ICES WGAGFM REPORT 2012

SCICOM Steering Group on Human Interactions on Ecosystems

ICES CM 2012/SSGHIE:12

Ref. SCICOM

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

2–4 May 2012

Derio, Spain

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. 2012. Report of the Working Group on the Application of Genetics in Fisheries and Mariculture (WGAGFM), 2-4 May 2012, Derio, Spain. ICES CM 2012/SSGHIE:12. 61 pp. https://doi.org/10.17895/ices.pub.8962

Contents

Executive summary

The Working Group on the Application of Genetics in Fisheries and Mariculture (WGAGFM) met in Derio in the Basque Country, Spain, 2–4 May 2012 and was chaired by Dorte Bekkevold. The meeting had 24 participants; 18 members and 6 PhD students, representing 11 countries altogether.

Discrimination of populations and stock units of fishes is one of the main requirements of fisheries management, and may be achieved with various tools, including genetic characterization. The ability to address complex and diverse questions in fisheries biology and management using molecular markers has improved greatly in recent years. Members met to discuss and review different molecular marker-based approaches to identifying evolutionary units and the origin of individuals and samples, including applications of relevance in an aquaculture context, followed by discussion about how methods can be implemented and integrated into fisheries and ecosystem management, such as via uptake into the data collection framework.

Following technological and bioinformatic developments to characterise genes and genomes and their association with specific biological traits, the repository of gene-associated markers for identifying individuals, populations and species is expanding rapidly. This has meant a shift in tool-of-choice and a dramatic increase in studies scanning moderate to large panels of genetic markers, demonstrating the utility of these markers in conservation and fisheries management, as well as in aquaculture breeding. Gene-associated markers may not only provide short-term "population tags" for describing and monitoring changes in stock distribution and for population assignment, but importantly they facilitate new insights into the nature and scale of adaptive diversity in wild fish populations – a key component of biodiversity that determines in part population resilience and ability to cope with environmental change. A presentation by one of the members highlighted an example of a recent application of such tools with broad scope for implementation in management, control and enforcement. The FishPopTrace project thus developed and applied panels of gene-associated markers for hake, sole, herring and cod, which can be implemented as a standard tool to identify source populations (assignment) and to verify the claimed population of origin from landed fish through to processed fish products, in the 'ocean to fork' sense.

In some cases genetic analysis using standard methods may not be powerful enough to adequately resolve stock relationships on demographic scales. Parasite/pathogens have been used as biological tags for stock assignment, but success was sometimes limited. Recent progress in genetic characterization of parasites/pathogens has now revealed unprecedented levels of resolution, thus providing a 'magnifying glass' to study their host's evolutionary history. Despite their substantial contribution to ecosystems, parasites have generally been understudied (and issues of their conservation generally ignored), and so has their applicability for identification of host stock structure. Although the scope for using population genetic data for parasites as proxies for determining host structure will strongly depend strongly on factors, which need be examined and evaluated carefully on a case-by-case basis, the evolution of the parasite–host biological system is likely to yield valuable information about host stock structure and dynamics.

Members discussed means to improve integration of genetic information into fisheries management under a reformed Common Fisheries Policy, the Marine Strategy Framework Directive and the new Data Collection Framework 2014-2020. Despite the

huge potential for implementation, European examples are still few and mainly target local and short-term management applications. For genetic data implementation into the DCF efficiently it is of importance to at a political level initiate an informative mutual dialogue with relevant stake-holders such as DGMARE, ICES Stock Assessment Working Groups and experts in fisheries genetics. It is therefore recommended that the ICES secretariat initiate and organise a workshop defining guidelines for the integration of genetic data and information in support of the implementation of EU policies such as the MSFD and the reformed CFP. Participants should include genetic experts as well as experts involved in the coordination and implementations of the MSFD and CFP.

1 Opening of the meeting

The Working Group on the Application of Genetics in Fisheries and Mariculture (WGAGFM) met in Derio, in the Basque Country 2–4 May 2012. The Terms of Reference (ToR) were decided by ICES Science Committee in Gdansk, Poland, in 2011. Dr Dorte Bekkevold (Denmark) chaired the meeting, which opened at 09:00 on Wednesday, 2 May and closed at 13.30, Friday, 4 May 2012.

1.1 Attendance

The meeting had 24 participants; 18 members and 6 PhD students, representing 11 countries (Belgium, Canada, Denmark, Germany, Iceland, Italy, Norway, Republic of Ireland, Russian Federation, Spain and United Kingdom; Annex 2).

1.2 Venue

The meeting was held at the AZTI-Tecnalia, Parque Tecnológico de Bizkaia, Astondo Bidea - Edificio 609, 48160 Derio (Bizkaia), and was jointly hosted by the University of the Basque Country (UPV/EHU) and the AZTI Foundation. The WG wish to express their appreciation to hosts Dr. Andone Estonba, Dr. Aitor Albaina Vivanco (both UPV/EHU) and Dr. Naiara Rodriguez-Ezpeleta (AZTI-Technalia), as well as the rest of the staff at the AZTI foundation for their hospitality and kind and efficient assistance throughout.

1.3 Meeting format

WGAGFM has an established framework for completing its ToR. Prior to the meeting, small *ad hoc* working groups, under the leadership of one or two persons, are established to prepare position papers related to specific issues in the Terms of Reference. The leader(s) of each 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(s) is responsible for preparing the final report text from their sessions. Prior to the meeting an agenda is circulated to all members.

2 Adoption of the agenda

2.1 ToR a) Review the potential for using parasites, microbes and viruses as "magnifying glass" for fish stock characterisation

Filip A.M. Volckaert, Anna K. Daníelsdóttir, Tine Huyse, Stanley D. King, Jackie Lighten, Martin Llewellyn and Estibaliz López de Abechuco

2.1.1 Rationale

Parasitism is arguably the dominant life-style in ecosystems: each host species is associated with at least one parasite (Marcogliese, 2005). Although commonly thought of as negligible, parasite biomass within an ecosystem can be substantial and comparable to, if not higher than, that of the top predators (Kuris *et al*., 2008). Despite this substantial contribution to ecosystems, parasites have been understudied and underestimated (Marcogliese, 2004).

Previous work utilizing parasites and microbes in an ecological context has shown that we are able to study cryptic host processes indirectly though their parasitic fauna, because of the intimate relationship that has evolved between parasite and host (Nieberding and Olivieri, 2007). Because of this close relationship, the substantial role in ecosystem function and the potential to limit population size and range (Ricklefs, 2011), parasite diversity and abundance (alpha diversity) may be used as a marker for overall ecosystem health (persistence, productivity, organization and resilience) (Marcogliese, 2005). Through a comparative method this approach can be adapted to utilize parasites as an indirect marker for individual contemporary population health assessments (Hudson *et al*., 2006).

We are also better able to characterize history, evolution, and demographics of these populations (stocks) by comparing their genomic diversity (Avise, 2000). Migration, adaptation and random processes leave an imprint in the genome of populations and clades, dating back as far as the time of speciation of a taxon. Thus if we investigate these processes within a parasitic species in close association to a particular host, the faster mutation rates of the parasite may provide a more resolved genealogy that reflects cryptic population processes of the host species (Wirth *et al*., 2005; Nieberding and Olivieri, 2007). It may therefore allow characterization of host stocks at a higher level of resolution. For example the "young" age of most northern marine habitats provides inadequate time (~<12Ka since last glacial maximum) for population differentiation in some species, especially those with recruitment times of a few years or more. This is especially true when considering that Hauser and Carvalho (2008) estimate that in large populations, several thousand generations are required to achieve detectable differentiation (e.g., $F_{st} \sim 0.002$) even under complete isolation (migration = 0). The use of parasites in a comparative population genetics approach may be the answer to detecting the limited structure that has accrued in host species over this time. Unfortunately, as discussed below, research focusing on the comparative population genetics of parasites and their hosts, especially marine species, is sparse (Criscione *et al*., 2005; Prugnolle *et al*., 2005).

2.1.2 Choosing a parasite appropriate to the scale of host evolution

The taxomomic and life history diversity present among parasitic organisms provides access to multiple lines of investigation around their hosts' ecology and evolutionary history. Indeed, in some cases parasite genetics can be even more informative than that of the host (Criscione *et al*., 2006). However, not all parasites will be equally informative, and parasite choice must be carefully tailored to the research question at hand. As such, in order to achieve the desired "magnifying glass" effect, a thorough understanding of the life history of the parasite and taxonomic scale of study, in conjunction with a focused research question, is essential to elucidate cryptic patterns of host evolutionary history and demographic processes. Crucially, parasite proxies for host evolutionary dynamics are potentially informative at multiple taxonomic scales including the species level (i.e. phylogenetically), the population level (i.e. phylogeographically) and the family level (i.e. demographically).

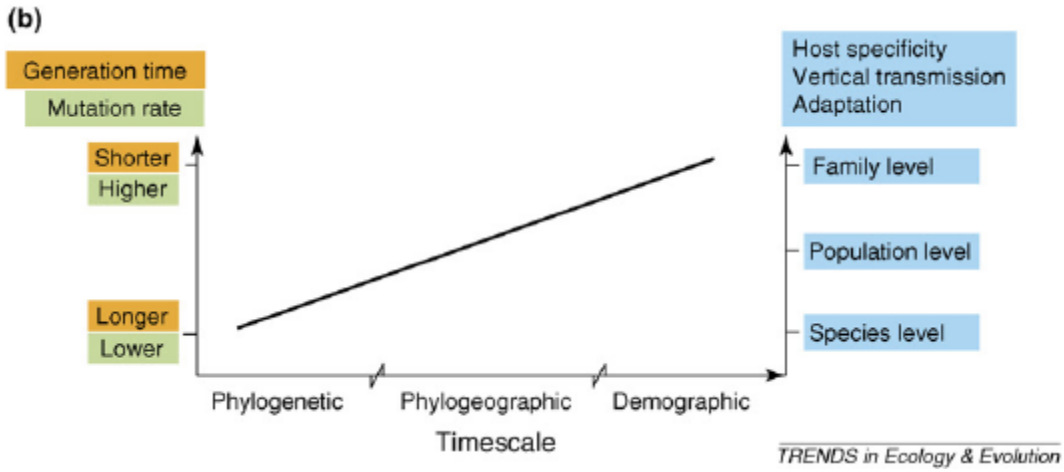

Figure 1: Relationships between 'proxy' parasite traits and the different timescales. (b) Variation in the traits of 'proxy' parasites according to the considered time scale. 'Proxy' parasites that are useful at the phylogenetic scale display longer generation times and lower mutation rates than do those useful at the demographic scale. Moreover, host-specificity, vertical transmission and adaptation to the host might be limited to the species level for parasite taxa useful at the phylogenetic scale, whereas those useful at the demographic scale should be specific, vertically transmitted and adapted to the host family level (Nieberding and Olivieri, 2007).

A principal concern among those interested in improving the resolution of host intra-species population scale analyses (like stock assessment) is the mutational turnover differential between parasite and host. Nieberding *et al*. (2004) suggested that at the population scale (and the family scale) parasites have the most "magnifying" potential, and prospective parasites should exhibit a short generation time and thus a higher rate of mutation that their host.

Fortunately, most parasitic species undergo multiple generations during the typical lifespan of their host. In addition, effective population size is likely to be critical. In general, parasite species typified by relatively small effective population sizes may be preferable. In combination with potentially small effective population sizes, and population sub-structuring (infra-population), elevated rates of genetic drift may promote rapid divergence between multiple hierarchical levels of parasite populations (inter-, and intra-population) (Nieberding and Olivieri, 2007).

Parasite host specificity will strongly influence a researcher's choice of proxy. Where long-term persistence on hosts and high host specificity is a feature of parasite life history, these patterns of divergence will yield more accurate information regarding host ecology and evolutionary history than the converse (Whiteman and Parker, 2005; Rannala and Michalakis, 2002). However, selecting parasite proxies based on mode of transmission and/or life cycle (direct or complex) should be judged on a case-by-case basis, as each can garner informative data in different circumstances. For example, conflated signals of host processes may arise if a parasite goes through numerous mobile hosts during an indirect life cycle. Nonetheless, previous studies have demonstrated that with careful consideration, a parasitic species with an indirect life cycle may accurately reveal cryptic host population patterns (*e.g.* population assignment - Criscione *et al*., 2006). For non-parasitic proxies (i.e. microbes) non-epidemic dispersal is preferable in order to preserve small-scale structure (Falush *et al*., 2003). In addition to the close association between host and parasite, temporal congruence between both genealogies reduces the potential for misinterpreting conflated genetic signals

within the parasite that do not reflect processes of the current host (*e.g.* ancestral polymorphism) (Whiteman and Parker, 2005).

Life-history differences between parasitic organisms of population genetic importance exist beyond host specificity. Many parasite species possess alternative reproductive strategies that do not follow classical patterns of Mendelian inheritance (*e.g.* asexual reproduction through selfing and cloning, as well as polyploidy). Careful consideration must be taken to assess whether the reproductive strategy may disrupt congruence between genealogies of the host and parasite.

Finally, in addition to elucidating demographic processes within fish populations, ectoparasites (*e.g.* copepods) that share environmental stressors with their host may potentially be used to explore adaptive traits at the genetic level – salinity, temperature, etc. This could be of interest in fish stock assessment, although it remains unexplored so far.

2.1.3 Parasitic organisms as proxies for host population genetic diversity from across the tree of life

Convincing cases on the application of parasites and microbes to improve the resolution of population structure are available from diverse areas of the tree of life (Nieber ding and Olivieri, 2007).

1) An emblematic case is the commensal human bacterium *Helicobacter pylor*i whose genome traced global evolution (Falush *et al*., 2003) and regional cultural evolution of the host. While the study confirmed patterns of migration seen in human it became really useful in the latter study of two human populations who had coexisted for more than 1000 years but remained isolated for cultural reasons (Wirth *et al*., 2004). Genetic analysis revealed three ancestral groups within the Muslim community going back to the ancestral Europe1 group, and the Buddhist community corresponded with a mosaic of ancestral East Asian and ancestral Europe1. These data fit with the recent history of the region. Viral markers such as the human immunodeficiency virus (HIV) and human T-cell lymphotrophic virus (HTLV-I) have also been used (Wirth *et al*., 2005). Host-specificity over a time period of relevance has been the key to detect the pattern.

2) The global population of lion (*Panthera leo*) was studied with maternally and biparentally inherited genetic markers and the lion feline immunodeficiency virus (FIV_{Ple}). Across the range these genetic and viral markers indicated the presence of discrete units despite the high potential for dispersal. For example, genetic patterns of the FIV_{Ple} revealed three distinct populations within the Greater Serengeti Ecosystem, which were undetectable through examination of host genetics alone (Antunes *et al*., 2008). The success of the study is related to the non-epidemic and population-specific infection of the virus occurring within a time frame (Pleistocene) relevant for discrimination.

3) Whale lice (*Cyamus spp*.) are closely associated with their hosts, the right whales (*Eubalaena spp.*), and require close contact between host individuals to facilitate transmission. They migrate between hosts when these are in close contact and represent large populations (with higher levels of genetic variation). Genetic characterisation of populations in three whale lice species infecting individual right whale species showed a close concordance

with patterns seen in the host genetics, as well as inferring cryptic patterns of migration in the host (Kaliszewska *et al*., 2005).

4) Aphid symbionts, namely the bacterium *Buchnera aphidicola* co-speciated with their host and because of their faster evolution they allowed for a better understanding of the host evolution (Jousselin *et al*., 2009).

5) Marine plankton populations are notoriously difficult to differentiate given their highly dynamic nature of species oscillations and chaotic behavior (Huisman and Weissing, 1999). Viruses seem to be good indicators of populations-specific differences. For example, two environments harbour different populations of the prasinophycee *Ostreococcus tauri* judging from the community of DNA viruses (Bellec *et al*., 2010).

2.1.4 Examples on the use of parasite population genetics in fish population characterisation

One of the potential applications of parasite genotypes is the assignment of the host to the source population, and may be particularly important in the management of migratory fish stocks. For example, in a study of the steelhead trout, *Oncorhynchus mykiss*, the genetics of an infecting trematode, *Plagioporus shawi,* allowed discrimination of host populations to rivers separated by as little as approximately 50 km, where patterns of host microsatellite diversity fail to resolve such fine-scale population assignment (Criscione *et al*., 2006). The use of parasites as genetic markers has also shown a broad range of applications in the identification of population structure. Anisakis nematodes have proven to be useful as a biological tag to separate and characterise populations of marine species from cephalopods to whales (Kuhn *et al*., 2011), including fish species such as herring *Clupea harengus* (Cross *et al*., 2007), horse mackerel *Trachurus trachurus* (Mattiucci *et al*., 2008; Mattiucci and Nascetti, 2008) and Pacific sardine *Sardinops sagax* (Baldwin *et al*., 2011). They are highly specific to their very mobile final host (marine mammals), and use a range of intermediate and paratenic hosts (from crustaceans and chaetognaths to fish). The use of Anisakis and other parasites in this context could have important implications in fisheries management.

Parasite dynamics of a single host have traditionally been analysed at the parasite community level; they have proven to be effective to determine changes over time (MacKenzie and Abaunza, 1998). However, specific host-parasite relationships may also provide information about the life-history of the host or its migration pattern (Criscione *et al*., 2006). Parasite invasion pathways affect the distribution and extent of genetic variation within and among its introduced populations. The evaluation of these genetic invasion pathways of parasites could help to know how parasites invade and become established in new geographic regions (Miura *et al*., 2006). This could help in understanding population dynamics, particularly in a context of global change (Rohr *et al*., 2011). Unfortunately, there is little knowledge about the complex life cycles of parasites, especially for species that sequentially parasitise different host species or when the parasite-host assemblage fluctuates due to the migratory behaviour of the host species.

Perhaps one of the most compelling examples of the importance of this application comes from Atlantic salmon, *Salmo salar*. Wild populations of Norwegian Atlantic salmon were decimated by the introduction of a foreign strain of the ectoparasite *Gyrodactylus salaris* (Meinila *et al*., 2004). To its native host populations (Baltic Atlantic salmon), this strain of *G. salaris* is relatively harmless, but when it was introduced to Norwegian Atlantic salmon as the result of the import of contaminated aquaculture broodstock, it effectively destroyed populations in over 40 rivers (Johnsen and Jen-

sen, 1991; Meinila *et al*., 2002) causing over $670 million US in damages and rising by ~ $57 million per year (Hansen *et al*., 2005; Bakke *et al*., 2007). Upon closer genetic inspection, it was revealed that this species was actually a complex of at least six strains, a result of reduced gene flow due to isolation (Meinila *et al*., 2004). This disaster has shown the detrimental effects that introduced parasites can have on their host.

