SALSEA-MERGE FP7-ENV-2007-1
Grant Agreement No 212529

Work Package 1

Deliverables - D 1.1

Integration Strategy Report and database structure (D 1.1)

(Month 6)

SUMMARY

The aim of Work Package 1 of the SALSEA-MERGE project was to develop a genetic methodology for the identification of European Atlantic salmon caught at sea to their region, and ideally river, of origin. Existing information on molecular genetic markers was reviewed in detail and the potential of different types of existing and new genetic variation for assignment considered. It was concluded that currently available microsatellite loci were likely to provide a sufficient basis for a substantive useful within continent regional assignment capacity. General considerations and a specific power analysis of existing data sets indicated that the previously identified "Virginia Panel" of 15 loci minus one, referred to as the "SALSEA-14" would provide a sufficient suite for regional assignment, at least at broad to medium geographical scales. This would also be optimal as it represented the majority of the loci for which there was maximal overlap in use among different research groups, maximising the amount of existing data. As such, it kept to minimum the amount of work that would be required for the integration and extension of existing genetic data sets for European stocks to create a unified database for assignment. The objective was achieved by a unique, large-scale transnational collaborative endeavour involving 12 European genetic research laboratories undertaking a progressive series of integrated work activities. These activities lead to the creation of a GRAASP (**G**enetically-based **R**egional **A**ssignment of **A**tlantic **S**almon **P**rotocol). It was also recognized, as set out in the initial project design, the loci selected for the initial GRAASP development were unlikely to provide the most cost-effective, accurate or highest resolution molecular tool. It was agreed that for the purposes of genetic assignments within the SALSEA-Merge programme, given time constraints, the SALSEA-14 based GRAASP would be applied within the SALSEA-Merge programme on an initial set of archival and contemporary survey cruise samples of post-smolts, demonstrating the power of GRAASP and delivering initial insights into the marine biology of European Atlantic salmon. However, it was also agreed that the value of exploring the potential of other microsatellites and of mitochondrial and nuclear SNPs remained in so far as it would potentially provide the basis for the enhancement of the GRAASP, either during or after the project. Improved cost-effectiveness and higher resolution assignment would expedite its routine application in studies of the marine ecology of Atlantic salmon in the NE Atlantic Ocean.

Background

Methods for identifying the region or river/tributary origin of fish using DNA profiling have advanced significantly in the last five years and are now widely applied to salmonid stocks in the Pacific Northwest (Shaklee *et al.,* 1999). These programmes demonstrate the power of these methods, for a number of salmonid and other freshwater and marine species. The basic elements needed to develop DNA methods for Atlantic salmon in the North Atlantic exist. Available DNA markers e.g. microsatellites can already resolve fine scale resolution of genetic structure and with existing statistical techniques and appropriate base line information can assign individuals to population of origin within regions both using mixed stock analysis (MSA) and individual assignment (IA) approaches (e.g. the Baltic - Koljonen *et al.* 2007). This strongly supports the view that a DNA based approach to the identification of the origin of marine samples of Atlantic salmon could be exploited to significantly advance our understanding of migration and distribution patterns.

A number of classes of DNA markers show strong evidence for potential to provide the resolution needed for regional assignment. Existing work shows point mutations in mtDNA with highly restricted regional distributions that could be informative for some regional groups (e.g. Verspoor *et al.*, 2002, 2006; Makhrov et al. 2005). However, further work is needed to identify a suite of markers to comprehensively cover the European range of the salmon and confirm their diagnostic potential. Existing work also shows regionally restricted distributions of microsatellite alleles and varying levels of regional differentiations among existing microsatellite loci of which ~1700 have been identified in Atlantic salmon. Again, those best suited for regional discrimination remains to be established.

There is considerable potential for identifying single nucleotide polymorphisms (SNPs) for the loci of protein coding genes, (McMeel et al., 2001) with the capacity to contribute regional assignment (Rengmark et al., 2006); in salmon MEP-2*, which shows a strong regional differentiation across Europe, would be a particularly valuable locus for which to develop a DNA typing method. However, screening of randomly identified nuclear SNPs, could also identify further markers with strong regional differentiation. A recent pilot study carried out by University College Cork, Ireland and Biotrove in the US has shown the diagnostic potential of SNP" s in Atlantic salmon research (Coughlan et al. 2007).

Building on existing molecular genetic information on the Atlantic salmon and recent advances in genetic stock identification (GSI) methodology, it was intended to identify a suite of molecular markers capable of assigning fish to region of origin to resolve structuring of river stocks into regional stock groups and allow Individual Assignment (IA) and Mixed Stock Analysis (MSA) of salmon in mixed stock fisheries. To achieve this a robust, comprehensive genetic baseline of salmon stocks for the suite of loci, encompassing in the major European salmon rivers, will be created. Genetic markers

represent an advance over conventional tagging in that genetic assessment can potentially allow most, if not all, wild fish to be assigned to region or river of natal origin. This obviates the need to produce and tag hatchery fish as surrogates for wild populations, reducing the need for large scale tagging and recovery programmes.

The specific objectives of the work package were:

- integration of existing and new genetic data from across the European range of the species into an optimal database to support the identification of the region or river/tributary of origin of fish captured in the North Atlantic
- a practical electronic database, that integrates and extends existing microsatellite and mtDNA data to provide required baseline information
- a suite of molecular markers which can identify salmon to region or river of origin
- optimisation and validation of the database and the assignment methodology

Reported here is the strategy for the integration of the various elements of the R &D programme under Work Package 1 (Appendix 1).

Review of Background Genetic Data

An up to date review of existing studies of molecular genetic variation in Atlantic salmon existing was carried out, starting in 2008 at the SALSEA-SALMAN II Symposium and Workshop in Paris, France, following indication that SALSEA-Merge was to be funded. This was necessary given substantive new developments over the previous year with regard to new classes of variation (e.g. SNPs – Single Nucleotide Polymorphisms) and available data on European salmon populations for existing classes of variation (e.g microsatellites). The review was continued and concluded at the SALASEA-Merge Genetics meeting in Stansted, UK, August 2008.

At the Stanstead meeting, an integration strategy to realise a functional **G**enetically-based **R**egional **A**ssignment of **A**tlantic **S**almon **P**rotocol (GRAASP) was agreed. This integration strategy had two development facets that needed to be defined:

- choosing a suitable suite molecular markers for assignment given those available
- creation of a single unified comprehensive baseline of information for these markers

Molecular Markers

Following a series of presentations and a qualitative review of available information, the assembled expertise in Atlantic salmon molecular genetics at the two meetings concluded that at present there was no potential for improving assignment power using currently available restriction enzyme detected mitochondrial DNA variation. The potential for beneficial use of mtDNA polymorphisms had been demonstrated with respect to assignment to continent of origin (Gilbey et al. 2006). However, it was agreed that this class of variation might have potential to complement and enhance assignment methods based on nuclear loci in the future, when better, more cost-effective methods for resolving mitochondrial single nuclelotide polymorphisms (SNPs) became available. Existing knowledge of mtDNA variation was insufficient to evaluate its full potential as they were based on a limited screening of a small number of restriction fragment length polymorphisms. This represented a small fraction of the mtDNA genome. Thus the potential of mtDNA variation as a source of regional molecular markers remained to be assessed and required more informative sequencing based approaches that could screen a larger portion of this molecule for useful polymorphisms. Little population work had been carried out on nuclear SNPs suggesting that these markers, though considerable work was underway to develop this group of molecular markers (e.g Ryynänen and Primmer 2006; Ryynänen et al. 2007).

A meta-analysis carried out in 2004 (Garcia de Leaniz et al., unpublished manuscript; Appendix II), considered in the context of more recent information, formed the basis of the review of microsatellite data. This encompassed work done by 41 research groups and reported in 150 studies. The trans-range river coverage of the most extensively used, reliably typed microsatellites in 2004 is given in Table 1; data for each microsatellites was available for <200 rivers; for Europe, excluding the Baltic region, data was available for <100 rivers most all microsatellites in the table with most rivers analysed for less than 6 of the 15 identified as the "best" set at the Virginia meeting (i.e. the "Virginia" Panel of microsatellites). The median number of loci per population screened was 9 (Figure 1). However, an update of the available data in 2008 at the Stansted meeting (not shown) found that substantial, but regionally patchy, progress had been in increasing coverage, particularly for the Virginia Panel microsatellites, through programmes such as the Baltic Sea GSI work (Koljonen et al. 2007), the EU funded ASAP project (Griffiths et al. 2010) and the Irish National Atlantic Salmon Genetic Stock Identification (GSI) project, which commenced in 2006/2007 (unpublished).

An assessment of the assignment power of the existing data was carried out found that 14-16 microsatellite markers was a sufficient basis for regional assignment as well as the most cost-effective and efficient approach to exploit existing data toward this end (Figure 2). This shows the assignment success increases only slowly with further loci beyond 14-16 loci but to increase if the number of alleles in the loci used can be increased. There was also some variation among loci independent of allele number.

The review of molecular markers in Atlantic salmon, with respect to the potential for assignment and the extent of available existing information, concluded that, within the context of technological constraints and partner laboratory analysis capabilities, the molecular marker-based assignment protocol development should

- be restricted to microsatellite markers given their high levels of variability and the relative abundance of existing information compared to other classes of genetic loci
- build on the information base for a set of microsatellite loci (Virginia Panel) identified by salmon researchers in 2004 as optimal for population genetic studies of the species

Table 1 Trans-range coverage in Atlantic salmon up until August 2004 of the most extensively used, reliably typed microsatellites.

Locus name	No. repeat bases	Max Size range	Max no. Alleles	No. Rivers screened
Ssa197	4	135-279	30	186
Ssa 202	4	200-320	18	166
Ssa171	4	193-285	32	126
SSOSL85	2	176-228	20	122
SSOSL311	2	112-192	33	108
SSOSL438	2	110-151	17	108
Ssa85	2	110-181	27	102
SSsp2210	4	104-164	15	98
SSsp2216	4	202-305	18	98
SSsp1605	4	222-254	14	97
Ssa 14	2	138-152	4	80
Ssa 289	2	107-132	8	80
SsaF43	2	99-131	13	71
SSsp2201	4	259-371	21	68
Ssa20-19	2	78-111	9	66
SSOSL417	2	155-213	19	62
SSspG7	4	112-214	24	57
SsaD-30	2	217-247	10	56
SSLEEI84	2	171-229	28	52
SSLEEN82	2	209-233	13	52
Ssa13-37	4	112-126	4	50
SsaD486	4	162-210	7	35
SsaD144	4	112-298	35	34
SstrP15Str15INRA	2	210-220	6	33
SSOSL25	2	150-180	?	29
SSsp3016	4	70-130	?	1

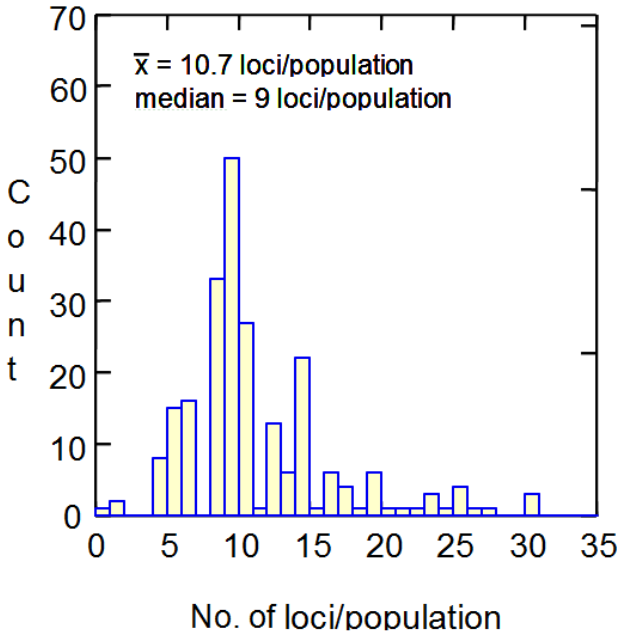

Figure 1 The distribution of numbers of studies relative to the number of loci/population based on a review of work done up to 2004 by 41 research groups and reported in 150 studies.

European Atlantic salmon stocks

Analysis of the literature (e.g. WWF 2001) and the NASCO (North Atlantic Conservation Organisation) database showed that there are ~2300 European rivers known to have salmon (Figure 3); the size and productivity of the rivers in respect of Atlantic salmon varies widely based on angling catches from <10 per year and a few kilometres long, to many tens of thousands of fish and many hundreds of kilometres in length. These rivers span 28^{o} latitude and 32^{o} longitude. Detailed data on microsatellite variation outside the Baltic is available for only one region in Europe – Ireland (National Atlantic Salmon Genetic Stock Identification (GSI) project, with reasonable coverage for rivers elsewhere in southwestern Europe (EU funded ASAP project - Griffiths et al. 2010) as well as Northwest Russia (Tonteri et al. 2006) and Denmark/SW Sweden (Nielsen et al. 1999; unpublished). However, there is incomplete and limited overlap of the microsatellites used and relatively little published or unpublished information exists for two of the largest salmon producing areas, Scotland and Norway, as well as Iceland. As well, for those microsatellites used in common, there has been no inter-calibration and integration of data sets (except in SW Europe – Griffiths et al. 2010).

This review of specific work on microsatellite variation in European Atlantic salmon stocks identified major data gaps that would have to be filled. This made clear the integration challenges and technical issues underlying assemblage of a baseline of

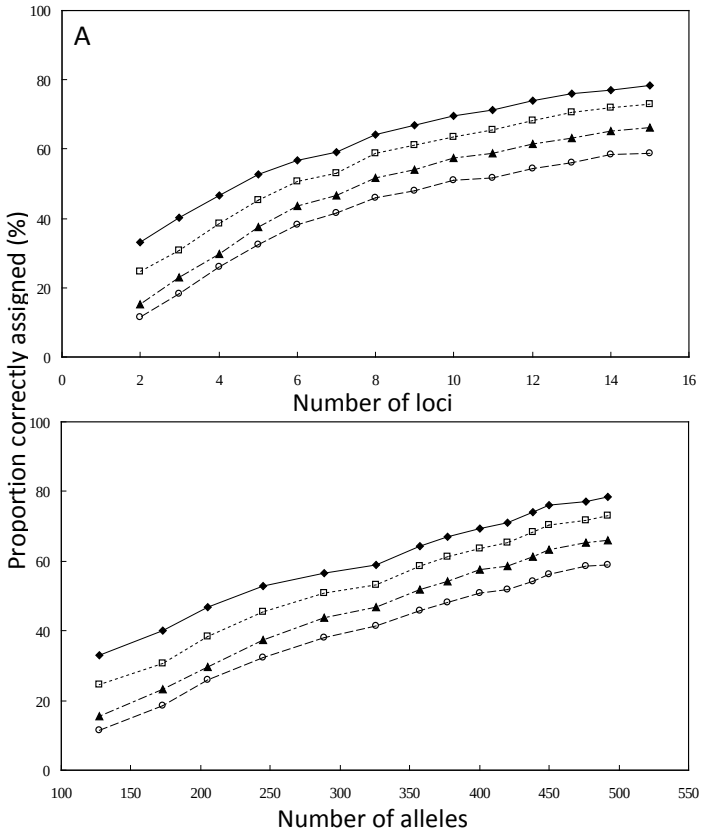

Figure 2 Relationship between A) number of loci, and B) number of alleles, and assignment success based on individual assignment simulations using existing data from Irish rivers: to Region (River Board; solid line, filled diamonds), District (Fishery District; short dashed line, open squares), river (short/long dashed line, filled triangles), and exact match to population (long dashed line, open circles).

information on microsatellite variation in European Atlantic salmon stocks to achieve a comprehensive GSI assignment to region/river of origin. It was felt that to be successful and achieve acceptable levels of assignment accuracy, it would be necessary to achieve baseline coverage of rivers spanning the full European distributional range and representing a high proportion of European production. Existing recreational and commercial catch data indicated that to achieve ~85% coverage of salmon production, it would be necessary to have a single data base of genetic information on the 450 most productive rivers. To achieve this required:

- extending existing river analyses to the full set of 14+ microsatellite loci
- expanding geographical coverage
- integrating existing and new data bases produced by national laboratories by inter-calibration of genotypes and standardization of genotype nomenclature

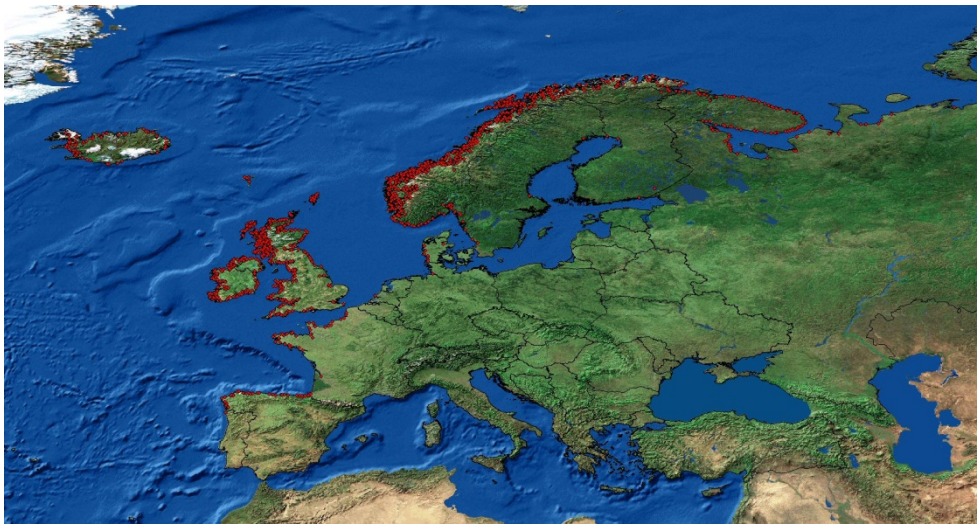

Figure 3 Map of the distribution of mouths of rivers in Europe with spawning populations of anadromous Atlantic salmon.

