;
- 6) When 5) is delivered, database developer and host should establish contact and establish a practical, workable strategy for development and implementation of the database;
- 7) The Atlantic salmon meta-data set developed as part of the NASCO Salmon at Sea initiative under the EU SALSEA MERGE project be used as a development template for a prototype fish population genetics database;
- 8) Work on 3) to 7) should be progressed, facilitated and monitored by the WGAGFM members leading this ToR;
- 9) Progress should be reported to the 2009 WGAGFM by the ToR leaders.

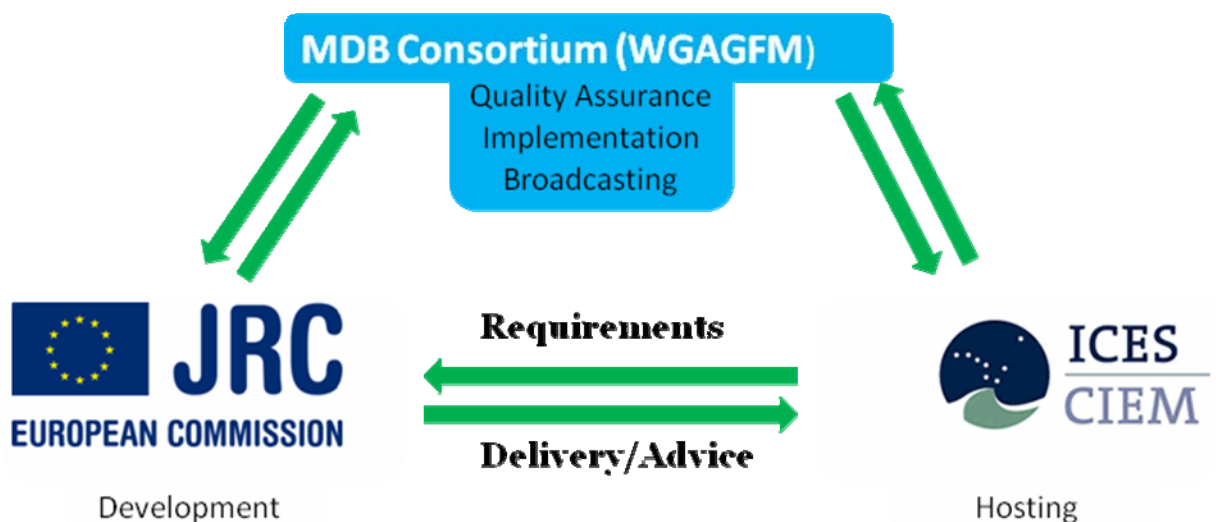


Figure 1. Proposed activity structure during the fish population genetics meta-database (FPG-MDB).

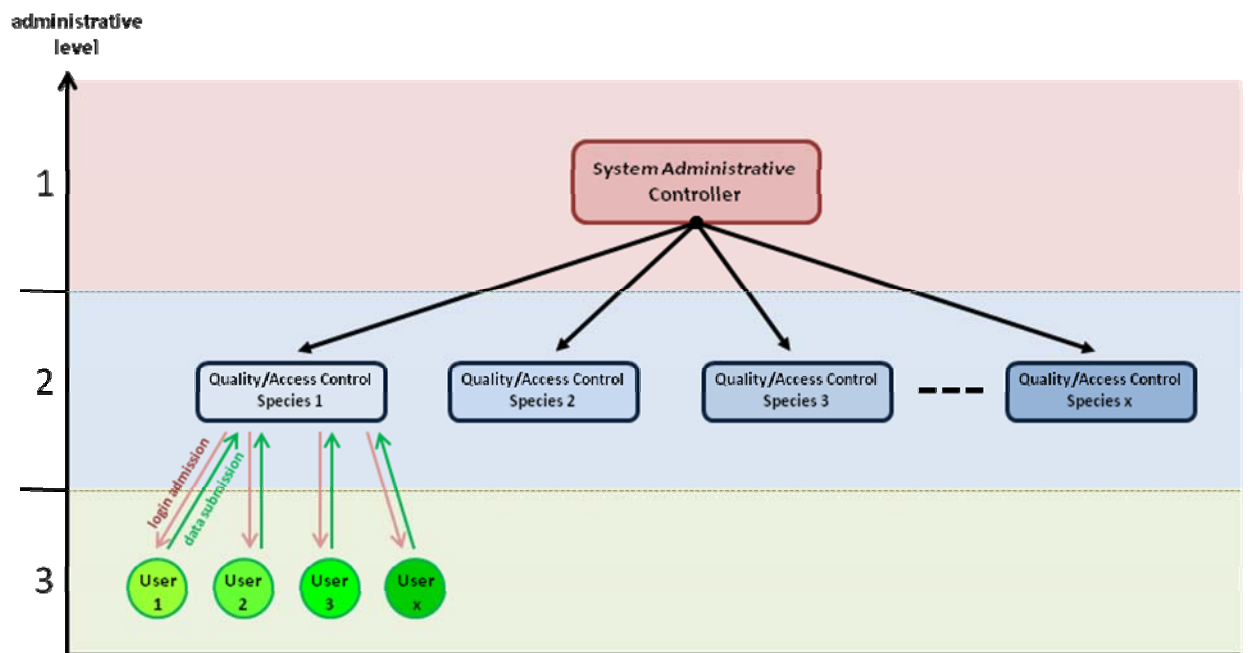


Figure 2. Schematic illustrating the allocation of administrative rights with regard to the proposed fish population genetics meta-database. One level 1 administrator would be nominated by the WGAGFM for a yet to be determined time period. This administrator would have access and administrative rights for all parts of the database. Level 2 administrators would also be assigned by the WGAGFM for a determined time period. Their administrative rights would be restricted to a defined set of data (e.g. species specific). Their major task would be data quality assurance and to provide end-users applying for database access with admissions. At level 3, end-users would be permitted access to the database by level two administrators for data upload.

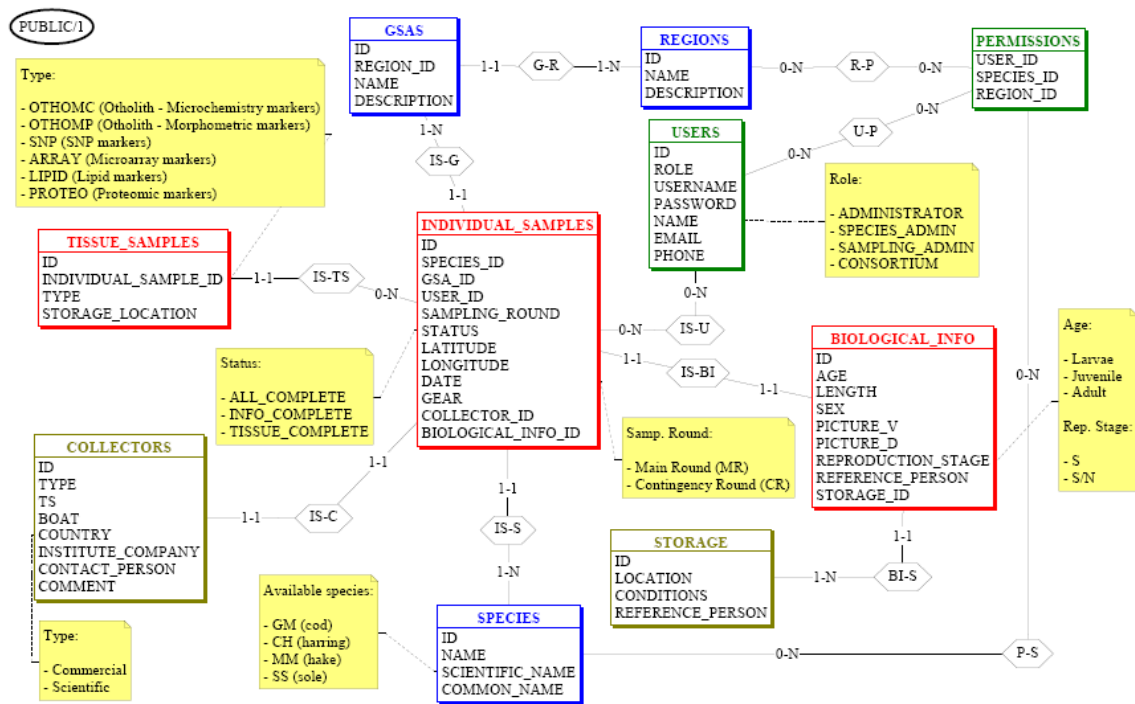


Figure 3. Example of an Entity Relationship Scheme (ERS) (here for a fish specimen collection database), illustrating database content and architecture (kind permission of Alessandro Ghigi; European Commission Joint Research Centre). As pointed out in the text an ERS should be developed as one of the first steps of the database developments. This must be preceded by the identification of data categories and their relationship, by the WGAGFM.

5 ToR d): Review progress for optimising the storing of otoliths and scales for DNA analysis

M.C. Cross⁵*, T.F. Cross, A-B. Florin, R. Hanel, R. Kirby and M. O'Sullivan

5.1 Introduction

Archived material (such as scales and otoliths) is proving to be a useful source of DNA for molecular genetics in fisheries and aquaculture, but this was not the original purpose of these collections. Beginning with Reibisch's observations of otolith annuli in 1899, mineralised fish tissues, like otoliths and scales (but also cephalopod statoliths (Zumholz *et al.*, 2007) and other invertebrate samples) have been collected and stored for biochronology. An informal survey of 30 fisheries laboratories, estimated a minimum of 800,000 otoliths were used for ageing worldwide in 1999 (Campana and Thorrold, 2001). Exact age determination of fish is one of the most important elements in the study of their population dynamics (Rodríguez Mendoza, 2006). Incremental rings provide information on age, and the widths of daily, as well as annual rings, reflect individual growth patterns. Besides age and growth determination, otoliths and to a lesser extent scales, have been the object of studies in many different fields including species and stock identification, environmental reconstruction of fish habitats (Campana, 1999, 2005; Brazner *et al.*, 2004; Humphreys *et al.*, 2005; Popper *et al.*, 2005) and to provide information on the ambient environmental conditions by trace element analysis. Microchemical studies of otoliths and other biogenic aragonites are based on a variety of analytical techniques, including particle induced X-ray emission PIXE (Elfman *et al.*, 1999), Synchrotron X-ray Fluorescence Analysis SYXRF (Tsukamoto *et al.*, 1998; Zumholz *et al.*, 2007), electron probe microanalysis EPMA (Daverat *et al.*, 2005), solution-based inductively-coupled plasma mass spectrometry ICP-MS and Laser Ablation ICP-MS (Arai *et al.*, 2006). Furthermore, these high-resolution analytical techniques allow the information to be related to age and growth based on the ring structure. Thus, otoliths and also scales can be regarded as archives recording individual life histories.