2.1.5 Examples on the use of parasite, microbe and virus population genetics in open ocean aquaculture

Interest has been growing to genetically profile populations of viruses (Castro-Nallar *et al*., 2011), bacteria and parasites (*e.g*., salmon lice - Yasuike *et al*., 2012; Glover, pers. comm.) to mitigate any detrimental effects of introducing non-native strains of infectious agents to natural environments through aquaculture. Pathogen dynamics have been affected considerably by aquaculture; it includes artificial human mediated transfer between fish populations, often between different geographical regions. In the case of aquaculture in open systems introduced pathogens can potentially interact with the natural environment. For example, parasite genetics revealed that the major outbreak of the infectious salmon anemia virus (ISAV) in Chile originated from the introduction of fish originating from Norwegian stocks, causing massive mortalities of local populations of cultured Atlantic salmon (Castro-Nallar *et al*., 2011). Hayward *et al*. (2001) showed that the ITS rRNA region of *Gyrodactylus anguillae* (Platyhelminthes, direct life cycle) which infects eels (*Anguilla spp*.), is identical across three different continents suggesting a recent introduction through international eel trade.

In one case it has been possible to identify cryptic demographic structure with an invasive parasite introduced through aquaculture. Infection of European eel (*Anguilla anguilla)* with an exotic nematode parasite, *Anguillicola crassus,* revealed a parasite population barrier in the western English Channel, indicating spatial separation between eels captured east and west of here (Wielgoss *et al*., 2008, 2010).

The interactions of the parasite and microbial community between aquaculture facilities and natural environments are undeniably considerable, with fish farms having a measurable impact on the natural parasite community (Krkošek *et al*., 2011). Transmission from natural populations to aquaculture facilities is particularly prevalent, where stressed and immuno-compromised fish become susceptible to infection and may become foci of infections and facilitate explosion in numbers of parasites. Susceptibility to infection in such cases may be tightly correlated with genetic composition of host and parasite, and the extent of shared evolutionary history between strains of parasite and host species.

2.1.6 Use of symbiotic gut and ectoderm microbes as population tags

The analysis of genetic variation of bacterial communities as host population tags is a new and promising research field. Human population structure has already been successfully identified with stomach bacteria (Wirth *et al*., 2004) and the gut microbiome (Qin *et al*., 2010), although protocol standardization and methodological pitfalls may influence interpretation (Fonseca *et al*., 2012). Lately, several cases of interest to fisheries have been published. The microbial community of epidermal mucus of North Atlantic cod (*Gadus morhua*) showed a diversity of bacteria with distinct differences at a scale of hundreds of kilometers (Baltic, Icelandic, and North Seas) based on rDNA (Wilson *et al*., 2008). No temporal variation was found at two of the three locations (Baltic and Icelandic waters). In a study based on bacterial profiles from the outer and mouth mucus and the surrounding seawater of whiting (*Merlangius merlangus*) from the Irish and Celtic seas, Smith *et al*. (2009) were able to assign posi-

tions of fish within 79.7 to 57.3 km of their known harvest location. In a study at a smaller scale, the microbial community on the gills of young and adult plaice (*Pleuronectes platessa*) showed on a mudflat in the Wadden Sea temporal (sub-adult and adult) and spatial (within kilometers) differences between catch sites (Wegner *et al*., in press).

Higgins *et al*. (2010) found that multiple biological markers, including bacterial communities, were more powerful than by any single technique to classify Atlantic cod to their population of origin. This rationale was emphasized by Baldwin *et al*. (2012), who suggest the integration of fish-parasite based techniques for future fish stock assessments. It is particularly the case for pelagic fish species in which population structure may be particularly difficult to detect, as well as to resolve fish population structure over small geographic areas.

2.1.7 Advantages and disadvantages, criteria and problems of using parasites and microbes as a "magnifying glass" for stock structure:

Advantages

- Parasites and microbes may provide high resolution of population processes in comparison to host genetics at different scales (population identification, migration, refuges, cryptic processes, …)
- Integration of parasites within a comparative framework may be cost effective in relation to information obtained (*e.g.* compared to Next Generation sequencing of host)
- Fast analysis when operational
- Availability of complementary information to the stock assessment;
- Spatio-temporal information on parasites can be integrated in epidemiology and indicators of global change (climate, ecosystem health, …)
- Non-lethal field sampling and preservation in ethanol

Disadvantages

- Parasites are patchily distributed among and within host populations.
- There is a need for considerable biological information and an understanding of parasite and host life history, including reproductive strategies (Begg and Waldman, 1999)
- Parasite taxonomic expertise are required for quality control
- The potential for false genetic signals might lead to the wrong interpretations
- A "silver-bullet" approach across host species is not applicable, and there is a need for a case by case approach (compare with common microsat or SNP marker used in genotyping)
- Seasonal population dynamics of parasites affect availability
- Parasites may be highly sensitive to pollution and hence may quickly vanish; they may require conservation status (Whiteman and Parker, 2005)
- The running cost is higher than morphological characterisation
- Potential logistical constraints increase cost effectiveness (for example amount of DNA)

- Patterns of parasite population differentiation are only a proxy for host population characterization, and thus an indirect indication of cryptic host processes.

Conceptual and practical criteria for parasites and microbes:

- Strict host specificity, namely a close association of host and parasite over the evolutionary period of interest
- Monoxenous parasites preferred to avoid signal conflation via intermediate host activities (*e.g. Gyrodactylus* group of flatworms – Huyse and Volckaert, 2005)
- Parasite species should be abundant
- Parasite species should infect individuals across the full host range
- Parasite species should exhibit low gene flow between populations independent of the host
- Life-history traits relate to a short generation time/short-lived/high mutation rate
- Case-specific selection of taxa
- Sampling: temporal-spatial sampling has to have a resolution appropriate to the host, and might focus on the presumed breeding population of the host; the representative number of individuals is more important at the component (population) level than at the infra-population (individual). The reason is that infra-population is strongly determined by infection history, transmission dynamics and reproductive strategy.

Criteria for microbes:

- Sampling is more relaxed as it occurs at the assemblage/community level
- Sampling should be consistent in time and space (Wilson *et al*., 2008)
- Sampling may target the ectoderm or the intestine

2.1.8 Conclusions

1) It is unfortunate that the evolutionary history and population dynamics of fish parasites and microbes are understudied. As a consequence there is a serious knowledge gap, which seems to have widened because of the relatively small number of parasite experts.

2) However, the potential to use parasites and microbial communities in fish stock assessment is high. Judging from the few cases documented and from the latest developments in parasitology, microbiology and genomics, parasites are a good proxy for stock assessment.

3) The evolution of the parasite/microbe – host biological system harbors great advantages to study populations and stocks (including the higher resolution of the host population structure). However, disadvantages appear at various levels. Therefore clear criteria have to be established when selecting a parasite or microbial community for population delineation.

4) Parasites are often the first organisms to disappear in polluted/unhealthy ecosystems. Despite their low acceptance in human society, they merit equal conservation status as their free-living hosts in order to guarantee ecosystem health and good use in fish stock management.

2.1.9 WGAGFM recommends

1) To, given that parasite population genetics can be a proxy for identifying host fish populations (including farmed and native groups), make good use of it, when appropriate for the research question addressed. This requires promoting interdisciplinary interaction between fish biologists, fisheries scientist, ecologists, evolutionary biologists, parasitologists, bacteriologists and virologists in order to enhance parasite supported stock.

2) To integrate molecular species identification (*e.g.* DNA barcoding) with population genetic characterization of the parasite to facilitate throughput in stock identification.

2.1.10 References

Avise, J. 2000. Phylogeography: The history and formation of species. Harvard University Press.

Bakke, T., Harris, P. and Cable, J. 2007. The biology of gyrodactylid monogeneans: the "Russian-doll Killers". Advances in Parasitology, 64: 161–376.

Baldwin, R. E., Rew, M. B., Johansson, M. J., Banks, M. A., and Jacobson, K. C. 2011. Population structure of three species of Anisakis nematodes recovered from Pacific Sardines (*Sardinops sagax*) distributed throughout the California current system. Journal of Parasitology, 97: 545–554.

Baldwin, R. E., Banks, M. A., and Jacobson, K. C. 2012. Integrating fish and parasite data as a holistic solution for identifying the elusive stock structure of Pacific sardines (*Sardinops sagax*). Reviews in Fish Biology and Fisheries, 1: 137–156.

Begg, G. A. and Waldman, J. R. 1999. An holistic approach to fish stock identification. Fisheries Research, 43: 35–44.

Bellec, L., Grimsley N., Derelle E. *et al*. 2010. Abundance, spatial distribution and genetic diversity of *Ostreococcus tauri* viruses in two different environments. Environmental Microbiology Reports, 2: 313–321.

Bjorkevoll, I., Olsen, R. L., and Skjerdal, O. T. 2003. Origin and spoilage potential of the microbiota dominating genus *Psychrobacter* in sterile rehydrated salt-cured and dried salt-cured cod (*Gadus morhua*). International Journal of Food Microbiology 84: 175–187.

Criscione, C., Poulin, R., and Blouin, M. 2005. Molecular ecology of parasites: elucidating ecological and microevolutionary processes. Molecular Ecology, 14: 2247–2257.

Criscione, C., Cooper, B., and Blouin, M. 2006. Parasite genotypes identify source populations of migratory fish more accurately than fish genotypes. Ecology, 87: 823–828.

Cross, M. A., Collins, C., Campbell, N., Watts, P. C., Chubb, J. C., Cunningham, C. O., Hatfield, E. M. C., and MacKenzie K. 2007. Levels of intra-host and temporal sequence variation in a large CO1 sub-units from *Anisakis simplex sensu stricto (*Rudolphi 1809) (Nematoda: Anisakisdae): implications for fisheries management. Marine Biology, 151: 695–702.

Falush, D., Wirth, T., Linz, B, Pritchard, J. K., Stephens, M., Kidd, M., Blaser, M. J., Graham, D. Y., Vacher, S., Perez-Perez, G. I., Yamaoka, Y., Megraud, F., Otto, K., Reichard, U., Katzowitsch, E, Wang, X., Achtman, M, and Suerbaum, S. Traces of human migrations in *Helicobacter pylori* populations. Science, 299: 1582–1585.

Fonseca, V. G., Nichols, B., Lallias, D., Quince, C., Carvalho, G. R., Power, D. M., and Creer S. 2012. Sample richness and genetic diversity as drivers of chimera formation in nSSU metagenetic analyses. Nucleic Acids Research, 2012**:** 1–9.

Haakon, H. Bachmann, L., and Bakke, T. A. 2003. Mitochondrial DNA variation of *Gyrodactylus* spp. (Monogenea: Gyrodactylidae) populations infecting Atlantic salmon, grayling, and rainbow trout in Norway and Sweden. International Journal for Parasitology, 33: 1471–1478.

Hansen, H., Olstad, K., Bakke, T., and Bachmann, L. 2005. Species, subspecies or strains: the taxonomy of *Gyrodactylus salaris* and *G. thymalli* (Monogenea) Revisited. Proceedings of the IX European Multicolloquium of Parasitology. Volume 2, 229–234. July 18–23, Valencia, Spain (ISBN:88-7587-143-4).

Hauser, L. and Carvalho, G. 2008. Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. Fish and Fisheries, 9: 333–362.

Hovda, M. B., Lunestad, B. T., Sivertsvik, M., and Rosnes, J. T. 2007. Characterisation of the bacterial flora of modified atmosphere packaged farmed Atlantic cod (*Gadus morhua*) by PCR-DGGE of conserved 16 S rRNA gene regions. International Journal of Food Microbiology, 117: 68–75.

Hudson, P., Dobson, A., and Lafferty, K. 2006. Is a healthy ecosystem one that is rich in parasites? Trends in Ecology and Evolution, 7: 381–385.

Huisman, J., and Weissink F. J. 1999. Biodiversity of plankton by species oscillations and chaos. Nature, 204: 407–410.

Huyse, T., and Volckaert, F. 2005. Comparing host and parasite phylogenies: *Gyrodactylus* flatworms jumping from goby to goby. Systematic Biology, 54: 710–718.

Johnsen, B., and Jensen, A. 1991. The *Gyrodactylus* story in Norway. Aquaculture, 98: 289–302.

Josselin, E., Desdevises Y., and Coeur d'acier A. 2009. Fine-scale cospeciation between *Brachycaudus* and *Buchnera aphidicola*: bacterial genome helps define species and evolutionary relationships in aphids. Proceedings of the Royal Society of London B, 276: 187–196.

Kaliszewska, Z. A., Seger, J., Rowntree, V. J., Knowlton, A. R. and Marshalltilas, K., Patenaude, N. J., Rivarola, M., Schaeff, C. M., Sironi, M., Smith, W. A., Yamada, T. K., Barco, S. G., Benegas, R., Best, P. B.; Brown, M. W.; Brownell Jr, R. L., Harcourt, R., and Carribero, A. 2005. Population histories of right whales (Cetacea: *Eubalaena*) inferred from mitochondrial sequence diversities and divergences of their whale lice (Amphipoda:*Cyamus*). Publications, Agencies and Staff of the U.S. Department of Commerce. Paper 88. http://digitalcommons.unl.edu/usdeptcommercepub/88

Kijewska, A., Dzido, J., Shukhgalter, O., and Rokicki, J. 2009. Anisakid parasites of fishes caught on the African shelf. Journal of Parasitology, 95: 639–645.

Kuris, A. M., Hechinger, R. F., Shaw, J. C., Whitney, K. L., Aguirre-Macedo, L., Boch, C. A., Dobson, A. P., Dunham, E. J., Fredensborg, B. L., Huspeni, T. C., Lorda, J., Mababa, L., Mancini, F. T., Mora, A. B., Pickering, M., Talhouk, N. L., Torchin, M. E. and Lafferty, K. 2008. Ecosystem energetic implications of parasite and free-living biomass in three estuaries. Nature Letters, 454: 515–518.

Krkosek, M., Connors, B. M., Ford, H., Peacock, S., Mages, P., Ford, J. S., Morton, A., Volpe, J. P., Hilborn, R., Dill, L. M. and Lewis, M. A. 2011. Fish farms, parasites and predators: implication for salmon population dynamics. Ecological Applications, 21: 897–914.

Kuhn, T., Garcia-Marquez, J. and Klimpel, S. 2011. Adaptive radiation within marine Anisakid nematodes: A zoogeographical modeling of cosmopolitan, zoonotic parasites. PloS ONE **6**: e28642.

Lester, R. J. G. and MacKenzie, K. 2009. The use and abuse of parasites as stock markers for fish. Fisheries Research, 97: 1–2.

MacKenzie, K., and Abaunza, P. 1998. Parasites as biological tags for stock discrimination of marine fish: a guide to procedures and methods. Fisheries Research, 38: 45–56.

Marcogliese, D. J. 2004. Parasites: small players with crucial roles in the ecological theater. Ecohealth 1: 151–164.

Marcogliese, D. 2005. Parasites of the superorganism: Are they indicators of ecosystem health? International Journal for Parasitology, 35: 705–716.

Mattiucci S., Farina V., Campbel N., *et al*. 2008. Anisakis spp. larvae (Nematoda : Anisakidae) from Atlantic horse mackerel: Their genetic identification and use as biological tags for host stock characterization. Fisheries Research, 89: 146–151.

Mattiucci, S. and Nascetti, G. 2008. Advances and trends in the molecular systematics of Anisakis nematodes, with implications for their evolutionary ecology and host-parasite coevolutionary processes. Advances in Parasitology, 66: 47–148.

Mejia-Madrid, H. H., Vazques-Dominguez, E., And Perez-Ponce de Leon. 2007. Phylogeography and freshwater basins in central Mexico: recent history as revealed by the fish parasite Rhabdochona lichtenfelsi (Nematoda). Journal of Biogeography, 34: 787–801.

Meinila, M., Kuusela, J., Ziętara, M., and Lumme, J. 2002. Primers for amplifying ~820 bp of highly polymorphic mitochongdrial CO1 gene of *Gyrodactylus salaris*. Hereditas, 137: 72–74.

Meinila, M., Kuusela, J., Ziętara, M., and Lumme, J. 2004. Initial steps in speciation by geographical isolation and host switch in the salmonid pathogen *Gyrodactlyus salaris* (Monogenea). International Journal for Parasitology, 34: 515–526.

Miura, O., Torchin, M. E., Kuris, A. M., Hechinger, R. F., and Chiba, S. 2006. Introduced cryptic species of parasites exhibit different invasion pathways. Proceedings of the National Academy of Sciences of the USA, 103: 19818–19823.

Nieberding, C. M., and Olivieri, I. 2007. Parasites: proxies for host genealogy and ecology? Trends in Ecology and Evolution, 22:156–165.

Nieberding, C., Morand S., Libois R., and Michaux J. R. 2004. A parasite reveals cryptic phylogeographic history of its host. Proceedings of the Royal Society of London B, 271: 2559–2568.

Nieberding, C., and Morand, S. 2006. Comparative phylogeography: the use of parasites for insights into host history. Ed. by S. Morand, B. Krasnov and R. Poulin. I Micromammals and macroparasites. Springer, New York, pp 277–293.

Prugnolle, F., Roze, D., Theron, A., and De Meeus, T. 2005. F-statistics under alternation of sexual and asexual reproduction: a model and data from schistosomes (platyhelminth parasites). Molecular Ecology, 14: 1355–1365.

Rannala, B., and Michalakis, Y. 2003. Population genetics and cospeciation: from process to pattern. *In* Tangled Trees: Phylogeny, Cospeciation and Coevolution. Ed. by R. D. M. Chicago: The University of Chicago Press, Pp 120-143.

Reynisson, E., Lauzon, H. L., Magnusson, H., Jonsdottir, R., Olafsdottir, G., Marteinsson, V. and Hreggviosson, G. O. 2009. Bacterial composition and succession during storage of North-Atlantic cod (*Gadus morhua*) as superchilled temperatures. BMC Microbiology, 9: 250.

Ricklefs, R. E. 2011. A biogeographical persepective on ecological systems: some personal reflections. Journal of Biogeography, 38: 2045–2056.

Rohr, J. R., Dobson A. P., Johnson P. T. J., *et al.* 2011. Frontiers in climate change-disease research. Trends in Ecology and Evolution, 26: 270–277.

Smith C. J., Danilowicz B. S., and Meijer W. G. 2009. Bacteria associated with the mucus layer of *Merlangius merlangius* (whiting) as biological tags to determine harvest location. Canadian Journal of Fisheries and Aquatic Sciences, 66: 713–716.

Wegner K. M., Shama L. N. S., Kellnreitner F., and Pockberger M. (in press) Diversity of immune genes and associated microbes of European plaice *Pleuronectes platessa*. Estuarine and Coastal Shelf Science.

Whiteman, N. K. and Parker, P. G. 2005. Using parasites to infer host population history: a new rationale for parasite conservation. Animal Conservation, 8: 175–181.

Wielgoss, S., Taraschewski, H., Meyer, A., and Wirth, T. 2008. Population structure of the parasitic nematode *Anguillicola crassus*, an invader of declining North Atlantic eel stocks. Molecular Ecology, 17: 3478–3495.