Integration Strategy

The work package sub-task delineation set out in the DoW (Appendix I) was reviewed at the Stansted Genetics subgroup meeting. It was agreed that, overall, the basic strategy defined at the time of the project proposal submission provided a good starting point for development of a GRAASP: **G**enetically-based **R**egional **A**ssignment of **A**tlantic **S**almon **P**rotocol. However, in light of work carried out since then, two modifications to the strategy set out in the DoW were agreed. The first was to focus in the first instance on the development of a GRAASP based on microsatellites and to exclude consideration of existing mtDNA polymorphisms. This was agreed due to the clear evidence that microsatellites alone and to focus on the microsatellites in the Virginia panel of 15 microsatellites minus SsaD486, SsaD486 being for the most part only variable in North American populations. It was felt that these 14 (ie the SALSEA-14; see Table 2) could provide a high level of assignment success and accuracy, and there was little evidence for potentially useful mtDNA variation in the polymorphisms that had so far been identified. The second, given the apparent power of a microsatellite based tool for regional assignment, despite recent developments in nuclear SNPs, to go straight for the development of a comprehensive and sufficient microsatellite baseline for assignment based on the SALSEA-14. Thus the main focus would be on the integration of the existing and new genetic information generated by the 12 contributing partner laboratories into a single trans-European data base for the assignment of Atlantic salmon to region/river of origin. A flow chart of the integration strategy is set out in Figure 4.

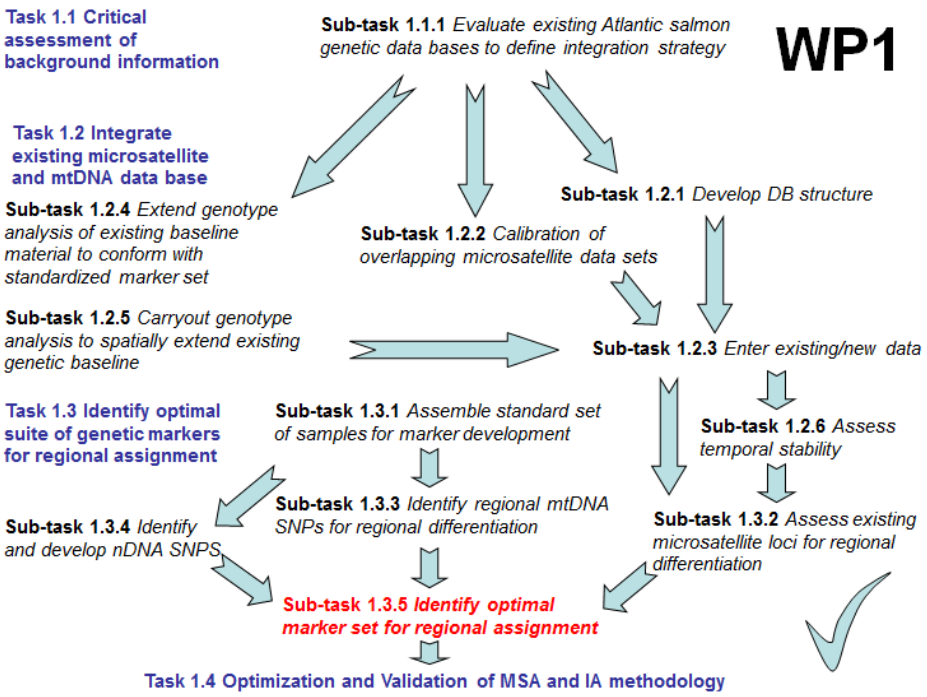

Figure 4 Agreed operational flow chart for integration strategy

References

Coughlan, J. et al. 2007

Griffiths, A. M., Machado-Schiaffino, G., Dillane, E., Coughlan, J., Horreo, J. L., Bowkett, A. E., Minting, P., Toms, S., Roche, W., Gargan, P., McGinnity, P., Cross, T., Bright, D., Garcia-Vazquez, E. & Stevens, J.R. (2010) Genetic stock identification of Atlantic salmon (Salmo salar) populations in the southern part of the European range. BMC Genetics, 11: 31. 7

Koljonen, M-L., King, T.L. and Nielsen, E.E. (2007). Genetic Identification of Individuals and Populations. Chapter 9, In *The Atlantic Salmon: Genetics, Conservation & Management* (Verspoor, E., Stradmeyer, L & Nielsen, J.L., eds). Blackwell Publishing, Oxford. pp 270-298.

Makhrov, A., Verspoor, E., Artamonova, V.S. and O'Sullivan, M, (2005) Atlantic salmon colonization of the Russian Arctic coast: pioneers from North America Journal of Fish Biology 67, Suppl A: 68-79.

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

Shaklee, J.B., Beacham, T.B., 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

Verspoor, E. O'Sullivan, M., Arnold, A.L, Knox, D. & Amiro, P.G. (2002). Restricted matrilineal gene flow & historical population fragmentation in Atlantic salmon (*Salmo salar* L.) within the Bay of Fundy, eastern Canada. **Heredity** 89: 465-472.

Verspoor, E. O'Sullivan, M. Arnold, A.M. Knox, D. Curry, A. Lacroix, G. & Amiro, P. (2006). The Nature and Distribution of Genetic Variation at the Mitochondrial Nd1 Gene of the Atlantic Salmon (*Salmo salar* L.) Within and Among Rivers Associated With The Bay Of Fundy and the Southern Uplands of Nova Scotia. **FRS Research Services Internal Report No** 18/05.

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

Ryynänen HJ and Primmer CR (2006) Single nucleotide polymorphism (SNP) discovery in duplicated genomes: intron-primed exon-crossing (IPEC) as a strategy for avoiding amplification of duplicated loci in Atlantic salmon (Salmo salar) and other salmonid fishes. BMC Genomics 7: 192

Ryynänen H, Tonteri A, Vasemägi A, Primmer CR (2007) A comparison of the efficiency of single nucleotide polymorphisms (SNPs) and microsatellites for the estimation of population and conservation genetic parameters in north European Atlantic salmon (Salmo salar) populations. Journal of Heredity, 98, 692-704.

Tonteri A, Titov S, Veselov A, Zubchenko A, Koskinen MT, Lesbarrères D, Kaluzhin S, Bakhmet I, Lumme J and Primmer CR (2005) Phylogeography of anadromous and non-anadromous Atlantic salmon (*Salmo salar*) from northern Europe Annales Zoologici Fennici 42: 1-22

APPENDIX I

Work Package 1 Task Description

Task 1.1 Critical Assessment of background genetic information

Sub-task 1.1.1 *Evaluate existing Atlantic salmon genetic data bases to define integration strategy*

An up-dated catalogue of existing data bases containing genetic information on microsatellite and mitochondrial DNA locus variation will be assembled, with a view to the requirements for their integration into a single, unified trans-European data base for genetic identification of the origin of European salmon captured in marine waters. Primary consideration will be given to identifying the degree of commonality of the markers sets in different data bases held in different regions and laboratories, differences in sampling approaches underlying the data bases (e.g. spatial intensity, sample size, etc), need for further baseline information to get optimal regional coverage, fill geographic gaps, etc. This will be achieved by assembling a meta-data base through correspondence with known active or past research groups. This meta-data base will then be used as the basis for convening a workshop to define a practical, optimal integration strategy which will form the basis for task 2.1. A catalogue of biological material available for the development of baseline information, including the analysis of temporal stability of stock differentiation. Locus and baseline GAPS will be identified.

Participants: Lead – 3,8; other 1,2, 5-13

Task 1.2 Integrate existing microsatellite and mtDNA data base

Sub-task 1.2.1 *Develop DB structure*

The database strategy will be translated into an electronic database format implemented using standard database software, to be identified in the strategy developed in subtask 1.1.1. Entry of data into the base, and data access and review, will be web-enabled to allow remote, rapid and universal access by consortium members. An IPR agreement will be developed and agreed with consortium partners to ensure IPR but maximize availability of individual existing data sets to other researchers and to the science community generally.

Participants: Lead – 1; other

Sub-task 1.2.2 *Calibration of overlapping microsatellite data sets*

Existing methodologies will be used to standardize and calibrate exising microsatellite and mtDNA databases, to make it possible to align them so that they can be combined for pan-European analyses. This will be achieved by the exchange of standard sets of samples and overlapping of sample analyses and the development of a standard variant nomenclature. The nominated participants will travel between laboratories to assess the specific integration challenges and define appropriate solutions. This will include the development of computer macros to translate existing nomenclatures into the agreed standard nomenclature and data storage formats. The calibration report will set out how databases will be linked.

Participants: Lead - 5,6; other 1

Sub-task 1.2.3 *Enter existing data*

Participants, whom own most of the existing information on microsatellite and mtDNA variation in European Atlantic salmon stocks, will input this information, once calibrated and standardized (see Subtask 1.2.2) into the web-based electronic database constructed as part of subtask 1.2.1

Participants: Lead – other 1; 3-13

Sub-task 1.2.4 *Extend genotype analysis of existing baseline material to conform with standardized marker set*

Existing regional European microsatellite and mtDNA databases are only partially overlapping with regard to the genetic

loci screened. Consortium laboratories will undertake further genetic analysis to extend existing data sets to include the full set of agreed markers, and as identified in Subtask 1.1.1, using their archival DNA or tissues. This data will then be entered into the Trans-European genotype database.
Participants: Lead – 2; other 3-12

Sub-task 1.2.5 *Carryout genotype analysis to spatially extend existing genetic baseline*
Existing regional microsatellite and mtDNA databases only partially cover the European range of the species and there are significant spatial gaps, which need to be filled to provide the required geographical coverage for accurate assignment. Consortium laboratories will undertake further genetic analysis to spatial extend existing data sets for the full set of agreed markers, and as identified in Subtask 1.1.1, using newly collected or archival DNA or tissues. Much of the material will be obtained from routine management electrofishing and trapping surveys. This data will then be entered into the Trans-European genotype database.
Participants: Lead – 2; other 1, 3-7, 10,11,13
NB: When European database for the agreed standard set of genetic markers is completed with information from the subtasks 1.2.2 to 1.2.4, it will be lodged in electronic form via ICES as part of WP5, with access according to an agreed IPR protocol agreed among consortium members.

Sub-task 1.2.6 *Assess temporal stability*
An assessment of a number of river populations will be undertaken in order to determine their genetic stability and thus provide an indication of the integrity of the newly established baseline in the long-term and to make recommendations on required sampling frequencies.
Participants: Lead – 9; other 1,3-13

Task 1.3 Identify optimal suite of genetic markers for regional assignment
Sub-task 1.3.1 *Assemble standard set of samples for marker development*
A representative subset of river/regional samples spanning the European range of the salmon will be assembled in order to provide material for identifying region specific genetic markers. This will be achieved by defining a number of agreed selection criteria, which are developed out of consideration of existing information on regional variation and European geography. Three replicate sample sets, one for each lead group in sub-tasks 1.3.2, 1.3.3., 1.3.4 will be created from a fish specifically collected for this purpose. The sets will be made by sub-sampling aliquots of tissue from ethanol preserved whole fish, with the remainder of the carcasses deposited for future use in a institution specialising in maintaining reference collections such as the British Museum. This collection will then be available as reference baseline DNA for future genetic studies. An agreed protocol of future access to this material will be agreed with the institution where the sample collection is deposited.
Participants: Lead – 2; other 1-13

Sub-task 1.3.2 *Assess existing microsatellite loci for regional differentiation*
A subset of 400 of the over 1700 microsatellite loci identified in Atlantic to date will be selected based on an assessment of their suitability for population genetics work. Based on screening of the reference collection assembled in Subtask 1.3.1, a suite of 8-12 multiplexable microsatellite loci which give high resolution regional assignment potential will be chosen and optimal conditions for genotyping established.
Participants: Lead – 1; other 4,11

Sub-task 1.3.3 *Identify regional mtDNA SNPs for regional differentiation*
mtDNA from the D-loop, ND1, Cytb and other gene regions will be sequenced to identify single nucleotide polymorphism (SNPs). Based on screening of the reference collection assembled in Subtask 1.3.1, a suite of SNPs which provide useful regional assignment capability will be chosen and optimal conditions for typing established.
Participants: Lead – 3; other 4,8

Sub-task 1.3.4 *Identify and develop nDNA SNPS*

Known nDNA sequences identified in electronic databases such as GENBANK and cGRASP will be reviewed and the most promising selected for screening for SNPs. Based on screening of the reference collection assembled in Subtask 1.3.1, a suite of 100 SNPs which provide useful regional assignment capability will be chosen and optimal conditions for typing established.

Participants: Lead – 6; other 2,7

Task 1.4 Optimisation and Validation of MSA and IA methodology

Sub-task 1.4.1 *Assembly of blind test samples from baseline river and non-baseline rivers*

Two sets of samples of 200 fish will be assembled, one will be used as a known test sample for optimization and the other will be used as a blind test sample for validation. The samples will be assembled by a consortium member independent of the partners involved in the optimization and validation, and supplied to the appropriate partner as the start of subtask 1.4.2. Also provided will be data for weighting of samples in baseline by productivity potential and catch, data simulations, assess effect of temporal variation.

Participants: Lead – 2; other 1, 3-13

Sub-task 1.4.2 *Optimisation and Validation of MSA and IA methodology*

The samples provided under subtask 1.4.1 will be genetically typed for the standard set of microsatellite and mtDNA loci, as well as for the newly developed regional markers from subtasks 1.3.2, 1.3.3, and 1.3.4. The data will then be used with baseline data from Task 1.2 and 1.3 to run simulations and optimisations using standard available genetic assignment software e.g. STRUCTURE, BAYES, SPAM. Findings will be reported to and discussed by all partners at a WP1 workshop. The output to be used in WP3 Task 3.1. The most parsimonious set of markers giving optimal MSA and IA will be identified and used for defining the MSA/IA baseline selection of loci and the set of loci to be used for screening marine samples in WP3.

Participants: Lead – 3; 1, 6, 7

Deliverables from WP1

D 1.1 - Report on integration strategy and data base structure (month 3)

D 1.2 - Report on new regional markers (month 20)

D 1.3 – Web-based Trans-European genotype database for Atlantic salmon GSI (month 24)

Appendix II

Microsatellite DNA loci for Genetic Stock Identification (GSI) of Atlantic salmon: choice, challenges and opportunities

C. García de Leániz[1], S. Consuegra[2] & E. Verspoor[3]

[1]University of Wales Swansea, Department of Biological Sciences

Swansea, SA2 8PP, Wales, UK

[2]Gatty Marine Institute, University of St. Andrews,

St Andrews, Fife KY16 8LB, Scotland, UK

[3]FRS Freshwater Laboratory, Pitlochry, Perthshire,

Scotland, PH16 5LB, UK

Running title: *Genetic Stock Identification of Atlantic salmon using microsatellites*

Contents

Abstract

The development of highly variable DNA microsatellite loci and new, more powerful assignment methods has made it possible to achieve, for the first time, the fine scale resolution necessary to study Atlantic salmon populations in detail and to assign individuals to their population of origin through Genetic Stock Identification (GSI). However, to realise its full potential, GSI requires the use of an optimal, common suite of loci, as well as appropriate base line allele frequency data. Here we review information from over 150 studies to examine the characteristics of different markers currently being used across the species' range, to review the desired properties of microsatellite loci for successful stock identification, and to examine the scope and limitations of GSI in Atlantic salmon. Most loci have been used to screen a relatively small number of populations, often within relatively narrow geographical areas, and this constitutes the first obstacle for successful GSI in Atlantic salmon. A second obstacle lies in the limited number of loci screened per population, although with the development of newer, easier to score markers, more comprehensive studies based on a larger number of loci are beginning to emerge. Finally, for GSI to be useful, the degree of temporal stability in microsatellite frequencies needs to be assessed, since the amount of temporal variation within populations may exceed the amount of variation between populations. Thus before microsatellite-based GSI becomes a viable proposition for routine stock identification in mixed stock fisheries several conditions need to be met: (1) a common, optimal suite of microsatellite needs to be agreed and tried across the species' range, (2) allele scoring needs to be validated across different laboratories, (3)

sample sizes and baseline data needs to be increased to include most of the existing populations, and (4) the degree of temporal stability in allelic frequencies needs to be assessed. In addition, we note that endangered populations may be difficult to detect in mixed stock samples due to their small size and relatively low contribution, and thus a precautionary, holistic approach to stock assignment is recommended.

Key words (max 6): microsatellites, genetic stock identification, GSI, Atlantic salmon, temporal stability, mixed stock fisheries, conservation genetics

Introduction

Sustainable management of anadromous salmonids such as the Atlantic salmon (*Salmo salar*) requires a clear understanding of how the species are structured into reproductively distinct populations, and the extent to which different stocks intermingle in the sea and are affected by exploitation (Shaklee et al., 1999; Koljonen, 2001). Unfortunately, and despite many years of research (Hansen and Quinn, 1998; Mills, 1999), the marine ecology of the Atlantic salmon remains poorly known (Shearer, 1992; Mills, 1993). Physical tagging has shown that salmon from different regions migrate to different feeding areas in the sea, and are exploited differentially in coastal and high seas fisheries (Jákupsstovu, 1988; Jensen, 1988; Hansen, 1993; Erkinaro et al., 1999; Holm et al., 2000; Kallio-Nyberg et al., 2000; Jacobsen et al., 2001; Økland et al., 2001; Jutila et al., 2003). However, the extent to which different populations contribute to a given fishery, or how their contribution varied annually or seasonally, is not known. Indeed, nowhere are such shortcomings more evident than in mixed-stock fisheries (Shaklee et al., 1990), where lack of proper identification means that fishing pressure can at times be applied indiscriminately to different salmon populations, regardless of their abundance or conservation status. Even for the most thoroughly studied of the salmon fisheries, that off West Greenland, existing knowledge is often only capable of estimating the relative contribution of North American and European salmon (Reddin and Friedland, 1999); very little is known about the contribution of fish originating from individual rivers, or even from different regions. With the possible exception of the

Baltic region (Koljonen et al., 1999; Koljonen, 2001; Koljonen, 2004), current knowledge on the contribution of different stocks to Atlantic salmon fisheries (Faroese, Irish Sea, Norwegian Sea) remains largely inadequate (Jacobsen et al., 2001).