Archived collections of scales and otoliths are often the only source of samples for retrospective research on fish populations and, as noted above, they are also an invaluable source of DNA. Scales and otoliths will also continue to be collected by fisheries biologists for ageing studies. In addition, field personnel are very familiar with collecting these materials and collection of additional material is often restricted by time constraints. While it is recognised that non-mineralised tissue would be the best source of DNA in the future, such as a fin clip stored appropriately, the present report concentrates on DNA in scales and otoliths.

5.2 Traditional Methods

Traditionally, scales were removed from fish, air dried in paper envelopes then stored together in groups. When required for ageing, scales were soaked in water or sodium peroxide. Alternatively, scales were sometimes mounted in resin or impressions made on plastic ribbon. It is likely that all this archive material, including plastic ribbon, could provide samples for genetic analysis. In the case of otoliths these were also allowed to air-dry in paper envelopes. Prior to reading otoliths were

⁵ m.cross@ucc.ie

cleaned, burnt, cracked or ground. Thus untreated otoliths are a much more reliable source of DNA (Heath *et al.*, 2007).

The issue of cross contamination between specimens when sampling scales or otoliths is more serious when DNA extraction is envisaged because it is easy to carry small amounts of soft tissue or mucous between specimens unless precautions are taken (instruments are cleaned (wiped) thoroughly between samples). It is also likely that sample processing stations will be contaminated by slime and tissues from a number of fish.

5.3 Location of DNA

In scales, DNA is contained in the overlying dermal tissue and to a much lesser extent within the scale matrix. Therefore, DNA may be extracted from unwashed and mounted scales (Mitchell *et al.*, 2008) and in some cases from scales preserved in formalin (Hua Yue and Orban, 2001). In the case of otoliths, the structure is mostly inorganic (Paabo *et al.*, 2004). Thus DNA is extracted largely from the tissue left adhering to the surface of the otolith after removal from the fish. The quality and quantity of DNA extracted depends on how well the scales or otoliths have been collected (no contamination) and stored.

5.4 Previous Genetic Studies

Numerous studies exist whereby archived scales and otoliths have been used as a source of DNA. The results of such temporal studies can serve to answer questions in relation to natural (e.g. bottlenecks, geographic expansion/contraction of a species range) and anthropogenic (e.g. over-fishing, restocking) effects. Hutchinson *et al.* 2003 used archived cod (*Gadus morhua*) otoliths for large-scale temporal genetic analysis after successful extraction of nuclear DNA from otoliths dating from 1954, stored in damp warehouse conditions. The extracts were amplified using the polymerase chain reaction (PCR) and screened for variation at three nuclear microsatellite DNA loci. Loci that yielded small products (less than 220 base pairs (bp)) were selected owing to the degraded nature of the recovered DNA. This study illustrates how mineralised tissue may still be a viable source of genetic information despite unsuitable storage conditions. Martinez *et al.* 2001 used archived Atlantic salmon (*Salmo salar*) scale samples to evaluate the effect of restocking policies on natural populations. Microsatellite alleles up to 232 bp long were successfully amplified after the DNA had been extracted from two to four unwashed scales using a rapid Chelex-based extraction method (Estoup *et al.*, 1996). These are just two of the many examples where archived mineralised tissues have yielded genetic information. Other examples include Nielsen *et al.* (1999, 2007), Adcock *et al.* 2000, Hoarau *et al.* 2005 and Poulsen *et al.* 2006. These studies used available archival materials. Methods of minimizing deterioration of DNA quantity and quality in archived or newly collected materials would be advantageous for future studies.

5.5 Degradation of DNA

DNA is very susceptible to degradation by hydrolysis, radiation and naturally occurring enzymes. This often results in the break-up of DNA into fragments consisting of between 100 and 400 bp. Probably the most important factors in the long-term storage of DNA are the presence of water, specifically the amount of 'available water', and enzyme cofactors, typically metal cations (Rollo, 1998; Hofreiter *et al.*, 2001; Paabo *et al.*, 2004) Since DNA is acidic, it can catalyze its own depurination (hydrolysis of purine moieties) in the presence of water. In addition

deoxyribonuclease (DNase) enzymes, like most proteins, require water to maintain their conformation, and so their activity declines as the amount of available water is reduced. Humidity also encourages the growth of bacteria and fungi. Therefore archived samples in paper envelopes should be stored in the driest conditions possible and protected from direct sunlight to favour the preservation of DNA (Rivers and Ardren, 1998; Nielsen *et al.*, 1999).

5.6 Optimal Methods

Two situations must be considered; treatment of archival collections to ensure minimum future DNA degradation and, recommendations in relation to new collections. However, the subsequent optimization of storage methods pertains to both situations. Sub-optimal storage conditions of archived material in paper envelopes and inevitable degradation of DNA will lead to low quality and quantity of DNA, which, due to the sensitivity of the molecular techniques, makes cross contamination in the laboratory (as opposed to cross contamination at sampling, as discussed above) a particular threat (Nielsen *et al.*, 1999; Hutchinson *et al.*, 1999; Hofreiter *et al.*, 2001; Gilbert *et al.*, 2005; Wandeler *et al.*, 2007). DNA that has degraded into small fragments leads to fewer template molecules to extract, and these can be overwhelmed by the amplification of fresh contaminant DNA during PCR. Specific care should be taken when handling such samples. Gloves should be worn to minimise contamination by human DNA, and any instruments used need to be cleaned to prevent cross-contamination between samples.

5.6.1 Archival collections

So far, little attention has been given to the storage of these important collections as a source of DNA, although we know of steps being taken in a Swedish institute where new facilities have been provided to store scales and otoliths in a controlled-environment; 45 +/- 8% humidity and 21 +/- 2°C (Mecklenburg *et al.*, 2004). It is recognised that one of the most common sources of contamination is by fungi, and this can be detected by simple visual inspection. Prior to continued storage, archived samples should therefore be checked and then inspected annually thereafter. In the event of fungal contamination, samples may be dried at 65°C for at least twenty minutes. Whether obviously contaminated (requiring heat treatment) or not, all sample envelopes should subsequently be spread out in a single layer in a dry environment for a day or two prior to re-storage. Samples should then be stored in the driest possible environment taking into consideration that humid, dark environments favour fungal growth. If this material is considered as being of high value, transferring it to 70% ethanol or immediate DNA extraction and storage should be considered (De Salle *et al.*, 1993; Yoshinaga *et al.*, 1997; Rohland *et al.*, 2004; Hunter *et al.*, 2008, and see also ICES WGAGFM 2007 report ToR c), though there may be merit to keeping both the original scales or otoliths and also extracted DNA—see below.

5.6.2 New Collections

Best practice for new collections of scales and otoliths should ensure rapid drying of material to minimise biological degradation. Due to the sensitivity of molecular techniques, relatively little material is required. Therefore, scales should be spread thinly within the envelope when being collected (Nielsen *et al.*, 1999), and then the envelopes laid out to air-dry individually as soon as possible after collection prior to accumulating for storage. Acid free paper envelopes are recommended (Monica Akerlund, Stockholm Natural History Museum, pers. comm.). Subsequently the

storage conditions described above to minimise further DNA degradation, should be applied. It is worth considering a dual approach, where DNA is immediately extracted from scales or otoliths using the most appropriate techniques, but that aliquots of the mineralised tissues are also stored under optimal conditions. Setting aside scales or otoliths for future use in this way will allow for the advent of new improved DNA extraction technologies (Rivers and Ardren, 1998; Hutchinson *et al.*, 2003; Rainbow-in press).

5.7 Multiple Uses

In the past, otoliths have been collected primarily for ageing, and the preparation methods involved may reduce DNA quantity and quality. Thus, consideration should be given to the order in which otoliths are used in various laboratory studies to allow for initial DNA extraction prior to subsequent experimentation. Furthermore, Heath *et al.* 2007 reported that DNA extraction from the otoliths of common coral trout (*Plectropomus leopardus*) had no significant effect on their integral structure and subsequently allowed age estimation. In the case of scales, there is usually enough material collected to allow sampling exclusively for DNA. Many of the scales collected from an individual fish may consist of replacement scales and may not be suitable for ageing, and thus will be available for other techniques including DNA extraction.

5.8 Conclusions

As stated above many archived collections exist in fisheries laboratories and research institutes around the world and in some cases may be known only to a few people. As their value extends to the whole scientific community a means of informing researchers in this field of the existence of these collections would be very valuable. According to Rivers and Ardren 1998, the age of samples suitable for DNA extraction should not be an issue, as sufficient DNA has been extracted from sub-fossil mammalian bones up to 47,000 years old. This depends entirely on the conditions in which the tissue was preserved. With the development of new technologies in the future, optimal preservation of archival collections will facilitate further valuable research for conservation and sustainable exploitation of fish stocks.

5.9 Recommendations

- ICES should request information on all scale and otolith collections held by fisheries laboratories, institutes and universities. This information should be accessible via the proposed ICES genetic database.
- If disposal of archived material is envisaged, the scientific community should be notified via the ICES network, so that alternative storage can be arranged if there is sufficient interest.
- Archived material should be inspected on a regular basis and subsequently stored in controlled conditions. It should be borne in mind that most fungi prefer dark humid environments. Thus dry storage conditions are vital. Samples contaminated with fungi may be heated to 65°C for twenty minutes to halt current fungal growth.
- For new scale or otolith collections, every precaution should be taken to prevent cross-contamination. There is a much greater risk of contamination due to the sensitivity of molecular techniques.
- Preferably a non-mineralised tissue should be sampled in the future, (e.g. a fin clip), solely for genetic analysis, due to the clear benefits of molecular

- If the only samples being taken in future surveys are scales or otoliths collected in the traditional way, then the material should be placed in a thin layer within the envelope to aid rapid drying. Both scales and otoliths should be stored as recommended for archived material.