Wielgoss, A., Tscharntke, T., Buchori, D., Fiala, B., and Clough, Y. 2010. Temperature and a dominant dolichoderine and species affect ant diversity in Indonesian cacao plantations. Agriculture, Ecosystems and Environment, 135: 253–259.

Williams, H., MacKenzie, K., and McCarthy, A. 1992. Parasites as biological indicators of the population biology, migration, diet, and phylogenetics of fish. Reviews in Fish Biology and Fisheries, 2: 144–176.

Wilson, B., Danilowicz, B. S. and Meijer, W. G. 2008. The diversity of bacterial communities associated with Atlantic cod *Gadus morhua*. Microbial Ecology, 55: 425-434.

Wirth T., Meyer A. and Achtman M. 2005. Deciphering host migrations and origins by means of their microbes. Molecular Ecology, 14: 3289–3306.

Yasuike, M., Leong, J., Jantzen, S. G., von Schalburg, Kristian, R., Nilsen, F., Jones, S. R. M., Koop, B. 2012. Genomic resources for sea lice: analysis of ESTs and mitochondrial genomes. Marine Biotechnology, 14: 155–166.

Zietara, M.S., Rokicka, M., Stojanovski, S., and Lumme, J. 2010. Introgression of distant mitochondria into the genome of *Gyrodactylus salaris*: nuclear and mitochondrial markers are necessary to identify parasite strains. Acta Parasitologica, 55: 20–28.

2.2 ToR b) Review on the use of adaptive SNPs and other adaptive markers for genetic identification of populations (breeding stocks)

Tom Cross, Aitor Albaina, Ian Bradbury, Jens Carlsson, Gary Carvalho, Geir Dahle, John Gilbey, Sarah Helyar, Phil McGinnity, Paulo Prodohl, Johan Robbens, José A. Sánchez, Niklas Tysklind, Daria Zelenina

2.2.1 Preamble and context

In parallel with the escalation in advanced genomic technologies and associated bioinformatics pipelines to characterise genes and genomes and their association with specific biological traits and disease, there is an expanding repository of gene-associated markers for identifying individuals, populations and species. Increasingly, studies scanning moderate to large panels of loci are implicating a subset of these markers in local adaptation and demonstrating the utility of these markers in conservation and management. The notion of using so-called "adaptive" markers that potentially influence survival and/or reproductive success ("fitness variation") in fisheries management is not new. For example, Sick (1965) examined the distribution of haemoglobin gene frequencies in cod populations. Although, the author did not elucidate the underlying basis of functional differences among the haemoglobin types, the study revealed frequency differences among populations from different areas, thereby generating a genetic basis for the recognition of different "races" of cod. Following the early application of molecular genetic markers to fisheries management (Ryman and Utter, 1987), the general approach was to use markers that were indicative of demographic processes only (e.g. population size, gene flow) – so-called "neutral" markers (see section 2). Typically however, a small proportion of

protein polymorphisms (allozymes) and more recently DNA- based population markers such as microsatellites, exhibited unusually high levels of genetic divergence among populations. In most cases such aberrant or "outlier" loci were removed from population genetic analyses because they violated theoretical assumptions of the neutral model of molecular evolution.

Here, we examine why and how the recent availability of large numbers of genetic markers potentially influenced by directional selection can provide powerful and informative tools in fisheries management. While we make reference to several marker types, we focus our attention on genetic differences scattered throughout the genome of all species, - "single nucleotide polymorphisms" (SNPs), given their abundance relative to other marker types. More formally, a SNP is a DNA sequence variation occurring when a single building block of DNA ("nucleotide"- A, T, C or G) – in the genome differs among individuals of a species. SNPs are commonly the result of single base substitution of one nucleotide by another, although some authors consider that single/small base insertions and deletions (indels) should also be considered as SNPs. Individuals in populations may share allele frequencies at many SNP and other loci, that together create a unique DNA pattern for that population, termed a "population signature".

While our ability to address complex and diverse questions in fisheries biology and management using molecular markers has improved greatly in recent years (reviewed in Hauser and Carvalho, 2008), the generic nature of the questions has remained fundamentally similar. Questions relating to the stock composition of catches, population assignment, traceability and spatial and temporal distribution of genetically discrete breeding stocks (often recently diverged), remain high priorities. (It is the "spawning group"- a group of fish spawning at a particular time and place and thus not interbreeding to any substantial degree with any other such group spawning in a different area or in the same area at a different time- that represents the fundamental evolutionary unit of biodiversity and hence management.) It is crucial that such units can be recognized, monitored and conserved. The fact that many such groups will exhibit heritable biological variation in fitness related life history traits including growth, size, fecundity and maturation, means that such intrinsic heterogeneity should be incorporated into baseline stock assessments. While neutral markers can provide a fundamental demographic framework for inferring evolutionary processes driving population persistence, potential for local adaptation and recovery of reduced stocks, they typically reveal little direct information on the biological significance of population differences. Adaptive markers, in contrast, may not only provide short-term "population tags" for describing and monitoring changes in stock distribution and for population assignment, but importantly they facilitate new insights into the nature and scale of adaptive diversity in wild fish populations – a key component of biodiversity that determines in part population resilience and ability to cope with environmental change (e.g. over-exploitation and climate change). Thus, markers under selection can add a new and relevant dimension to biodiversity - the genetic component of adaptive diversity, which demands a reconsideration of the most relevant spatial and temporal scales for population conservation.

2.2.2 Evolutionary background

Molecular markers, as used in fisheries and aquaculture, consist of proteins or sections of DNA which show heritable genetic variability (polymorphism) within a particular species and can be interpreted following Mendelian rules of inheritance (Carvalho and Pitcher, 1994). The different forms are referred to as "alleles" and dif-

ferent frequencies of these alleles can characterise or "tag" particular fish or invertebrate populations or stocks. Fundamentally, polymorphism occurs because of mutation at the DNA level, and is then moulded within individual populations by evolutionary forces, the most important of which are genetic drift (a random or stochastic process) and natural selection. There are several types of natural selection, but in the present context the focus is on directional selection. (Directional selection increases the frequency of a certain allele or alleles, i.e. confers reproductive advantage to individuals that possess particular allelic combinations over others.) An additional important evolutionary factor, often mediating the rate of genetic change, is the movement of individuals between populations which when accompanied by successful reproduction, is technically referred to as gene flow. Polymorphic genes, which are not influenced by natural selection, are referred to as neutral markers. This means that they occur in parts of the genome that are not affected by natural selection and that genetic drift is the only intra-population evolutionary force acting on them (disregarding mutation). In contrast, certain parts of the genome of a species are influenced by natural selection, in addition to genetic drift. Markers associated with these regions are referred to as selective or adaptive markers. Those markers influenced by directional selection usually have higher evolutionary rates (change much faster) than neutral markers. Thus, adaptive markers can be much more discriminatory of individuals from distinct populations than neutral markers, and can be of much greater utility in assignment or traceability studies (see below). Mathematical methods exist (as described below) to distinguish these two markers types, allowing for separate analysis, which is necessary not only because of the different management implications that can be derived for the two marker types but also because the statistical analytical framework for analysis of both marker types is often not compatible. In the 1960s, when the only molecular markers available were the so-called allozymes (polymorphic enzymes detected by protein electrophoresis), testing whether the marker was influenced by selection or not was typically complex and time-consuming, often requiring data on environmental correlates or biochemical investigations. Most protein markers, unless otherwise indicated were regarded as neutral (though since the underlying genes were coding for particular proteins - were so-called transcribed genes, we now might regard this assumption as doubtful). During the remainder of the 20th century and early years of the 21st, as DNA techniques increased in number and complexity, through mitochondrial and then nuclear DNA, most markers appeared to behave as neutral, with certain exceptions such as the Major Histocompatibility Complex (MHC, deEyto *et al*., 2007), but systematic direct testing for selection was rare. With the advent of huge numbers of molecular markers, chiefly SNPs, indels and microsatellites, from genomic studies and Next Generation Sequencing (NGS), it is now evident that a number of them do not fit the neutral model of molecular evolution. An increasing number of studies are now reporting on "outliers markers", which are often characterised by higher levels of inter population divergence as measured by standard population structuring estimators (e.g. F_{ST}). This observation has prompted a renewed interest in analytical methods, which can be used to assess whether particular markers are potentially influenced by directional selection. Over the past few years, statistical procedures have been developed to test whether the extent and distribution of genetic variants fits models of neutrality or selection (see below).

2.2.3 Utility of adaptive markers: a critique

Despite the fact that some markers show large population differences that exceed neutral expectations (are so-called "outliers"), selection was rarely measured in the

past, as noted above. Outlier loci identification is often based on positive test results using one of several available approaches (see below). How "adaptive" individual loci are likely to be will ultimately require evaluation on a case-by-case basis with careful mapping and annotation of associated genes. We will refer to these outliers as "adaptive" here, but recognize the biological significance of such diversity is often unknown initially. It is nonetheless of paramount importance to test for marker 'neutrality' prior to use (i.e. for exploring population structure) by using outlier tests as outlined below and recognize the limitations of available tests in definitively establishing a link to selection or adaptation.

As these "outlier tests" rely on direct comparisons among loci and across populations, processes which influence locus specific rates of divergence may themselves or in concert act to bias and increase the rate of false positives (i.e., Type one error). Population structure, particularly associated with small isolated populations, is the most commonly cited cause of false positives in testing for adaptive loci (e.g., Foll and Gaggiotti 2008; Excoffier *et al.* 2009; see below for details on specific methods). Also differences in mutation rate of loci used may influence the rate of false positives (Nei and Maruyama 1975; Beaumont 2008). Accordingly this may increase the rate of false positives for balancing selection (e.g., Mäkinen *et al*. 2008) but perhaps also for directional selection (see Bradbury *et al*. 2011a). Finally, it also seems likely that bias associated with the ascertainment of adaptive markers may also be responsible for increasing the rate of false positives. Ascertainment bias may result if the panel used to ascertain markers such as SNPs is not representative of the geographical region in which the markers are applied and may manifest itself as a decrease in diversity in regions outside the ascertainment region. Bradbury *et al*. (2011b) observed significant decreases in diversity in SNPs in Atlantic cod isolated in the West Atlantic but genotyped range wide. In this case, a significant portion of SNPs fixed in the eastern Atlantic, test positive for selection and it is impossible to rule out ascertainment bias as a cause. Accordingly, the authors chose to restrict tests for selection to discrete geographic regions identified using putatively neutral loci and principle component analysis (Bradbury *et al*. 2010).

Given that adaptive loci may be influenced by selection, errors can result if the purpose of a study is to make inferences of demographic or historical population processes. In cases where demographic inferences are desired, such as when estimating gene flow or population size, markers under selection should be removed prior to analyses (Beaumont and Nichols, 1996). The inclusion of information for loci under directional selection themselves, or loci tightly linked to regions under selection, leads, as noted above, to violation of assumptions for most demographic or neutral population genetic models, and may bias results and cause erroneous inference about population demographic parameters (see discussion in Laval *et al*. 2010). In weakly structured species, the effect of just a few loci on overall patterns could be significant, but provided selected loci make up only a small proportion of the total marker number, biological inference is not generally expected to be severely biased (Luikart *et al*. 2003). On the other hand, markers under selection can be exploited for specific purposes, such as investigating population structure on ecological rather than evolutionary timescales (Waples and Gaggiotti, 2006), for increasing the power for traceability and the assignment of individuals to populations of origin (Nielsen *et al.* 2009, Nielsen *et al.* 2012), identifying candidate genes under selection (Bonin 2008; Brieuc and Naish, 2011; Hemmer- Hansen *et al*. 2011), and resolving the adaptive landscape for the identification of conservation units (Allendorf *et al*., 2010).

2.2.3.1 Temporal stability of signal from adaptive markers

Given the dynamic response of adaptive markers to selection the issue of temporal stability must be addressed. However, studies analyzing temporal stability in non-neutral (adaptive) markers are still scarce, maybe due to economic constraints, since genotyping multiple markers in multiple years and for a sufficient number of individuals, requires substantial effort and resources (but see case study on FishPopTrace below). However, the temporal stability of adaptive loci increases the likelihood that such loci are useful, when used in conjunction with neutral markers, as population markers (e.g. Waples, 1998; Nielsen *et al*., 2007). In this sense, one potential concern about adaptive markers is that the frequencies of alleles under selection may change more quickly because of environmental change (e.g. Ackerman *et al*., 2011). These authors also raised a question about how many generations it would take, on average, for adaptive SNP loci to demonstrate temporal allele frequency differences. Although further work is required, we present some case studies that reveal insights to such issues.

1) The presence of temporal instability in adaptive markers has recently been raised by Andre *et al*. 2011 when genotyping one microsatellite locus under selection (associated with salinity) in herring (*Clupea harengus*). They tested temporal stability by applying hierarchical analysis of molecular variance, with sampling years nested within locations. The authors used the locus-by-locus option in ARLEQUIN 3.1; statistical significance being obtained from 10 000 permutations. Further, they divided all fish into year classes based on otolith ageing, then tested for genetic heterogeneity among cohorts with n≥20 individuals, within and among locations. They found that samples from the North Sea clustered by year of sampling and concluded that this could be a year-class effect. This highlighted the influence temporal differences among cohorts may have on inferences drawn from adaptive markers.

2) In contrast, other studies indicate stability in adaptive markers over periods of several decades. Poulsen *et al*. (2011), analyzed 92 gene-associated single-nucleotide polymorphism (SNP) markers in Atlantic cod (*Gadus morhua*) from several sampling sites within the North Sea and adjacent areas, and tested these loci for temporal stability including long- and short-term temporally replicated samples (up to 38 years apart), from a subset of populations. Of this panel, they observed three loci that showed signatures of directional selection and highly elevated levels of genetic differentiation. The analysis of the historical samples revealed long-term temporally stable patterns of both neutral and adaptive divergence among some populations, interpreted as indicating long-term temporal adaptive stability driven by strong local selection. Moreover, adaptive loci did not vary clinally with either latitude or longitude indicating that outlier signals were associated with one or perhaps a few population samples, supporting a strong role for local selection over neutral isolation by distance

3) Similar conclusions of temporal stability where drawn by Ackerman *et al*. (2011). Their study identified four SNP loci from a panel of 42 loci, as candidates for directional selection in sockeye salmon (*Oncorhynchus nerka*) from the Copper River and adjacent coastal drainages in south-central Alaska. They evaluated the information content of the four adaptive loci showing that such loci improved the ability to identify the origin of individual fish and to estimate the composition of Pacific salmon populations

in mixed fisheries. A total of 18 sampling locations were sampled in multiple years. For all eight locations, temporal collections (spanning one to four generations) failed to demonstrate any significant departures from homogeneity (α = 0.05), thus confirming temporal stability of the adaptive SNPs allele frequencies.

Thus, some of these illustrative studies indicate that across ecological time scales of decades, that adaptive markers can show sufficient temporal stability to yield spatially meaningful biological differentiation. However, it should be emphasized that where reference databases are established, it is advisable to re-sample representative sites at a frequency determined by biological (e.g. frequent extreme fluctuations in population abundance) or local environmental factors (e.g. rapidly changing temperature or over-exploitation).

2.2.4 General strategies and data analysis:

The identification of panels of adaptive loci can be achieved by two general approaches. In the first, a large panel of loci (100s-1000s) is examined and those under selection identified. In the second approach candidate genes or functional loci only are directly targeted. Both approaches are commonly reported in the literature but the suitability will depend on the availability of genetic resources for a given species and the understanding of the selective landscape. In general, for non-model marine organisms, genome scans for adaptive markers are often the most suitable and informative, since existing genomic resources are often limited, though this may not always be the case (e.g., Nielsen *et al*. 2009).

Approach 1 – Genome scans for the identification of adaptive markers

The search for signatures of selection in molecular data has had a long tradition in evolutionary biology (e.g. reviewed by Helyar *et al*., 2011). Although discarded when facing population genetics questions such as describing migration or demographics, the utility of outlier loci for origin (stock) assignment purposes is very high and they usually outperform neutral loci in this context (e.g. Russello *et al*., 2012).

For molecular markers there are basically two main groups of methods to detect outlier loci. Arguably, the classic method is the one developed by Beaumont and Nichols (1996), based on the F_{ST} outlier approach, which is readily implemented in the software LOSITAN - selection detection workbench (Antao *et al*., 2008). The main problem with the Beaumont and Nichols (1996) approach is that there is an assumption that populations are at drift-migration equilibrium, which is unrealistic in most natural situations. Two Bayesian methods which account for this problem are BAYESFST (Beaumont and Balding, 2004) and BayeScan (Foll and Gaggiotti, 2008) which identify loci potentially under selection through estimates of locus effects on F_{ST} (Narum and Hess, 2011). These two methods are based on the same regression model but differ in the way that the effect of selection is inferred. Both programs have been widely applied, but they have also recently been found to be vulnerable to complex population structure scenarios, such as when populations are hierarchically structured, leading to correlated allele frequencies among samples (Excoffier *et al.* 2009). In such situations the Arlequin 3.5 software may be more appropriate (Excoffier *et al.* 2009) as the implementation of a hierarchical model results in higher variance between simulated neutral loci and thus leads to a more conservative estimate of the number of outlier loci (Excoffier *et al*., 2009).

In general, having too few samples can substantially reduce the statistical power of these methods (Foll and Gaggiotti, 2008), so they detect only extreme outlier loci, missing many potential candidate loci. In contrast, too many samples can also bias results, resulting in increased false positive rates (Excoffier *et al*., 2009). This bias could be reduced through analyzing balanced sub-sets of samples, i.e. using a similar number of samples from each of a number of populations or groups of populations identified through other approaches, such as clustering methods. Again, the important thing is to have clarity in the question that is being addressed. If the goal is to identify sets of markers with high discriminatory power among different populations/groups of populations, then in principle it does not matter if a detected outlier is truly subject to selection, or if it is a false positive, provided that the signal is temporally stable. In this case, the outlier detection can be viewed as an explorative preliminary exercise supporting downstream analyses. However, if evolutionary or demographic processes are being investigated, the inclusion of loci under selection may influence results substantially so careful attention should be paid to the design of the strategy to identify outlier loci.

Although the identification of candidate loci under selection (outlier loci) has been mainly performed in the few last years with LOSITAN software, recently several papers showing a lower rate of false positives in BayeScan compared with LOSITAN, has made the former the software-of-use (e.g. Narum and Hess, 2011). However, a lot of technical questions still remain, especially about the minimum numbers of individual markers that should be tested. In this sense, a recent paper applying simulations (Landguth *et al*., 2012) has concluded that amplifying more (and more variable) loci is likely to increase the power of landscape genetic inferences more than increasing number of individuals. Moreover, the advent of NGS techniques yielding hundreds-to-thousands of markers, along with cheaper and more reliable high-throughput genotyping technologies (e.g. Garvin *et al*., 2010, Ekblom and Galindo, 2011, Nielsen *et al*., 2011) will soon address this technical constraint to the acquisition of a sufficient number of samples covering most of the species distribution.