Differences among salmon populations in migratory behaviour (Kallio-Nyberg et al., 1999; Jacobsen et al., 2001), coupled with seasonal fluctuations in oceanic conditions (Friedland et al., 2000; Holm et al., 2000; Friedland et al., 2003), can be expected to result in substantial stock differences in vulnerability to fisheries and in marine survival between populations (Reddin et al., 1997; Kallio-Nyberg et al., 1999; Kallio-Nyberg et al., 2000; Kallio-Nyberg et al., 2004). This is most certainly the case in relation to Atlantic and Baltic stocks in Europe (Hansen, 1993; Hansen and Quinn, 1998), but differences are also evident at a finer scale. For example, stocks from the inner Bay of Fundy (unlike those in the Outer Bay) generally do not migrate to West Greenland and show a much higher marine mortality than stocks elsewhere in the region (Jessop, 1976; Mills, 1989). The extent to which other stocks experience different marine conditions is largely unknown, though at the very least, most stocks will differ in migratory routes if only because their natal rivers differ in geographical location, e.g. salmon from southern and northern Europe migrating to West Greenland will follow different, though perhaps partially overlapping, outward and return paths.

The impact of intercepting fisheries on different salmon populations can only be ascertained if the migrations and contributions of each stock component are known. To do so, extensive surveys are required, and fish catches need to identified to the population, river, or at least, region of origin. Given the large number of salmon rivers

(WWF, 2001), and the fact that some watersheds are likely to contain multiple genetic populations (e.g., Spidle et al., 2001; Verspoor et al., 2002), the identification of fish by means of physical tagging becomes economically, as well as logistically, impractical. Rather, it was long recognised that the only practical solution for identifying stocks in mixed stock fisheries rested in the use of biological markers or 'natural tags', either inherent or acquired by fish as a consequence of their origin. Initial attempts to discriminate between salmon stocks focused on non-genetic markers, such as variation in parasite loads (Nyman and Pippy, 1972; Beverley-Burton and Pippy, 1978), differences in meristics and morphological characters (Claytor and MacCrimmon, 1988; Claytor et al., 1991), variation in growth patterns (de Pontual and Prouzet, 1988; Friedland and Reddin, 1994; Cadrin and Friedland, 1999), and microchemical analysis of hard parts (Lapi and Mulligan, 1981; Mulligan et al., 1983; Campana et al., 2000). However, success was limited and confined largely to identification of continent of origin or, at best, to discrimination of regional stock groups (Fournier et al., 1984; Reddin and Friedland, 1999).

Advances in molecular biology in the early 80's facilitated the use of genetic tools in fisheries management and the development of mixed-stock analysis (Carvalho and Hauser, 1995; Ward, 2000). Protein polymorphisms were the first molecular markers to show useful stock differentiation in Atlantic salmon (e.g., Ståhl, 1981; Verspoor, 1986; Ståhl, 1987), but their use was limited in practise by the need for fresh or freshly-frozen tissue and the need to sacrifice the fish (reviewed by Verspoor et al., 2005). With the identification of more informative gene loci, the focus then turned to DNA markers,

initially to mitochondrial DNA (mtDNA) variation (Davidson et al., 1989; Bermingham et al., 1991), and later to analysis of variation at minisatellite (Ferguson et al., 1995) and microsatellite loci (McConnell et al., 1995; McConnell et al., 1997) found in nuclear DNA (nDNA). Nuclear DNA markers generally show higher, more useful numbers of polymorphic loci and levels of allelic diversity, which make them better suited for stock discrimination (Carvalho and Hauser, 1995; O'Connell and Wright, 1997; Ward, 2000).

Microsatellite genes consist of repeat units of 2-5 DNA bases which are believed to be mostly neutral to selection (Hancock, 1999) and generally afford a high level of resolution, thus being ideally suited for the differentiation of populations (reviewed in (O'Connell and Wright, 1997). Nowadays, genetic screening of salmon populations with microsatellites is cheap, accurate, and fast (O'Reilly et al., 1996; Paterson et al., 2004). Furthermore, samples can be screened non-destructively using small quantities of DNA either from fresh tissue or dried scales (Nielsen et al., 1999). The latter has opened up the possibility of using archived scale collections for historical analysis (e.g., Nielsen et al., 1997; Säisä et al., 2003; Consuegra et al., 2005), and to sample extant populations from remote parts of the species' range. Thus the development of highly variable DNA microsatellite markers, coupled with new and powerful statistical techniques, has made it possible to achieve, for the first time, the fine scale resolution necessary to study Atlantic salmon populations in detail and to assign individuals to their population of origin.

Genetic Stock Identification (GSI) - the use of genetic markers for identifying the contribution of different populations in mixed stock fisheries, as pioneered by (Beacham

et al., 1985; Beacham et al., 1985) for Pacific salmon – is still rudimentary when applied to Atlantic salmon. Baseline data is only available for a limited number of populations, often screened with different sets of markers (Koljonen and McKinnell, 1996; Verspoor, 1997; King et al., 2001). To realise the full potential of GSI of Atlantic salmon requires the identification of an optimal suite of markers and appropriate base line information on river stocks and populations. With the advent of cheaper, faster technologies, the development of different molecular markers in Atlantic salmon has increased exponentially (Figure 1), to the point where there are now c. 1,700 microsatellite sequences available for use. Obviously this poses a logistic problem since a prerequisite for successful stock identification is the adoption of the *same* set of markers. At present it is not known which microsatellites are being used, how reliable they are, how many populations are being screened, or how realistic the scope is for sharing a common database. Clearly, considerable work still needs to be undertaken before salmon populations can be routinely managed on a population by population basis. Thus, our objectives on this paper were threefold: (1) to examine the characteristics, extent, and reliability of the different microsatellites currently being used across the species' range, (2) to review the desired characteristics of microsatellite markers, sample size requirements, and sampling considerations for successful genetic stock identification in Atlantic salmon, and (3) to assess the feasibility, scope and limitations of microsatellite variation for genetic stock identification of Atlantic salmon populations.

Methods

Cross-table questionnaires were sent to 41 key research groups working on genetic stock discrimination in Atlantic salmon. For each screened population we asked for information on its name and geographical location, type of samples (juveniles or adults), number of years that the population had been screened, and source reference for the study. Likewise, for each marker used, we compiled information on locus name, number of repeat bases and motif, allele size, number of alleles, reliability (detection of null alleles, incidence of stuttering), GenBank/primer sequences, and source reference.

Data was compiled for 150 studies, comprising both published and unpublished data sets. This allowed to piece together a database of 229 anadromous and land-locked Atlantic salmon populations from both Europe (n = 159) and N. America (n = 70) screened with 158 different loci across the entire species' range (Figure 2). This forms the basis of the comparative study presented below. The full database is available from the authors upon request.

Criteria for choosing microsatellite loci for GSI of Atlantic salmon

Type of nuclear markers

Five types of nuclear DNA markers have been used for population differentiation in Atlantic salmon: minisatellites (Taggart et al., 1995; Mjølnerød et al., 1997; Goodier and Davidson, 1998), microsatellites (e.g., King et al., 2001), SNPs (Ryynänen and Primmer, 2004), and more recently, expressed sequence tags (EST; Vasemägi et al., 2005) and amplified fragment length polymorphisms (AFLP; Moen et al., 2004). The advantages and disadvantages of different nuclear markers for stock discrimination are discussed in

O'Connell and Wright (1997), Campbell et al. (2003) and Vasemägi et al. (2005). Being more variable, microsatellite loci require smaller sample sizes than minisatellites or allozymes to achieve comparable levels of precision (Galvin et al., 1995; O'Reilly and Wright, 1995), and largely for this reason appear to be the marker of choice for GSI in Atlantic salmon, being by far the most extensively used type of markers (n = 77) in the studies reviewed.

Usage of different microsatellite loci

A prerequisite for successful stock identification is the adoption of the *same* set of markers. However, inspection of the database indicates that of the 77 microsatellite loci considered, none has been used to screen *all* 229 populations studied so far (Figure 3), thus negating the first prerequisite for successful GSI (i.e. a common set of markers).

On average, each locus was used to screen just over 32 populations (median = 8 populations per locus). The most widely used microsatellite loci are Ssa197 and Ssa202 with 186 and 166 populations screened, respectively, followed by Ssa171 (126 populations), SSOSL85 (122), and SSOSL311 and SSOSL438, both with 108 populations. These include the first microsatellites that were developed (McConnell et al., 1995; Slettan et al., 1995; O´Reilly et al., 1996) and have, thus, been longest in use. Most other microsatellites have been used to screen between 20 and 40 populations, often within relatively small geographical areas. Clearly, for GSI to be properly implemented in Atlantic salmon some form of agreement must be reached on a common set of microsatellite loci to be used.

Number of loci

Determining the optimum number of loci for successful stock identification can be problematic (e.g., Campbell et al., 2003; Guinand et al., 2004) since the optimum number of loci will largely depend on the variability and size of the samples to be assigned (Manel et al., 2004). For example, Bernatchez and Duchesne (2000) found in simulations that increasing the number of loci had the greatest effect on the accuracy of the assignment tests for GSI, whereas increasing allelic diversity, and thus differentiation, was most influential in the case of parentage assignment. On the other hand, Koskinen et al. (2004) concluded using resampling techniques that even a modest increase in the number of microsatellites resulted in a substantial improvement in discrimination power in European grayling. O'Connell and Wright (1997) suggested that more than 5 loci need to be screened to obtain reasonable confidence limits when addressing population genetics questions. Cornuet et al. (1999) examined the influence of the number of loci on the performance of the Bayesian assignment method and concluded that approximately 10 microsatellite loci were needed to obtain 100% correct classification under various scenarios. Letcher and King (1999) showed that 4 loci and 7 alleles per locus were sufficient to correctly assign offspring to parental family, but that 10 loci and 10 alleles per locus were required to for correct assignment to grandparents. In farmed Atlantic salmon (Norris et al., 2000) found that 8 highly variable microsatellite loci provided accurate (>95%) assignment to parental pair in a family selection program. In Pacific salmonids, genetic stock identification is carried out successfully with 8-15 loci,

depending on the species (Beacham et al., 1999; Beacham et al., 2000; Nelson et al., 2003; Withler et al., 2004). In the case of the much less variable isozymes, 51 loci are screened for GSI of Pacific salmon across the Pacific Rim (Habicht et al., 2001), which can perhaps also serve as an indication on the number of allozyme markers that would need to be screened in Atlantic salmon.

Despite the difficulty in determining an optimum number of loci suitable for all studies, some general recommendations can still be made. As a rough rule of thumb, it may concluded that c. 10-15 microsatellites with moderate allele diversity are probably adequate for GSI of Atlantic salmon under most conditions (King et al., 2001; Koljonen et al., 2002; Koljonen, 2004; Skaala et al., 2004), though it is probably wise to first test the discrimination power of the chosen loci in a subsample (e.g., Banks et al., 2003). Use of a greater number of loci, say 20 loci or more, are probably only advantageous over small spatial scales, when population differentiation is low, and provided sample sizes are large. On the other hand, if the purpose of the study is solely the discrimination of populations over very large spatial scales, for example of European and North American origin, then the screening of samples with 2 or 3 carefully chosen microsatellites will provide the necessary accuracy under most conditions (e.g., McConnell et al., 1995; King et al., 2001; Koljonen et al., 2002; Wennevik et al., 2004). Indeed, even just one microsatellite locus combined with a mtDNA polymorphism will suffice for discriminating between salmon from the two continents (Gilbey et al., 2005)

Inspection of the database indicates that Atlantic salmon populations have been screened with widely different number of markers, from a minimum of 1 microsatellite

locus (2 populations) to a maximum of 104 different markers including minisatellites, microsatellites, SNPs and ESTs (1 population). However, most populations have been screened with a relatively small number of microsatellite loci (mean = 10.7 loci, median = 9 loci, Figure 4), which falls short of the recommended number. With the development of more, easier to score markers, more comprehensive studies based on a larger number of loci are beginning to appear (Tonteri et al., 2005; Vasemägi et al., 2005a,b).

Number of alleles

Power to discriminate between stocks depends ultimately on the relative amount of genetic variation within and among populations (Smouse and Chevillon, 1998; Smith et al., 2001). Therefore, markers with a large number of alleles generally offer greater discrimination power than those with more limited variation. However, too many alleles will also require larger sample sizes, whereas too few alleles may not afford enough differentiation. Thus there is a trade-off for stock discrimination between sample size, number of loci, and allele diversity (O'Connell and Wright, 1997). Smith et al. (2001) recommend that populations should be sampled with loci having between 5 and 10 alleles, while Smouse and Chevillon (1998) concluded that accurate assignments may be achieved by having as few as 4 loci provided there are more than 8 alleles per locus.

In general, employing additional loci is preferred over choosing more polymorphic loci (Neff et al., 2000; Norris et al., 2000; Villanueva et al., 2002), and simulations have shown that 6-10 alleles per locus may be adequate for most population assignment

studies (Bernatchez & Duchesne, 2000). However, criteria for optimal parental assignment (Letcher and King, 1999; Reynolds and Templin, 2004) will generally differ from criteria for optimal mixed stock analysis (King et al., 2001). In general it can be concluded that choosing microsatellite loci with a moderate number of alleles per locus (5-15) would probably work best for GSI of Atlantic salmon over wide geographical scales (e.g., Spidle et al., 2001; Spidle et al., 2003) and when sample sizes per population may be limited (c. 50). The more polymorphic loci are probably best applied for stock discrimination over smaller geographical areas (Shaklee et al., 1999; Beacham et al., 2004) and with larger sample sizes (e.g., Reynolds and Templin, 2004). In any event, it is probably wise to first test the discrimination of the chosen loci in a subsample (Banks et al., 2003), or to conduct a priori power analysis (Reynolds and Templin, 2004).

Sample size

Although sample size will largely be dictated by the purpose of the study, O'Connell and Wright (1997) recommend that a minimum of 50 individuals should be sampled per population for loci having between 5 and 10 alleles. Larger sample sizes will normally be required for more variable loci, as depicted in Figure 5. Genetic stock identification of Pacific salmon is based on samples of 28-1192 individuals per population (typically 50-100) in the case of allozymes (Habicht et al., 2001) and 200+ individuals per population in the case of microsatellites (Withler et al., 2004). Simulations have shown that increasing the sample size increases the reliability of assignments (Withler et al., 2004),

with GSI accuracy improving substantially up to a sample size of 50 fish per population (Winans et al., 2004).

In general, baseline population sample sizes should also be 50 or greater for accurate population assignments based on genetic data (Shaklee et al., 1999; Guinand et al., 2004). In the case of Pacific salmon, the allozyme-based GSI program relies on allelic frequencies computed for 280 different sites across the Pacific Rim, from Washington State to Russia (Habicht et al., 2001). Data for 229 populations of Atlantic salmon (Figure 6) indicates that the average sample size was 65.3 fish per population (median = 60 fish/population) though for many populations sample size fell short of the minimum suggested above.

Base repeats

Most of the microsatellite loci used for stock identification in Atlantic salmon are still dinucleotide repeat polymorphisms (n = 112 or 78% of microsatellite loci), probably because they were the first to be developed (McConnell et al., 1995; McConnell et al., 1995; Slettan et al., 1995; O'Reilly et al., 1996; Sánchez et al., 1996; Slettan et al., 1996). However there is an increasing number of tetranucleotide loci being used (n = 31 or 22% of cases), which are generally easier to screen and less prone to scoring errors (O'Reilly et al., 1996; Paterson et al., 2004; King et al., 2005), which must be taken into account when choosing a suite of markers for GSI (O'Connell and Wright, 1997).

In general, tetranucleotide microsatellites display higher allele diversity (mean = 25.1 range = 4-55) than that provided by dinucleotides (mean = 16.0, range = 2-46), which

makes the former more powerful for stock discrimination, other things being equal. On the other hand, dinucleotides have generally a shorter fragment size than tetranucleotides (see below), which may make them more suitable for studies involving degraded DNA (e.g. analysis of archived scale collections), when only small DNA fragments may amplify successfully.

Fragment size and suitability for multiplexing

The majority (52%) of dinucleotide microsatellites employed to discriminate between Atlantic salmon populations have a minimum fragment size that lies in the range 100-159 bp, whereas the size of fragments amplified with tetranucleotide microsatellites are generally larger than 220 bp. Consideration to fragment size can be important when dealing with old or degraded samples, such as scales or bones remains, where fragment size is recommended to be under 150 bp, or even below 120 bp (O'Connell and Wright, 1997; Nielsen et al., 1999). In this context, the development of new SNPs and ESTs of small fragment size (Moran, 2002; Ryynänen and Primmer, 2004; Vasemägi et al., 2005) should prove particularly useful for studies based on archival samples. However, the need for non-overlapping fragment sizes when multiplexing (an important factor in keeping the costs of genetic typing down) will also dictate the suite of microsatellites which are suitable (O'Reilly et al., 1996; Neff et al., 2000). Thus, for most GSI applications a combination of non-overlapping loci with long and short fragment sizes will need to be used.