5.10 References

- Adcock, G.J., Ramirez, J.H.B., Hauser, L., Smith, P., and Carvalho, G.R. 2000. Screening of DNA polymorphisms in samples of archived scales from New Zealand snapper. *Journal of Fish Biology* 56, 1283–1287.
- Arai, T., and Takafumi, H. 2006. Differences in the Trace Element Deposition in Otoliths between Marine- and Freshwater-resident Japanese Eels, *Anguilla japonica*, as determined by Laser Ablation ICPMS. *Environmental Biology of Fish* 75, 173–182.
- Brazner, J.C., Campana, S.E., Tanner, D.K., and Schram, S.T. 2004. Reconstructing habitat use and wetland nursery origin of yellow perch from Lake Superior using otolith elemental analysis. *Journal of Great Lakes Research* 30, 492–507.
- Campana, S.E. 1999. Chemistry and composition of fish otoliths: pathways, mechanisms and applications. *Marine Ecology Progress Series* 188, 263–297.
- Campana, S.E. 2005. Otolith science entering the 21st century. *Marine and Freshwater Research* 56, 485–495.
- Campana, S.E., and Thorrold, S.R. 2001. Otoliths, increments, and elements: keys to a comprehensive understanding of fish populations? *Canadian Journal of Fisheries and Aquatic Sciences* 58, 30–38.
- Daverat, F., Tomas, J., Lahaye, M., Palmer, M., and Elie, P. 2005. Tracking continental habitat shifts of eels using otolith Sr/Ca ratios: validation and application to the coastal, estuarine and riverine eels of the Gironde-Garonne-Dordogne watershed. *Marine and Freshwater Research* 56, 619–627.
- DeSalle, R., Williams, A.K., and George, M. 1993. Isolation and characterization of animal mitochondrial DNA. In: *Molecular evolution: producing the biochemical data*. E.A. Zimmer., T.J. White, R.L. Cann and A.C. Wilson AC (eds.). *Methods in Enzymology* 224, 176–204. Academic Press.
- Elfman, M., Limburg, K.E., Kristiansson, P., Malmqvist, K., and Pallon, J. 1999. Application of micro-PIXE to fish life history analyses: trace element analysis of otoliths. *Nuclear Instruments and Methods in Physics Research Section B* 150, 272–286.
- Estoup, E., Largiader, C.R., Perrot, E., and Chourrout, D. 1996. Rapid one-tube extraction for reliable PCR detection of fish polymorphic markers and transgenes. *Molecular Marine Biology and Biotechnology* 5, 295–298.
- Gilbert, T. M. P., Bandelt, H.J., Hofreiter, M., and Barnes, I. 2005. Assessing ancient DNA studies. *Trends in Ecology and Evolution*: 20, No.10: 541–544.
- Heath, D., Lou, D.C., Ovenden, J., and Begg, G.A. 2007. Effect of DNA extraction on ageing success in coral trout (*Plectropomus leopardus*) otoliths. *Journal of Fish Biology*, 302–307.
- Hoarau, G., Boon, E., Jongma, D.N., Ferber, S., Palsson, J., Van der Veer, H.W., Rijnsdorp, A.D., Wytze, T.S., and Olsen, J.L. 2005. Low effective population size and evidence for inbreeding in an overexploited flatfish, plaice (*Pleuronectes platessa* L.). *Proceedings of the Royal Society B* 272, 497–503.
- Hofreiter, M., Serre, D., Noinar, H., Kuch, M., and Paabo, P. 2001. Ancient DNA. *Nature Reviews*, 353–359.

- Hua Yue, G., and Orban, L. 2001. Rapid isolation of DNA from fresh and preserved fish scales for polymerase chain reaction. *Marine Biotechnology* 3, 199–204.
- Humphreys, R.L.Jr., Campana, S.E., and DeMartini, E.E. 2005. Otolith elemental fingerprints of juvenile Pacific swordfish *Xiphias gladius*. *Journal of Fish Biology* 66, 1660–1670.
- Hunter, S.J., Goodall, T.L., Walsh, K.A., Owen, R., and Day, J.C. 2008. Nondestructive DNA extraction from blackflies (Diptera: Simuliidae): retaining voucher specimens for DNA barcoding projects. *Molecular Ecology Resources* 8, 56–61.
- Hutchinson, W.F., Carvalho, G.R., and Rogers, S.I. 1999. A non-destructive technique for the recovery of DNA from dried fish otoliths for subsequent molecular genetic analysis. *Molecular Ecology* 8, 891–894.
- Hutchinson, W.F., Oosterhout, C. van., Rogers, S.I., and Carvalho, G.R. 2003. Temporal analysis of archived samples indicates marked genetic changes in declining North Sea cod (*Gadus morhua*). *Proceedings of the Royal Society, B* 270: 2125–2132.
- Martinez, J.L., Dumas, J., Beall, E., and Garcia-Vazquez, E. 2001. Assessing introgression of foreign strains in wild Atlantic salmon populations: variation in microsatellites assessed in historic scale collections. *Freshwater Biology*, 46: 835–844.
- Mecklenburg, M.F., Tumosa, C.S., and Pride, A. 2004. Preserving legacy buildings. *ASHRAE Journal*, 41: 18–23.
- Mitchell, D., McAllister, P., Stick, K., and Hauser, L. 2008. Sperm contamination in archived and contemporary herring samples. *Molecular Ecology Resources*, 8: 50–55.
- 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: 265–276.
- Nielsen, E.E., MacKenzie, B.R., Magnussen, E., and Meldrup, D. 2007. Historical analysis of Pan I in Atlantic cod (*Gadus morhua*): temporal stability of allele frequencies in the southeastern part of the species distribution. *Canadian Journal of Fisheries and Aquatic Sciences*, 64: 1448–1455.
- Paabo, S., Poinar, H., Serre, D., Jaenicke-Despres, V., Hebler, J., Rohland, N., Kuch, M., Krause, J., Vigilant, L., and Hofreiter, M. 2004. Genetic analyses from ancient DNA. *Annual Review of Genetics*, 38: 645–679.
- Popper, A.N., Ramcharitar, J., and Campana, S.E. 2005. Why otoliths? Insights from inner ear physiology and fisheries biology. *Marine and Freshwater Research* 56, 497–504.
- Poulsen, N.A., Nielsen, E.E., Schierup, M.H., Loeschcke, V., and Gronkjaer, P. 2006. Long-term stability and effective population size in North Sea and Baltic Sea cod (*Gadus morhua*). *Molecular Ecology* 15, 321–331.
- Rainbow, P S. Marine Biological collections in the 21st century. *Zoologica Scripta*-in press.
- Reibisch, J. 1899. Über die Eizahl bei *Pleuronectes platessa* und die Altersbestimmung dieser Form aus den Otolithen. *Wissensch Meeresuntersuchungen, Kiel* 4, 231–248.
- Rivers, P.J., and Ardren, W.R. 1998. The Value of Archives. *Fisheries Research* 23, 6–9.
- Rodríguez Mendoza, R.P. 2006. Otoliths and their applications in fishery science. *Ribarstvo* 64, 89–102.
- Rohland, N., Siedel, H., and Hofreiter, M. 2004. Nondestructive DNA extraction method for mitochondrial DNA analyses of museum specimens. *Biotechniques*, 36: 814–821.
- Rollo, R. 1998. Ancient DNA: Problems and Perspectives for Molecular Microbial Palaeoecology. *Advances in Molecular Ecology*. G.R.Carvalho (ed). IOS Press, 133–149pp.
- Tsukamoto, K., Nakai, I., and Tesch, W.V. 1998. Do all freshwater eels migrate? *Nature*, 396: 635–636.

- Wandeler, P., Hoeck, P.E., and Kelle, L.F. 2007. Back to the future: museum specimens in population genetics. *Trends in Ecology and Evolution*, 22: 634–642.
- Yoshinaga, K., Yoshioka, H., Kurosaki, H., Hirasawa, M., Uritani, M., and Hasegawa, K. 1997. Protection by trehalose of DNA from radiation damage. *Bioscience Biotechnology and Biochemistry*, 61: 160–161.
- Zumholz, K., Hansteen, T.H., Piatkowski, I.U., and Croot, P.L. 2007. Influence of temperature and salinity on the trace element incorporation into statoliths of the common cuttlefish (*Sepia officinalis*). *Marine Biology*, 151: 1321–1330.

6 ToR e): To evaluate prospects for the use of genetic monitoring for evaluating conservation status, intraspecific biodiversity, and stock “health” in fishes.

Eric Verspoor, Ellen Kenchington, Philip McGinnity, Niklas Tysklind, Eef Cauwelier and Oleg Vasin

6.1 Rationale

Fisheries around the world are in crisis (e.g. Beddington *et al.*, 2007) and there is increasing recognition of the importance of monitoring, not only “stock” abundance to regulate fisheries exploitation, but also the inherent (i.e. genetic) spatial and temporal biological diversity within fish species (Hedgcock *et al.*, 2007), both in support of fisheries management and fish conservation. The effective management of marine fisheries remains an on-going challenge beset by the lack of essential insights for the effective regulation of exploitation.

In their recent review, Schwartz *et al.* (2006) conclude that few management programmes “...take full advantage of the potential afforded by molecular genetic markers, which can provide information relevant to both ecological and evolutionary time frames, while costing less and being more sensitive and reliable than traditional monitoring approaches.” With the notable exception of Pacific Salmon (e.g. Beacham *et al.*, 2004), there are only a few examples of the routine application of genetic tools in fisheries management.

Perceptions, such as those of Walters and Martell (2004), that “Some newer and very expensive activities, such as genetic stock identification ...have produced some apparently useful results ...but suspiciously little contribution to the actual practice of management.”, remain widespread. This is undoubtedly conditioned to some extent by the usefulness and cost of gaining genetic information using molecular markers. However, there is also an institutional inertia and resistance to new approaches and paradigms.

The last decade has seen the scope for the cost-effective application of genetic monitoring increase dramatically as new, more technologies are exploited (Schwartz *et al.*, 2006), increasing the potential for gaining important and useful demographic and evolutionary insights. This arises from new classes of markers, such as microsatellite loci and SNPs (single nucleotide polymorphisms– e.g. Morin *et al.*, 2004), as well as new screening (e.g. capillary electrophoresis and DNA microchips – e.g. Hardenbol *et al.*, 2005) and statistical methodologies (e.g. Bayesian methods – Beaumont, 2003). The potential for deriving useful insights from the analysis of genetic variation in marine fishes covered by the ICES remit is considered here.

6.2 Breeding populations and fisheries

Harvesting of fish biomass needs to be sustainable i.e. to be carried out so as to maintain the viability of the fish stocks targeted and their ability to produce an exploitable recruitment surplus. Fundamentally, this means maintaining the breeding capacity of a stock, and in the longer term, its adaptive diversity for coping with environmental change. These concerns are fundamentally genetic issues.

Breeding populations or aggregates were recognised by early fisheries scientists as being of fundamental importance in fisheries recruitment and, therefore, to fisheries management (see Berst and Simon, 1981). However, despite early recognition of their

importance, identifying and managing breeding populations has proved difficult in practise, particularly in marine fish as their logistically challenging environment makes it difficult to define their spatial and temporal dynamics. Yet spatial or temporal structuring into distinct breeding groups (genetic populations), where it occurs, will be of fundamental importance to local species recruitment in marine fisheries where they differ inherently in their distribution in time or space, or in characteristics such as size at age, age of maturation, or spawning success. As such the conservation of breeding populations, and the heritable adaptive diversity they represent, will be central to the maintenance of healthy fish stocks, and fisheries management must be focused on assessing and monitoring their viability and capacity for fisheries exploitation. Currently, in the absence of detailed population genetic insights, management of many species is focused on geographical stocks, often defined on the basis of oceanographic regions or arbitrary map co-ordinates (e.g. "ICES rectangles"). Such stocks may represent part or all of one or more breeding populations.