Once the requirements cited above are fulfilled, a conservative and maybe optimum strategy could be to discard, for demography purposes, only loci that consistently appeared to be under directional selection, as denoted by being candidate loci in both BayeScan and LOSITAN, because the two approaches use different assumptions and algorithms (e.g., Richter-Boix, 2011). While BayeScan is generally considered as the best approach, discrepancy between the outputs of different outlier loci detection softwares is still under debate (e.g. Manel *et al*., 2009; Pérez-Figueroa *et al*., 2010; Gomez-Uchida *et al*., 2011; Narum and Hess, 2011; Nunes *et al*., 2011; Russello, 2012).

Finally, in a recent review, Le Corre and Kremer (2012) provided several case studies where the integration of various external information (phenotypic, transcriptomic, functional genomic information, etc.) with outlier detection methods, led to the successful identification of important selected genes. These authors illustrated the need for more integrated approaches for detecting selected loci, suggesting shifting from testing individual markers to multilocus approaches and integrating knowledge about candidate genes and phenotypic data, as well as ecological and environmental data.

Approach 2 – The use of candidate loci

Developments in sequencing technologies, the increasing availability of information relating to functional loci and even complete genomes of non-model species, now allow individual markers to be targeted as embedded in functional loci, or linked to

functional traits and hence potentially under selection. This technique can be a powerful way of linking genotype to phenotype and to further link both of these to the environment. In heterogeneous environments, this may result in directional selection pressures upon the loci and result in high discriminatory power. Further, such loci may be of particular use for examining the genetic composition and adaptive responses of populations to changes in environmental conditions. In general, the investigation of candidate loci has yielded great success for the identification of adaptive markers. For example, Nielsen *et al*. (2011) reported a significantly higher proportion of candidate genes showing evidence of selection (30%) than randomly selectively SNPs (6%), highlighting the utility of this approach. In this particular case, candidate genes where chosen from publicly available databases with functions expected to be associated with temperature, growth and reproduction. In the near future, with increased success in identification, candidate genes allow exploration of functional relationships. This contrasts with adaptive markers identified using genome scans, where the ability to annotate and explore functionality of loci may be limited (e.g. <5%; Bradbury *et al*., 2010).

2.2.5 Case Studies

The following reviews six studies where adaptive - or presumably adaptive - markers have been applied in a management and/or conservation context.

2.2.5.1 Case study: Traceability of fish populations and fish products – FishPopTrace

Several tools are available to understand the extent to which fish populations interbreed and to trace back the geographic origin of landed fish. These include physical external tags, natural tags and genetic markers. However, once a fish enters the food supply chain, several tools become less suitable. Cooking excludes the use of external features, as only the fillet in its processed state is available. Tools for monitoring natural populations and application to fisheries enforcement should therefore meet stringent criteria: they should mirror population identity and stability over ecological (environmental isolation) and evolutionary (limited interbreeding) scales. Traceability tools should be available throughout the food supply chain from capture to a customer's plate (from ocean to fork) and should be amenable to forensic validation for use in a court of law.

Analyses of SNPs can reach hitherto unprecedented levels of group identification, rendering them optimal tools in fundamental biology, conservation and traceability. In addition, the identification of SNPs not only responsive to changes unrelated to environmental differences (neutral SNPs), but also to natural selection (adaptive SNPs), greatly improves power of assignment.

FishPopTrace – a recently completed EU Framework 7 project (http://fishpoptrace.jrc.ec.europa.eu/) – demonstrates the application of SNP markers to the mapping and traceability of fish populations, with an initial focus on four marine fish species: cod, herring, hake and sole. The core approach was the development of a "SNP chip" for each of three species: sole, hake, and herring. (Canadian researchers had already created a SNP chip for cod to assist in aquaculture research). These devices enabled testing the identity of 1536 possible SNPs for each group of individuals from a specific population.

A common concern for European consumers is the source of Atlantic cod, *Gadus morhua*. Fish from the Baltic are worth less because they tend to have lower quality flesh and higher levels of contaminants. The cod team used its SNP chip to examine, blind-

ly, samples from both locations. By looking at 20 SNPs, the researchers correctly identified the origin of each individual fish. With just 10 SNPs, 96% of the unknown samples were still correctly identified.

The SNP chip for sole (*Solea solea*) also performed well. This flatfish is severely over fished in Europe. Only two of the twelve fishery areas within European waters are considered to be fished within safe biological limits. A key question is whether sole from the North Sea can be distinguished from populations in the Mediterranean. Just one SNP could reveal which sole was which, with 96% accuracy.

European hake (*Merluccius merluccius*) must be 27 cm long to be legally landed in the eastern Atlantic, while in the Mediterranean, vessels can catch hake that are only 20 cm in length. Fishing vessels in the Bay of Biscay are known to occasionally catch smaller fish, which are then misreported as originating in the Mediterranean. FishPopTrace has shown that just 10 SNPs can reveal the origin of hake with near-perfect accuracy.

The most challenging test case was perhaps Atlantic herring, (*Clupea harengus*), a geographically wide-spread and abundant species, with complex seasonal migratory behaviour. Herring within European waters typically display only minor and sometimes transient genetic differences among populations. By applying the SNP chip to herring, however, it was possible to accurately distinguish many entities, including those in the northeast Atlantic and North Sea, a goal important to a joint EU - Norwegian fishery management plan. The flexibility of combining differing numbers of SNPs allowed the identification of some herring groups at smaller scales, even around the United Kingdom, where there is substantial misreporting of catches.

Thus, it has been possible, by varying the numbers used on a SNP-chip, to assign individuals back to their source grouping across different geographic scales with high levels of certainty and reproducibility. Importantly, in the present context, it was those SNP markers showing evidence of selection that exhibited the highest level of genetic differentiation among target populations, and could therefore be grouped to generate a "minimum (and thus, cost-effective) SNP panel with maximum power. Such outputs are especially important, since previous types of genetic markers either detect levels of population differences that were too low for accurate assignment, or there were inherent difficulties in comparing data generated from different laboratories.

The species-specific panels of SNP markers can be implemented by control and enforcement authorities for essentially two purposes: i) as a standard tool to identify source populations (assignment) and ii) to verify the claimed population of origin from landed fish through to processed fish products, in the 'ocean to fork' sense. Implementation may be broad. In the first case it is envisaged that fisheries management will take greater account of biological structure as opposed to the current arbitrary geographic structure. It is also expected that the subtle differences between adjacent stocks will be based on adaptive markers.

2.2.5.2 Case study: Use of the adaptive marker Pan-1in management of spawning cod off the Norwegian coast

Until the mid 1970s, the local Norwegian Coastal cod (NCC) was managed as part of the highly migratory Northeast Arctic cod (NEAC) stock. Due to continued decline in survey results, ICES advised zero catch for NCC for the years 2004-2011, and recommended establishing a recovery plan to rebuild the NCC stocks. The rebuilding plan for the NCC, put in operation in 2011, aims at gradually reducing fishing mortality

until research surveys show results similar to the years 1995-1998. Fishing closure of targeted spawning grounds in the spawning season has been one regulation tool in the plan. So far seasonal closure has been in operation in the two main spawning areas for NEAC; Henningsværstraumen (since 2005) and Borgundfjorden (since 2009). These two spawning sites are also spawning sites for the local populations of NCC. The intention is to monitor the surrounding areas outside these closures, and, based on the fraction of NEAC outside the closures, the managers will decide on opening or closure.

Catches and survey indices have routinely been estimated by distinguishing between NCC and NEAC based on the visual inspection of the otoliths, a method requiring skilled readers. Cod from gillnet catches are commonly landed without heads and guts, making it impossible to read otoliths. A DNA method was introduced in 2005 enabling sampling of beheaded landings and analysis within 24 hours. The Pantophysin locus (*PanI*) exhibits particularly large differences in allele frequencies between samples of NEAC collected in the Barents Sea and the NCC samples from coastal areas of Norway. The method enables managers to calculate the fraction of NEAC in landings within 24 hours, and therefore react in real time to changes in the stock composition

The area round the Lofoten Islands (approx 200 km^2) is closed for all fishing activity from January 1 to June 30, with the exception of rod and hand line fishing. The Directorate of Fisheries, who regulates and monitors the fishing activity in Norwegian waters, would consider opening the area for gillnet fisheries, for all vessels smaller than 15 meters, if the fraction of NEAC in the commercial catches outside the closed area exceeded a preset fraction level of approximately 0.7. This occurred in 2011, when the fraction of NEAC in the catches exceeded 0.9 and the area was opened (Figures 1 and 2).

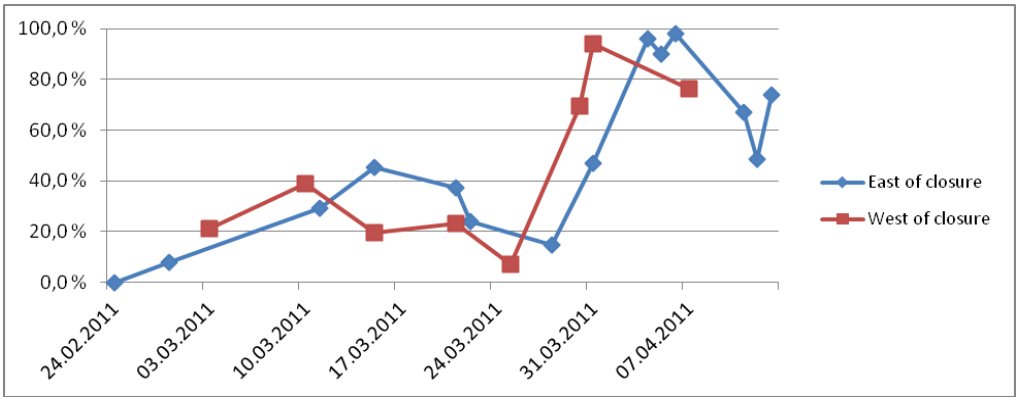

Figure 1. Fraction of Northeast arctic cod in the landings from catches around the closure in the Lofoten Islands.

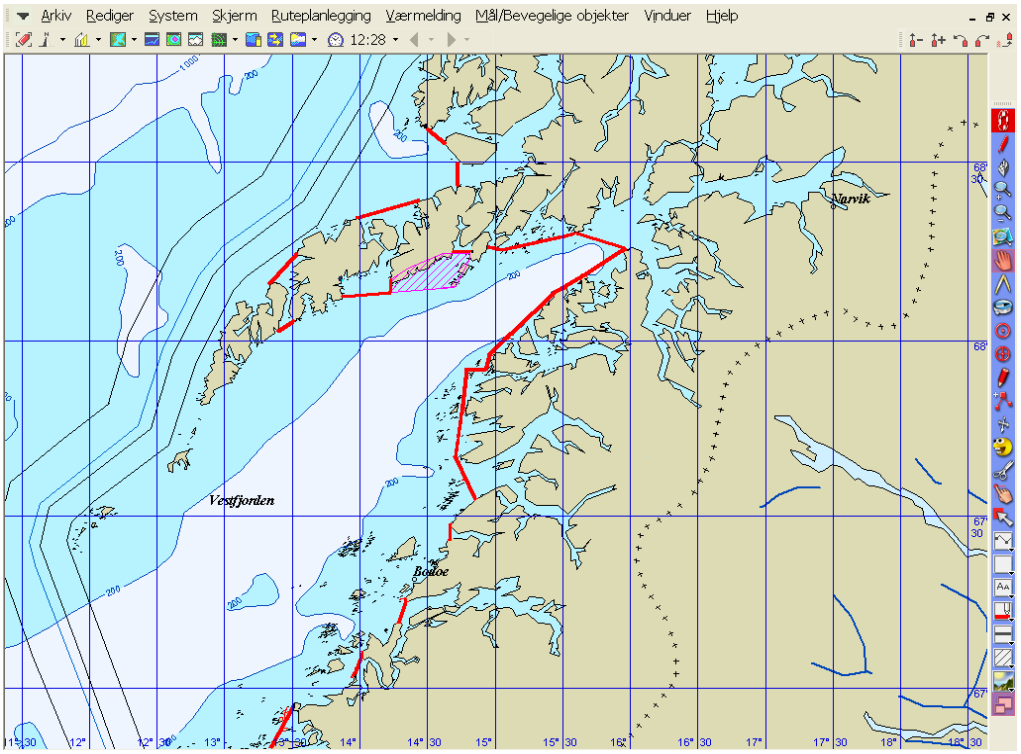

Figure 2. The Lofoten area in northern Norway, where the shaded area indicates the closed area south of the Lofoten Islands.

2.2.5.3 Case study: Barents Sea cod and Pan I – example of cautionary use of adaptive markers

When one creates a genetic baseline, samples from spawning sites should be included. The expectation then is that the genetic characteristics of fish throughout the whole habitat of the population will be similar; but it may happen that the ratios of allele frequencies of adaptive marker are not the same in different locations of the area, which a particular population inhabits. For example, all northeast Arctic cod are considered to spawn in the same area off the Lofoten Islands. However the comparison of PanI marker allele frequencies in the cod samples collected during the feeding period in different parts of the Barents Sea revealed clinal variation in allele frequencies from the northwest to the southeast (Markeenko *et al.*, 2012). Moreover, in most cases this effect is temporally stable. The putative reason for this phenomenon is that the Pantophysin marker is "very adaptive" and during the feeding period, cod gravitates towards different localities according to water temperature and/or salinity (or something else). Thus, if we include the Panthophysin locus in a panel of markers and explore the traceability of cod caught in the southern part of the Barents Sea, we may obtain misleading assignments because of the discrepancy of F frequencies between the reference sample and the sample(s) being examined.

2.2.5.4 Case study: Adaptive markers and sockeye salmon (*Oncorhynchus nerka*) fishery management

Conservation of Pacific salmon (*Oncorhynchus* sp.) stocks is complicated by the presence of both stocks that are healthy and stocks that are depressed, within the same mixed stock fisheries. As such, there has been a need for stock-specific fisheries management. Coded wire tags, parasites and scale patterns have all been used for this purpose with varying degrees of success, as has Genetic Stock Identification (GSI)

using a variety of molecular markers (Shaklee *et al*., 1999). Demographic structure of sockeye salmon has been identified and utilised for GSI purposes at a hierarchy of scales. Using neutral microsatellites alone the salmon cluster into 19 geographically based regions relating to individual river basins and groups of small lakes on both the British Columbia mainland coasts and Vancouver Island (Miller *et al*., 2002; Beacham *et al*., 2006). However, using the adaptive class II MHC locus, a second level of structure is revealed based on individual lake systems, with the allele frequency differentiation among such systems exceeding neutral microsatellite variation by an order of magnitude (Miller *et al*., 2002). This increased discriminatory power has also been reflected in the power of such adaptive loci in performing GSI. A comparison of stock identification using lake stocks from across the species Pacific range showed that the MHC locus was more effective at stock identification than 13 of the 14 microsatellite loci presently utilized (Beacham *et al*., 2005). Thus GSI in sockeye salmon has utilised both the regional structure revealed by neutral microsatellites and the smaller scale demographic differences revealed by the adaptive and more discriminatory MHC locus.

2.2.5.5 Case study: Evolutionary significant units

In addition to the utility of adaptive markers for population or individual assignment, they may also resolve the genetic component of adaptive diversity. As diverse adaptive portfolios of populations have been linked to fisheries stability, adaptive markers can provide a means of resolving conservation units in exploited species. Such conservation units are often defined as discrete and evolutionarily significant units, where "significant" means that the population is important to the evolutionary legacy of the species as a whole and if lost would likely not be replaced over ecological time scales (COSEWIC, 2010). Discreteness may refer to genetic isolation, habitat discontinuity, or ecological isolation. Significance may refer to deep phylogenetic divergence (e.g. glacial races), adaptive (e.g. life history variation), or ecological uniqueness, and its inclusion in the definition reflects the opinion that isolation in and of itself is not deemed sufficient for designation. In this context, the use of loci which reflect local adaptation may be particularly informative and help define units of conservation. For example, in Atlantic cod in Canadian waters, initial conservation assessments were based on one unit. Using microsatellite loci both Bentzen *et al*. (1996) and Ruzzante *et al*. (1998) then reported significant structure which was largely driven by a single locus, Gmo132. This locus has since been shown (Nielsen *et al*., 2006) to have elevated divergence associated with hitch-hiking selection (i.e. linkage to a gene under selection). Similar observations of elevated divergence associated with the Pantophysin locus have been made in Canadian waters, though the structure is much lower than observed in the eastern Atlantic (Beacham *et al*. 2002). Recently Bradbury *et al*. (2010) examined 1641 expressed SNPs in cod from 19 locations throughout Canadian and adjacent waters. This work identified a suite of loci as potentially experiencing selection associated with ocean temperature. These findings directly resulted in an increase of the number of ESUs recognized in Canadian waters, a revision which better reflects the adaptive diversity present in the species.

2.2.5.6 Case Study: Understanding local adaptation and environmental change impacts in salmonids

Anthropogenic climate change and its impacts are currently a major worldwide concern at many levels (Kerr, 2007), and impact in the oceans, such as increase in mean sea temperature and change in seasonal shifts to date are well recorded (Burrows *et al*., 2011). Climate change is predicted to alter the productivity, population sizes and

migration patterns, and result in extended invasions, extirpations and substitutions of many marine commercially exploited species (Cheung *et al*. 2009). Hence, understanding the mechanisms of how climate change might influence natural populations is crucial for the development of effective monitoring techniques and a predictive framework that will help us cope with climate change impacts (Hansen *et al*., 2012).

Although selection and genetic adaptation have been extensively tested in laboratory organisms, understanding the molecular mechanisms underlying environmental adaptation patterns of natural non-model organisms is hindered by the complexity of variables involved and the lack of individual pedigree and species genome information. Nevertheless considerable progress has been made by comparing both neutral and non-neutral markers and evaluating their association with environmental variables. Hansen *et al*. (2012) reviewed the use of genetic markers to evaluate adaptive responses to environmental change and produced criteria for demonstrating adaptive genetic change, which are summarised as: (i) suitable genetic variation exits; (ii) the monitored genes are relevant to the specific environmental stress; (iii) genes are analysed over time; (iv) selection is tested; (v) shifts in allele frequencies coincide with changes expected in response to the environmental change and (vi) adaptive genetic change is not inferred when replacement by a different genetic population has occurred instead.