Temporal stability

For GSI to be useful, the degree of temporal stability in microsatellite frequencies must first be assessed (Tessier and Bernatchez, 1999), since genetic variation among years within populations may equal (or even exceed) genetic variation between populations (Garant et al., 2000; Heath et al., 2002), thus severely curtailing the scope for successful stock discrimination. Microsatellite allelic frequencies have been shown to temporally stable in some Pacific salmonids (Beacham et al., 2004), even in bottlenecked populations (Withler et al., 2000), but the results for other stocks are not so clear (Heath et al., 2002). In Atlantic salmon, several authors have reported temporal stability in landlocked (Tessier and Bernatchez, 1999; Potvin and Bernatchez, 2001) and anadromous populations (Beacham and Dempson, 1998; Nielsen et al., 1999; Säisä et al., 2003), but other studies have shown significant shifts in allele frequencies over various time scales (Nielsen et al., 1997; Garant et al., 2000; Consuegra et al., 2005). This suggests that temporal stability should not be assumed, particularly in small or depressed populations where the likelihood of random genetic drift is greatest.

In the case of anadromous Atlantic salmon with overlapping generations, several years of microsatellite data will normally be needed to test for temporal changes in allelic frequencies since at least two complete cohorts need to be screened in order to properly test for temporal stability. Depending on the age structure of the populations, 2-4 years of microsatellite data may be the minimum required for southern populations, and up to 6 years and longer for more northerly stocks. However, inspection of the

database (Figure 7) indicates that the majority of the 229 populations have been screened for only 1 year (67%), and only in 26 populations (11%) is there enough data to examine the degree of temporal stability in microsatellite frequencies. Clearly, longer studies need to be undertaken to assess the temporal stability of Atlantic salmon populations and the real scope for GSI based on microsatellite frequencies. Moreover, samples should be stratified by year classes to test for homogeneity in allelic frequencies between cohorts whenever possible (e.g., Consuegra et al., 2005), and the amount of genetic differentiation between populations should be reported relative to the observed degree of annual variation within populations (Beacham et al., 1999; Beacham et al., 2000; Withler et al., 2000; Beacham et al., 2004)

Origin of samples

In genetic studies of wild salmonids, sampling of breeding adults is generally preferred over sampling of juveniles (Beacham et al., 2004), as it lessens the possibility of non-random representation of families (Hansen et al., 1997) and other sources of bias, such as differential juvenile to adult mortality or contributions of non-native fish from hatcheries. However, sampling of adults may not always be possible and it may also pose problems if the origin of fish is uncertain. For example, fish may stray temporarily into other rivers, making it important to sample adults at spawning time in spawning areas. In this respect, as juveniles can generally be safely attributed to local spawning, sampling of wild juveniles may be preferred under some conditions. In any event, when sampling of juveniles is the chosen option, samples should be obtained from wild fish,

and not from fish in hatcheries, as analysis of cultured fish may lead to erroneous allele frequency estimates (e.g. Säisä et al., 2003). Sampling should be carried out over a wide an area as possible to capture the true genetic structure of the breeding population, and not merely the distribution of spawning redds in the area. Moreover, given the limited post-hatching dispersal of Atlantic salmon alevins (García de Leániz et al., 2000; Webb et al., 2001), it is probably wise to avoid sampling alevins and 0+ fry during their first summer, concentrating instead on parr and smolts, as older juveniles are more likely to be thoroughly distributed than younger ones. Considering what is known about the effects of domestication, sampling juveniles from hatcheries to infer the genetic structure of wild populations is not recommended.

Inspection of the GSI salmon database indicates that the majority of the populations sampled for microsatellite variation are anadromous (96% of the samples), and only a few (8 populations, 4% of samples) consist of land-locked and non-migratory stocks. Samples are composed predominantly (60%) from juveniles, rather than from breeding or returning adults (Figure 8). For a relatively small number of populations (n = 34) samples are available from both adults and juveniles (the origin of samples for a further 13 populations is not known). Considering what is known about the non-random distribution of juveniles in streams (e.g., Mjølnerød et al., 1999; Olsén et al., 2004) and gene flow between populations (Consuegra et al., 2005), the stage of development should probably be taken into account when interpreting microsatellite frequencies for stock discrimination. Failure to sample the population adequately will tend to

underestimate the degree of genetic variation, potentially introducing serious errors in the estimated degree of population differentiation and on the accuracy of stock assignments.

Reliability

Several factors determine the reliability of microsatellite loci for stock identification, of which differential mutation rates, facility of scoring, stuttering, and presence of null alleles are perhaps the most important (O'Connell and Wright, 1997). Together these can affect the accuracy and reliability of assignments (O'Reilly et al., 1998), which makes it essential to validate the scoring of loci among all laboratories involved in GSI ((White and Shaklee, 1991). There may important differences in scoring microsatellite alleles between different laboratories (e.g., O'Reilly et al., 1996; Wennevik et al., 2004), particularly if they employ different systems.

Mutations seem to be rare in Atlantic salmon, and neither mutations nor scoring errors are constant among loci (O'Reilly et al., 1998). Fishes tend to have longer, more variable microsatellites than most other vertebrates (Slettan et al., 1993; Brooker et al., 1994; Slettan et al., 1997), with mutation rates varying significantly among loci depending on microsatellite length (Neff and Gross, 2001). Differences in mutation rates among loci can introduce bias in the estimation of population divergence (Olsen et al., 2004), and thus on the accuracy of genetic stock identification. In particular, highly polymorphic microsatellite loci yield significantly lower F_{ST} estimates than less variable loci (Olsen et

al., 2004), and this will also need to be taken into account when choosing a suite of microsatellite loci for GSI.

Stuttering and the presence of null alleles can limit the utility of genetic markers as they can curtail the amount of detected genetic variation or, worse, introduce errors during sequencing (O'Reilly et al., 1996; O'Connell and Wright, 1997). Such problems can be associated with poor quality DNA, but also with the design of the primers, with size homoplasy, and with identity by state rather than by descent (Estoup and Angers, 1998). Fortunately, such problems appear to be rare. Of all the loci examined, null alleles were reported for nine loci only: SSsp2213 and SSsp2215 (Paterson et al., 2004); Est12, Est22, Est46, Est76, Est103, and Est125, (Vasemägi et al., 2005); and MHCI (Grimholt et al., 2002) which may need some optimisation. Similarly, significant stuttering was reported only for six loci in some populations: μ60 (Estoup et al., 1993), SSOSL25 and SSOSL417 (Slettan et al., 1995), Ssa28 (Skaala et al., 2004), Ssa85 (O´Reilly et al., 1996) and SsaF43 (Sánchez et al., 1996). Statistical procedures can be applied for correcting the presence of null alleles in the samples (e.g., Brookfield, 1996), but it is probably best to avoid loci with null alleles for GSI purposes.

Complementary markers for stock identification

The current tendency is to rely increasingly on molecular markers alone for stock identification. However, a more holistic, multidisciplinary approach based on the combined use of genetic and non-genetic markers (Table 1) can enhance discriminatory power and improve accuracy in stock assignments (Begg and Waldman, 1999; Waldman,

1999; Cadrin, 2000). This would be particularly useful in the analysis of mixed-stock fisheries, where baseline genetic data may be fragmentary or incomplete (Seeb et al., 1998; Habicht et al., 2001), and where Bayesian assignment methods that make use of prior information can greatly improve accuracy (Baudouin et al., 2004; Koljonen, 2004). For example, smolt age varies predictably among anadromous populations (Metcalfe and Thorpe, 1990) and this can be used in conjunction with genetic markers to increase the precision of assignments in stock discrimination (Koljonen and Pella, 1997). Likewise, trace elements and stable isotopes that leave a permanent chemical signature in the otoliths of fishes can be used as natural tags to discriminate between populations (e.g. Thorrold et al., 1998; Campana and Thorrold, 2001). In Pacific salmon, elemental analysis of otoliths has been used to assess homing accuracy (e.g., Quinn et al., 1999) but the same approach could also be used to aid discrimination in mixed-stock analysis.

Combining genetic markers with different modes of inheritance (e.g. maternal vs. bi-parental) or different response to selection (e.g. selected vs. neutral) can help to discriminate genetically similar stocks subjected to different selective pressures (Tessier et al., 1995; Landry and Bernatchez, 2001; Tonteri et al., 2005; Vasemägi et al., 2005). In this respect, recent developments in microarray technology have greatly simplified the screening of mtDNA variation in salmonids (e.g., Moriya et al., 2004), making it feasible to compare sequences relatively cheaply. Indeed, the combined use of mtDNA variation with a single microsatellite marker provides a more rapid and cost-effective method of discriminating North American and European Atlantic salmon than it is possible using microsatellites alone (Gilbey et al., 2005). In Pacific salmon, allozyme and mtDNA data

are still used for GSI alongside data for microsatellite loci, as the former databases are more extensive (Seeb et al., 1998; Habicht et al., 2001). Moreover, for some populations allozymes provide better, more accurate identification than microsatellites (Winans et al., 2004). Thus, whenever possible, genetic estimates of stock composition should be compared to those derived from other markers including physical tags (e.g., Brodziak et al., 1992; Koljonen and McKinnell, 1996) .

Discriminatory power and implications for conservation

The power to detect specific populations, in mixed stocks samples, declines as their contribution declines (Reynolds and Templin, 2004a). Additional bias may be introduced when comparing mixture estimates among samples separated either temporally or spatially (Reynolds and Templin, 2004b). These limitations have a number of important implications for the conservation of threatened stocks (Koljonen, 1995; Nelson et al., 2003), since their typically small size may render them undetectable in mixed-stock fisheries. First, because the consequences of failing to detect an endangered population in mixed stock samples are normally much worse than the consequences of making an erroneous contribution when none exists (Figure 9), assignment methods should be adjusted to minimize the likelihood of Type I errors (declaring a hypothesis false when it is actually true), possibly at the expense of inflating Type II errors (failure to falsify a hypothesis when it is actually false). Secondly, inferences from molecular markers alone should not be taken as conclusive evidence for lack of contribution of small populations in mixed stock samples; where available, ancillary information from physical tagging or

other markers should be taken into account before drawing firm conclusions regarding the impact of a given fishery on endangered stocks. Finally, when the purpose of GSI is to afford greater protection to vulnerable populations, assignment methods should bear in mind the conservation status of all potential populations since anthropogenic disturbances to salmon stocks tend to clumped spatially (Frissell, 1993; Reeves et al., 1995).

Summary and Conclusions

The use of microsatellites as genetic tags has considerable potential for improving understanding of the ocean migrations of Atlantic salmon, the extent of population mixing at sea, and the marine exploitation of populations from different geographical areas. Microsatellites, hence, can extend knowledge gained initially from physical tagging, and more recently from other, often less powerful genetic markers such as allozymes, mtDNA or mini satellites. Their use will help to assess the effect of mixed stock fisheries and allow for a more rational, effective way of managing and protecting wild salmon populations.

As the abundance of wild salmon declines and aquaculture production increases, it is likely that microsatellite loci will increasingly be used to assess the introgression of farmed fish into natural populations, or to monitor the effects of supplemental breeding programs. Traditionally, the effects of stocking had been assessed with the aid of allozyme markers (e.g., García de Leániz et al., 1989; Verspoor and García de Leániz,

1997), but discrimination was generally limited to a few polymorphisms. The use of microsatellite loci, in combination with new assignment methods in mixed stock analysis (e.g.(Manel et al., 2004) (Koljonen, 2004), now offers the possibility of accurately assessing the seasonal contribution of wild and cultured fish originating from different hatcheries (e.g., Kallio-Nyberg and Koljonen, 1997; Skaala et al., 2004; Vasemägi et al., 2005). Microsatellites should also prove invaluable in developing more stringent containment procedures designed to minimize escapes from salmon farms, and in assessing the merits of stocking, sea ranching, and supplemental breeding programs in comparison to other restoration strategies. Indeed, the use of microsatellites as genetic tags may for some purposes even replace the need for conventional tagging studies, or augment such studies by providing information on all fish captured, and not just from the recapture of tagged individuals (Ruzzante et al., 2001).

Development of a GSI database for Atlantic salmon

However, before microsatellites can be used effectively as natural tags, a database containing allele frequencies for each population needs to be developed, shared and maintained by the different laboratories working on identification of Atlantic salmon populations, similar to the one already in place for Pacific salmon (Habicht et al., 2001). The main challenges regarding the development and implementation of such database are summarized below:

1. *Validation and Reliability*. Allele scoring needs to be replicated and validated across different laboratories involved in GSI of Atlantic salmon; quality control guidelines and best-practice protocols may need to be developed to minimise scoring errors.
2. *Number and type of microsatellite loci.* For successful GSI, a trade off exists between sample size, number of loci and allele diversity. A suite of 10-15 loci with moderate allele diversity, ease of scoring and capacity for multiplexing are probably best for most applications.
3. *Sample sizes*. The minimum recommended sample size is 50 fish per year class and population, but sample sizes in excess of 100 fish/year class/population may be required under some conditions. There may often be a trade-off between the number of populations sampled and the number of individuals sampled per population, given that available resources will generally be limited.
4. *Base-line data.* Ideally, the database should encompass base line allele frequencies for as many populations as possible, including wild natural populations, supplementation hatcheries, and commercial fish farms. This would allow for the greatest power and flexibility in meeting various demands (forensic, food traceability, GSI in mixed stocks fisheries, monitoring of restoration programs, etc..)
5. *Temporal stability*. The degree of temporal stability in allelic frequencies needs to be assessed, and the ratio between population variation and annual variation

reported. Populations normally need to be screened over several years, and the results stratified by year class (or migration year).

6. *Origin of samples*. It is essential that samples are representative of the population, which is best met by sampling breeding adults rather than juveniles or just adults returning to a river. If sampling of juveniles is the only option, then older parr are preferred over alevins and fry, as the former are less likely to be biased by non-random distribution of families.
7. *Complementary information*. The inclusion in the database of complementary data on genetic and non-genetic markers, such as allozymes, mtDNA, life history traits, parasite loads, trace elements, etc. should be considered as it can greatly improve the discrimination power of the assignment tests, particularly in the case of Bayesian-based methods.
8. *Precautionary approach to GSI.* Endangered populations may be difficult to detect in mixed stock samples due to their small size and relatively low contribution. Thus, a precautionary approach based on minimising Type I assignment errors and using a variety of methods for stock identification of small populations is advocated.

Choice of an optimal suite of microsatellite loci for GSI

The disparity of markers used by different laboratories, and the ever increasing development of new microsatellite loci, make it is essential to choose a common suite of c. 15 microsatellite loci, that is agreed and implemented by different laboratories. To

capitalize on the information already available, it is recommended that the chosen microsatellite panel includes the most common loci (Table 2), and meets the following suitability criteria (Table 3):

(a) good discrimination power across different geographical scales

(b) moderate allele diversity

(c) be easy to score

(d) amenable to multiplexing

(e) free of stuttering and null alleles

(f) sizeable baseline data

(g) selectively neutral

(h) low mutation rate

(i) capable of being used on fish scales and archival material

Acknowledgements

We thank the following for responding to our questionnaire and helping us build the Atlantic salmon GSI database: P. Moran, J.L. Campos-Parada, E. Garcia-Vazquez, F. Ayllon, J.A. Sanchez-Prado, D. Bernardo, B. Jordan, E. Dillane, P. McGinnity, K. Kohlmann, C. Primmer, A. Tonteri, A. Vasemägi, M.-L. Koljonen, K. Hindar, V. Wennevik, E. Nielsen, B. Adams, T. King, and P. O'Reilly. Financial support for this study

was provided by The Atlantic Salmon Trust and the Westcountry Rivers Trust through the ASAP project. Their help is gratefully acknowledged.

References

Banks, M. A., Eichert, W. and Olsen, J. B. (2003) Which genetic loci have greater population assignment power? *Bioinformatics* **19**(11), 1436-1438.

Baudouin, L., Piry, S. and Cornuet, J. M. (2004) Analytical Bayesian approach for assigning individuals to populations. *J Hered* **95**(3), 217-224.

Beacham, T. D. and Dempson, J. B. (1998) Population structure of Atlantic salmon from the Conne River, Newfoundland as determined from microsatellite DNA. *J. Fish Biol.* **52**, 665-676.

Beacham, T. D., Le, K. D. and Candy, J. R. (2004) Population structure and stock identification of steelhead trout (*Oncorhynchus mykiss*) in British Columbia and the Columbia River based on microsatellite variation. *Env. Biol. Fishes* **69**, 95-109.

Beacham, T. D., Pollard, S. and Le, K. D. (1999) Population structure and stock identification of steelhead in southern British Columbia, Washington, and the Columbia river based on microsatellite DNA variation. *Trans. Am. Fish. Soc.* **128** (6), 1068-1084.

Beacham, T. D., Pollard, S. and Le, K. D. (2000) Microsatellite DNA population structure and stock identification of steelhead trout (*Oncorhynchus mykiss*) in the Nass and Skeena Rivers in northern British Columbia. *Marine Biotechnology* **2** (6), 587-600.

Beacham, T. D., Withler, R. E. and Gould, A. P. (1985) Biochemical genetic stock identification of chum salmon (*Oncorhynchus keta*) in Southern British Columbia. *Can. J. Fish. Aquat. Sci.* **42** (3), 437-448.

Beacham, T. D., Withler, R. E. and Gould, A. P. (1985) Biochemical genetic stock identification of Pink salmon (*Oncorhynchus gorbuscha*) in Southern British Columbia and Puget Sound. *Can. J. Fish. Aquat. Sci.* **42**, 1474-1483.

Begg, G. A. and Waldman, J. R. (1999) An holistic approach to fish stock identification. *Fisheries Research* **43**(1-3), 35-44.

Bermingham, E., Forbes, S. H., Friedland, K. and Pla, C. (1991) Discrimination between Atlantic salmon (Salmo salar) of North American and European origin using restriction analyses of mitochondrial DNA. *Can. J. Fish. Aquat. Sci.* **48**, 884-893.

Bernatchez, L. and Duchesne, P. (2000) Individual-based genotype analysis in studies of parentage and population assignment: how many loci, how many alleles? *Can. J. Fish. Aquat. Sci.* **57**, 1-12.