Ideally, management should be specifically focused on breeding populations rather than on arbitrary spatial clusters of conspecifics. The importance of doing so is increasing, as more and more species undergo local and regional declines and the fisheries they support collapse (e.g. Grand Banks Atlantic cod - Smedbol and Wroblewski, 2002). Reversing these declines and developing sustainable fisheries represents a growing challenge, one which can be met by adopting population-sensitive approaches within overall ecosystem-based management (e.g. Policansky and Magnuson, 1998).

6.3 Structuring into breeding populations

Most marine fish species are widely distributed across a range of geographical and environmental conditions. Depending on their inherent biological character, habitat fragmentation combined with homing to natal spawning areas can variously lead to structuring into multiple, distinct breeding aggregations. To the extent that these are temporally coherent and interbreeding is more or less constrained over time, aggregations will evolve distinct gene pools, reinforced in some cases by selective processes associated with environmental heterogeneity and adaptation to local environmental conditions. This can lead to a wide range of potential population structuring scenarios.

The issue of connectivity of breeding groups (=genetic populations or sub-populations) has recently reviewed by Hedgecock *et al.* (2007). Its extent will depend on the interaction of evolved biological characteristics of the species with specific environmental conditions. Biological factors of particular importance will be the basic breeding strategy of a species (broadcast spawning vs. pair-mating vs. brood guarding) and, where it is disposed to disperse or migrate, on the strength and scope for natal homing. Important environmental factors will be the physical or hydrographic fragmentation and isolation of spawning areas, and its temporal stability, as well as their spatial and temporal environmental differentiation with regard to physiochemical factors which affect growth, survival and reproductive success. The latter may lead to differential selection on the heritable variation for these traits found in all species and, where genetic exchange is limited, lead to adaptive genetic differentiation among populations. Environmental variations may also lead to locally varying population abundance which may influence levels of dispersal and genetic exchange. The collective interaction of these factors over time will determine the nature and extent of contemporary structuring into populations.

Most of the 34 fin and shell fish species of interest to ICES have been studied with regard to genetic population structuring (Table 1) but for most species understanding of population structuring remains rudimentary or local. Indeed, for many species even rudimentary biogeographical understanding of species ranges, overall distribution within range, and distribution of spawning areas (Figure 1) is still more or less unknown. Even for the best studied species such as the Atlantic salmon, Atlantic cod and Atlantic herring, understanding is far from complete, though for Atlantic salmon this incompleteness relates largely to the distribution of spawning within river systems and in many areas this type of information is already available.

From the species which have been studied, however, it is clear that the full gamut of population structuring scenarios is likely to exist, among as well as within species. At one extreme, is the generally high degree of structuring seen in the anadromous Atlantic salmon (King *et al.*, 2007; Vaha *et al.*, 2008), and at the other extreme the apparent absence of structuring in the catadromous eel (e.g. Dannewitz *et al.*, 2005; Maes *et al.*, 2006). In other species it would appear that in some areas structuring may encompass populations which are more or less completely reproductively isolated (e.g. herring – Bekkevold *et al.*, 2005; cod – Pogson and Fevolden, 2003) to populations that may be transient and only exist in more or less isolation for relatively short periods of time in evolutionary terms (e.g. herring – Jørgensen *et al.*, 2005; cod – Smedbol and Wroblewski, 2002), though these times may still be highly significant from a management perspective (years, decades or centuries). Potentially at least populations within species may exhibit a range of continuous or ephemeral genetic connectivity ranging from no genetic exchange to rates that approach effective panmixia.

The nature and extent of structuring of species stocks into distinct breeding populations, together with their degree of connectivity, will have important implications for both short and long-term stock demographics and for species conservation requirements. It will also have important implications for the way that fisheries can best be managed to achieve sustainability in the face of exploitation pressures and environmental change (e.g. global warming). This needs to be understood along with the demographic and reproductive status of the breeding populations of which fish stocks are composed. In many respects, this can be most effectively achieved by the application of genetic assessment and monitoring methods.

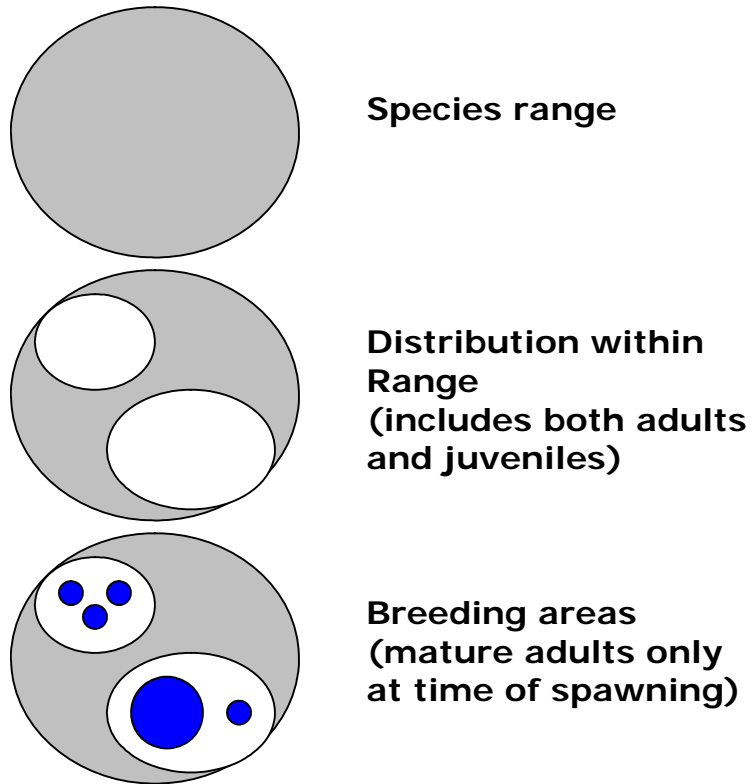


Figure 1. A schematic of the different levels of biogeographical information relevant to understanding structuring of fish into genetic populations.

Table 1. List of species for which ICES undertakes assessments with associated level of information available to provide genetic advice. Categories summarizing the state of knowledge on population structure refer to Figure 1 (1=range known; 2=distribution within range known; 3=breeding areas known). The number of crosses reflects the relative number of population genetic studies undertaken for each species (range 0-5). Categories summarizing the geographic scale of existing population genetic models are: N=none, L=local, R=range. Table is based on rough assessments of existing information sources by WGAGFM members.

AL CLASSIFICATION	COMMON NAME	SPECIES	BIOLOGY	POPULATION GENETIC DATA	POPULATION GENETIC MODELLING
Demersal Fish					
Gadoids	Cod	<i>Gadus morhua</i>	3	+++++	L, R,G
	Haddock	<i>Melanogrammus aeglefinus</i>	+	++	R
	Saithe/Pollock	<i>Pollachius virens</i>	+	+	R
	Pollack	<i>Pollachius pollachius</i>	+	+	R
	Whiting	<i>Merlangius merlangus</i>	+	++	R,G
	Norway Pout	<i>Trisopterus esmarkii</i>	+	+	R
	Hake	<i>Merluccius merluccius</i>	+	+++	L,R,G
Flatfish	Plaice	<i>Pleuronectes platessa</i>	3	++	L, R
	Sole	<i>Solea solea</i>	3	+++	L, R
	Dab	<i>Limanda limanda</i>	3	in progress	N
	Lemon Sole	<i>Microstomus kitt</i>	+		N
	Witch Flounder	<i>Glyptocephalus cynoglossus</i>	2	+	N
	Megrim	<i>Lepidorhombus whiffiagonis</i>	2	+	N
	Turbot	<i>Psetta maxima</i>	2	++	L
	Brill	<i>Scophthalmus rhombus</i>	2	+	N
	Flounder	<i>Platichthys flesus</i>	2	++	L, R
	Greenland Halibut	<i>Reinhardtius hippoglossoides</i>	+	+	L
	Halibut	<i>Hippoglossus hippoglossus</i>	3	++	L, R
Monkfish	Anglerfish	<i>Lophius budegassa</i>	+	+	N
	Anglerfish	<i>Lophius piscatorius</i>	+	++	N
Redfish	Beaked Redfish	<i>Sebastes mentella</i>	+	++	L,R
	Golden Redfish	<i>Sebastes marinus</i>	+	++	L,R
	Acadian Redfish	<i>Sebastes fasciatus</i>	+	++	L,R
Crustaceans	Nephrops	<i>Nephrops norvegicus</i>	+	+++	Yes
	Northern prawn	<i>Pandalus borealis</i>	+	+++	Yes
Pelagic Fish	Mackerel	<i>Scomber scombrus</i>	+	+	R
	Horse Mackerel	<i>Trachurus trachurus</i>	+	+++	Yes
	Herring	<i>Clupea harengus</i>	+	+++	Yes
	Capelin	<i>Mallotus villosus</i>	+	+++	
	Sandeel	<i>Ammodytes marinus</i>	+	+	No
	Blue Whiting	<i>Micromesistius poutassou</i>	+	+++	Yes
	Sprat	<i>Sprattus sprattus</i>	2	++	R
	European Anchovy	<i>Engraulis encrasicolus</i>	+	+++	Yes
	Sardine	<i>Sardina pilchardus</i>	+	+++	Yes
	Swordfish	<i>Xiphias gladius</i>	+	+++	Yes
	Bluefin Tuna	<i>Thunnus thynnus</i>	+	+++	Yes
Deepwater Fish	Blue Ling	<i>Molva dypterygia</i>	?	No	No

AL CLASSIFICATION	COMMON NAME	SPECIES	BIOLOGY	POPULATION GENETIC DATA	POPULATION GENETIC MODELLING
	Ling	<i>Molva molva</i>	?	No	No
	Tusk	<i>Brosme brosme</i>	+	+	No
	Alfonsinos	<i>Beryx spp.</i>	+	++	Yes
	Argentine	<i>Argentina silus</i>	?	No	No
	Black Scabbardfish	<i>Aphanopus carbo</i>	?	++	No
	Greater forkbeard	<i>Phycis blennoides</i>	?	No	No
	Orange Roughy	<i>Hoplostethus atlanticus</i>	+	++	Yes
	Roundnose Grenadier	<i>Coryphaenoides rupestris</i>	+	++	Yes
	Blackspot Sea Bream	<i>Pagellus bogaraveo</i>	+	++	Yes
Diadramous Fish	Salmon	<i>Salmo salar</i>	3	+++++++	L,R
	Sea Trout	<i>Salmo trutta</i>	3	++	L,R
		*(<i>B. splendens</i> ; <i>B. decadactylus</i>)			

6.4 Genetic assessment and monitoring

A wide range of genetic metrics have been developed within the conceptual framework of general population genetics theory (Weir 1996; Balding *et al.*, 2003). At the centre of this framework is the concept of distinct breeding assemblages known as demes or genetic populations, which constitute local groups of conspecific individuals within which interbreeding is more or less random but among which interbreeding is constrained. Each such population, by virtue of its reproductive isolation, represents a distinct pool of genetic material (=gene pool). Where a species contains many such biological groups they are also often referred to in the literature as sub-populations; in this conceptualization they are seen to exist as subunits of the overall species population. The conceptual framework of population genetics makes it possible to draw inferences on breeding populations from the analysis of the distribution of genetic variation and how this change over time (genetic monitoring), and has the potential to provide valuable insights for fisheries management. These insights relate to the structuring of stocks into breeding populations and the extent of their reproductive connectivity, as well as to their demographic and evolutionary status.