It is important to note that while the previously discussed approaches (LOSITAN, BAYESCAN, ARLEQUIN) identify adaptive loci as outliers which show atypically high differentiation, the signals of environmental selection on selected loci may not necessarily appear as high differentiation (Narum and Hess, 2011). Hence, another way of detecting loci potentially under selection is pursuing the expectation that they will show allele distributions which are incongruent with patterns seen in other parts of the genome, which will show patterns resulting from evolutionary forces such as migration and drift. This approach was employed by Joost *et al*. (2008) to develop a Spatial Analysis Method (SAM), where correlations between allele frequencies and environmental variables are tested through univariate regression models. Significant associations between loci and environmental variables detected through the SAM method then need to be tested against neutral population structure to avoid targeting outlier loci, representative of population structure but not necessarily associated with the target environmental variable (Excoffier *et al*., 2009). It is worth noting that loci detected with the SAM method are not necessarily FST outliers (Narum *et al*. 2010). Despite the attractiveness of detecting adaptive markers associated with known variables, the main limitation of SAM approach is the appropriate choice and collection of pertinent environmental variables.

A similar approach was employed by Narum *et al*. (2010) who identified six SNP loci strongly associated to climatic regimes in redband trout (*Oncorhynchus mykiss*), by comparing montane and desert populations. Putatively neutral markers yield expected demographic patterns: desert populations were more isolated from each other than montane ones, which seemed to experience stronger gene-flow, with an isolation-by-distance pattern; Conversely, SNPs identified to be associated with temperature or precipitation regimes, not only showed allele frequencies strongly correlated with specific climatic variables, but showed a two-fold increase in levels of differentiation among populations experiencing variable extremes (despite neutral structure), and showed significant isolation-by-temperature across populations. Such work demonstrates the combined utility of correlation and outlier based tests for adaptive marker identification.

In a similar study evaluating the genetic component of the propensity for anadromy in *O. mykiss* populations, Narum *et al*. (2011) found three SNPs significantly associated with anadromy, while controlling for neutral population structure and other markers associated with environmental variables (temperature, elevation, upstream distance, etc.). These SNPs were then used to construct a model to predict propensity for anadromy in other local populations, which further corroborates their association with the trait of interest.

Although the cases reviewed here (Narum *et al*., 2010, 2011) do not yet fulfil all the criteria stipulated by Hansen *et al*. (2012) (i.e. temporal analysis of allele shifts), they lay the necessary framework for the routine monitoring of markers associated to traits of interest to evaluate anthropogenic impacts such as climate change, and to formulate predictions of the potential for adaptation of specific populations.

2.2.6 The WGAGFM recommend

- That markers under directional selection continue to be identified and employed in analysis as such markers have been shown to yield informative insights on both the scale and dynamics of populations and in identifying potential underlying drivers.
- That markers associated with functional variation be used to explore the predictive power to study the impacts of selective pressures including fisheries, climate change, and pollution.
- That based on the binary nature and ease of cross laboratory calibration when considering SNP databases, that the potential to combine datasets over larger spatial scales across labs be explored.

2.2.7 Action list

- That biological inferences from adaptive markers be interpreted within a demographic framework (e.g. based on neutral markers) since population-level demographic change will underpin long-term processes that influence, in part, the resilience and recovery of exploited individuals.
- A comparative analysis of neutral, selective, combinations and subsets of markers is undertaken to maximise the power to detect signals of differentiation across a range of geographic scales.
- As with neutral markers, when adaptive markers are employed, temporal stability should be periodically checked (as selective pressures may change over time).

2.2.8 References

Ackerman, M. W., Habicht, C., and Seeb, L.W. 2011. Single-Nucleotide Polymorphisms (SNPs) under Diversifying Selection Provide Increased Accuracy and Precision in Mixed-Stock Analyses of Sockeye Salmon from the Copper River, Alaska. Transactions of the American Fisheries Society, 140: 865–881.

Allendorf, F. W., Hohenlohe, P. A., and Luikart, G. 2010. Genomics and the future of conservation genetics. Nature Reviews Genetics, 11: 697–709.

Andre, C., Larsson, L. C., Laikre, L., Bekkevold, D., Brigham, J., Carvalho, G. R., Dahlgren, T. G., Hutchinson, W. F., Mariani, S., Mudde, K., Ruzzante, D. E., and Ryman, N. 2011. Detecting population structure in a high gene-flow species, Atlantic herring (*Clupea harengus*): direct, simultaneous evaluation of neutral vs putatively selected loci. Heredity, 106: 270–280.

Antao, T., Lopes, A., Lopes, R. J., Beja-Pereira, A., and Luikart, G. 2008. LOSITAN: A workbench to detect molecular adaptation based on a FST-outlier method. BMC Bioinformatics, 9:323–327.

Beacham, T. D., Brattey, J., Le, K. D., Miller, K. M., and Withler, R. E. 2002. Multiple stock structure of Atlantic cod (*Gadus morhua*) off Newfoundland and Labrador determined from genetic variation. ICES Journal of Marine Science, 59: 650–665.

Beacham, T. D., Candy, J. R., McIntosh, B., MacConnachie, C., Tabata, A., Kaukinen, K., Deng, L., Miller, K.M. and Withler, R.E. 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., McIntosh, B., MacConnachie, C., Miller, K. M., Withler, R. E. and Varnavskaya, N. 2006. Pacific Rim Population Structure of Sockeye Salmon as Determined from Microsatellite Analysis. Transactions of the American Fisheries Society, 135: 174–187.

Beaumont, M. A. 2008. Selection and sticklebacks. Molecular Ecology, 17: 3425–3427.

Beaumont, M. A. and Balding, D. J. 2004. Identifying adaptive genetic divergence among populations from genome scans. Molecular Ecology, 13: 969–980.

Beaumont, M. A. and Nichols. R. A. 1996. Evaluating loci for use in the genetic analysis of population structure. Proceedings of the Royal Society B, 263: 1619–1626.

Bentzen, P., Taggart, C. T., Ruzzante, D. E., and Cook, D. 1996. Microsatellite polymorphism and the population structure of Atlantic cod (*Gadus morhua*) in the northwest Atlantic. Canadian Journal of Fisheries and Aquatic Sciences, 53: 2706–2721.

Bonin, A. 2008. Population genomics: a new generation of genome scans to bridge the gap with functional genomics. Molecular Ecology, 17: 3583–3584.

Bradbury, I. R, Coulson, M. W., Campana, S. E., Paterson, I. G. and Bentzen, P. 2011a. Contemporary nuclear and mitochondrial genetic clines in a north temperate estuarine fish reflect Pleistocene vicariance. Marine Ecology Progress Series, 438: 207–218

Bradbury, I. R., Hubert, S., Higgins, B., Bowman, S., Paterson, I., Snelgrove, P. V. R., Morris, C., Gregory, R., Hardie, D., Hutchings, J. and Bentzen, P. 2011b. Evaluating SNP ascertainment bias and its impact on population assignment in Atlantic cod, *Gadus morhua*. Molecular Ecology Resources Suppl.1: 218–225.

Bradbury, I. R., Hubert, S., Higgins, B., Bowman, S., Paterson, I., Snelgrove, P. V. R., Morris, C., Gregory, R., Hardie, D., Hutchings, J. and Bentzen, P. 2010. Parallel adaptive evolution of Atlantic cod in the eastern and western Atlantic Ocean in response to ocean temperature. Proceedings of the Royal Society of London Biological Sciences B, 277: 3725–3734.

Brieuc, M. S. O., and Naish, K. A. 2011. Detecting signatures of positive selection in partial sequences generated on a large scale: pitfalls, procedures and resources. Molecular Ecology Resources, 11: 172–183.

Burrows, M. T., Schoeman, D. S., and Buckley, L. B. 2011. Tha pace of shifting climate in marine and terrestrial ecosystems. Science, 334: 652–655.

Carvalho, G. R. and Pitcher, T. J. (eds.) 1994. Molecular genetics in fisheries. Reviews in Fish Biology and Fisheries: 4.

Cheung, W. W. L., Lam, V. W. Y., Sarmiento, J. L., Kearney. K., Watson, R., and Pauly, D. 2009. Projecting global marine biodiversity impacts under climate change scenarios. Fish and Fisheries, 10: 235–351.

deEyto, E., McGinnity, P., Consuegra, S., Coughlan, J., Tufto, J., Farrell, K., Megens, H.-J., Jordan, W., Cross T., and Stet, R.J.M. 2007. Natural selection acts on Atlantic salmon MHC variability in the wild. Proceedings of the Royal Society B, 274 : 861–869.

Ekblom, R., and Galindo, J. 2011. Applications of next generation sequencing in molecular ecology of non-model organisms. Heredity, 107: 1–15.

Excoffier, L., Hofer, T., and Foll, M. 2009. Detecting loci under selection in a hierarchically structured population. Heredity, 103: 285–298.

Foll, M. and Gaggiotti, O. E. 2008. A genome scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. Genetics, 180: 977–993.

Garvin, M. R., Saitoh, K. and Gharrett, A. J. 2010. Application of single nucleotide polymorphisms to non-model species: a technical review. Molecular Ecology Resources, 10: 915–934.

Gomez-Uchida, D., Seeb, J. E., Smith, M. J., Habicht, C., Quinn, T. P. and Seeb, L. W. 2011. Single nucleotide polymorphisms unravel hierarchical divergence and signatures of selection among Alaskan sockeye salmon (*Oncorhynchus nerka*) populations. BMC Evolutionary Biology, 11: 48.

Hansen, M. M., Olivieri, I., Waller, D. M. and Nielsen, E. E. 2012. Monitoring adaptive genetic responses to environmental change. Molecular Ecology, 21: 1311–1329.

Hauser, L. and Carvalho, G. R. 2008. Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. Fish and Fisheries, 9: 333–362.

Helyar, S. J., Hemmer-Hansen, J., Bekkevold, D., Taylor, M. I., Ogden, R., Limborg, M. T., Cariani, A., Maes, G. E., Diopere, E., Carvalho, G. R. and Nielsen, E. E. 2011. Application of SNPs for population genetics of non-model organism: new opportunities and challenges. Molecular Ecology Resources, 11(Suppl. 1): 123–136.

Hemmer-Hansen, J., Nielsen, E. E, Meldrup, D. and Mittelholzer, C. 2011. Identification of single nucleotide polymorphisms in candidate genes for growth and reproduction in a nonmodel organism; the Atlantic cod, *Gadus morhua*. Molecular Ecology Resources, 11 Supplement s1: 71–80.

Joost, S., Kalbermatten, M., and Bonin, A. 2008. Spatial analysis method (SAM): a software tool combining molecular and environmental data to identify candidate loci for selection. Molecular Ecology Resources, 8: 957–960.

Kerr, R. A. 2007. Global warming is changing the world. Science, 316: 188–190.

Landguth, E. L., Fedy, B. C., Oyler-McCance, S. J., Garey, A. L., Emel, S. L., Mumma, M., Wagner, Fortin, M. J., and Cushman, S. A. 2012. Effects of sample size, number of markers, and allelic richness on the detection of spatial genetic pattern. Molecular Ecology Resources, 12: 276–284.

Laval, G., Patin, E., Barreiro, L. B., and Quintana-Murci, L. 2010. Formulating a historical and demographic model of recent human evolution based on resequencing data from noncoding regions. PLoS ONE 5: e10284.

Le Corre, V., and Kremer, A. 2012. The genetic differentiation at quantitative trait loci under local adaptation. Molecular Ecology, 21: 1548–1566.

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

Makeenko, G. A., Volkov, A. A., Mugue, N. S., and Zelenina, D. A. 2012. Spatial-temporal variability of PanI locus allele frequencies in Atlantic cod from the open part of the Barents sea. Russian Journal of Genetics – in press.

Mäkinen, H. S., Cano, J. M. and Merilä, J. 2008. Identifying footprints of directional and balancing selection in marine and freshwater three-spined stickleback (*Gasterosteus aculeatus*) populations. Molecular Ecology, 17: 3565–3582.

Manel, S., Conord, C. and Després, L. 2009. Genome scan to assess the respective role of host-plant and environmental constraints on the adaptation of a widespread insect. BMC Evolutionary Biology, 9: 288.

Miller, K. M., Beacham, T.D., and Withler, R.E.. 2002. Use of Neutral and Adaptive Genetic Variation in the Management and Conservation of Pacific Salmon. ICES CM 2002/U:17.

Narum, S. R. and Hess, J. E. 2011. Comparison of Fst outlier tests for SNP loci under selection. Molecular Ecology Resources, 11: 184-194.

Narum, S. R., Campbell, N. R., Kozfkay, C. C., and Meyer, K. A. 2010. Adaptation of redband trout in desert and montane environments. Molecular Ecology, 19: 4622-4637.

Narum, S. R., Zendt, J. S., Frederiksen, C., Campbell, N., Matala, A., and Sharp, W. R. 2011. Candidate genetic markers associated with anadromy in *Oncorhynchus mykiss* of the Klickitat River. Transactions of the American Fisheries Society, 140: 843–854.

Nei, M., and Maruyama, T. 1975. Lewontin-Krakauer neutral genes. Genetics, 80: 396.

Nielsen, E. E., Cariani, A., Mac Aoidh, E., Maes, G. E., Milano, I., Ogden, R., Taylor, M. I., Hemmer-Hansen, J., Babbucci, M., Bargelloni, L., Bekkevold, D., Diopere, E., Grenfell, L., Helyar, S., Limborg, M.T., Martinsohn, J. T., McEwing, R., Panitz, F., Patarnello, T., Tinti, F., Van Houdt, J., Volckaert, F. A. M., Waples, R. S., FishPopTrace consortium, and Carvalho, G. R. 2012. Gene-associated markers provide tools for tackling illegal fishing and false eco-certification. Nature Communications- in press.

Nielsen, E. E., Hansen, M. M., and 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.

Nielsen, E. E., Hemmer-Hansen, J., Poulsen, N. A., Loeschcke, V., Moen, T., Johansen, T., Mittelholzer, C., Taranger, G. L., Ogden, R., and Carvalho, G. R. 2009. Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (*Gadus morhua*). BMC Evolutionary Biology, 9: 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.

Nielsen, R., Paul, J. S., Albrechtsen, A., and Song, Y. S. 2011. Genotype and SNP calling from next-generation sequencing data. Nature Reviews Genetics, 12: 443–451.

Nunes, V. L., Beaumont, M. A., Butlin, R. K and Paulo, O. S. 2011. Multiple approaches to detect outliers in a genome scan for selection in ocellated lizards (*Lacerta lepida*) along an environmental gradient. Molecular Ecology, 20: 193–205.

Pérez-Figueroa, A., García-Pereira, M. J., Saura, M., Rolán-Álvarez, E., and Caballero, A. 2010. Comparing three different methods to detect selective loci using dominant markers. Journal of Evolutionary Biology, 23: 2267-2276.

Poulsen, N. A., Hemmer-Hansen, J., Loeschcke, V., Carvalho, G. R. and Nielsen, E. E. 2011. Microgeographical population structure and adaptation in Atlantic cod *Gadus morhua*: spatio-temporal insights from gene-associated DNA markers. Marine Ecology Progress Series, 436: 231–243.

Richter-Boix, A., Quintela, M., Segelbacher, G., and Laurila, A. 2011. Genetic analysis of differentiation among breeding ponds reveals a candidate gene for local adaptation in *Rana arvalis*. Molecular Ecology, 20: 1582–1600.

Russello, M. A., Kirk, S. L., Frazer K. K., and Askey, P. J. 2012. Detection of outlier loci and their utility for fisheries management. Evolutionary Applications, 5: 39–52.

Ruzzante, D. E., Taggart, C. T., and Cook, D. 1998. A nuclear DNA basis for shelf- and bank-scale population structure in NW Atlantic cod (*Gadus morhua*): Labrador to Georges Bank. Molecular Ecology, 7: 1663–1680.

Ryman, N. and Utter, F. 1987. Population Genetics and Fishery Management. Seattle, WA: University of Washington Press.

Shaklee, J. B., Beacham, T. D., Seeb, L. and White, B. A. 1999. Managing fisheries using genetic data: case studies from four species of Pacific salmon. Fisheries Research, 43: 45–78.

Sick, K. 1965 Haemoglobin polymorphism of cod in the North Sea and the North Atlantic Ocean. Hereditas, 54: 19–48.

Waples, R. S. 1998. Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. Journal of Heredity, 89: 438–450.

Waples, R. S. and Gaggiotti, O. 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Molecular Ecology, 15: 1419–1439.

2.3 ToR c) Continuing assessment of the SNP-technology

Geir Dahle, Naiara Rodriguez-Ezpeleta

2.3.1 Preamble

Technological and analytical developments in Single Nucleotide Polymorphism analysis, with special focus of applications in fish and fisheries are discussed. Due to rapidly increasing numbers of applications using RAD-based approaches, this technology receives special attention.

2.3.2 State of the art

Nowadays different reduced-complexity approaches to genome-wide data are available at a reasonable price. Of the different types of analyses, the **R**estriction site **A**ssociated **D**NA (RAD) sequencing on the next-generation sequencers seems to be one of the most promising approaches to producing genomic data at population level in non-model organisms (Miller *et al*., 2007; Baird *et al* 2008). RAD sequencing is a method to identify thousands of informative SNPs using NGS (**N**ext **G**eneration **S**equencing) suitable for genotype-phenotype association mapping, scaffolding genome assemblies through linkage mapping, QTL analysis, hybridization and gene flow analysis, phylogeography, and population genetics. The method involves cutting the genome with one or more restriction enzymes and then sequencing the ends of the resulting fragments. The fragments from one individual are ligated to a modified adapter containing a unique identifying sequence. The fragments from many individuals can therefore be pooled together and sequenced in one operation. The resulting reads can be separated bioinformatically by identifying the unique identifying sequence at the start of each read. Using this method many SNPs can be identified in a single experiment. The number of studies using RAD sequencing to identifying SNPs are growing fast and in a variety of species; barley (Poland *et al*., 2011), artichoke (Scaglione *et al*., 2012), eggplants (Barchi *et al*., 2011), stickleback (Hohenlohne *et al*., 2010), trout (Hohenlohe *et al*., 2011).

2.3.3 High-throughput versus low-throughput

We discriminate between low or medium-throughput genotyping, i.e. a genome wide analysis on a low number of individuals, and high-throughput genotyping, i.e. genotyping selected SNPs in a large number of individuals. The latter uses the knowledge

acquired through a genome wide analysis (low-throughput) of a species, combined with annotation and validation of a large number of SNPs. From these data a panel of targeted SNPs is selected for any specific task. This method has been utilized in discriminating between farmed and wild salmon in Norway. Karlsson *et al* (2011) selected SNPs from a total panel of 7000. They found that the precision of assignment increased with the number of loci, and was close to 100 % using only the 60 highest-ranked loci of the 7000 on the original SNP-chip. This selection enables less expensive high-throughput genotyping.

2.3.4 Ascertainment bias

Ascertainment bias is a systematic geographically correlated reduction in estimates of diversity caused by the discovery process of any molecular marker, and it has been identified as a significant impediment to the widespread utility of large SNP panels (Helyar *et al*. 2011). Theoretically, ascertainment biases could influence any analysis or inference based on SNP allele frequency when the SNPs are isolated from a limited sample but applied in a larger geographical context. The ascertainment panel used for development of SNPs should therefore be carefully matched with the type of question their application warrants.