Beverley-Burton, M. and Pippy, J. H. C. (1978) Distribution, prevalence and mean numbers of larval *Anisakis simplex* (Nematoda: Ascaridoidea) in Atlantic salmon, *Salmo*

salar L., and their use as biological indicators of host stocks. *Env. Biol. Fishes* **3**(2), 211-222.

Brodziak, J., Bentley, B., Bartley, D., Gall, G. A. E., Gomulkiewicz, R. and Mangel, M. (1992) Tests of genetic stock identification using coded wire tagged fish. *Can. J. Fish. Aquat. Sci.* **49** (7), 1507-1517.

Brooker, A. L., Cook, D., Bentzen, P., Wright, J. M. and Doyle, R. W. (1994) Organization of microsatellites differs between mammals and cold-water teleost fishes. *Can. J. Fish. Aquat. Sci.* **51**, 1959-1966.

Brookfield, J. F. Y. (1996) A simple new method for estimating null allele frequency from heterozygote deficiency. *Molecular Ecology* **5**, 453-455.

Cadrin, S. X. (2000) Advances in morphometric identification of fishery stocks. *Rev. Fish Biol. Fish.* **10**(1), 91-112.

Cadrin, S. X. and Friedland, K. D. (1999) The utillity of image processing techniques for morphometric analysis and stock identification. *Fisheries Research* **43**, 129-139.

Campana, S. E., Chouinard, G. A., Hanson, J. M., Fréchet, A. and Brattey, J. (2000) Otolith elemental fingerprints as biological tracers of fish stocks. *Fisheries Research* **46**, 343-357.

Campana, S. E. and Thorrold, S. R. (2001) Otoliths, increments and elements: keys to a comprehensive understanding of fish populations? *Can. J. Fish. Aquat. Sci.* **58**, 30-38.

Campbell, D., Duchesne, P. and Bernatchez, L. (2003) AFLP utility for population assignment studies: analytical investigation and empirical comparison with microsatellites. *Molecular Ecology* **12**, 1979-1991.

Carvalho, G. R. and Hauser, L. (1995). Molecular genetics and the stock concept in fisheries. In Molecular Genetics in Fisheries: 55-79. Carvalho, G. R. and Pitcher, T. J. (Eds.). London: Chapman & Hall.

Claytor, R. R. and MacCrimmon, H. R. (1988) Morphometric and meristic variability among North American Atlantic salmon (*Salmo salar*). *Can. J. Zool.* **66**, 310-317.

Claytor, R. R., MacCrimmon, H. R. and Gots, B. L. (1991) Continental and ecological variance-components of European and North American Atlantic salmon (*Salmo salar*) phenotypes. *Biol. J. Linn. Soc.* **44**, 203-229.

Consuegra, S., Verspoor, E., Knox, D. and García de Leániz, C. (2005) Asymmetric gene flow and the evolutionary maintenance of genetic diversity in small, peripheral Atlantic salmon populations. *Cons. Genetics*.

Cornuet, J. M., Piry, S., Luikart, G., Estoup, A. and Solignac, M. (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics* **153**(4), 1989-2000.

Craik, J. C. A. and Harvey, S. M. (1987) A biochemical method for distinguishing between wild and farmed salmonid fishes by their carotenoid pigmentation. *Journal of the Forensic Science Society* **27**, 47-55.

Craik, J. C. A., Harvey, S. M., Jákupsstovu, S. H. and Shearer, W. M. (1987) Identification of farmed and artificially reared Atlantic salmon among the catch of the wild salmon fishery of the Faroes. *Ices C M* **M:26**, 1-5.

Davidson, W. S., Birt, T. P. and Green, J. M. (1989) A review of genetic variation in Atlantic salmon, Salmo salar L., and its importance for stock identification, enhancement programmes and aquaculture. *J. Fish Biol.* **34**, 547-560.

de Pontual, H. and Prouzet, P. (1988) Numerical analysis of scale morphology to discriminate between atlantic salmon stocks. *Aquat Living Resour* **1**, 17-27.

Dempson, J. B. and Power, M. (2004) Use of stable isotopes to distinguish farmed from wild Atlantic salmon, Salmo salar. *Ecol Freshw Fish* **13**, 176-184.

Einum, S., Fleming, I. A., Cote, I. M. and Reynolds, J. D. (2003) Population stability in salmon species: effects of population size and female reproductive allocation. *J Anim Ecol* **72**(5), 811-821.

Erkinaro, J., Økland, F., Moen, K. and Niemelä, E. (1999) Return migration of the Atlantic salmon in the Tana River: distribution and exploitation of radiotagged multi-sea-winter salmon. *Boreal Environmental Research* **4**, 115-124.

Estoup, A. and Angers, B. (1998). Microsatellites and minisatellites for Molecular Ecology: theoretical and empirical considerations. In Advances in Molecular Ecology: 55-85. Carvalho, G. R. (Ed.). IOS Press.

Estoup, A., Presa, P., Krieg, F., Vaiman, D. and Guyomard, R. (1993) $(CT)_n$ and $(GT)_n$ microsatellites: a new class of genetic markers for *Salmo trutta* L. (brown trout). *Heredity* **71**, 488-496.

Ferguson, A., Taggart, J. B., Prodöhl, P. A., McMeel, O., Thompson, C., Stone, C., McGinnity, P. and Hynes, R. A. (1995) The application of molecular markers to the study and conservation of fish populations, with special reference to Salmo. *J. Fish Biol.* **47**(Suppl. A), 103-126.

Fournier, D. A., Friedland, K. D., Beacham, T. D., Riddell, B. E. and Busack, C. A. (1984) Estimating stock composition in mixed stock fisheries using morphometric, meristic, and electrophoretic characters. *Can. J. Fish. Aquat. Sci.* **41**, 400-408.

Friedland, K., Esteves, C., Hansen, L. P. and Lund, R. A. (1994) Discrimination of Norwegian farmed, ranched and wild-origin Atlantic salmon, Salmo salar L., by image processing. *Fish Mgmt Ecol* **1**, 117-128.

Friedland, K. D., Hansen, L. P., Dunkley, D. A. and MacLean, J. C. (2000) Linkage between ocean climate, post-smolt growth, and survival of Atlantic salmon (Salmo salar L.) in the North Sea area. *ICES J. Mar. Sci.* **57**, 419-429.

Friedland, K. D. and Reddin, D. G. (1994) Use of otolith morphology in stock discriminations of Atlantic salmon (*Salmo salar*). *Can. J. Fish. Aquat. Sci.* **51**, 91-98.

Friedland, K. D. and Reddin, D. G. (2000) Gowth patterns of Labrador Sea Atlantic salmon postsmolts and the temporal scale of recruitment synchrony for North American salmon stocks. *Can. J. Fish. Aquat. Sci.* **57**, 1181-1189.

Friedland, K. D., Reddin, D. G. and Castonguay, M. (2003) Ocean thermal conditions in the post-smolt nursery of North American Atlantic salmon. *ICES J. Mar. Sci.* **60**(2), 343-355.

Frissell, C. A. (1993) Topology of extinction and endangerment of native fishes in the Pacific Northwest and California (U.S.A.). *Cons. Biol.* **7**(2), 342-354.

Galvin, P., McKinnell, S., Taggart, J. B., Ferguson, A., O'Farrell, M. and Cross, T. F. (1995) Genetic stock identification Atlantic salmon using single locus minisatellite DNA profiles. *J. Fish Biol.* **47** (Suppl A.), 186-199.

Garant, D., Dodson, J. J. and Bernatchez, L. (2000) Ecological determinants and temporal stability of the within-river population structure in Atlantic salmon (Salmo salar L.). *Molecular Ecology* **9**, 615-628.

García de Leániz, C., Fraser, N. and Huntingford, F. A. (2000) Variability in performance in wild Atlantic salmon, Salmo salar L., fry from a single redd. *Fish Mgmt Ecol* **7**, 489-502.

García de Leániz, C., Verspoor, E. and Hawkins, A. D. (1989) Genetic determination of the contribution of stocked and wild Atlantic salmon, *Salmo salar* L., to the angling fisheries in two Spanish rivers. *J. Fish Biol.* **35**(Suppl. A), 261-270.

Gilbey, J., Knox, D., O'Sullivan, M. and Verspoor, E. (2005) Novel DNA markers for rapid, accurate and cost-effective discrimination of continental origin of Atlantic salmon (Salmo salar L.). *ICES J. Mar Sci.*

Goodier, J. L. and Davidson, W. S. (1998) Characterization of novel minisatellite repeat loci in Atlantic salmon (*Salmo salar*) and their phylogenetic disttribution. *J Mol Evol* **46**, 245-255.

Grimholt, U., Drabløs, F., Jørgensen, S. M., Høyheim, B. and Stet, R. J. M. (2002) The major histocompatibility class I locus in Atlantic salmon (Salmo salar L.): polymorphism, linkage analysis and protein modelling. *Immunogenetics* **54**, 570-581.

Guinand, B., Scribner, K. T., Topchy, A., Page, K. S., Punch, W. and Burnham-Curtis, M. K. (2004) Sampling issues affecting accuracy of likelihood-based classification using genetical data. *Env. Biol. Fishes* **69**, 245-259.

Habicht, C., Templin, W. D., Guthrie, C. M., III, Wilmot, R. L., Winans, G. A., Iwamoto, E. M., Seeb, J. E. and Seeb, L. (2001). Status report for genetic stock identification studies of Pacific Rim sockeye salmon. 21.

Hancock, J. M. (1999). Microsatellites and other simple sequences: genomic context and mutational mechanisms. In Microsatellites. Evolution and Applications: 1-9. Goldstein, D. B. and Schlötterer, C. (Eds.). Oxford: Oxford University Press.

Hansen, L. P. (1993). Movements and migration of salmon at sea. In Salmon in the Sea and New Enhancement Strategies: 26-39. Mills, D. (Ed.). London: Fishing News Books.

Hansen, L. P. and Quinn, T. P. (1998) The marine phase of the Atlantic salmon (*Salmo salar*) life cycle, with comparisons to Pacific salmon. *Can. J. Fish. Aquat. Sci.* **55**, 104-118.

Hansen, M. M., Nielsen, E. E. and Mensberg, K. D. (1997) The problem of sampling families rather than populations: relatedness among individuals in samples of juvenile brown trout Salmo trutta L. *Molecular Ecology* **6**, 469-474.

Heath, D. D., Busch, C., Kelly, J. and Atagi, D. Y. (2002) Temporal change in genetic structure and effective population size in steelhead trout (Oncorhynchus mykiss). *Molecular Ecology* **11**, 197-214.

Holm, M., Holst, J. C. and Hansen, L. P. (2000) Spatial and temporal distribution of post-smolts of Atlantic salmon (Salmo salar L.) in the Norwegian Sea and adjacent areas. *ICES J. Mar. Sci.* **57**, 955-964.

Jacobsen, J. A., Lund, R. A., Hansen, L. P. and O´Maoileidigh, N. (2001) Seasonal differences in the origin of Atlantic salmon (Salmo salar L.) in the Norwegian Sea based on estimates from age structure and tag recaptures. *Fisheries Research* **52**, 169-177.

Jákupsstovu, S. H. (1988). Exploitation and migration of salmon in Faroese waters. In Atlantic Salmon: Planning for the Future: 458-482. Mills, D. and Piggins, D. (Eds.). London: Croom-Helm.

Jensen, J. M. (1988). Exploitation and migration of salmon on the high seas, in relation to Greenland. In Atlantic Salmon: Planning for the Future: 438-457. Mills, D. and Piggins, D. (Eds.). London: Croom-Helm.

Jessop, B. M. (1976). Distribution and timing of tag recoveries from native and nonnative Atlantic salmon (Salmo salar) released into Big Salmon River, New Brunswick. *J. Fish. Res. Canada* **33**, 829-833.

Jutila, E., Jokikokko, E., Kallio-Nyberg, I., Saloniemi, I. and Pasanen, P. (2003) Differences in sea migration between wild and reared Atlantic salmon (Salmo salar L.) in the Baltic Sea. *Fisheries Research* **60**, 333-343.

Kallio-Nyberg, I., Jutila, E., Saloniemi, I. and Jokikokko, E. (2004) Association between environmental factors, smolt size and the survival of wild and reared Atlantic salmon from the Simojoki River in the Baltic Sea. *J. Fish Biol.* **65**, 122-134.

Kallio-Nyberg, I. and Koljonen, M.-L. (1997) The genetic consequence of hatchery-rearing on life-history traits of the Atlantic salmon (Salmo salar L.): a comparative analysis of sea-ranched salmon with wild and reared parents. *Aquaculture* **153**, 207-224.

Kallio-Nyberg, I. and Koljonen, M.-L. (1999) Sea migration patterns in the Atlantic salmon: a comparative study of two stocks and their hybrids. *Boreal Environment Research* **4**, 163-174.

Kallio-Nyberg, I., Koljonen, M.-L. and Saloniemi, I. (2000) Effect of maternal and paternal line on spatial and temporal marine distribution in Atlantic salmon. *Animal Behaviour* **60**, 377-384.

Kallio-Nyberg, I., Peltonen, H. and Rita, H. (1999) Effects of stock-specific and environmental factors on the feeding migration of Atlantic salmon (Salmo salar) in the Baltic Sea. *Can. J. Fish. Aquat. Sci.* **56**, 853-861.

Kennedy, B. P., Folt, C. L., Blum, J. D. and Pagechamberlain, C. (1997) Natural isotope markers in salmon. *Nature* **387**, 766-767.

King, T. L., Eackles, M. S. and Letcher, B. H. (2005) Microsatellite DNA markers for the study of Atlantic salmon (Salmo salar) kinship, population structure, and mixed-fishery analysis. *Mol. Ecol. Notes* **5**, 130.

King, T. L., Kalinowski, S. T., Schill, W. B., Spidle, A. P. and Lubinski, B. A. (2001) Population structure of Atlantic salmon (Salmo salar L.): a range-wide perspective from microsatellite DNA variation. *Molecular Ecology* **10**, 807-821.

Klaue, A., Kennedy, B., Blum, J., Folt, C. and Lohmann, K. (2000) Sr isotope markers in otolith growth increments of Atlantic salmon. *Journal of Conference Abstracts* **5**(2), 590-590.

Koljonen, M.-L. (1995) Distinguishing between resident and migrating Atlantic salmon (*Salmo salar*) stocks by genetic stock composition analysis. *Can. J. Fish. Aquat. Sci.* **52** (4), 665-674.

Koljonen, M.-L. (2001) Conservation goals and fisheries management units for Atlantic salmon in the Baltic Sea area. *J. Fish Biol.* **59**(Suppl. A), 269-288.

Koljonen, M.-L. (2004) Changes in stock composition of annual Atlantic salmon catches in the Baltic Sea on basis of DNA-microsatellite data and Bayesian estimation. ICES C M 2004/Stock Identification Methods/EE:08.

Koljonen, M.-L., Jansson, H., Paaver, T., Vasin, O. and Koskiniemi, J. (1999) Phylogeographic lineages and differentiation pattern of Atlantic salmon (Salmo salar) in the Baltic Sea with management implications. *Can. J. Fish. Aquat. Sci.* **56**, 1766-1780.

Koljonen, M.-L. and McKinnell, S. (1996) Assessing seasonal changes in stock composition of Atlantic salmon catches in the Baltic Sea with genetic stock identification. *J. Fish Biol.* **49**, 998-1018.

Koljonen, M.-L. and Pella, J. J. (1997) The advantage of using smolt age with allozymes for assessing wild stock contributions to Atlantic salmon catches in the Baltic Sea. *ICES J. Mar. Sci.* **54**, 1015-1030.

Koljonen, M.-L., Tähtinen, J., Säisä, M. and Koskiniemi, J. (2002) Maintenance of genetic diversity of Atlantic salmon (Salmo salar) by captive breeding programmes and the geographic distribution of microsatellite variation. *Aquaculture* **212**, 69-92.

Koljonen, M. and McKinnell, S. (1996) Assessing seasonal changes in stock composition of Atlantic salmon catches in the Baltic Sea with genetic stock identification. *J Fish Biol* **49**, 998-1018.

Koskinen, M. T., Hirvonen, H., Landry, P. and Primmer, C. R. (2004) The benefits of increasing the number of microsatellites utilized in genetic population studies: an empirical perspective. *Hereditas* **141**, 61-67.

L'Abée-Lund, J. H., Vøllestad, L. A. and Beldring, S. (2004) Spatial and temporal variation in the grilse proportion of Atlantic salmon in Norwegian rivers. *Trans Am Fish Soc* **133**, 743-761.

Landry, C. and Bernatchez, L. (2001) Comparative analysis of population structure across environments and geographical scales at major histocompatibility complex and microsatellite loci in Atlantic salmon (Salmo salar). *Molecular Ecology* **10**, 2525-2539.

Lapi, L. A. and Mulligan, T. J. (1981) Salmon stock identification using a microanalytic technique to measure elements present in the freshwater growth region of scales. *Can. J. Fish. Aquat. Sci.* **38**, 744-751.

Letcher, B. H. and King, T. L. (1999) Targeted stock identification using multilocus genotype `familyprinting´. *Fisheries Research* **43**, 99-111.

MacCrimmon, H. R. and Claytor, R. R. (1985) Meristic and morphometric identity of Baltic stocks of Atlantic salmon (*Salmo salar*). *Can. J. Zool.* **63**, 2032-2037.

MacCrimmon, H. R. and Claytor, R. R. (1986) Possible use of taxonomic characters to identity Newfoundland and Scottish stocks of Atlantic salmon, *Salmo salar* L. *Aquaculture and Fisheries Management* **17**, 1-17.

MacFarlane, A. J., Sutherland, D. L., Goff, T. R. and Gilmour, K. M. (2002) In vivo photostabilization of tetracycline hydrochloride by reduced glutathione for use an ecological marker in Salmo salar. *Aquaculture* **213**(1-4), 253-236.

Manel, S., Gaggiotti, O. E. and Waples, R. S. (2004) Assignment methods: matching biological questions with appropriate techniques. *TREE*.