Despite early recognition of their importance to fisheries recruitment and management, identifying and managing breeding populations has proved difficult, particularly for marine species. The lack until recently of simple, cost-effective methods for studying and monitoring genetic variation have, until recently, made it almost impossible to integrate population genetic considerations into practical species conservation and management. As a result the field of population genetics developed over the last century largely as a theoretical discipline, supported by genetic analyses whose scope was highly constrained and based largely on laborious methods such as the microscopic analysis of chromosome structure of model organisms such as *Drosophila*. Building on advances in DNA screening over the last two decades, it is becoming increasingly possible to achieve this essential integration and the general potential that now exists for applying genetic methods to biological monitoring in species conservation and management has been reviewed by Schwartz *et al.* (2006).

Identifying Individuals, Populations and Species

Species

Biodiversity is fundamentally genetic diversity and is easily assessed at the species and ecosystem level where underlying genetic diversity leads to unambiguous morphological differentiation among different species. This area applies to both fin fish and shell fish, and their diseases and pathogens. However, it is more difficult to assess where well defined heritable differences are lacking as species designations become more arbitrary. This may be the case at some or all life history stages and provide little scope for assessing levels of interspecific hybridisation and genetic introgression, or identifying cryptic species, whose evolutionary divergence is only unambiguously resolvable at the molecular level (e.g. Hyde *et al.*, 2008).

True evolutionarily and biologically distinct species can be expected to have evolved heritable diagnostic differences in many regions of their nuclear and mitochondrial genomes. Thus in most cases where morphological differentiation is not possible or reliable, rapid cost-effective diagnostic molecular markers will be able to be identified e.g. restriction fragment digestion of amplified DNA regions (e.g. Pendas *et al.*, 1995; Rocha-Olivares, 1998). The key issue, as with any diagnostic marker, be they morphological or molecular, is to be able to demonstrate that it is indeed diagnostic. If so, diagnostic nuclear markers can be used to identify first generation hybrids which will be heterozygous for the marker used. Levels of introgression (i.e. the proportion of individuals of one species, or of its genome, containing gene variants derived from another species can also be assessed by sampling. However, not all offspring in post-F1 generations will inherit species diagnostic markers, and the probability of inheritance of a marker needs to be taken into account in estimating introgression levels. By using more diagnostic markers, the accuracy of estimation of rates and degrees of introgression can be increased (Campton, 1987).

Populations

Understanding how fish and pathogen species are structured into breeding populations arguably poses one of the biggest biological challenges facing fisheries managers. This encompasses long standing phylogenetic biodiversity associated with long-term genetic isolation, often encompassing large groups of breeding populations, and more evolutionarily ephemeral short and long term metapopulation structuring within phylogeographic groups. The nature of the structuring that exists will be reflected in the nature and distribution of genetic diversity within and among individuals. However, identifying informative variation and devising appropriate sampling strategies for discriminating the wide range of possible population structuring scenarios is not a trivial problem. Developing a complete understanding, particularly where evolutionary divergence is limited and transient can be expected to require an iterative process of sample collection and analysis, and careful collective consideration of genetic, demographic and environmental information. Indeed, where evolutionary divergence is transient and short-term within a metapopulation dynamic, standard genetic metrics such as F_{ST} will often prove to be relatively uninformative. However, in population studies, considerable care must be taken as the absence of evidence in many cases will not be evidence of absence as the absence of evidence will often be a reflection of the limitation of the genetic loci and metrics used.

In principal, at least, population structure can be resolved by analysing the level of genetic differentiation with regard to molecular variation either among samples or

among individuals. At the sample level, genetic distance metrics such as F_{ST} or others (see chapters in Balding *et al.*, 2003) are generally used to establish whether the level of genetic differentiation is inconsistent with samples having derived from a single genetic population. Differentiation can arise variously due to genetic drift, new local mutations, and selectively driven local adaptation. Using individual information, an assessment is made to determine whether more than one distinct cluster of related individuals is observed as expected if they belong to different more or less distinct breeding populations (e.g. Pritchard *et al.*, 2000). For both approaches, it is essential that observed differentiation is stable across generations rather than due to within-generation selection among different temporal or spatial groups derived from a single breeding population.

The degree of genetic differentiation, and hence the ability to resolve structuring, depends on a number of interacting variables. These include the length and degree of isolation, generation time, effective population size, mutation rate, and degree of adaptive selection. In general, differentiation increases the longer and greater the degree of reproductive isolation, and with shorter generation times and stronger local adaptation. At loci unaffected by selection, it is inversely proportional to the historical numbers of breeders (N_e - effective size of populations) and, assuming the same level of genetic exchange, the same level of differentiation takes longer to achieve in larger populations than in smaller populations. Thus it is easiest to detect structure where species have a short generation time and populations are small and locally adapted, have been isolated for a long time, and have no genetic exchange. It is hardest to resolve structuring in species with a long generation time and composed of large, recently established populations with little adaptive differentiation and relatively high levels of continuous, or sporadic, genetic exchange.

The degree of genetic differentiation observed will also vary depending on the molecular variation considered. At loci subject to genetic drift, the level of genetic differentiation will vary due to chance with some showing higher levels than others. Also, in general, loci that have higher mutation rates can be expected to differentiate more quickly among populations. Furthermore, other things equal, in large populations, loci subject to adaptive differentiation will be relatively more differentiated than those subject to genetic drift as genetic drift decreases with increasing population size. Thus “neutral loci” can be expected to be more useful in identifying population structuring where populations are small and loci involved in adaptive differentiation more useful where populations are large.

The analysis of variation at loci showing differentiation among breeding populations can be used to determine the proportional contribution of different breeding populations to samples of fish taken in fisheries or captured as part of ecological studies into the temporal and spatial dynamics of stocks at sea. This methodology, known as mixed stock analysis (MSA), is routinely used in the management of Pacific salmon fisheries and in studies of their marine ecology (e.g. Beacham *et al.*, 2005, 2006), and has recently been reviewed by Koljonen *et al.* (2007). It has also been explored for some marine species (e.g. Ruzzante *et al.*, 2000; Jónsdóttir *et al.*, 2007).

Individuals

Where structuring into breeding populations occurs, variation at loci showing differentiation can be used to determine the population membership of individual fish. The ability to assign individuals successfully will depend on a number of factors: level of genetic differentiation, level of polymorphism at loci used, the number of loci, the size of baseline samples, the number of populations, and the statistical method

used (Koljonen *et al.*, 2007). In general the power of individual assignment (IA) decreases as F_{ST} gets smaller. Thus it is likely to be more useful in situations where populations are both smaller as well as smaller in number, particularly when using hypervariable loci such as microsatellite DNA. To date the use of IA has been limited largely to species such as salmonids where breeding populations are highly differentiated (e.g. Beacham *et al.*, 2005, 2006), though its use has been explored in some marine species (e.g. *Sebastes* - Roques *et al.*, 1999; *Clupea* - Jørgensen *et al.*, 2005). In situations where IA is possible, it can be used to study the spatial boundaries of populations and estimate the degree of population mixing, and to provide insights into population dynamic processes such as dispersal and migration (Schwartz *et al.*, 2006).

Monitoring Population Status

Analysis of genetic variation can, in principal, be used to provide insight into the short or long-term biological status of populations (Swartz *et al.*, 2006), either by using genetic variation to tag individuals, or by providing genetic metrics that can be monitored for changes. However, the routine application of genetic markers to fisheries management remains largely confined to Pacific salmon fisheries (e.g. Beacham *et al.*, 2005, 2006) and relates largely to insights gained from MSA and IA analyses. There is potential for the application of genetic markers in mark and recapture studies to assess abundance and vital rates (mortality/survival), exploiting the fact that each individual has a unique genotype (Lukacs and Burnham, 2005). However, in practise, this application is likely to be restricted to small populations where there is a reasonably high probability of recapture/detection of "marked" individuals. It is also likely to be further limited to species where physical tagging is difficult but biological material can still be obtained.

The lack of wider routine application, particularly to most other marine fish species, is likely to be attributable to two factors. Firstly, as discussed above, understanding of population structuring in most species is limited and the use of genetic metrics is in most cases dependent on having such an understanding. Secondly, the estimation of most genetic metrics is based on evaluating the distorting effects on the distribution of genetic variation of genetic drift, mutation and migration. Unfortunately, the effect decreases as the number of breeders increases and therefore, in general, increasingly large sample sizes appear to be needed to obtain equivalently accurate and reliable estimates as population size increases. Thus, in general, monitoring populations using genetic methods and metrics becomes increasingly difficult as the size of the breeding populations monitored increases (Figure 2). The only exception is metrics such as reaction norm that measure genetically-based phenotypic change (e.g. Barot *et al.*, 2004). Here the changes measured are due to selection rather than genetic drift, a force which has the potential to be increasingly strong, and therefore more easily detected, in larger populations.

The numbers of spawners contributing to the next generation is a key management metric in fish conservation and establishing sustainable levels of exploitation, but direct counts are difficult and proxy estimates of spawning escapement and biomass are generally employed. This demographic variable is generally difficult to estimate. One genetic metric that is being increasingly seen as useful in this regard is N_e – the effective breeding size of populations, or its related derivative, N_b – the effective number of breeders. In a population with no genetically effective immigrants, this is related to the historical number of breeders contributing to each generation and their relative reproductive success. In so far as genetic exchange increases, N_e also

increases for a given number of breeders. Estimates of N_e can be obtained by a number of different methods based on the analysis of linkage disequilibrium across loci, departures from Hardy-Weinberg expectations, and changes in allelic diversity over time (Waples, 2005). However, obtaining accurate estimates of N_e with tight confidence limits is problematic and appears to be a direct function of sample sizes, at least for some methods (Waples, 2006). As such employing N_e as a monitoring metric is potentially practical for small to medium sized populations, where sample sizes of 50 to a few hundred will provide useful estimates for detecting changes. It is unlikely to prove to be a useful monitoring metric for large breeding populations except where changes in numbers of breeders are in terms of orders of magnitude. New methods need to be developed to calculate N_e values more accurately where confidence limits are not a function of N_e itself, if this metric is to be able to be more widely applied.