2.3.5 Validation of SNPs

SNPs are the most versatile class of functionally associated genetic markers, and allow assessment of the correlation between genetic and phenotypic variation. However, some of the SNP candidates identified in initial rounds of sequencing are likely to be sequencing "artefacts". Validation of the SNP markers is therefore necessary. Because of cost issues, validation is often based on a subset of the initially detected SNPs. Deep sequence coverage will detect SNPs with higher probability of being validated. Validation can be performed with various techniques, such as primer extension, hybridization, ligation, PCR amplification, and restriction enzyme digestion. Salmonid fishes have experienced several whole genome duplication events (Danzmann *et al* 2008). These duplicated genomes contain paralogous sequence variants (PSVs) which are readily mistaken for SNPs. A PSV is created when there is a base pair difference between the sequences of two paralogs, but the substitution does not segregate within either paralogue. Multisite variation (MSV) is another source of variation in polyploid genomes. MSVs are polymorphic and possibly informative; however, approaches to detect MSVs are limited, making analysis difficult. There has however been developed a software for identifying MSV on the salmon Illumina SNP array (Giskehaug *et al* 2010), increasing the number of usable polymorphic markers by 35%.

2.3.6 Bioinformatics (handling and analysis of data)

The preponderance of RAD sequencing and GBS data comes from the Illumina platform instruments such as the GAIIx or the HiSeq 2000. The uniqueness of the data generated by these instruments is the number of files created and their size. Even after pre-processing, the primary data produced in a single sequencing run can reach 1 terabyte, and, because the evolution of sequencing technologies is far from static, this number is expected to increase in a near future (Sexton, 2012). Thus, compared to previous eras in population genetics in which data generation was the limiting factor, the challenge now is not the data generation, but the storage, handling and analysis of the information obtained (Pennisi, 2011). Thus, as all other applications of these

new technologies, the SNP discovery and genotyping using next generation sequencing data requires unprecedented needs in infrastructure, software and personnel.

2.3.6.1 Infrastructure

High throughput sequencing datasets can range from occupying a few to hundreds of gigabytes per sample, implying high requirement of disk storage, memory and computing power for the downstream analyses, and often needing supercomputing centres or cluster facilities. A basic computational starting point consists on a server with 16 cores, 48 gigabytes of RAM and a 10 terabytes of disk space, but as sequencing throughput increases this needs to be scaled up through a larger server or a computer cluster. If the sequencing throughput does not justify such a large scale computing set up for a single lab and/or the home institution is not equipped with an appropriate IT infrastructure, other alternatives exist, such as the use of cloud-computing (e.g. the Elastic Compute Cloud from Amazon), which allows scientists to virtually rent both storage and processing power, by accessing servers as they need them. The downside of this option is that it requires moving data from researchers to 'the cloud' back and forth, which given file sizes, is not trivial (Rodríguez-Ezpeleta *et al*. 2012).

2.3.6.2 Software

There are hundreds of commercial and open-source softwares available that provide some aspect of analysing next generation sequencing data, but none addresses the usual concern when it comes to high-throughput data analysis, there is no 'Swiss army knife'-type software that covers all possible biological questions and combinations of experiment designs and data types (Rodríguez-Ezpeleta *et al*. 2012). In general, command line open-source customizable programs are preferred by expert bioinformaticians, but may be difficult to use for biologists that lack computing experience. In these cases, commercial easy-to-use software packages may be a solution. These tools have some lack of flexibility and configurability and can lead to the temptation of simply applying a preconfigured workflow ("black box") without fully considering or understanding whether each of the steps is appropriate for the particular project´s objectives and dataset (Paszkiewicz and Studholme, 2012). In all cases, a careful documentation of the analysis steps required for a given application, which often involves choosing among tens of available software and an extensive set of parameters for each step, is necessary. New algorithms are continuously emerging, adding increasing complexity to choosing to optimal analysis approach.

The software pipeline Stacks (Catchen *et al*., 2011) has recently been developed for building genetic maps and for identifying thousands of SNPs from RAD-Tag NGS data, usable in phylogeography and population structure studies. Although Stacks can generate several summary statistics and compute population genetic measures such as F_{is} within populations and F_{st} between populations, additional downstream analyses using as input data the identified SNPs are required to fully understand the population structure of the species of interest. Given the large amount of SNPs to be used development of new or adaptation of existing software to handle thousands of loci is required (Helyar *et al*., 2011).

2.3.6.3 Personnel

To fully capitalise on next generation sequencing data for SNP identification and genotyping, a considerable degree of bioinformatics expertise is required. The bioinformatics pipeline is however only one of the major bottlenecks of the field. There-

fore, along with technology considerations, it is additionally critical to have a well-trained cadre of bioinformatics specialists operating in close cooperation with the research group. The tasks of these bioinformatics experts include, but are not limited to, participating in the experimental design, optimizing data analysis pipelines in the parallel computing environment, automating bulk transfers of large volumes of data, filtering data and assigning biological significance, making decisions and intervening in the course of analysis, and interacting with investigators to suggests analysis strategies and methods. Although in high demand, few individuals that are proficient at both 'wet' laboratory based disciplines and 'dry' computational methods are currently available (Sexton, 2012).

2.3.6.4 Strategy

A common workflow of a population genomic study using SNPs detected by NGS is illustrated in Figure 1. Steps inside the dotted line are the ones that can be considered new in experiments using NGS in comparison to experiments using more standard techniques. A company founded by the inventors of RAD sequencing, Floragenex (http://www.floragenex.com/), makes this technique commercially available providing ready-to-use genotypes following the reception of purified DNA, and some institutions, such as the University of Edinburgh (http://genepool.bio.ed.ac.uk/), provide RAD sequencing services in a collaborative framework. In both cases, RAD-Tag library preparation, Illumina sequencing (including quality control assessment) and bioinformatics analyses to identify genotypes are performed. Subcontracting implies a tight cooperation with the expert wet-lab and bioinformatics team that will generate and analyse the data so that experiment design and interpretation of the data takes data generation and analysis issues into account. Another alternative is to perform some or all of these steps in-house. This allows more control over the type of RAD tags obtained (different species may require different restriction enzyme combinations, different coverage, etc.) and over the bioinformatics analysis to identify SNPs. Still, handling the large files obtained from next generation sequencing experiments in-house is only possible if the above requirements of infrastructure, software and personnel are fulfilled. Additionally, the downstream population genetic structure analyses will also require some degree of appropriate infrastructure and computational knowledge due to the large amount of data (thousands of SNPs) to be handled.

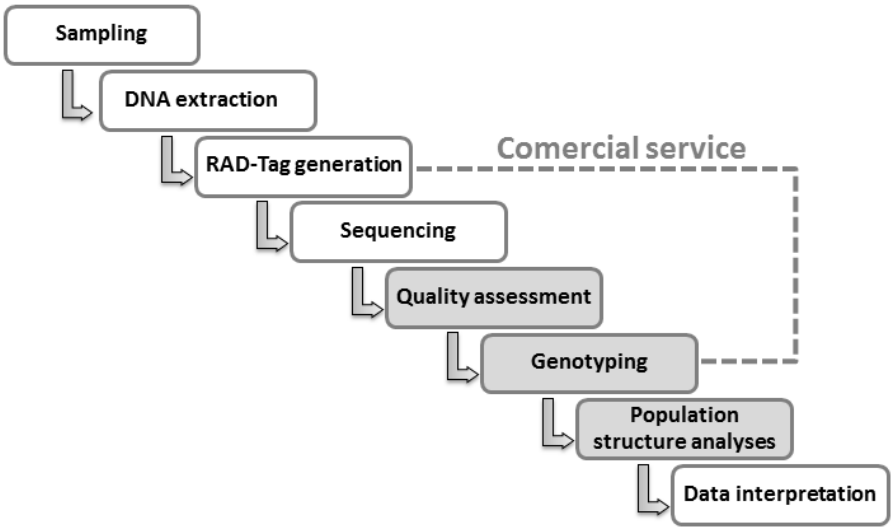

Figure 1: Workflow of a population genomic study using SNPs detected by next generation sequencing. Steps between the dotted line (all or a subset of them) can be subcontracted to companies such as Floragenex. Steps in grey require bioinformatics knowledge.

2.3.6.5 Cost issues

Compared to standard techniques, RAD sequencing is an extremely cost-effective approach for SNP discovery and genotyping (e.g. genotyping more than 10,000 SNPs in 200 individuals costs approximately 60,000 Euros if done commercially and even cheaper if done in house). Still, while the cost for SNP identification and genotyping is steadily decreasing as new advances in NGS technologies arose, this is still not the case for SNP validation, especially when large numbers of SNPs are concerned. Microchip based technology is still expensive for population based studies, and is currently restricted to commercial and/or large research centres. There are many low-to-medium throughput options for SNP genotyping available, with no apparent consensus as for the best system (i.e. comparison of performance in relation to technical genotyping issues: including genotyping errors and missing allele call; etc.). Technology continues to move fast and deciding on a particular platform may be difficult (but see Fluidigm system for an interesting option).

2.3.7 Recommendation

That SNP technology due to the constant technological developments is assessed and discussed by WGAGFM members on a continuous basis.

2.3.8 References

Baird, N., Etter, P., Atwood, T., Currey, M., Shiver, A., Lewis, Z., Selker, E., Cresko, W., and Johnson, E. 2008. Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. Plos One, 3(10):e3376.

Barchi, L., Lanteri, S., Portis, E., Acquadro, A., Valè, G., Toppino, L., and Rotino, G. L. 2011. Identification of SNP and SSR markers in eggplant using RAD tag sequencing. BMC Genomics, 12:304 doi:10.1186/1471-2164-12-304.

Catchen, J., Amores, A., Hohenlohe, P., Cresko, W., and Postlethwait, J. 2011. Stacks: building and genotyping loci de novo from short-read sequences. G3: Genes, Genomes, Genetics: 1: 171–182.

Danzmann, R. G., *et al*. 2008. Distribution of ancestral proto-Actinopterygian chromosome arms within the genomes of 4Rderivative salmonid fishes (Rainbow trout and Atlantic salmon), BMC Genomics, 9, 557.

Gidskehaug, L., Kent, M., Hayes, B. J., and Lien, S. 2010. Genotype calling and mapping of ultisite variants using an Atlantic salmon iSelect SNP-array. Bioinformatics 27(3): doi:10.1093/bioinformatics/btq673

Helyar, S. J., Hemmer-Hansen, J., Bekkevold, D., Taylor, M. I., Ogden, R., Limborg, M. T., Cariani, A., Maes, G. E., Diopere, E., Garvalho, G. R., and Nielsen, E. E. 2011. Application of SNPs for population genetics of non-model organisms: new opportunities and challenges. Molecular Ecology Resources, 11 (Suppl 1): 1–14.

Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A., Cresko, W. A. 2010. Population Genomics of Parallel Adaptation in Threespine Stickleback using Sequenced RAD Tags. PLoS Genetics 6(2): e1000862. doi:10.1371/journal.pgen.1000862.

Hohenlohe, P. A., Amish, S. J., Catchen, J. M., Allendorf, F. W., Luikart, G. 2011. Next-generation RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and westslope cutthroat trout. Molecular Ecology Resources, 11: 117–122.

Karlsson, S., Moen, T., Lien, S., Glover, K. A., and Hindar, K. 2011. Generic genetic differences between farmed and wild Atlantic salmon identified from a 7K SNP-chip. Molecular Ecology Resources (2011) 11 (Suppl. 1), 247–253.

Miller, M., Dunham, J., Amores, A., Cresko, W., and Johnson, E. 2007. Rapid and costeffective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. Genome Res, 17: 240–248.

Paszkiewicz and Studholme. 2012. High-througput sequencing data analysis software: current state and future developments. Ed. by Rodríguez-Ezpeleta *et al*. *In* Bioinformatics for High Throughput Sequencing, Springer: pp 231–248.

Pennisi, E. 2011. Human genome 10th anniversary. Will computers crash genomics? Science: 331: 666–668.

Poland, J. A., Brown, P. J., Sorrells, M. E., Jannink, J.-L. 2012. Development of High-Density Genetic Maps for Barley and Wheat Using a Novel Two-Enzyme Genotyping-by-Sequencing Approach. PLoS ONE 7(2):e32253.doi:10.1371/journal.pone.0032253.

Rodríguez-Ezpeleta N, and Aransay, A. M. 2012. Introduction. Ed. By Rodríguez-Ezpeleta *et al*. *In* Bioinformatics for High Throughput Sequencing, Springer: pp 1–9.

Scaglione, D., Acquadro, A., Portis, E., Tirone, M., Knapp, S. J., and Lanteri, S. 2012. RAD tag sequencing as a source of SNP markers in *Cynara cardunculus* L. BMC Genomics, 13:3 doi:10.1186/1471-2164-13-3

Sexton, D. 2012. Computational infrastructure and basic data analysis for high-throughput sequencing. Ed. by Rodríguez-Ezpeleta *et al*. *In* Bioinformatics for High Throughput Sequencing, Springer: pp 55–65.

2.4 ToR d) Contribute to the development of the scientific understanding of sustainability of aquaculture by collaborating with WGMASC and WGEIM and planning for the ASC 2012 theme session on "Genetic impact of aquaculture on wild populations"

All WGAGFM participants attending meeting

2.4.1 Context

WGAGFM has a long history of providing advice to ICES on genetic aspects of aquaculture, both with regards to use of molecular approaches to set up and optimize breeding designs, as well as assessments of potential effects of aquaculture on ecosystems and fish and shellfish resources. Steps have been taken to develop collaboration between WGAGFM and the new EG "Sustainability of aquaculture", and the scope for the WGAGFM chair and potentially other interested members of WGAGFM to participate in a meeting with the new WG during the next ASC in Bergen is being examined. WGAGFM has identified several areas of potential synergies emanating from collaboration between EGs. These include work on 1) the identification of mariculture issues and advisory needs with respect to status and sustainability of living marine resources in the face of mariculture, and 2) application of genomic marker based approaches to optimisation of breeding programs (e.g. using inference about intra-specific biodiversity associated with traits of interest), both in established and novel mariculture species. Other, yet unidentified, areas of potential synergies may also apply. At time of writing formal collaboration with the EG (integrating WGMASC and WGEIM) has not been established. Yet, WGAGFM agree that interaction is likely to be beneficial, while at the same time it requires careful consideration how such interaction could and should take place. Resources for contributing to WGs and attending EG meetings are contributed by individual member countries' institutions' own budgets. This is a significant limiting factor for the extent to which members are able to contribute to multiple EGs and has to be taken into consideration, when drawing up specific plans to facilitate and ensure interaction. One way of integrating activities could be for the two groups to exchange comments on respective relevant ToR resolutions and recommendations. Shared ToR could also be developed. WGAGFM suggest producing a commented list of areas of expertise held within WGAGFM that should be of relevance to work carried out by the "Sustainability of aquaculture" EG. The list could subsequently be submitted for the attention of the "Sustainability of aquaculture" EG. The list should be updated annually by WGAGFM.

WGAGFM have not collaborated with WGMASC/WGEIM on the planning of the ASC 2012 theme session on "Genetic impact of aquaculture on wild populations". However, the scope for the theme session falls within the field of expertise of WGAGFM, and individual WGAGFM members have submitted abstracts to the session.

2.4.2 WGAGFM recommends

- That EG chairs (of the integrated WGMASC/WGEIM and WGAGFM) explore the scope for interaction during the ASC in Bergen 2012 and hereunder discuss the scope for exchanging ToR resolutions with the aim to provide comments on respective ToR, where appropriate.

- That ICES SCICOM SSGHIE compile and disseminate a portfolio of expertise from individual EGs to facilitate future collaborations.

2.5 Term of Reference e) Evaluate potential for collaboration with other EGs in relation to the ICES Science Plan and report on how such cooperation has been achieved in practical terms (e.g. joint meetings, back-to-back meetings, communication between EG chairs, having representatives from own EG attend other EG meetings).

All WGAGFM participants attending meeting

2.5.1 Context

WGAGFM discussed the potential for collaboration and agree that there should be scope for collaboration between WGAGFM and several other EGs, both under the SSGHIE as well as between WGAGFM and EGs under other SCICOM steering groups. It was acknowledged that it is practically difficult for individual members to engage in regular interaction with all relevant EGs (apart from what is accomplished via *ad hoc* discussions between members from different EGs), but also that the aim should be for the widest possible levels of collaboration. In some cases, joint meetings could contribute to the collaboration, but it is also important that annual meetings allow WGAGFM members time to discuss specialized issues, which may not be attainable to the same extent during multi-disciplinary meetings. Back-to-back meetings may in some cases be a useful approach but are at the same time practically difficult to plan and require increased resource allocation from participating members. Having representatives from own EG attend other EG meetings could be useful, but also requires close coordination of the respective EG activities. It is envisaged that the advertised shift to multi-annual ToR may improve the integration of work among EGs, as it gives EG chairs and members the opportunity to follow respective EG work developments over a longer time frame, hence enabling better scope for planning joint activities. To aid visibility of the issues discussed and capacity held within individual EGs, portfolio of expertise could be compiled and disseminated for each EG.

2.5.2 WGAGFM recommends

That it remains an aim to develop collaborations and sharing of knowledge between EGs, both between those under SSGHIE and between WGAGFM and EGs under other steering groups.

That ICES SCICOM scientific steering groups compile and disseminate a portfolio of expertise from individual EGs to facilitate future collaborations.

2.6 ToR f) Contribute to the improved integration of genetic information into fisheries management under a reformed Common Fisheries Policy the Marine Strategy Framework Directive and the new Data Collection Framework 2014–2020

Jann Martinsohn, Reinhold Hanel, Tom Cross and Dorte Bekkevold

2.6.1 Abstract

While genetic approaches in a number of examples have proven their value for the management of marine living resources and conservation measures, the inclusion of

genetic data for such purposes is not yet routinely applied, and integration of genetic information in marine fishery management schemes has been slow.

The ICES WGAGFM has repeatedly and consistently advocated the improved integration of genetic data and information into fisheries management and conservation, also under the Common Fisheries Policy (CFP) remit. For example the WGAGFM recommended on various occasions the building of a "metadatabase for fish(eries) genetic data compilation". Additionally in its 2011 report the WGAGFM suggested to explore opportunities for the integration of genetic data into fisheries management resulting from the European Union Data Collection Framework Regulation. This has also been discussed in March 2012 during an Expert Working Group meeting of the Scientific, Technical and Economic Committee for Fisheries (STECF), together with ICES one of the major advisory bodies to the European Commission, during which ICES was also represented.

This ToR intends to reflect on opportunities arising from current EU policy developments, such as the implementation of the Marine Strategy Framework Directive, the ongoing CFP reform and the envisioned new DCF (2014–2020), for a better integration of genetic approaches and concepts into fisheries management and conservation measures, and to come forward with tangible recommendations to reach this goal.