McConnell, S., Hamilton, L., Morris, D., Cook, D., Paquet, D., Bentzen, P. and Wright, J. (1995) Isolation of salmonid microsatellite loci and their application to the population genetics of Canadian east coast stocks of Atlantic salmon. *Aquaculture* **137**(1-4), 19-30.

McConnell, S. K., O'Reilly, P., Hamilton, L., Wright, J. M. and Bentzen, P. (1995) Polymorphic microsatellite loci from Atlantic salmon (Salmo salar): genetic differentiation of North America and European populations. *Can. J. Fish. Aquat. Sci.* **52**, 1863-1872.

McConnell, S. K. J., Ruzzante, D. E., O'Reilly, P. T., Hamilton, L. and Wright, J. M. (1997) Microsatellite loci reveal highly significant genetic differentiation among Atlantic salmon (*Salmo salar* L.) stocks from the east coast of Canada. *Molecular Ecology* **6**(11), 1075-1089.

Metcalfe, N. B. and Thorpe, J. E. (1990) Determinants of geographical variation in the age of seaward-migrating salmon, Salmo salar. *J Anim Ecol* **59**, 135-145.

Mills, D. (1989). Ecology and Management of Atlantic salmon. London: Chapman & Hall.

Mills, D., (ed.). (1993). Salmon in the Sea and New Enhancement Strategies. London: Fishing News Books.

Mills, D., (ed.). (1999). The Ocean Life of Atlantic Salmon. Oxford, UK: Fishing News Books.

Mjølnerød, I. B., Refseth, U. H. and Hindar, K. (1999) Spatial association of genetically similar Atlantic salmon juveniles and sex bias in spatial patterns in a river. *J. Fish Biol.* **55**, 1-8.

Mjølnerød, I. B., Refseth, U. H., Karlsen, E., Balstad, T., Jakobsen, K. S. and Hindar, K. (1997) Genetic differences between two wild and one farmed population of Atlantic salmon (*Salmo salar*) revealed by three classes of genetic markers. *Hereditas* **127**, 239-248.

Moen, K., Høyheim, B., Munck, H. and Gomez-Raya, L. (2004) A linkage map of Atlantic salmon (*Salmo salar*) reveals an uncommonly large difference in recombination rate between the sexes. *Anim. Gen.* **35**(2), 81-92.

Moran, P. (2002) Current conservation genetics: building an ecological approach to the synthesis of molecular and quantitative genetic methods. *Ecol Freshw Fish* **11**(1), 30-55.

Moriya, S., Urawa, S., Suzuki, O., Urano, A. and Abe, S. (2004) DNA microarray for rapid detection of mitochondrial DNA happlotypes of chum salmon. *Marine Biotechnology* **6**(5), 430-434.

Mortensen, D. G., Wertheimer, A. C., Maselko, J. M. and Taylor, S. G. (2002) Survival and straying of Auke Creek, Alaska, pink salmon marked with coded wire tags and thermally induced otolith marks. *Trans. Am. Fish. Soc.* **131** (1), 14-26.

Mulligan, T. J., Lapi, L. A., Kieser, R., Yamada, S. B. and Duewer, D. L. (1983) Salmon stock identification based on elemental composition of vertebrae. *Can. J. Fish. Aquat. Sci.* **40**, 215-229.

Muncy, R. J., Parker, N. C. and Poston, H. A. (1990) Inorganic chemical marks induced in fish. *Am Fish Soc Symp* **7**, 541-546.

Neff, B. D., Fu, P. and Gross, M. R. (2000) Microsatellite multiplexing in fish. *Trans. Am. Fish. Soc.* **129** (2), 584-593.

Neff, B. D. and Gross, M. R. (2001) Microsatellite evolution in vertebrates: inferences from AC dinucleotide repeats. *Evolution* **55** (9), 1717-1733.

Neff, B. D., Repka, J. and Gross, M. R. (2000) Parentage analysis with incomplete sampling of candidate parents and offspring. *Molecular Ecology* **9**, 515-528.

Nelson, R. J., Wood, C. C., Cooper, G., Smith, C. and Koop, B. (2003) Population structure of sockeye salmon of the central coast of British Columbia: implications for recovery planning. *North American Journal of Fisheries Management* **23**(3), 703-720.

Nielsen, E. E., Hansen, M. M. and Loeschcke, V. (1997) Analysis of microsatellite DNA from old scale samples of Atlantic salmon Salmo salar: a comparison of genetic composition over 60 years. *Molecular Ecology* **6**, 487-492.

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

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

Norris, A. T., Bradley, D. G. and Cunningham, E. P. (2000) Parentage and relatedness determination in farmed Atlantic salmon (*Salmo salar*) using microsatellite markers. *Aquaculture* **182**(1-2), 73-83.

Nyman, O. L. and Pippy, J. H. C. (1972) Differences in Atlantic salmon, *Salmo salar*, from North America and Europe. *Journal of the Fisheries Research Board of Canada* **29**, 179-185.

O'Connell, M. and Wright, J. M. (1997) Microsatellite DNA in fishes. *Rev. Fish Biol. Fish.* **7**, 331-363.

O'Reilly, P. and Wright, J. M. (1995) The evolving technology of DNA fingerprinting and its application to fisheries and aquaculture. *J. Fish Biol.* **47**, 29-55.

O'Reilly, P. T., Hamilton, L. C., McConnell, S. K. and Wright, J. M. (1996) Rapid analysis of genetic variation in Atlantic salmon (Salmo salar) by PCR multiplexing of dinucleotide and tetranucleotide microsatellites. *Can. J. Fish. Aquat. Sci.* **53**, 2292-2298.

O'Reilly, P. T., Herbinger, C. and Wright, J. M. (1998) Analysis of parentage determination in Atlantic salmon (Salmo salar) using microsatellites. *Anim. Gen.* **29**, 363-370.

O´Reilly, P. T., Hamilton, L. C., McConnell, S. K. and Wright, J. M. (1996) Rapid analysis of genetic variation in Atlantic salmon (Salmo salar) by PCR multiplexing of dinucleotide and tetranucleotide microsatellites. *Can J Fish Aquat Sci* **53**, 2292-2298.

Økland, F., Erkinaro, J., Moen, K., Niemelä, E., Fiske, P., McKinley, R. S. and Thorstad, E. B. (2001) Return migration of Atlantic salmon in the River Tana: phases of migratory behaviour. *J. Fish Biol.* **59**, 862-874.

Olsen, J. B., Habicht, C., Reynolds, J. and Seeb, J. E. (2004) Moderately and highly polymorphic microsatellites provide discordant estimates of population divergence in sockeye salmon, Oncorhynchus nerka. *Env. Biol. Fishes* **69**, 261-273.

Olsén, K. H., Petersson, E., Ragnarsson, B., Lundqvist, H. and Järvi, T. (2004) Downstream migration in Atlantic salmon (Salmo salar) smolt sibling groups. *Can. J. Fish. Aquat. Sci.* **61**, 328-331.

Parrish, D. L., Behnke, R. J., Gephard, S., McCormick, S. D. and Reeves, G. H. (1998) Why aren´t there more Atlantic salmon (*Salmo salar*)? *Can. J. Fish. Aquat. Sci.* **55**, 281-287.

Paterson, S., Piertney, S. B., Knox, D., Gilbey, J. and Verspoor, E. (2004) Characterization and PCR multiplexing of novel highly variable tetranucleotide Atlantic salmon (Salmo salar L.) microsatellites. *Mol. Ecol. Notes* **4**, 160-162.

Poole, W. R., Webb, J. H., Matthews, M. A. and Youngson, A. F. (2000) Occurrence of canthaxanthin in Atlantic salmon, Salmo salar L., fry in Irish rivers as a indicator of escaped farmed salmon. *Fish Mgmt Ecol* **7**(5), 377-385.

Potvin, C. and Bernatchez, L. (2001) Lacustrine spatial distribution of landlocked Atlantic salmon populations assessed across generations by multilocus individual assignment and mixed-stock analyses. *Molecular Ecology* **10**, 2375-2388.

Presa, P. and Guyomard, R. (1996) Conservation of microsatellites in three species of salmonids. *J. Fish Biol.* **49**, 1326-1329.

Quinn, T. P., Volk, E. C. and Hendry, A. P. (1999) Natural otolith microstructure patterns reveal precise homing to natal incubation sites by sockeye salmon (*Oncorhynchus nerka*). *Can. J. Zool.* **77**, 766-775.

Reddin, D. G. and Friedland, K. D. (1999) A history of identification to continent of origin of Atlantic salmon (*Salmo salar* L.) at West Greenland, 1969-1997. *Fisheries Research* **43**, 221-235.

Reddin, D. G., Helbig, J., Thomas, A., Whitehouse, B. G. and Friedland, K. D. (1997). Survival of Atlantic salmon (*Salmo salar* L.) and its relation to marine climate. In

Managing Wild Atlantic Salmon. New Challenges- New Techniques: 24-49. Whoriskey, F. G., Jr. and Whelan, K. F. (Eds.). Atlantic Salmon Federation.

Reddin, D. G., Stansbury, D. E. and Short, P. B. (1988) Continent of origin of Atlantic salmon (*Salmo salar* L.) at West Greenland. *Journal du Conseil* **44**, 180-188.

Reeves, G. H., Benda, L. E., Burnett, K. M., Bisson, P. A. and Sedell, J. R. (1995) A disturbance-based ecosystem approach to maintaining and restoring freshwater habitats of evolutionarily significant units of anadromous salmonids in the Pacific northwest. *Am Fish Soc Symp* **17**, 334-349.

Reynolds, J. H. and Templin, W. D. (2004) Comparing mixture estimates by parametric bootstrapping likelihood ratios. *J. Agri. Biol. Env. Stat.* **9**(1), 57-74.

Reynolds, J. H. and Templin, W. D. (2004) Detecting specific populations in mixtures. *Env. Biol. Fishes* **69**, 233-243.

Ruzzante, D. E., Hansen, M. M. and Meldrup, D. (2001) Distribution of individual inbreeding coefficients, relatedness and influence of stocking on native anadromous brown trout (Salmo trutta) population structure. *Molecular Ecology* **10**, 2107-2128.

Ryynänen, H. and Primmer, C. R. (2004) Distribution of genetic variation in the growth hormone 1 gene in Atlantic salmon (*Salmo salar*) populations from Europe and North America. *Molecular Ecology* **13**, 3857-3869.

Säisä, M., Koljonen, M.-L. and Tähtinen, J. (2003) Genetic changes in Atlantic salmon stocks since historical times and the effective population size of a long-term captive breeding programme. *Cons. Genetics* **4**, 613-627.

Sánchez, J. A., Clabby, C., Ramos, D., Blanco, G., Flavin, F., Vázquez, E. and Powell, R. (1996) Protein and microsatellite single locus variability in Salmo salar L. (Atlantic salmon). *Heredity* **77**(4), 423-432.

Sánchez, J. A., Clabby, C., Ramos, D., Blanco, G., Flavin, F., Vázquez, E. and Powell, R. (1996) Protein and microsatellite single locus variability in Salmo salar L. (Atlantic salmon). *Heredity* **77**, 423-432.

Schaffer, W. M. and Elson, P. F. (1975) The adaptive significance of variations in life history among local populations of Atlantic salmon in north America. *Ecology* **56**, 577-590.

Seeb, J. E., Habicht, C., Olsen, J. B., Bentzen, P. and Seeb, L. W. (1998). Allozyme, mtDNA, and microsatellite variants describe structure of populations of pink and sockeye salmon in Alaska.

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**(1-3), 45-78.

Shaklee, J. B., Phelps, S. R. and Salini, J. (1990). Analysis of fish stock structure and mixed-stock fisheries by the electrophoretic characterization of allelic isozymes. In

Electrophoretic and Isoelectric Focusing Techniques in Fisheries Management: 173-196. Whitmore, D. H. (Ed.). Boca Raton: CRC Press, Inc.

Shearer, W. M. (1992). The Atlantic Salmon. Natural History, Exploitation and Future Management. London: Fishing News Books.

Skaala, Ø., Høyheim, B., Glover, K. and Dahle, G. (2004) Microsatellite analysis in domesticated and wild Atlantic salmon (Salmo salar L.): allelic diversity and identification of individuals. *Aquaculture* **240**, 131-143.

Slettan, A., Olsaker, I. and Lie, Ø. (1993) Isolation and characterization of variable (GT)n repetitive sequences from Atlantic salmon, *Salmo salar* L. *Anim. Gen.* **24**, 195-197.

Slettan, A., Olsaker, I. and Lie, Ø. (1995) Atlantic salmon, Salmo salar, microsatellites at the SSOSL25, SSOSL85, SSOSL311, SSOSL417 loci. *Anim. Gen.* **26**, 281-282.

Slettan, A., Olsaker, I. and Lie, Ø. (1996) Polymorphic Atlantic salmon, Salmo salar L., microsatellites at the SSOSL438, SSOSL439 and SSOSL444 loci. *Anim. Gen.* **27**, 57-58.

Slettan, A., Olsaker, I. and Lie, Ø. (1997) Segregation studies and linkage analysis of Atlantic salmon microsatellites using haploid genetics. *Heredity* **78**, 620-627.

Smith, B. R., Herbinger, C. M. and Merry, H. R. (2001) Accurate partition of individuals into full-sib families from genetic data without parental information. *Genetics* **158**, 1329-1338.

Smouse, P. E. and Chevillon, C. (1998) Analytical aspects of population-specific DNA fingerprinting for individuals. *Journal of Heredity* **89**(2), 143-150.

Spidle, A. P., Kalinowski, S. T., Lubinski, B. A., Perkins, D. L., Beland, K. F., Kocik, J. F. and King, T. L. (2003) Population structure of Atlantic salmon in Maine with reference to populations from Atlantic Canada. *Trans Am Fish Soc* **132**(2), 196-209.

Spidle, A. P., Schill, W. B., Lubinski, B. A. and King, T. L. (2001) Fine-scale population structure in Atlantic salmon from Maine's Penobscot River drainage. *Cons. Genetics* **2**, 11-24.

Ståhl, G. (1981) Genetic differentiation among natural populations of Atlantic salmon (Salmo salar) in Northern Sweden. *Ecol Bull* **34**, 95-105.

Ståhl, G. (1987). Genetic population structure of Atlantic salmon. In Population Genetics and Fisheries Management: 121-140. Ryman, N. and Utter, F. (Eds.). Seattle: University of Washington Press.

Taggart, J. B., Verspoor, E., Galvin, P. T., Morán, P. and Ferguson, A. (1995) A minisatellite DNA marker for discriminating between European and North American Atlantic salmon (Salmo salar). *Can. J. Fish. Aquat. Sci.* **52**, 2305-2311.

Tessier, N. and Bernatchez, L. (1999) Stability of population structure and genetic diversity across generations assessed by microsatellites among sympatric populations of landlocked Atlantic salmon (Salmo salar L.). *Molecular Ecology* **8**, 169-179.

Tessier, N., Bernatchez, L., Presa, P. and Angers, B. (1995) Gene diversity analysis of mitochondrial DNA, microsatellites and allozymes in landlocked Atlantic salmon. *J. Fish Biol.* **47**, 156.

Thorrold, S. R., Campana, S. E., Jones, C. M., McLaren, J. W. and Lam, J. W. H. (1998) Trace element signatures in otoliths record natal river of juvenile American shad (*Alosa sapidissima*). *Limnol Oceanogr* **43**, 1826-1835.

Tonteri, A., Titov, S., Veselov, A., Zubchenko, A., Koskinen, M. T., LesbarrŠres, D., Kaluzchin, S., Bakhmet, I., Lumme, J. and Primmer, C. R. (2005) Phylogeography of anadromous and non-anadromous Atlantic salmon (*Salmo salar*) from northern Europe. **42** (1), 1-22.

Vasemägi, A., Gross, R., Paaver, T., Koljonen, M.-L. and Nilsson, J. (2005) Extensive immigration from compensatory hatchery releases into wild Atlantic salmon population in the Baltic sea: spatio-temporal analysis over 18 years. *Heredity*.

Vasemägi, A., Nilsson, J. and Primmer, C. R. (2005) Expressed sequence tag-linked microsatellites as a source of gene-associated polymorphisms for detecting signatures of divergent selection in Atlanticsalmon (Salmo salar L.). *Mol Biol Evol* **22**(4), 1067-1076.

Verspoor, E. (1986) Spatial correlation of transferrin allele frequencies in Atlantic salmon (Salmo salar) populations from North America. *Can. J. Fish. Aquat. Sci.* **43**, 1074-1078.

Verspoor, E. (1997) Genetic diversity among Atlantic salmon (Salmo salar L.) populations. *ICES Journal of Marine Science* **54**, 965-973.

Verspoor, E., Beardmore, J. A., Consuegra, S., Garcia de Leániz, C., Hindar, K., Jordan, W. C., Koljonen, M.-L., Mahkrov, A. A., Paaver, T., Sánchez, J. A., Skaala, Ø., Titov, S. and Cross, T. F. (2005) Genetic protein variation in the Atlantic salmon: population insights gained from 40 years of research. *Journal of Fish Biology* **67**(Suppl. A).

Verspoor, E. and García de Leániz, C. (1997) Stocking success of Scottish Atlantic salmon in two Spanish rivers. *J. Fish Biol.* **51**, 1265-1269.

Verspoor, E., O'Sullivan, M., Arnold, A. L., Knox, D. and Amiro, P. G. (2002) Restricted matrilineal gene flow and regional differentiation among Atlantic salmon (Salmo salar L.) populations within the Bay of Fundy, eastern Canada. *Heredity* **89**, 465-472.

Villanueva, B., Verspoor, E. and Visscher, P. M. (2002) Parental assignment in fish using microsatellite genetic markers with finite numbers of parents and offspring. *Anim. Gen.* **33**(1), 33-41.