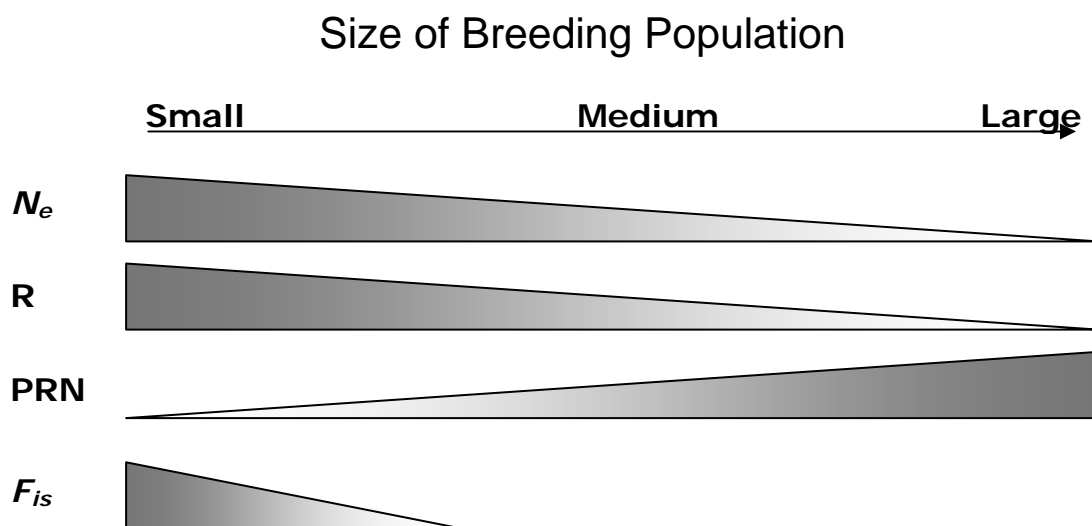


Figure 2. The ease with which some of the more typical genetic metrics can in principal be expected to be estimated and used for monitoring, relative to the size of the breeding population involved; population size range is from few individuals to hundreds of thousands; genetic metrics: N_e – effective breeding size of population, R – allelic richness (standardized allelic diversity), PRN – genetic probabilistic reaction norms for quantitative traits, F_{is} – coefficient of inbreeding.

N_e is affected by the breeding behaviour (e.g. life time monogamy, random mating – Waples, 2006) as well as the extent of genetic connectivity of populations i.e. m , the migration rate, and the joint analysis of these two parameters, for which theoretical approaches have been developed (e.g. Wang and Whitlock 2003) can be highly informative (e.g. Frazer *et al.*, 2007). However, again, the use of N_e and accurate estimation of m generally requires a good basic understanding of structuring into populations to ensure appropriate sampling is carried out.

In general, levels of genetic diversity will be a function of N_e ; as N_e increases, levels of genetic diversity can be expected to increase. In addition to using N_e as a proxy for genetic diversity, a number of other well-developed metrics are available for monitoring levels of genetic diversity within populations. Their routine and meaningful use in monitoring requires that sampling is random and encompasses

only the population of interest. Thus, again, an understanding of population structuring is essential if they are to be meaningfully employed. Also, ideally, loci should be representative markers of overall levels and types of genetic diversity in the genome so as to reflect overall genetic diversity. The size of sample required to obtain useful estimates depends on the number and frequency of allelic variants in a population and the required sample sizes to obtain robust estimates for monitoring population status will need to be assessed in each individual case.

The metric A is the absolute number of genetic variants (alleles) observed at a locus, or across loci. However, as increasingly variable loci are employed, the estimate of A becomes dependent on sample size and it is more informative to use the metric A_R or R (Allelic Richness) which is A standardized with respect to sample size. This provides a useful metric for assessing changes in diversity based on the average number of alleles detected in each generation for a given sample size. Other metrics that can be used are H (heterozygosity) and F_{IS} (coefficient of inbreeding). In general, the loss of alleles happens more rapidly than reductions in heterozygosity (Allendorf, 1986) and can be a more sensitive indicator of losses due to increased genetic drift associated with a decrease in the effective size of the breeding population (also known as a bottleneck). As population size increases, the effect of drift decreases and the magnitude of chance genetic change decreases, such that increasingly large samples are needed to detect change and routine monitoring becomes more difficult and less meaningful.

The random loss of genetic diversity will be of greatest in small populations where levels are relatively easy to monitor. However, significant changes can be expected when there is mass mortality and severe bottlenecking of large populations occurs. To detect such changes estimates of A_r are required before and after the mass mortality which may be quite difficult. For example, in the case of plaice, increase of fishing pressure in the 1950s, combined with breeding aggregation fishing in the North Sea triggered a rise in the number of inbred individuals, a situation that continues to the present (Hoarau *et al.*, 2005). This can be expected to be associated with a loss of molecular genetic diversity.

6.5 Conclusions

- Genetic metrics have considerable potential for application to the monitoring of fish populations, particularly in helping to define structuring into breeding populations, but also to provide demographic and biodiversity insights.
- For most commercial fish species covered by the ICES remit, the understanding of structuring into breeding populations is limited and further biological and genetical studies are required to understand this aspect of their biology which is fundamental to effective fisheries management and fish conservation.
- The full scope for using genetic metrics in marine fish species to monitor population status remains unclear and will depend, among other things, on the way they are structured into breeding populations, the sizes of the populations and their degree of interconnectivity.
- The scope for routine application of genetic metrics is currently largely restricted to anadromous species such as salmon, which are composed of large numbers of relatively small breeding populations.

- More work is needed to assess the potential for developing analytical methods for practical estimation of demographically informative genetic metrics such as N_e in large, relatively highly interconnected populations, as are likely to exist for many marine species.
- Genetic metrics can be usefully used in conjunction with existing demographic metrics derived using methods, such as tagging.

6.6 References

- Allendorf, F. W. 1986. Genetic drift and the loss of alleles versus heterozygosity. *Zoological Biology*, 5: 181–190.
- Balding, D., Bishop, M., and Cannings, C., eds. *Handbook of Statistical Genetics*, 2nd Edition. New York: John Wiley & Sons, Ltd.
- Barot, S., Heino, M., O'Brien, L., and Dieckmann, U. 2004. Long-term trend in the maturation reaction norm of two cod stocks. *Ecological Applications*, 14: 1257–1271.
- Beacham, T.D., Candy, J.R., McIntosh, B., MacConnachie, C., Tabata, A., Kaukinen, K., Deng, L., and Varnavskaya, N. 2005. Estimation of stock composition and individual identification of sockeye salmon on a Pacific Rim basis using microsatellite and major histocompatibility complex variation. *Transactions of the American Fisheries Society*, 134: 1124–1146.
- Beacham, T.D., Candy, J.R., Jonsen, K.L., Supernault, J., Wetklo, M., Deng, L., Miller, K.M., and Varnavskaya, N. (2006) Estimation of stock composition and individual identification of Chinook salmon across the pacific rim by use of microsatellite variation. *Transactions of the American Fisheries Society*, 135: 861–888.
- Beaumont, A. 2003. Conservation Genetics. p 751–792 In Balding D Bishop M, and Cannings C, eds. *Handbook of Statistical Genetics*, 2nd Edition. New York: John Wiley & Sons, Ltd.
- Beddington, J. R. Agnew, D. J., and Clark C. W. 2007. Current Problems in the Management of Marine Fisheries Science, 316: 1713–1716.
- Bekkevold, D., André, C., Dahlgren, T.G., Clausen, L.A.W., Torstensen, E., Mosegaard, H., Carvalho, G.R., Christensen, T.B., Norlinder, E., and Ruzzante, D.E. 2005. Environmental correlates of population differentiation in Atlantic herring. *Evolution*, 59: 2656–2668.
- Berst, A., and Simon, R. 1981. Proceedings of the International Stock Concept Symposium. *Canadian Journal of Fisheries and Aquatic Sciences*, 38.
- Campton, D. E. 1987. Natural hybridization and introgression in fishes. *In Population Genetics and Fishery Management* (Ryman, N. & Utter, F., eds), pp. 161–192. Seattle, WA: Washington Sea Grant Program, University of Washington Press.
- Dannewitz, J., Maes, G.E., Johansson, L., Wickström, H., Volckaert, F.A.M., and Jarvi, T. 2005. Panmixia in the European eel: a matter of time. *Proceedings of the Royal Society of London Biological Sciences* 272:1129–1137.
- Fraser, D.J, Hansen, M.M., Ostergaard, S., Tessier, N., Legault, M., and Bernatchez, L. 2007. Comparative estimation of effective population sizes and temporal gene flow in two contrasting population systems. *Molecular Ecology*, 16: 3866–3889.
- Hardenbol, P., Yu, F., Belmont, J., Mackenzie, J., Bruckner, C., Brundage, T., Boudreau, A., Chow, S., Eberle, J., Erbilgin, A., *et al.* 2005. Highly multiplexed molecular inversion probe genotyping: over 10,000 targeted SNPs genotyped in a single tube assay *Genome Res.*, 15, 269–275.
- Hedgecock, D., Barber, P., and Edmands, S. 2007. Genetic approaches to measuring connectivity. *Oceanography*, 20: 70–79.
- Hoarau, G., Boon, E., Jongma, D.N., Ferber, S., Pálsson, J., Van der Veer, HW., Rijnsdorp, AD., Stam, W.T., Olsen, J.L. 2005. Low effective population size and evidence for inbreeding in