2.6.2 Rationale and background

Genetic data can provide valuable information for fisheries management and conservation measures, for example spatial and temporal trends in exploited fish stocks. While this has been shown in numerous scientific studies, demonstrated through the application of genetic approaches in fisheries management scenarios (Hauser, 2008 and see below), and acknowledged by eminent scientific advisory bodies (STECF, 2011) genetic analysis for fisheries management and conservation is far from being considered routinely. There are various reasons for this reluctant uptake of genetic information (discussed in Waples, 2008). Generally the current scientific infrastructure is unfavourable: With the exception of genetic species identification ("DNA-Barcoding" Ward, 2009), at present the generation of genetic data and information on marine organisms is generally restricted to specific research projects with limited funding and timeframes. This, and the lack of robust and central data collection structures for genetic and genomic data (Verspoor, 2010), relevant for management of living marine resources is currently greatly impeding the uptake of genetic information.

The use of genetic and genomic information for fisheries and aquaculture management as well as conservation is regularly addressed by ICES through the WGAGFM, and has also been subject during the 36th Plenary Meeting of the STECF (PLEN-11-01) where outcomes of the FP7 funded project FishPopTrace (http://fishpoptrace.jrc.europa.eu), which performed extensive population genetic analysis on marine fish, including on historical samples, were presented and documented in the resulting report {STECF, 2011}. Also in March 2012, during the STECF – EWG 12- 01 Meeting on the future DCF Multi Annual Plan (MAP) 2014-2020, opportunities for fisheries management under the CFP remit, provided by modern genetic and genomic analytical approaches were presented as well as emerging challenges concerning the collection of genetic data. It was discussed whether and how data resulting from genetic and genomic analysis and monitoring with relevance for fisheries management under the CFP remit, could be collected and included under the DCF remit (STECF Report EWG 12- 01 in preparation).

The currently ongoing developments in the EU marine-policy framework, such as the implementation of the Marine Strategy Framework Directive (MSFD; European Commission, 2008), the overhaul of the Common Fisheries Policy (CFP; European Commission, 2009) in combination with the creation of the European Maritime and Fisheries Fund (EMFF; European Commission, 2011) and the envisioned new Data Collection Framework 2014-2020 (STECF Report EWG 12- 01 in preparation) provide an excellent opportunity to re-launch the endeavour and develop sound strategies in order to improve the integration of genetic information into fisheries management, including aquaculture activity and also conservation. In the following, EU policies and legislation which can be supported by the use of genetic information are briefly introduced. Subsequently questions critical to sustainable fisheries management and conservation, which can be addressed by genetic approaches, are presented, followed by a brief critical reflection on opportunities and shortcomings. Finally a series of recommendations designed to improve the inclusion of genetic information in fisheries management, aquaculture activities and conservation are put forward.

2.6.3 Relevant current policy & RTD initiatives

The Marine Strategy Framework Directive (Directive 2008/56/EC)

The Marine Strategy Framework Directive (European Parliament, 2008) establishes a common framework and objectives for the protection and conservation of the marine environment. In order to achieve these common objectives, EU Member States (MS) will draw up and implement coherent management plans in each region, and subsequently monitor their application. By 15 October 2012, Member States are to complete and report to the European Commission their Initial Assessment (Article 8), determination of Good Environmental Status (GES) (Article 9) and establishment of environmental targets (Article 10).

ICES has provided scientific support to the European Commission as a background for the preparation of the Commission Decision on criteria and methodological standards on good environmental status of marine waters (European Commission, 2010). The resulting eight reports have been prepared by groups of independent experts coordinated by the European Commission Joint Research Centre (JRC) and ICES.

EU Member States must determine the "ecological status" of their waters on the basis of 11 descriptors, which are further specified by a list of attributes, criteria and indicators. **Genetic information** is explicitly mentioned as a substantial indicator: Population genetic structure of key species is defined as an indicator for Descriptor 1 (Biodiversity), and could also be used as an indicator in Descriptor 3 (Commercially Exploited Fish). In addition, genetically distinct forms of native species should be assessed. Such an assessment of genetically distinct forms also applies to Descriptor 2 (Non-Indigenous Species).

Once the process of selection of indicators nears completion there is a need for harmonization of assessment and reporting between MS. It may not be appropriate to apply indicators in the same way within and between regions. However, the raw data obtained function as the fundamental building blocks for assessment. These data need to be compatible, reproducible and quality assured on a pan-European scale. This means that sampling and sample processing must follow internationally agreed procedures, independent of subsequent data analysis. Within several Descriptors, international standard guidelines may exist for some, if not all, of the selected indicators (for example contaminants). For other Descriptors, such as Biodiversity, Non-indigenous species, Food webs and Sea-floor integrity, there is likely to be a paucity

of technical guidelines. This specifically applies to indicators based on genetic information. However, such indicators are essential and provide adequate tools to tackle a variety of questions on the environmental status of marine ecosystems. Priority should be given to matching the emerging needs of genetic information on a species, population and individual level with the availability of internationally approved technical guidelines/methodological standards. Where there is a lack of such guidelines, measures should be taken to ensure these are developed, within the timeframe relevant to the MSFD assessment process.

The Common Fisheries Policy Reform and the European Maritime and Fisheries Fund

The Common Fisheries Policy (CFP) is currently under reform, aiming to bring fish stocks back to sustainable levels. In its proposal for a regulation on the Common Fisheries Policy (European Commission, 2011), the European Commission emphasizes the need for a fundamental overhaul and identifies a series of measures to improve the current situation. Examples are to ensure productivity of fish stocks and to maximise long-term yield, multi-annual plans governed by ecosystem approach, a ban on discards, new marketing standards and clearer labelling, regionalisation of fisheries management, a better framework for aquaculture, up-to-date information on state of living marine resources.

The CFP reform is accompanied by the introduction of a new financial instrument, the European Maritime and Fisheries Fund (EMFF; European Commission, 2011) which will replace the current European Fisheries Fund (EFF). The EMFF, as part of the EU's multi-annual financial framework for 2014-2020, is designed such that it can help deliver the objectives of the reformed Common Fisheries Policy (CFP) but it also supports the implementation of the EU Integrated Maritime Policy (IMP; European Commission, 2007). Importantly it also foresees measures for data collection and scientific advice. In its accompanying MEMO (European Commission, 2011, the European Commission emphasizes that scientists and researchers in the fields of marine environment, climate change, coastal protection, *etc*. should benefit from the EMFF. This should be further explored with respect to potential opportunities for the application of genetic approaches and concepts to fisheries management and conservation

As shown below by a list of fisheries and conservation relevant questions, genetic and genomic approaches can very well deliver great support to both the MSFD and the CFP. An opportunity will be missed if scientists and stakeholders do not find ways to integrate genetic information better and on a broader scale into fisheries management and conservation measures.

2.6.4 Marine Knowledge 2020 and the Data Collection Framework 2014–2020

It is generally acknowledged that marine data is generated in a highly fragmented way by many stakeholders from diverse disciplines. This leads to data dispersal and loss, which does greatly impede a coherent approach to marine research, causes redundant work on marine-related topics and has also a significant economic impact (European Commission, 2010). Also the field of fisheries genetics suffers from the lack of central databases compiling genetic data on marine organisms, which contributes to the slow uptake of genetic information into fisheries management and conservation (Verspoor, 2010).

Marine Knowledge 2020 is an EU initiative bringing together marine data from different sources with the aim of helping researchers, but also the industry and public authorities, to make more effective use of datasets to improve our understanding of

the marine realm and to develop new products and services. A key aspect is to compile data from multiple sources to serve a wide array of end users. Following this approach, users can benefit by access to data that was collected for a certain purpose or in the context of a specific project but which is reusable for other purposes. The different users will process the data and transform it into information and knowledge in different ways for their particular aims. The same line of argument was also used by the WGAGFM group when advocating the creation of a meta-database for fish and shellfish genetic data (Verspoor, 2010). The EU takes several measures to pursue this goal, such as supporting Member States' efforts on marine observations & data collection, development and implementation of EU policies to improve accessibility of public data & information (e.g. INSPIRE; European Parliament, 2007) and ultimately to transform the current fragmented arrangement of systems into one interconnected and interoperable structure. Important pillars of Marine Knowledge 2020 are the Global Monitoring for Environment and Security (GMES; European Commission, 2009), the European Marine Observation and Data Network (EMODnet; DG MARE, 2010) and the EU Fisheries Data Collection Framework (DCF; European Council, 2008). In the long run it appears that particularly EMODnet and the DCF could incorporate genetic data on marine organisms and commercially exploited fish respectively. The feasibility of including fish(eries) genetic data in the DCF has been discussed in March 2012 during an Expert Working Group meeting of the Scientific, Technical and Economic Committee for Fisheries (STECF), during which ICES was also represented. It was agreed that genetic and genomic information can help to elucidate fisheries management relevant questions and provide support to the Common Fisheries Policy. It was also concluded that the current DCF already provides for the possibility to collect data sets which are currently not routinely included such as is the case for genetic/genomic data under its remit, on specific request by end-users. To this end, also studies addressing the use of any specific analytical approaches and technologies, species-specific issues and regions can be carried out and co-funded under the DCF provisions.

Acknowledging the need to be able to accommodate opportunities arising from technology advancements and new data needs, the STECF DCF Expert Working Group endorsed to maintain this level of flexibility, also in the future DCF 2014-2020 (STECF Report EWG 12- 01 in preparation).

With respect to fisheries genetic data an opportunity might also emerge from FishFrame, a web based data-warehouse application (www.FishFrame.org) linking between stored nationally raw data and the aggregated data used in the assessment process, which is currently hosted by ICES and co-financed by the European Commission. While the variables included in FishFrame should satisfy all data needs for most assessment models including fishery based assessment models (Degel, 2006), currently genetic data is not included. It should be explored whether the integration of genetic data in FishFrame is feasible.

2.6.5 The CFP Control and IUU Regulations

Illegal, Unregulated and Unreported (IUU) fishing and fraud along the supply chain (mislabelling of fish products) is vastly contributing to the predominant overexploitation of many fish stocks worldwide. These illegal activities have severe adverse effects, as they undermine sustainable fisheries, cause destruction of marine ecosystems, obstruct socioeconomic development, and impede consumer information and protection (Martinsohn, 2011).

The European Union has recently taken initiative to curb illegal activities in the fisheries sector and developed two major and complementing legal instruments: in January 2010, Council regulation (EC) No 1005/2008, - the 'IUU regulation' (European Council, 2008), entered into force, and in November 2009, Council regulation (EC) No 1224/2009 established a new Community control system (European Council, 2009).

Both regulations place emphasis on detailed catch documentation and traceability for fishery products 'from ocean to fork', that is, covering all stages of the supply chain from catch, to landing, transport, processing, and the markets. Traceability is generally acknowledged as being a highly powerful tool in support of monitoring, control and enforcement in the fisheries sector. However, currently it is mainly based on certificates accompanying goods, and labelling of products, both measures which are vulnerable to falsification. Here genetic analysis can greatly help to authenticate fish and fish products, even if highly processed (Martinsohn, 2011). It is noteworthy that the CFP control regulation explicitly refers to 'genetic analysis' in Article 13 - Modern Technologies (European Council, 2009).

2.6.6 Genetic applications in support of fisheries management & conservation

Apart from fisheries control and enforcement (see above) a number of other issues or questions of relevance to fisheries management and conservation can be addressed using genetic approaches. These include 1) monitoring of biodiversity (among and within species), 2) identifying biologically relevant management units and identifying origin of individuals and mixed samples ('genetic stock identification' GSI, and 'mixed-stock analysis' MSA), 3) determining exploitation rates of individual populations and management units (monitored with GSI/MSA), 4) monitoring effects and scale of environmental change, 5) monitoring trophic interactions, e.g. from bar coding analysis of stomach contents and environmental samples, 6) genetic detection of stock sex- and age ratios, 7) monitoring of population sizes, 8) monitoring presence of escapees from aquaculture, and their interaction with wild stocks, 9) monitoring effects of stock-augmentation through releases/stocking. For the time being, some of these applications are more readily available outside an academic environment than others. The scope of this report is not to provide a review of applications and is far from being exhaustive; however a few illustrative examples of applications that are already available to management (albeit not in all cases implemented) are given here.

2.6.6.1 Example: Atlantic Salmon Genetic Stock Identification

The Irish Atlantic salmon genetic stock identification database is constructed using genotypic data from approximately 11 000 individual juvenile salmon screened for variation at 15 microsatellite loci. Samples of juvenile stages fry and parr were collected from 117 rivers (out of 140 recognised salmon rivers in Ireland) from approximately 240 sampling locations, starting with locations of highest productivity. Approximately 25% of the samples represent temporal replicates. The sampling of juveniles was conducted by electrofishing with GIS aided targeting (largely conducted by a Central Fisheries Board (now Inland Fisheries Ireland) team led by Drs P. Gargan and W. Roche) which identified major spawning areas in each river. 99% of total Irish salmon production is represented by the areas sampled. Genotypes of Irish ranched and farmed strains are also incorporated. The data are used from a management perspective to associate rivers and/or populations into management units based on levels of differentiation. We also endeavour to identify rivers/populations which may have special conservation importance, based on levels of allelic diversity and

differentiation from neighbouring rivers. We also examine the extent of temporal stability in small rivers particularly and the level of differentiation of these from neighbouring large populations, to determine whether these small rivers form parts of single management units or in contrast, show high levels of population genetic integrity. Most importantly the data are used as a baseline for genetic stock identification (GSI) in Irish commercial and recreational fisheries. This baseline has been used to assess mixed stock fisheries in both inshore and offshore locations around Ireland by the Marine Institute and following special government Ministerial requests (see references below). Because of the novel method of targeted sampling, the extent of population coverage and the intensity of the genetic screening, the Irish database is recognised as a vital management tool, that is being emulated in other Atlantic salmon producing countries. (Source: Prof. T Cross & Dr J. Coughlan; University College Cork IE).

2.6.6.2 Example: Atlantic Herring Genetic Stock Identification

The EU FP7 funded project FishPopTrace (http://fishpoptrace.jrc.europa.eu) have established a genetic stock identification database using single nucleotide polymorphism (SNP) information obtained for individuals collected from 20 spawning locations in the Northeast Atlantic, focussing on major stocks in the area. Major population components were identified (Limborg *et al*. 2012) and the SNP marker data were used to construct genetic assays ("minimum assays with maximal power"; Nielsen *et al*., 2012) to identify populations of fish to a forensic level of validation. Results demonstrated how application of gene-associated markers will likely revolutionise origin assignment. It was e.g. possible to correctly assign on average 99% of a sample of herring between two areas (W North Sea and Norway) between which, at best, very weak levels of genetic differentiation was previously reported, which previously prevented successful GSI using genetic markers such as microsatellite DNA. It was moreover demonstrated that the method could be used throughout the food supply chain, representing tissue samples ranging from freshly caught to highly processed fish (e.g. smoked, pickled). Improving this technology and spearheading its application in Europe, such new tools provide a new way forward for managing fish resources, as they will revolutionise origin assignment and become a highly valuable tool, also for fighting illegal fishing and mislabelling. (Source: Dr. D. Bekkevold; DTU Aqua DK & The FishPopTrace Consortium).

2.6.6.3 Example: Norwegian Coastal Cod Management

In the northeast, Atlantic cod is divided into two main management units, namely northeast Arctic cod and coastal cod. In Norway both groups can co-occur within the same fjords. While Arctic cod and coastal cod have traditionally been identified by otolith classification, meanwhile genetic analysis using microsatellites and the Pan I locus, which is under selection, is used to distinguish between both groups and to guide management decisions. (Source: Dr. G. Dahle; Institute of Marine Research Bergen, NO)

2.6.7 Opportunities and Challenges

It is meanwhile generally acknowledged that genetic analysis can provide valuable support to fisheries management and also contribute to scientific advice feeding into fisheries policy making and governance. A major opportunity arises through the swift advancement in the field of genetics and genomics and the rapid drop in costs of genetic analysis, particularly DNA sequencing, which also greatly benefits the field

of fish(eries) genetics. This has been discussed extensively in recent literature (Waples, 2008; Hauser, 2008; Martinsohn, 2011).

When applying genetic approaches, such as Genetic Stock Identification (GSI), for fisheries management purposes, the main effort lies in establishing and maintaining baseline data that will allow biologically significant inference in space and time. However, once a set of markers that allow assessment on the desired spatial scale is determined, the issue of database maintenance by temporal updates on baseline genetic information (e.g. on a 5-10 year basis) is no different from requirements for other types of data. This should open a venue to tap into existing databases or data collection schemes, such as the DCF, which already compile and hosting fisheries relevant data and information.

However, as stated above, currently genetic information on marine fish is still to a large degree exclusively accommodated in the research and academic environment. Therefore, a dialogue on how to best take advantage of genetic data and information, how and where to collect/compile/disseminate genetic data, and also how to finance such an endeavour, is needed. Only such a multi-disciplinary and multi-stakeholder approach will guarantee that genetic data and information will be available to fisheries management, conservation and scientific advice such that the greatest possible benefit is created. Quite obviously the ICES is very well positioned to catalyse such a dialogue and to ensure its efficiency.

2.6.8 Conclusion

The WGAGFM agrees that the lack of genetic information integration in fisheries management and conservation poses a substantial underuse of available methods that needs to be highlighted at member state and EU levels. The compilation of available genetic information in support of EU policy should be facilitated. WGAGFM recommend that that member states, e.g., under the remit of MSFD and CFP, build the necessary capacity and infrastructure to enable genetic monitoring at national and regional levels. To reach this goal national governmental fisheries agencies should ensure the availability of genetic core competence (e.g. staff with expertise in population genetics and genomics). Successful implementation in management and fisheries control and enforcement, requires national and/or regional databases, which could be initiated and maintained under the DCF. Acknowledging the importance of the subject, the WGAFGFM aims to continue discussions on how to implement genetic data in management and enforcement across fish and shellfish species in years to come.

2.6.9 Recommendations

The ICES WGAGFM recommends that:

4) That SCICOM contribute to a process to support that member states build the necessary capacity and infrastructure to enable genetic monitoring at national and regional levels under the remit of the Marine Strategic Framework Directive and the Common Fisheries Policy.

5) The ICES secretariat initiates and organises a workshop defining guidelines for the integration of genetic data and information in the Data Collection Framework in support of the implementation of EU policies such as the Marine Strategic Framework Directive (MSFD) and the reformed Common Fisheries Policy (CFP). Participants should include genetic experts as well as experts involved in the coordination and implementation of the MSFD and CFP.

6) The availability and feasibility of using national fisheries databases (such as the ICES hosted FishFrame) for the hosting of genetic data is explored.

2.6.10 References

Degel, H., and Jansen, T. 2006. FishFrame: Fisheries and stock assessment data framework. ICES Report. I. C. f. t. E. o. t. S. (ICES). Copenhagen, International Council for the Exploration of the Sea (ICES). ICES CM 2006/M:02: 22.

DG MARE – European Commission. (2010). Marine Knowledge 2020. Brussels: 28. ISBN 978 92 79 16449 1.