Waldman, J. R. (1999) The importance of comparative studies in stock analysis. *Fisheries Research* **43**(1-3), 237-246.

Ward, R. D. (2000) Genetics in fisheries management. *Hydrobiologia* **420**, 191-201.

Webb, J. H., Fryer, R. J., Taggart, J. B., Thompson, C. E. and Youngson, A. F. (2001) Dispersion of Atlantic salmon (*Salmo salar*) fry from competing families as revealed by DNA profiling. *Can. J. Fish. Aquat. Sci.* **58**, 2386-2395.

Wennevik, V., Skaala, Ø., Titov, S., Studyonov, I. and Nævdal, G. (2004) Microsatellite variation in populations of Atlantic salmon from North Europe. *Env. Biol. Fishes* **69**, 143-152.

White, B. A. and Shaklee, J. B. (1991) Need for replicated electrophoretic analyses in multiagency genetic stock identification (GSI) programs - examples from a Pink salmon (*Oncorhynchus gorbuscha*) GSI fisheries study. *Can. J. Fish. Aquat. Sci.* **48** (8), 1396-1407.

Winans, G. A., Paquin, M. M., Van Doornik, D. M., Baker, B. M., Thornton, P., Rawding, D., Marshall, A., Moran, P. and Kalinowski, S. (2004) Genetic stock identification of steelhead in the Columbia River basin: an evaluation of different molecular markers. *North American Journal of Fisheries Management* **24**(2), 672-685.

Withler, R. E., Candy, J. R., Beacham, T. D. and Miller, K. M. (2004) Forensic DNA analysis of Pacific salmonid samples for species and stock identification. *Env. Biol. Fishes* **69**, 275-285.

Withler, R. E., Le, K. D., Nelson, R. J., Miller, K. M. and Beacham, T. D. (2000) Intact genetic structure and high levels of genetic diversity in bottlenecked sockeye salmon, *Oncorhynchus nerka*, populations of the Fraser River, British Columbia, Canada. *Can. J. Fish. Aquat. Sci.* **57**, 1985-1998.

WWF (2001). The status of Wild Atlantic Salmon: A river by river assessment. WWF.

Figure 1. Trends (cumulative numbers) in the number of papers containing the terms "Atlantic salmon" and "microsatellites" listed in ISI Web of Science during 1990-2005 and in the number of Atlantic salmon microsatellite sequences deposited in Genbank during the same period.

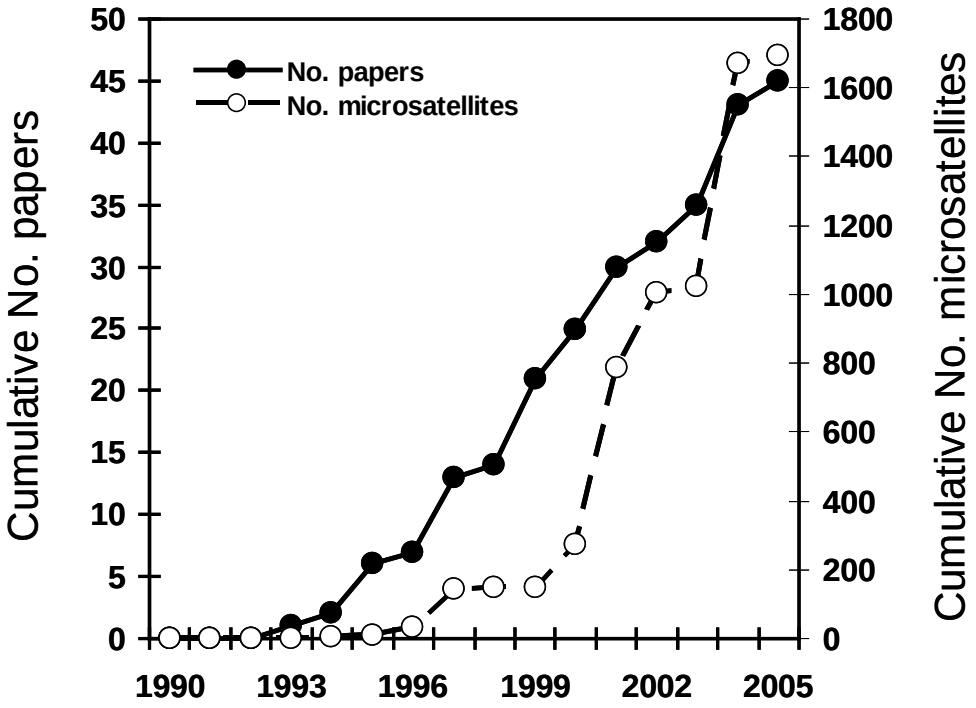

Figure 2. Map showing location of the Atlantic salmon populations screened for microsatellite variation and reviewed in the present study

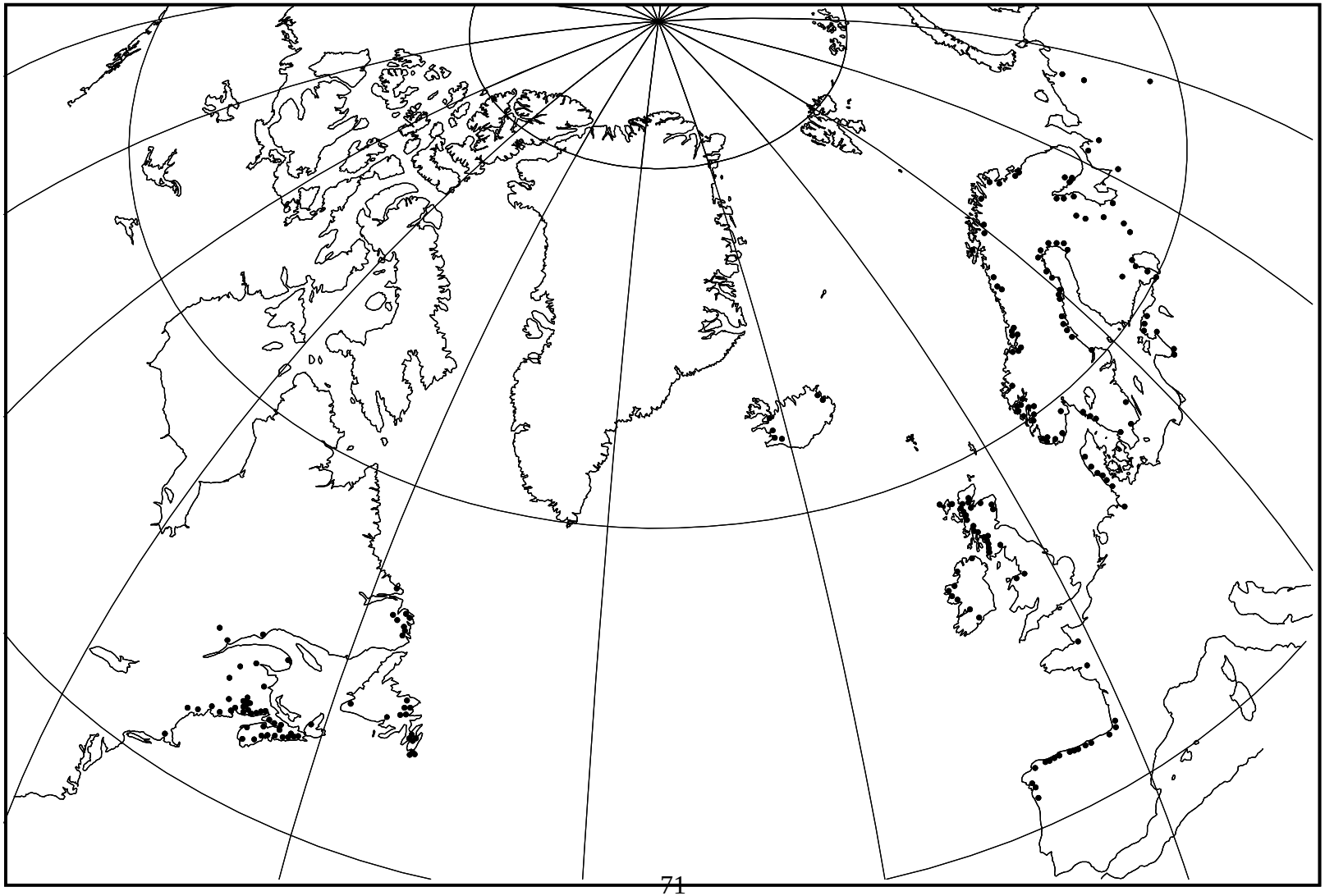

Figure 3. Number of different populations screened with the same microsatellite loci included in the Atlantic salmon GSI database.

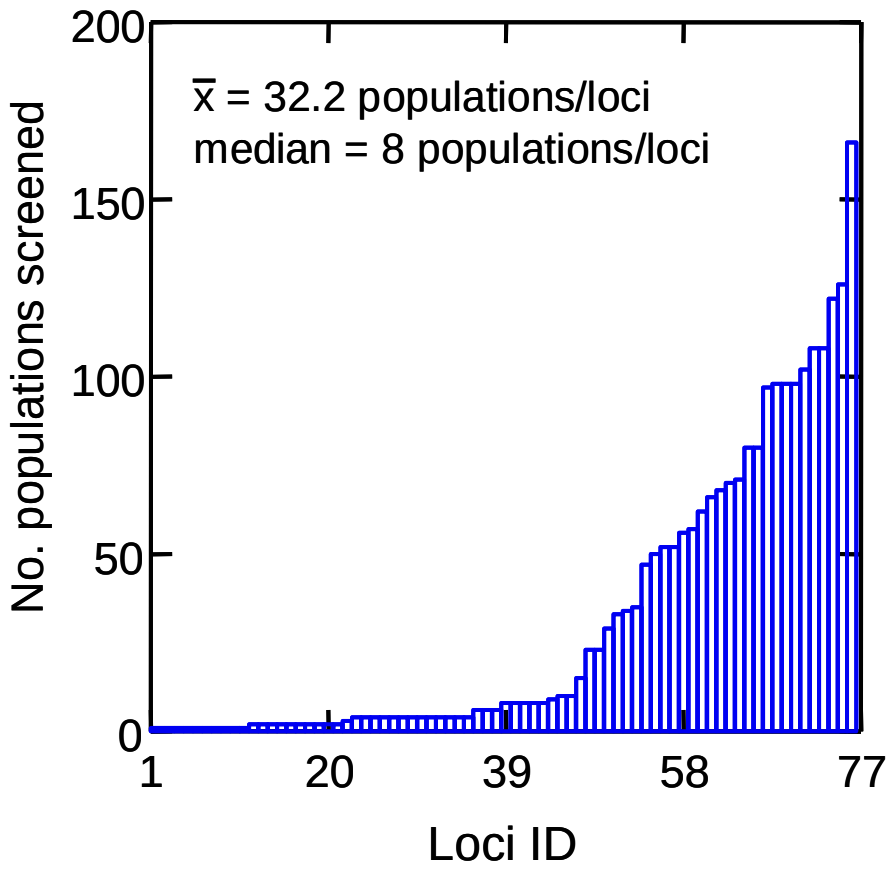

Figure 4. Number of microsatellite loci used to characterize the different populations included in the Atlantic salmon GSI database.

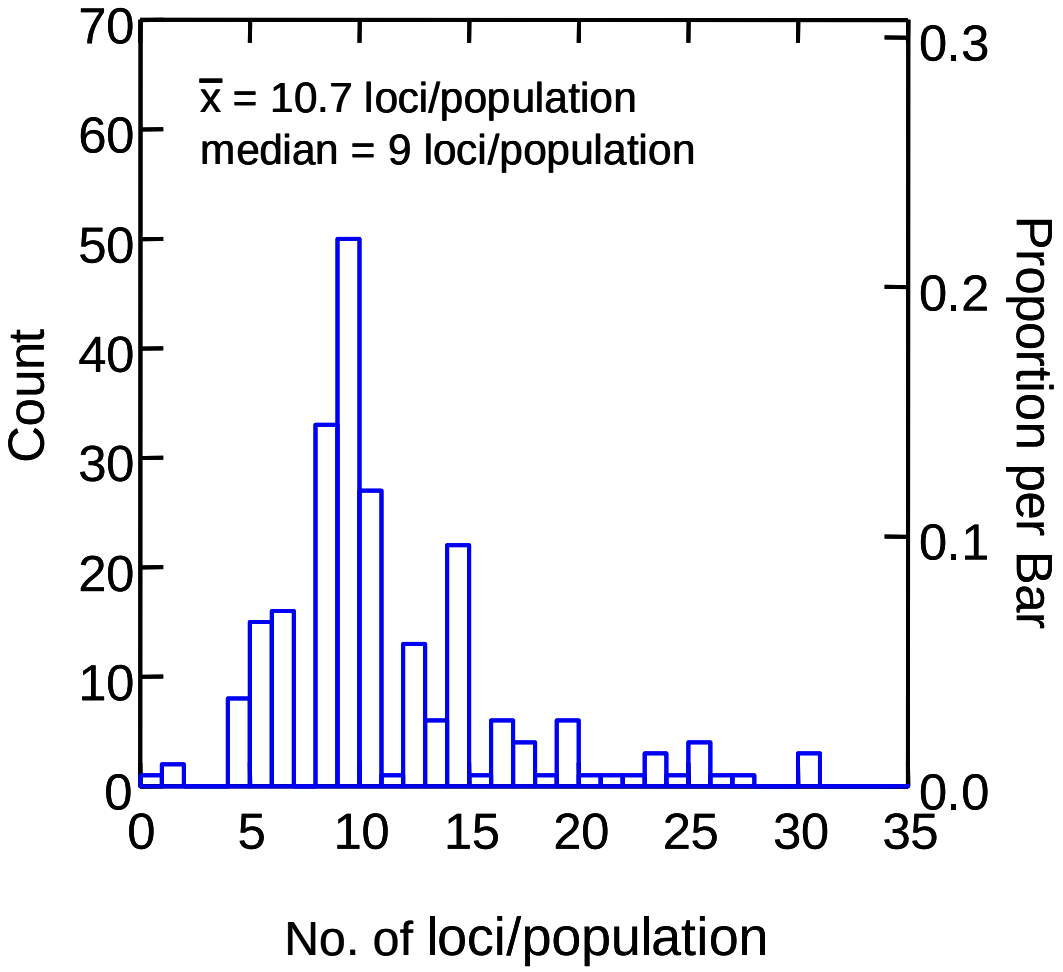

Figure 5. Influence of sample size, number of available loci and number of alleles per locus for choosing microsatellite loci for genetic stock identification of Atlantic salmon.

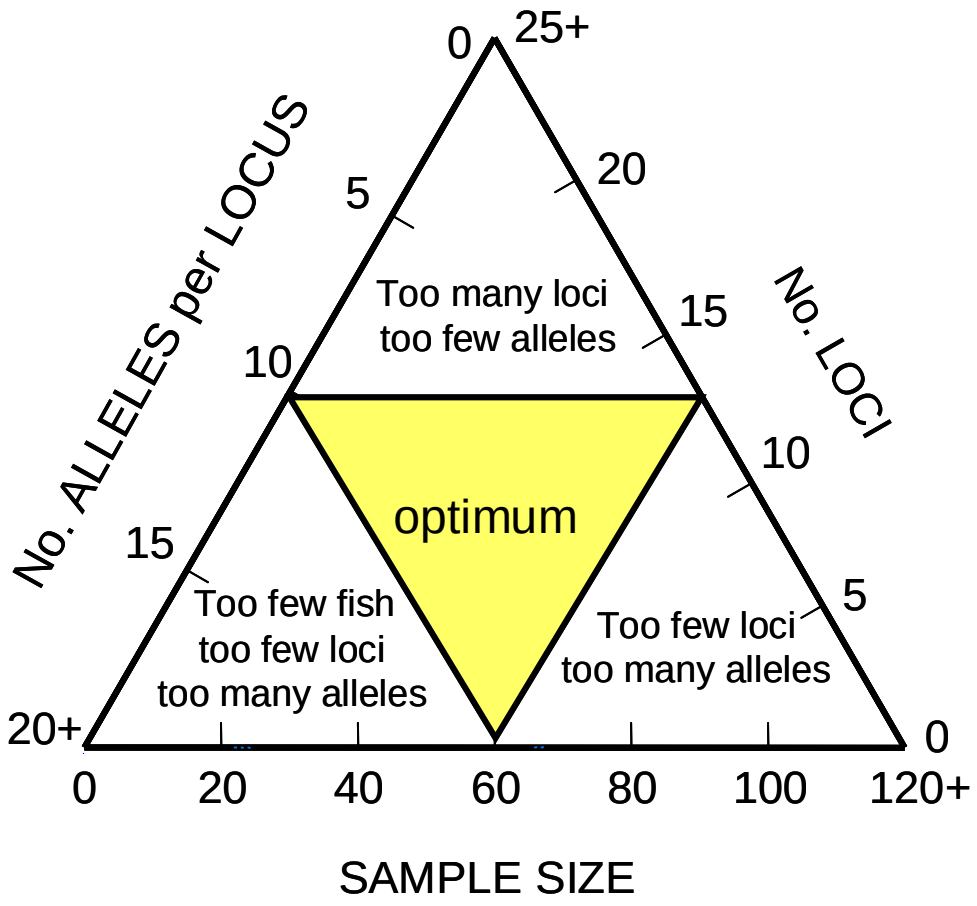

Figure 6. Number of fish screened per population in 77 populations included in the Atlantic salmon microsatellite GSI database.