- an overexploited flatfish, plaice (*Pleuronectes platessa* L.) Proceedings of the Royal Society B, 272, 497-503.
- Hyde, J.R., Kimbrell, C.A., Budrick, J.E., Lynn, E.A., and Vetter, R. D. 2008. Cryptic speciation in the vermilion rockfish (*Sebastes miniatus*) and the role of bathymetry in the speciation process. *Molecular Ecology*, 17:4, 1122–1136.
- Jónsdóttir, I.G., Marteinsdóttir, G., and Campana, S.E. 2007. Contribution of different spawning components to the mixed stock fishery for cod in Icelandic waters. *ICES Journal of Marine Science*, 64: 1749–1759.
- Jørgensen, H.B., Hansen M.M., and Loeschcke, V. 2005. Spring-spawning herring (*Clupea harengus* L.) in the southwestern Baltic Sea: do they form genetically distinct spawning waves? *ICES Journal of Marine Science*, 62: 1065–1075.
- King, T.L., Verspoor, E., Spidle, A.P., Gross, R., Phillips, R.B., Koljonen, M.-L., Sancehez, J.A., Morrison, C.L. 2007. Biodiversity and Population Structure. pp.119-171. In: *The Atlantic Salmon: Genetics, Conservation and Management* (edited by Verspoor, E., Stradmeyer, L., Nielsen, J.L.). Blackwell Publishing.
- Koljonen, M.-L., King, T.L., and Nielsen, E.E. 2007. Genetic Identification of Individuals and Populations, pp. 270–298. In: *The Atlantic Salmon: Genetics, Conservation and Management* (edited by Verspoor, E., Stradmeyer, L., Nielsen, J.L.). Blackwell Publishing, 119–171.
- Lukacs P.M., and Burnham, K.P. 2005. Review of capture-recapture methods applicable to noninvasive genetic sampling. *Molecular Ecology*, 14: 3909–3919.
- Maes, G.E., Pujolar, J.M., Hellemsans, B., and Volckaert, F.A.M. 2006. Evidence for isolation by time in the European eel (*Anguilla anguilla* L.) *Molecular Ecology*, 15: 2095–2107.
- Morin, P.A., G. Luikart, R.K., Wayne, and SNP Working Group 2004. SNPs in ecology, evolution and conservation. *Trends in Ecology and Evolution*, 19(4): 208–216.
- Pendas, A. M., Moran. P., Marlinez, J. L., and Garcia-Vazquez, E. 1995. Applications of 5S rDNA in Atlantic salmon, brown trout, and in Atlantic salmon x brown trout hybrid identification. *Molecular Ecology*, 4:275–276.
- Pogson, G. H., and Fevolden, S. 2003. Natural selection and the genetic differentiation of coastal and Arctic populations of the Atlantic cod in northern Norway: a test involving nucleotide sequence variation at the pantophysin (Pan I) locus. *Molecular Ecology*, 12: 63–74.
- Policansky, D., and Magnuson J.J. 1998. Genetics, Metapopulations, and Ecosystem Management of Fisheries. *Ecological Applications*, Vol. 8, No. 1, Supplement: Ecosystem Management for Sustainable Marine Fisheries, pp. S119–S123.
- Pritchard, J.K., Stephens, M., and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics*, 155:945–959.
- Rocha-Olivares, A. 1998. Multiplex haplotype-specific PCR: a new approach for species identification of the early life stages of rockfishes of the species-rich genus *Sebastes* Cuvier. *Journal of Experimental Marine Biology and Ecology*, 231: 279–290.
- Roques, S., Duchesne, P., and Bernatchez, L. 1999. Potential of microsatellites for individual assignment: the North Atlantic redfish (genus *Sebastes*) species complex as a case study. *Molecular Ecology*, 4:1703–1717.
- Ruzzante, D.E., Taggart, C.T., Lang, S., and Cook, D. 2000. Mixed-Stock Analysis of Atlantic Cod near the Gulf of St. Lawrence Based on Microsatellite DNA. *Ecological Applications*, 10: 1090–1109.
- Schwartz, M.K., Luikart, G., and Waples, R.S. 2007. Genetic monitoring: a promising tool for conservation and management. *Trends in Ecology and Evolution*, 22:25–33.

- Smedbol, R.K., and Wroblewski, J.S. 2002. Metapopulation theory and northern cod population structure: interdependency of subpopulations in recovery of a groundfish population. *Fisheries Research*, 55: 161–174.
- Vähä, J-P., Erkinaro, J., Niemelä, E., and Primmer C.R. 2008. Temporally stable genetic structure and low migration in an Atlantic salmon population complex: implications for conservation and management. *Evolutionary Applications*, 1: 137–154.
- Walters, C.J., and Martell, S.J.D. 2004. *Fisheries Ecology and Management*. Princeton University Press: Princeton. 399 pp.
- Waples, R.S. 2005. Genetic estimates of contemporary effective population size: to what time periods do the estimates apply? *Molecular Ecology*, 14: 3335–3352.
- Waples, R.S. 2006. A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics*, 7:167–184.
- Weir, B.S. 1996. *Genetic Data Analysis II*. Sunderland, MA: Sinauer Associates, Inc.

6.7 Recommendations

- 1) Base the monitoring of fish stocks on genetic as well as demographic metrics that are informative with regard to the status of their contributing breeding populations as this will help to provide insights needed for sustainable management
- 2) Build working models of local, regional and global breeding population structure in commercial fish species of interest to ICES, using existing information on their breeding biology and genetics, to help realise management based on breeding populations, and focus research directed at model improvement.
- 3) Incorporate genetic studies into existing fisheries research programmes to advance understanding of population structuring, and to help assess the benefits of using genetic metrics to assess population status.
- 4) Develop new and improved methods for estimating genetic metrics (e.g. N_e), particularly for large breeding populations and breeding populations with high connectivity, so as to make possible cost-effective monitoring of breeding populations
- 5) Establish strong direct links between individual species assessment working groups and the WGAGFM to facilitate 1-4 above.

7 Group Business

7.1 Draft Resolutions for 2009

The WG considered ToRs for the 2009 meeting and prepared a draft set for consideration by ICES (Annex 3). Five items were proposed.

7.2 Meeting Places in 2009 and 2010

In 2004 it was agreed that future meetings should plan for 3.5 days of business in order to allow for a full discussion of the ToRs. Over the past few years the ToRs are increasing in complexity and require more time for discussion in order to reach consensus on the report.

The WG finds it useful for planning purposes to determine meeting venues two years in advance. The 2009 meeting is provisionally planned for Poland at the invitation of R. Wenne. For 2010, Tom Cross from The University College Cork kindly offered to host the meeting.

Annex 1: List of participants

Name	Address	E-mail
Pierre Boudry	IFREMER La Tremblade Station P.O. Box 133 F-17390 La Tremblade France	pierre.boudry@ifremer.fr
Gary R. Carvalho	University of Wales, Bangor School of Biological Sciences LL57 2DG Bangor, Gwynedd United Kingdom	g.r.carvalho@bangor.ac.uk
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	t.cross@ucc.ie
Geir Dahle	Institute of Marine Research P.O. Box 1870 N-5817 Bergen Norway	Geir.Dahle@imr.no
Ann-Britt Florin	Swedish Board of Fisheries Institute of Coastal Research P.O. Box 109 SE-74071 Öregrund Sweden	ann-britt.florin@fiskeriverket.se
John Gilbey	Fisheries Research Services FRS Freshwater Fisheries Laboratory Faskally PH16 5LB Pitlochry United Kingdom	gilbeyj@marlab.ac.uk
Elizabeth Gosling	Galway-Mayo Institute of Technology Dublin Road Galway Ireland	elizabeth.gosling@gmit.ie
Reinhold Hanel	Leibniz-Institut für Meereswissenschaften Düsternbrooker Weg 20 D-24105 Kiel Germany	RHanel@ifm-geomar.de
Torild Johansen	Institute of Marine Research Institute of Marine Research Tromsø P.O. Box 6404 N-9294 Tromsø Norway	torild.johansen@imr.no

Name	Address	E-mail
Ellen L. Kenchington	Fisheries and Oceans Canada Bedford Institute of Oceanography P.O. Box 1006 Dartmouth NS B2Y 4A2 Canada	kenchington@mar.dfo-mpo.gc.ca
Anastasia Khrustaleva	Russian Federal Research Institute of Fisheries & Oceanography 17 Verkhne Krasnoselskaya RU-107140 Moscow Russian Federation	Khrustaleva@molgen.vniro.ru
Phillip McGinnity	The Marine Institute Rinville Co. Galway Oranmore Ireland	phil.mcginnity@marine.ie
Einar E. Nielsen (Chair)	The National Institute of Aquatic Resources Department of Inland Fisheries Vejlsovej 39 DK-8600 Silkeborg Denmark	een@dfu.min.dk
Martha O'Sullivan	Fisheries Research Services FRS Marine Laboratory P.O. Box 101 AB11 9DB Aberdeen United Kingdom	osullivanm@marlab.ac.uk
Paulo Prodohl	Queen s University Belfast School of Biology and Biochemistry 97 Lisburn Road BT9 7BL Belfast United Kingdom	p.prodohl@qub.ac.uk
Jochen Trautner	Johann Heinrich von Thünen- Institute, Institute for Fishery Ecology Palmaille 9 D-22767 Hamburg Germany	jochen.trautner@vti.bund.de
Niklas Tysklind	University of Wales, Bangor School of Biological Sciences LL57 2DG Bangor Gwynedd United Kingdom	osp231@bangor.ac.uk
Anti Vasemägi	University of Turku Department of Biology Vesilinnantie 5 FI-20014 Turku Finland	antvas@utu.fi
Olegs Vasins	Latvian Fish Resources Agency 8 Daugavgrivas Str. LV-1048 Riga Latvia	olegs.vasins@lzra.gov.lv

Name	Address	E-mail
Eric Verspoor	Fisheries Research Services FRS Marine Laboratory P.O. Box 101 AB11 9DB Aberdeen United Kingdom	verspoor@marlab.ac.uk
Anna Was	Galway-Mayo Institute of Technology Molecular Ecology Research Group Galway Ireland	anna.was@gmit.ie
Roman Wenne	Institute of Oceanology PO Box 68 PL-81-712 Sopot Poland	rwenne@cbmpan.gdynia.pl

Annex 2: Proposed Draft Resolutions for 2009

The **Working Group on the Application of Genetics in Fisheries and Mariculture [WGAGFM]** (Chair: E. Eg Nielsen, Denmark) will meet in Sopot, Poland from 1-3 April 2009 to:

- a) Report on progress with the establishment of a meta-database for genetic data on fish and shellfish genetics covered under the ICES remit;
- b) Review current status of the application of traceability methods in the fisheries sector based on genetics;
- c) Update and insights from the EU project SalseaMerge on establishment of a large scale genetic database for assigning individuals to population of origin;
- d) Assess the possibility for the development of an integrated global management model for Atlantic cod based on genetic information;
- e) To evaluate prospects for application of genetics/genomics to study and reduce the impact of fish and shellfish diseases in natural and cultured populations.

WGAGFM will report by 30 April 2009 to the attention of the Mariculture Committee.