European Commission. 2007. Communication from the Commission: An Integrated Maritime Policy for the European Union. Brussels, European Commission.

European Commission. 2009. Green Paper: Reform of the Common Fisheries Policy. COM(2009)163 final: 1.

European Commission. 2009. Proposal for a Regulation of the European parliament and of the Council on the European Earth observation programme (GMES) and its initial operations (2011–2013). Brussels: 42.

European Commission. 2010. Commission Decision of 1 September 2010 on criteria and methodological standards on good environmental status of marine waters (notified under document C(2010) 5956) Text with EEA relevance. 477/2010/EU. Brussels, Official Journal of the European L 232. 477/2010/EU: 11.

European Commission. 2010. Communication from the Commission to the European Parliament and the Council: Marine knowledge 2020 - marine data and observation for smart and sustainable growth. Brussels, European Commission: 14.

European Commission. 2011. Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the European Maritime and Fisheries Fund [repealing Council Regulation (EC) No 1198/2006 and Council Regulation(EC) No 861/2006 and Council Regulation No XXX/2011 on integrated maritime policy. Brussels.

European Commission. 2011. Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the Common Fisheries Policy. Brussels.

European Commission. 2011. Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the European Maritime and Fisheries Fund [repealing Council Regulation (EC) No 1198/2006 and Council Regulation(EC) No 861/2006 and Council Regulation No XXX/2011 on integrated maritime policy. Brussels.

European Commission. 2011. Questions and answers on the new European Maritime and Fisheries Fund (EMFF). MEMO. Brussels.

European Council. 2008. COUNCIL REGULATION (EC) No 1005/2008 of 29 September 2008 establishing a Community system to prevent, deter and eliminate illegal, unreported and unregulated fishing, amending Regulations (EEC) No 2847/93, (EC) No 1936/2001 and (EC) No 601/2004 and repealing Regulations (EC) No 1093/94 and (EC) No 1447/1999. Official Journal of the European Union L 286: 1.

European Council. 2008. Council Regulation(EC) No 199/2008 of 25 February 2008 concerning the establishment of a Community framework for the collection, management and use of data in the fisheries sector and support for scientific advice regarding the Common Fisheries Policy. Official Journal of the European Union L 60/1: 1.

European Council. 2009. COUNCIL REGULATION (EC) No 1224/2009 of 20 November 2009 establishing a Community control system for ensuring compliance with the rules of the common fisheries policy, amending Regulations (EC) No 847/96, (EC) No 2371/2002, (EC) No 811/2004, (EC) No 768/2005, (EC) No 2115/2005, (EC) No 2166/2005, (EC) No 388/2006, (EC) No 509/2007, (EC) No 676/2007, (EC) No 1098/2007, (EC) No 1300/2008, (EC) No

1342/2008 and repealing Regulations (EEC) No 2847/93, (EC) No 1627/94 and (EC) No 1966/2006. Official Journal of the European Union L 343/1(22/12/2009): 50.

European Parliament and Council. 2007. Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), Official Journal of the European Union. L 108:1.

European Parliament and Council. 2008. DIRECTIVE 2008/56/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 17 June 2008 establishing a framework for community action in the field of marine environmental policy (Marine Strategy Framework Directive). Official Journal of the European Union L(L164): 19.

Hauser, L., and Carvalho, G. R. 2008. Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. Fish and Fisheries, 9(4): 333.

Martinsohn, J. T. 2011. Deterring Illegal Activities in the Fisheries Sector: genetics, genomics, chemistry and forensics to fight IUU fishing and in support of fish product traceability. JRC Reference Reports. Luxemburg, European Commission: 72. EUR 24394 EN. ISBN 978-92-79-15905-3.

STECF. 2011. 36th Plenary Meeting Report of the Scientific, Technical and Economic Committee for Fisheries (PLEN-11-01). JRC Scientific and Technical Reports. Eds. J. Casey and H. Doerner. Ispra, European Commission.

Verspoor, E., A. Zanzi, *et al*. 2010. ToR d) Pursuing the establishment of a meta-database cataloguing molecular data in the field of fish and shellfish population genetics. Report of the Working Group on the Application of Genetics in Fisheries and Mariculture (WGAGFM) 2010. Cork, International Council for the Exploration of the Sea (ICES).

Waples, R. S., Punt, A. E. *et al*. (2008). "Integrating genetic data into management of marine resources: how can we do it better?" Fish and Fisheries, 9(4): 423.

Ward, R. D., Hanner, R. *et al*. 2009. "The campaign to DNA barcode all fishes, FISH-BOL." Journal of fish biology, 74(2): 329.

Annex 1: List of participants

Name	Address	Phone/Fax	Email
Naiara Rodríguez-Ezpeleta	Marine Research Division, Azti Technalia, Txatxarramendi ugartea z/g E-48395 Sukarrieta, Bizkaia	Phone : +34 667 174 514	nrodriguez@azti.es
Johan Robbens	Chemical Monitoring & Product technology Instituut voor Landbouw- en Visserij Onderzoek ILVO Visserij Ankerstraat 1 B-8400 Oostende, Belgium	Phone: + 32 (0)59 56.98.50 Fax: + 32 (0)59 33.06.29	johan.robbens@ilvo.vlaanderen.be
Jann Th. Martinsohn	Joint Research Centre (JRC) Institute for the Protection and Security of the Citizen (IPSC) JRC.G.4 - Maritime Affairs Via Enrico Fermi 2749 (TP 051) I-21027 Ispra (Va), Italy	Phone: +39 0332 78 6567 Fax: +39 0332 78 9658	jann.martinsohn@jrc.ec.europa.eu
Niklas Tysklind	University of Wales, Bangor School of Biological Sciences Bangor LL57 2DG Gwynedd UK		bssco8@bangor.ac.uk
Tom Cross	Aquaculture & Fisheries Development Centre, Environmental Research Institute, University College Cork, Cooperage Building, Distillery Fields, North Mall, Cork, Ireland	Phone: +353 (0)21 490 4652 Fax: +353 (0)21 490 4664	t.cross@ucc.ie

Geir Dahle	Institute of Marine Research PO Box 1870 N-5817 Bergen, Norway	Phone: +47 55 23 63 49 Fax: +47 55 23 63 79	geir.dahle@imr.no
Gary R.Carvalho	School of Biological Sciences, University of Bangor Environment Centre Wales Bangor, Gwynedd LL57 2UW UK	Phone: +44 (0)1248 382100 Fax: +44 (0)1248 371644	g.r.carvalho@bangor.ac.uk
Aitor Albaina Vivanco	Lab. Genetics, Fac. Science & Technology; Dpt. Genetics, Physical Anthropology & Animal Physiology; Univ. Basque Country Bº Sarriena s/n 48940 Leioa; Bilbao; Spain	Phone: +34-946015503 Fax: +34-946013145	aitoralbaina@hotmail.com
Andone Estonba	Lab. Genetics, Fac. Science & Technology; Dpt. Genetics, Physical Anthropology & Animal Physiology; Univ. Basque Country Bº Sarriena s/n 48940 Leioa; Bilbao; Spain		andone.estonba@ehu.es
Reinhold Hanel	Johann Heinrich von Thünen-Institute, Institute for Fishery Ecology Palmaille 9 D-22767 Hamburg Germany	Phone +49 40 38905 290 Fax+49 40 38905 261	reinhold.hanel@vti.bund.de
Dorte Bekkevold	Technical University of Denmark, Vejlsøvej 39, 8600 Silkeborg, Denmark	Phone: +45 35883130 Fax: + 45 35883150	db@aqua.dtu.dk
John Gilbey	Freshwater Laboratory, Faskally, Pitlochry, Perthshire, PH16 5LB UK	Phone: +44 (0)1224 876544 Fax: +44 (0)1796 473523	j.a.gilbey@marlab.ac.uk

Anna Kristin Danielsdottir	Food Safety & Environment Vinlandsleid 12 113 Reykjavik, Iceland	Phone: + 354422 5014 Fax: +354 422 5001	annak@matis.is
Sarah Helyar	Food Safety & Environment Vinlandsleid 12 113 Reykjavik, Iceland	Phone: +354 422 5000 Fax: +354 422 5001	Sarah.helyar@matis.is
Filip Volckaert	Katholieke Universiteit Leuven Laboratory of Animal Diversity and Systematics Ch. Deberiotstraat 32 - bus 2439 B-3000 Leuven, Belgium	Phone: +32 16 32 39 66 Fax: + 32 16 32 45 75	filip.volckaert@bio.kuleuven.be
Daria Zelenina	Russian Federal Research Institute for Fisheries and Oceanography, 107140, 17, V. Krasnoselskaya str., Moscow, Russian Federation		dzel67@mail.ru
Estíbaliz López de Abechuco	Marine Research Division, Azti Technalia, Txatxarramendi ugartea z/g E-48395 Sukarrieta, Bizkaia		elopez@azti.es
Ian Bradbury	Fisheries and Oceans Canada/Pêches et Océans Canada. 80 East White Hills Road, P.O. Box 5667 St. John's, NL, A1C 5X1, Canada	Phone (709) 772-3869, Fax. (709) 772-3578	ian.bradbury@dfo-mpo.gc.ca
Iratxe Montes	Lab. Genetics, Fac. Science & Technology; Dpt. Genetics, Physical Anthropology & Animal Physiology; Univ. Basque Country Bº Sarriena s/n 48940 Leioa; Bilbao; Spain		iratxe.montes@ehu.es

Jack Lighten	Department of Biology, Dalhousie University Halifax, NS, B3H 4J1, Canada	jackie.lighten@dal.ca
José A. Sánchez-Prado	Departamento de Biología funcional, Universidad de Oviedo, Spain	jafsp@correo.uniovi.es
Stanley King	Department of Biology, Dalhousie University Halifax, NS, B3H 4J1, Canada	Stanley.King@dal.ca
Unai Cotano (first day only)	Lab. Genetics, Fac. Science & Technology; Dpt. Genetics, Physical Anthropology & Animal Physiology; Univ. Basque Country Bº Sarriena s/n 48940 Leioa; Bilbao; Spain	
Urtzi Laconcha (first day only)	Lab. Genetics, Fac. Science & Technology; Dpt. Genetics, Physical Anthropology & Animal Physiology; Univ. Basque Country Bº Sarriena s/n 48940 Leioa; Bilbao; Spain	

Annex 2: Agenda

Wednesday 2nd

9.00 Welcome by local hosts Aitor Albaina Vivanco (University of the Basque Country, UPV/EHU) and Naiara Rodríguez-Ezpeleta (Marine Research Division, AZTI Foundation)

9.30 Welcome and updates from WG chair Dorte Bekkevold

9.45 - 12.30 Presentation and discussion of position papers for ToR a–f

a) *Produce a review on the potential for using parasites and pathogens as "magnifying glass" for fish stock characterisation*. Filip Volckaert (9:45-10:30)

10:30-11:00 – Coffee

b) *Produce a review and discussion on the use of adaptive SNP markers and other adaptive markers for genetic identification of populations (breeding stocks).* Tom Cross (11:00-11:45)

c) *Contribute to the continuing assessment of the SNP-technology.* Geir Dahle (11:45-12:30)

d) *Contribute to the development of the scientific understanding of sustainability of aquaculture by collaborating with WGMASC and WGEIM and planning for the ASC 2012 theme session on "Genetic impact of aquaculture on wild populations"*. Dorte Bekkevold (12:30-13:00)

13:00- 14:00 - Lunch

e) *Evaluate potential for collaboration with other EGs in relation to the ICES Science Plan and report on how such cooperation has been achieved in practical terms (e.g. joint meetings, back-to-back meetings, communication between EG chairs, having representatives from own EG attend other EG meetings).* Dorte Bekkevold (14:00-14:30)

f) *Contribute to the improved integration of genetic information into fisheries management under a reformed Common Fisheries Policy and the new Data Collection Framework 2014-2020.* Jann T. Martinsohn (14:30-15:30)

15.30 – 17.00 Formation of ToR working groups and parallel work sessions on ToR a-f.

17.00 – 18.00 Open session. Presentation of results, projects, management issues that you would like to share and discuss.

Thursday 3rd

9.00 Morning assembly w. updates on activities and practical information

9.15 – 13.00 Parallel work sessions on ToR a-f

13.00 - 14.00 Lunch

14.00 - 16.00 Work in groups on ToR a-f (continued)

16.30 – 17.00 Status of work in ToR groups – each ToR lead gives an update

17.00 – 18.00	Open session. Presentation of results, projects, management issues that you would like to share and discuss.

Friday 4th

9.00	Morning assembly
9.15 – 12.15	Presentation of ToR reports/recommendations
12.15 - 13.30	Suggestions for new ToR's for 2013, discussion of consequences of multi-annual management of SCICOM expert groups and future meeting venue.
13.30	End of meeting
13.30	Lunch

Annex 3: WGAGFM terms of reference for the next meeting

The **Working Group on Application of Genetics in Fisheries and Mariculture** (WGAGFM), chaired by Dorte Bekkevold, Denmark, will meet in Reykjavik, Iceland, 8-10 May 2012 to:

a) Produce a review of the identification and use of adaptive gene markers in shellfish aquaculture and for the genetic characterisation of wild populations;

b) Review and consider technological developments in fisheries forensics and management of exploited marine fishes with emphasis on contributions to sustainability and governance;

c) Produce a review on the use of metagenomics and metatranscriptomics as an approach for marine ecosystem management

d) Contribute to the continuing assessment of the SNP-technology.

WGAGFM will report by 31 May 2013 to the attention of the SSGHIE Committee.

Supporting Information

Priority	The current activities of this Group will lead ICES into issues related to the ecosystem affects of fisheries, especially with regard to the application of the Precautionary Approach. Consequently, these activities are considered to have a very high priority.

Scientific justification	Term of Reference a) There is an increasing pressure for sustainable aquaculture of many shellfish species in Europe and worldwide. There is evidence of local adaptation in many species of shellfish and locally adapted populations are often characterized by traits allowing them to survive and thrive under heterogeneous environmental conditions. Some such traits are also likely to be of interest in relation to mariculture production. The fast developing field of genomics offers a potential to obtain necessary markers for investigation of population structuring, genetic basis of unique traits, effective population sizes, interspecies hybridization, etc. This information is valuable, not only to ensure sustainable shellfish aquaculture but also from a marine ecosystem management/biodiversity conservation point of view, where the aim is to identify and protect local populations. The recent and continuing development of new genetic screening technologies has the potential to significantly aid the identification of adaptive markers. Thus it is timely to consider the issues pertaining to the use of such markers in relation to both aquaculture traits and wild population genetics applications. We will review the current information in this field. Term of Reference b) While there are various strategies employed to promote sustainability of exploited marine fish resources, issues relating to the enforcement of regulations, and more widely, the governance of marine fisheries, continues to present a significant impediment to reducing stock declines. Recent advances in technology, especially of DNA-based methods, together with applications at a species and population-level, provide robust and informative tools to tackle issues relating to illegality and consumer fraud. Therefore, a review of the need for an enhanced framework for law enforcement and prosecution, key technological developments, the stringency (and distinct features) of the forensic approach, followed by a synthesis of potential and actual applications to date is required. Term of Reference c) Metagenomics and metatranscriptomics consist respectively on sequencing the total DNA or RNA present in a given environmental sample to determine the species present and their metabolic activity. Given that virtually anything that contains living organisms can be sequenced, these techniques can be applied to a vast range of biological questions. In the marine environment, metagenomics and metatranscritpomics can be applied i) to assess temporal and spatial biodiversity and/or metabolic changes to monitor the effects of stressors such as climate change and human activity in a given community, ii) to identify the preys of a given fish by sequencing its stomach content, iii) to identify the parasites infecting a given fish population or iv) to assess water quality in aquaculture installations. There are certainly many more applications of these techniques that will help maintaining a good status for marine waters, habitats and resources, all within the MSFD. Therefore, a review of the potential of metagenomics and metatranscirptomics as cost-effective approaches for marine ecosystem management is required. Term of Reference d) Issues pertaining to ascertainment bias, cost, SNP choice, ease of analyses, screening platform, technical aspects related to genotyping, data management, and broader technological and statistical approaches should be further considered by members of this working group on an ongoing basis.
Resource requirements	The research programmes which provide the main input to this group are already underway, and resources are already committed. The additional resource required to undertake additional activities in the framework of this group is negligible.
Participants	The Group is normally attended by some 20–25 members and guests.
Secretariat facilities	None.
Financial	No financial implications.

Linkages to advisory committees	There are no obvious direct linkages with the advisory committees.
Linkages to other committees or groups	SIMWG, WGEVO, WGMASC, WGEIM
Linkages to other organizations	Linkage with the EC Joint Research Centre at Ispra, Italy.

Annex 4: Recommendations

Recommendation	Adressed to
1. WGAGFM recommend that given that parasite population genetics can be a proxy for identifying host fish populations (including farmed and native groups), to make good use of it, when appropriate for the research question addressed. This requires promoting interdisciplinary interaction between fish biologists, fisheries scientist, ecologists, evolutionary biologists, parasitologists, bacteriologists and virologists in order to enhance parasite supported stock identification.	SCICOM, WGPDMO
2. That genetic markers under directional selection continue to be identified and employed in genetic stock identification analysis as such markers have been shown to yield informative insights on both the scale and dynamics of populations and in identifying potential underlying drivers	SCICOM, SIMWG
3. That markers associated with functional variation be used to explore the predictive power to study the impacts of selective pressures including fisheries, climate change, and pollution.	SSGHIE, WGEVO
4. That based on the binary nature and ease of cross laboratory calibration when considering Single Nucleotide Polymorphism databases, that the potential to combine datasets over larger spatial scales across laboratories be explored	WGDIM
5. That SNP technology due to the constant technical developments is assessed and discussed by WGAGFM members on a continuous basis.	SCICOM
6. That WGMASC/WGEIM and WGAGFM chairs explore the scope for interaction during the ASC in Bergen 2012 and hereunder discuss the scope for exchanging Term of Reference resolutions with the aim to provide comments on respective Terms of Reference, where appropriate.	WGMASC/WGEIM
7. That SCICOM SSGHIE compile and disseminate a portfolio of expertise from individual Expert Groups to facilitate future collaboration between Expert Groups.	SCICOM
8. That the ICES secretariat initiate and organise a workshop defining guidelines for the integration of genetic data and information in the Data Collection Framework and information in support of the implementation of EU policies such as the Marine Strategic Framework Directive and the reformed Common Fisheries Policy. Participants should include genetic experts as well as experts involved in the coordination and implementations of the MSFD and CFP.	ICES Secretariat
9. That SCICOM contribute to a process to support that member states build the necessary capacity and infrastructure to enable genetic monitoring at national and regional levels under the remit of the Marine Strategic Framework Directive and the Common Fisheries Policy.	SCICOM, SISAM
10. That the availability and feasibility of using national fisheries databases (such as the ICES hosted FishFrame) for the hosting of genetic data is explored	ICES Data Centre