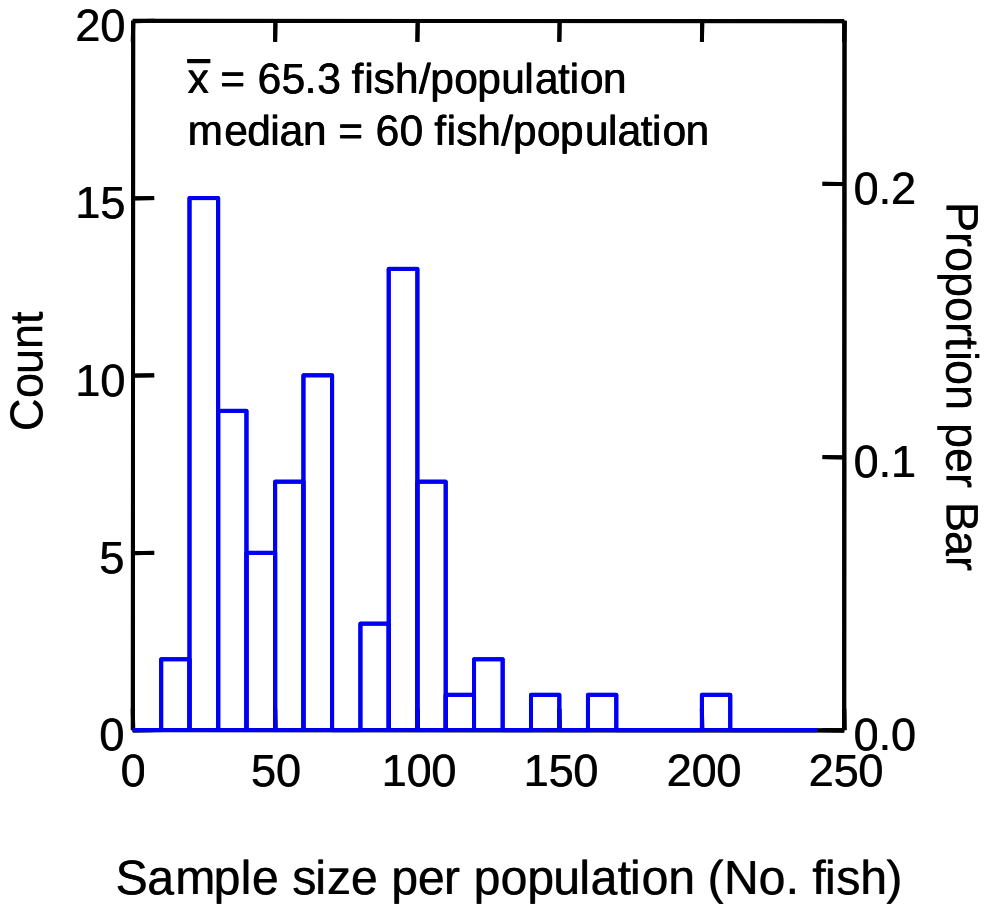

Figure 7. Number of years of data per population in the Atlantic salmon GSI database.

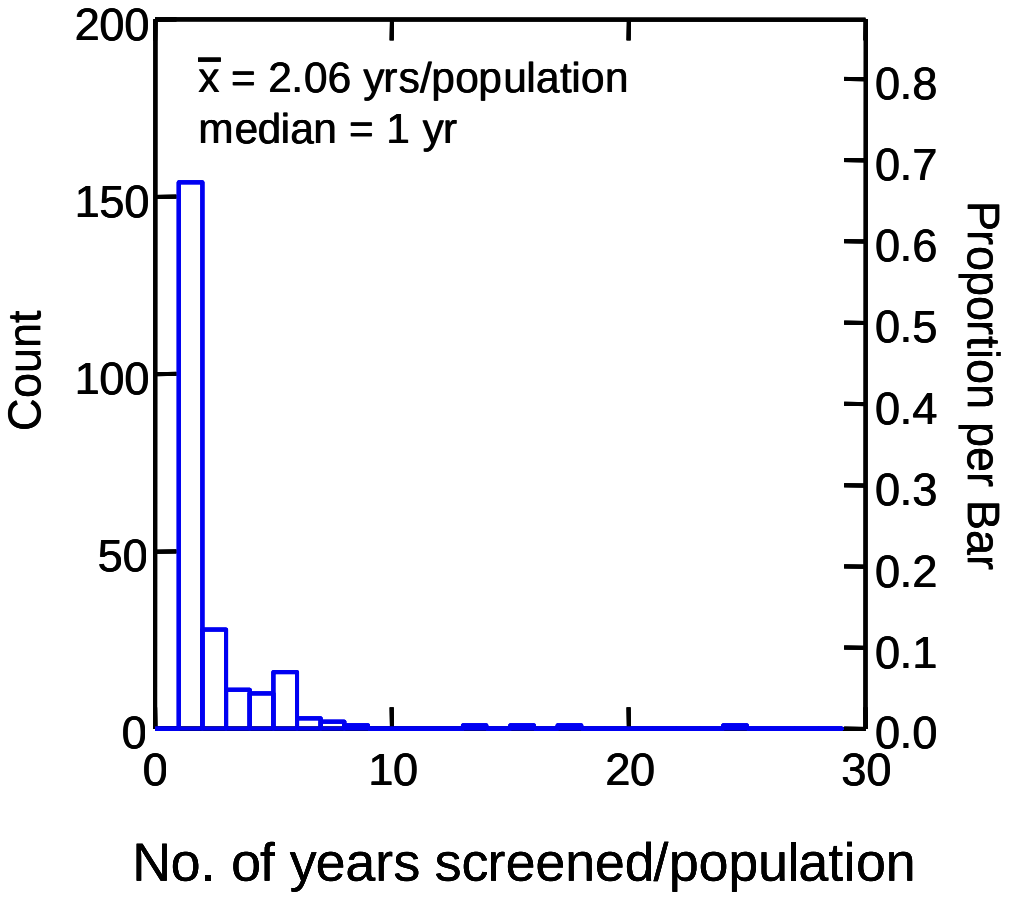

Figure 8. Origin of the samples used to screen the 229 populations of the Atlantic salmon microsatellite GSI database.

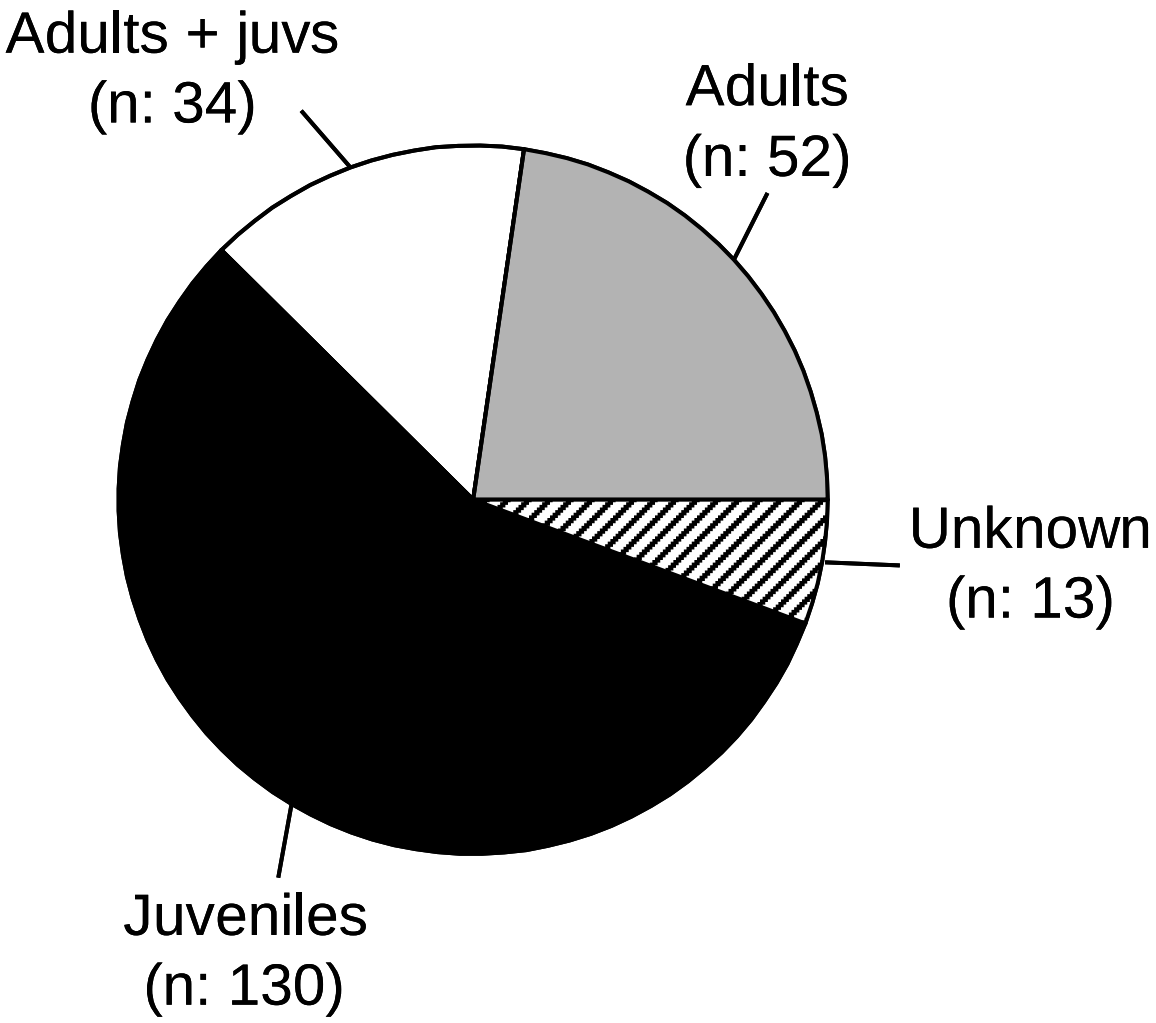

Figure 9. Diagram depicting the two types of qualitative errors in genetic stock identification of salmon populations (denoted by circles) in mixed stock fisheries: failing to detect the contribution of populations that are in fact present (Type I error), and wrongly detecting the contribution of populations that are in fact absent (Type II error). For conservation purposes, mixed stock analysis should be adjusted as to minimise Type I errors since the consequences of failing to detect the exploitation of small, endangered populations in the fishery (denoted by solid circles) are much worse than the consequences of inferring a mistaken contribution.

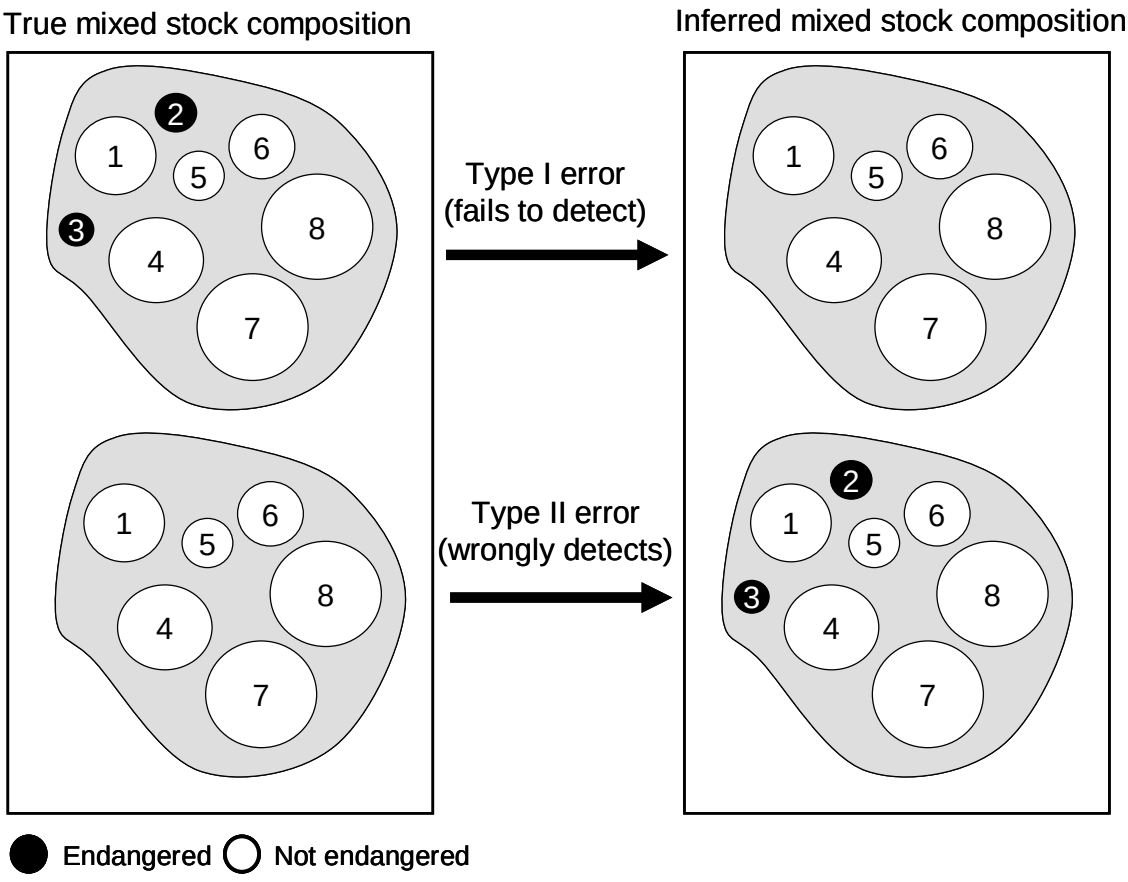

Table 1. Complementary sources of information that can be used in combination with genetic variation at microsatellite loci to increase assignment accuracy in stock identification.

Marker/Data	Reference
1. Additional molecular markers	
- mtDNA	Tonteri et al.(2005), Vasemägi et al. (2005), Gilbey et al. (2005)
- minisatellites	Ferguson et al. (1995), Galvin et al. (1995), Taggart et al. (1995)
- allozymes	Koljonen (1995), Koljonen and McKinnell (1996), Verspoor et al. (2005)
2. Biological tags	
- parasites	Nyman and Pippy (1972), Beverley-Burton and Pippy (1978)
- trace elements	Thorrold et al. (1998), Campana et al. (2000)
- stable isotopes	Kennedy et al. (1997), Klaue et al.(2000), Dempson and Power (2004)
- carotenoid pigments	Craik and Harvey (1987), Craik et al. (1987), Poole et al. (2000)
3. Age structure	Metcalfe and Thorpe (1990), Koljonen and Pella (1997), L'Abée-Lund et al.(2004)
4. Growth patterns in scales or otoliths	Schaffer and Elson (1975), Reddin and Friedland (1999), Friedland and Reddin (2000)
5. Scale and otolith morphology	de Pontual and Prouzet (1988), Reddin et al. (1988), Friedland et al. (1994), Friedland and Reddin (1994)
6. Body morphometrics and meristics	MacCrimmon and Claytor (1985), MacCrimmon and Claytor (1986), Claytor and MacCrimmon (1988), Claytor et al. (1991), Cadrin and Friedland (1999), Cadrin (2000)
7. Population size	Parrish et al. (1998), Einum et al. (2003)
8. Marine distribution	Kallio-Nyberg and Koljonen (1999), Jacobsen et al. (2001), Jutila et al. (2003)
9. Physical tags and induced marks	Muncy et al. (1990), Hansen and Quinn (1998), Jacobsen et al. (2001), MacFarlane et al. (2002), Mortensen et al .(2002)

Table 2. Compilation of the 25 most useful microsatellite loci for genetic stock identification in Atlantic salmon, based on the extent of usage and lack of reported problems with stuttering or null alleles.

Locus name	No. repeat bases	Max Size range	Max no. Alleles	Source
Ssa197	4	135-279	30	(O'Reilly et al., 1996)
Ssa 202	4	200-320	18	(O'Reilly et al., 1996)
Ssa171	4	193-285	32	(Slettan et al., 1995)
SSOSL85	2	176-228	20	(Slettan et al., 1995)
SSOSL311	2	112-192	33	(Slettan et al., 1996)
SSOSL438	2	110-151	17	(O'Reilly et al., 1996)
Ssa85	2	110-181	27	(Paterson et al., 2004)
SSsp2210	4	104-164	15	(Paterson et al., 2004)
SSsp2216	4	202-305	18	(Paterson et al., 2004)
SSsp1605	4	222-254	14	(McConnell et al., 1995)
Ssa 14	2	138-152	4	(McConnell et al., 1995)
Ssa 289	2	107-132	8	(Sánchez et al., 1996)
SsaF43	2	99-131	13	(Paterson et al., 2004)
SSsp2201	4	259-371	21	(Sánchez et al., 1996)
Ssa20-19	2	78-111	9	(Slettan et al., 1995)
SSOSL417	2	155-213	19	(Paterson et al., 2004)
SSspG7	4	112-214	24	(Sánchez et al., 1996)
SsaD-30	2	217-247	10	(Sánchez et al., 1996)
SSLEEI84	2	171-229	28	Unpublished.Genbank U86703
SSLEEN82	2	209-233	13	Unpublished. Genbank U86706
Ssa13-37	4	112-126	4	(King et al., 2005)
SsaD486	4	162-210	7	(King et al., 2005)
SsaD144	4	112-298	35	(Presa and Guyomard, 1996)
SstrP15 (Str15INRA*)	2	210-220	6	(Slettan et al., 1995)
SSOSL25	2	ca. 150-180		

Table 3. Some general suitability criteria for choosing microsatellite loci for genetic stock identification in Atlantic salmon

Criteria	Worst	Best
1. No loci	<6 or >20	10-15
2. No. alleles	<6 or >20	6-14
3. Discrimination power (F_{ST})	Low	High
4. Sample size (No. fish per population)	< 20	≥ 50
5. Base line data (no. pops screened)	Limited	Extensive
6. No. repeat bases	Dinucleotide	Tetranucleotide

7. Fragment size	< 90 bp or >300 bp	90-300 bp
8. Null alleles	Yes	No
9. Stuttering	Yes	No
10. Suitability for multiplexing	Low	High
11. Scoring	Difficult	Easy
12. Mutation rate	High	Low
13. Temporal stability	Unstable	Stable
14. Neutrality	Selected	Neutral