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	<p>Term of Reference a)</p> <p>The WGAGFM recommended in 2007 to establish a meta-database cataloguing existing data in the field of fish and shellfish genetics in order to counteract the increasing discrepancy between data generation and data storage, management and accessibility in the field of fish and shellfish genetics (see WGAGFM report 2007). It was argued that the existence of such a meta-database would significantly help to promote coherence across research programmes, enhance research progress, and facilitate the translation of results from fundamental research to support other stakeholders such as fisheries managers.</p> <p>After consultation of the ICES Data Centre (DC), and further discussion between the WGAGFM members, nine recommendations were forwarded aiming at streamlining and facilitating the development and implementation of the fish genetic meta-database. These recommendations will serve as a basis to monitor and scrutinize the progress of the database development. Results from our discussions with the ICES DC as well as our experience with the database prototype will be reported to the WGAGFM panel and serve as a basis for reviewing the database utility and to establish prospects.</p> <p>(Lead: J. Martinsohn, E. Verspoor)</p>
--------------------------	--

Term of Reference b)

The latest FAO report on the state of fisheries and aquaculture 2007 estimates that globally, 60% of marine fish stocks are at their maximum sustainable limits or overexploited. This disquieting level is even more alarming due to the growing serious global problem of Illegal, Unreported and Unregulated (IUU) fishing, a major impediment to the achievement of sustainable world fisheries. Worth between \$4bn and \$9bn per year, IUU fishing leads to major revenue losses, and effective measures to fight IUU fishing are still lacking. To reduce the market for illegal catches is of utmost importance, and the development of an effective traceability system to verify the species and origins of fish and

shellfish caught would greatly support such an approach, both for regulatory enforcement but also for consumer protection. In addition to species verification, a traceability system based on the identification of regional stocks is necessary to preclude fraudulent allocations. There is an urgent need to identify traceability markers that can be used throughout the food supply chain, from on-board samples, to processed product ("fish to fork"), and which exhibit minimal variance. Furthermore, it is essential that tools based on such markers are validated to internationally recognized forensic standards. Only under such stringent conditions can traceability data be used for legal enforcement and as evidence in a court of law. We will review briefly the range of genetic markers available for traceability of species and population identity, and then consider critically the potential contribution of molecular genetic approaches to traceability systems, with particular emphasis on the development of technologies that facilitate forensic validation.

(Lead: G. Carvalho, J. Martinsohn)

Term of Reference c)

Over the past two decades, an increasing proportion of North Atlantic salmon are dying at sea during their oceanic feeding migration. The specific reasons for the decline in this important species are as yet unknown. However, climate change is likely to be an important factor. In some rivers in the southern part of the salmon range, wild salmon now face extinction. This is in spite of unprecedented management measures to halt this decline. Arguably the greatest challenge in salmon conservation is to gain insight into the spatial and ecological use of the marine environment by different regional and river stocks, which are known to show variation in marine growth, condition, and survival. Salmon populations may migrate to different marine zones, whose environmental conditions may vary. To date it has been impossible to sample and identify the origin of sufficient numbers of wild salmon at sea to enable this vital question to be addressed. SALSEA-Merge will provide the basis for advancing our understanding of oceanic-scale, ecological and ecosystem processes. Such knowledge is fundamental to the future sustainable management of this key marine species. Through a partnership of 9 European nations the programme will deliver innovation in the areas of: genetic stock identification techniques, new genetic marker development, fine scale estimates of growth on a weekly and monthly basis, the use of novel high seas pelagic trawling technology and individual stock linked estimates of food and feeding patterns. In addition, the use of the three-dimensional Regional Ocean Modelling System, merging hydrography, oceanographic, genetic and ecological data, will deliver novel stock specific migration and distribution models. This widely supported project provides the basis for a comprehensive investigation into the problems facing salmon at sea. It will also act as an important model for understanding the factors affecting survival of many other important marine species. This ToR will report on progress made in the first year of this important ocean ecology project, particularly with regard to the application of genetics.

(Lead P. McGinnity, E. Verspoor).

Term of Reference d)

A lot of information has been gathered on the genetic population structure of commercially important marine fish. However, to date this knowledge has been poorly utilized in management. Even in cases where genetic data has shown that previously defined management areas conflict with the biology of the species, little or no action has been taken to change management practices. This is in contrast to the intention of e.g. the EU Common Fisheries Policy, which has identified the population the natural unit of evolutionary change, and as such, provides the genetic resources required for adaptive response to natural and man-made changes in the environment. In this ToR we aim at incorporating genetic information, in joint framework with previous ecological insights, to attempt the development a global management model for Atlantic cod. We will particularly focus on the identification

of geographical areas where management operates in conflict with knowledge on population structure and come up with suggestions for a new design of management areas. Also, we will identify areas with limited or lacking genetic information where more studies are warranted.
(Lead: E. Nielsen, E. Kenchington, T. Johansen).

Term of Reference e)

Pathogens and parasites negatively affect natural and cultured populations of fish and shellfish by induction of diseases. Wild populations can be reduced in numbers, which decreases fishery resources. Cultured populations are often more susceptible to diseases because of high rearing densities compared to wild populations and therefore undergo widespread mortality when infected. Economical implications of diseases are difficult to overestimate. Breeding of resistant stocks is often proposed to overcome such loss in aquaculture. Recently developed genomic methods can significantly increase our understanding of pathogen – host relationship, stress and immune response of the host. This knowledge will facilitate diagnostic, treatment and vaccination against diseases. Genomic approaches usually involve construction of cDNA libraries from various tissues of infected and non-infected organisms or by suppressive subtractive hybridisation (SSH) followed by sequencing of expressed sequence tags (ESTs). Microarray technology enables to identify up- and down- regulated fish and shellfish genes related to susceptibility or resistance to diseases. These candidate genes can be further characterised. A variety of molecular markers is available for identification of QTLs for disease resistance, which can ultimately be used in marker assisted selection (MAS) programmes.
(Lead: R. Wenne, P. Boudry)

Resource Requirements:	None required other than those provided by the host institute.
Participants:	The Group is normally attended by some 15–25 members and guests
Secretariat Facilities:	None required
Financial:	None required
Linkages To Advisory Committees:	ACOM.
Linkages To other Committees or Groups:	SIMWG , WGECO, WGMAFC, WGMASC
Linkages to other Organisations:	Linkage with the EC Joint Research Centre at Ispra, Italy.

Annex 3: Recommendations

RECOMMENDATION	FOR FOLLOW UP BY:
ToR a)	
1) Prior to inter-laboratory use of SNPs the performance of different platforms available should be evaluated to minimize potential bias in presently unknown error-rates.	Fisheries geneticists
2) SNP discovery should be based on a panel of individuals that optimises the “ascertainment width” (geographic coverage) and “ascertainment depth” (number of chromosomes and individuals per location) to avoid ascertainment bias.	Fisheries geneticists
3) Use SNPs in addition to other markers, not as a replacement.	Fisheries geneticists
4) SNPs should be the preferred markers for establishing long-term genetic datasets and subsequent genetic monitoring.	Fisheries geneticists
5) As with other markers the extent to which the employed SNPs are influenced by selection (and other evolutionary forces) should be determined.	Fisheries geneticists
6) Applying a combination of selected and neutral SNPs is possible and is likely to yield novel information on population structure.	Fisheries geneticists
7) The development and application of new analytical tools should be a research priority for examining large datasets such as generated by SNP studies.	Fisheries geneticists

ToR b)	
1) 4. QTL studies should be supported in both in wild and farmed aquatic species as they are one of the most direct ways to understand the genetic basis of phenotypic variation, linking classic quantitative genetic and genomic studies.	ICES
2) QTL studies should not be restricted to MAS. The development of QTL studies should be supported as they can also contribute to a better understanding of the genetic architecture of adaptive traits of interest to fisheries and their management.	Aquaculture geneticists
3) To aid identification of QTL in a wider variety of aquatic species, the current development of genomic resources - notably linkage and physical maps, EST and BAC libraries and whole genome sequences - should be encouraged.	Aquaculture geneticists
4) The development of statistical methods and software adapted to aquatic species should be supported to facilitate the development of linkage maps and to identify QTLs.	Aquaculture geneticists
5) The development and maintenance of divergent lines, segregating progenies, or other biological material of interest for QTL mapping should be encouraged.	Aquaculture geneticists
ToR c)	
1) 6. The development and implementation of a web-based fish population genetic meta-database, under the responsibility of WGAGFM, within the remit of ICES and in collaboration with the European Commission, as proposed in the WGAGFM report 2007, is progressed;	WGAGFM, ICES data centre
2) The meta-database serve as a portal cataloguing relevant primary information with respect to generated population genetic data, primary and secondary research reports, available biological samples, and point to the locations and contact points from which to get such data, samples and extended information;	WGAGFM, ICES data centre
3) The WGAGFM agree before September 2008 which types of meta data to include and the database structure, and the functionalities of the interface providing access to the data as well as about other features to be made available;	WGAGFM
4) The WGAGFM agree access regulation and control by database end-users as well as strategies to ensure data quality and validity before September 2008;	WGAGFM
5) Decisions regarding 3) and 4) the WGAGFM report to the database developer and host before September 2008;	WGAGFM
6) When 5) is delivered, database developer and	WGAGFM, ICES data centre

host should establish contact and establish a practical, workable strategy for development and implementation of the database;	WGAGFM
7) The Atlantic salmon meta-data set developed as part of the NASCO Salmon at Sea initiative under the EU SALSEA MERGE project be used as a development template for a prototype fish population genetics database;	WGAGFM
8) Work on 3) to 7) should be progressed, facilitated and monitored by the WGAGFM members leading this ToR;	WGAGFM
9) Progress should be reported to the 2009 WGAGFM by the ToR leaders.	
ToR d)	
1) ICES should request information on all scale and otolith collections held by fisheries laboratories, institutes and universities. This information should be accessible via the proposed ICES genetic database.	ICES
2) If disposal of archived material is envisaged, the scientific community should be notified via the ICES network, so that alternative storage can be arranged if there is sufficient interest.	ICES
3) Archived material should be inspected on a regular basis and subsequently stored in controlled conditions. It should be borne in mind that most fungi prefer dark humid environments. Thus dry storage conditions are vital. Samples contaminated with fungi may be heated to 65°C for twenty minutes to halt current fungal growth.	ICES
4) For new scale or otolith collections, every precaution should be taken to prevent cross-contamination. There is a much greater risk of contamination due to the sensitivity of molecular techniques.	ICES
5) Preferably a non-mineralised tissue should be sampled in the future, (e.g. a fin clip), solely for genetic analysis, due to the clear benefits of molecular approaches in fisheries biology. This is also important because of the competing uses for scale and otolith material.	ICES
6) If the only samples being taken in future surveys are scales or otoliths collected in the traditional way, then the material should be placed in a thin layer within the envelope to aid rapid drying. Both scales and otoliths should be stored as recommended for archived material.	
ToR e)	
1) Base the monitoring of fish stocks on genetic as well as demographic metrics which are informative with regard to the status of their	ICES

contributing breeding populations as this will help to provide insights needed for sustainable management.

- | | |
|---|-------------------------------|
| 2) Build working models of local, regional and global breeding population structure in commercial fish species of interest to ICES, using existing information on their breeding biology and genetics, to help realise management based on breeding populations and focus research directed at model improvement. | ICES |
| 3) Incorporate genetic studies into existing fisheries research programmes to advance understanding of population structuring and to help assess the benefits of using genetic metrics to assess population status. | ICES
Fisheries geneticists |
| 4) Develop new and improved methods for estimating genetic metrics (e.g. N_e), particularly for large breeding populations, and breeding populations with high connectivity, so as to make possible cost-effective monitoring of breeding populations. | ICES |
| 5) Establish strong direct links between individual species assessment working groups and the WGAGFM to facilitate 1-4 above. | |
-