International Atlantic Salmon Research Board

SAG(15)4

Enhancement of a North American Atlantic salmon genetic baseline for individual and stock identification and application of the baseline to historical scales collected at West Greenland

SAG(15)4

Enhancement of a North American Atlantic salmon genetic baseline for individual and stock identification and application of the baseline to historical scales collected at West Greenland

Note: Not to be cited without permission from the author.

April 27, 2015

Project Leader: Ian Bradbury, Research Scientist, Fisheries and Oceans Canada, St. John's, Newfoundland, (709) 772-3869, Ian.bradbury@dfo-mpo.gc.ca

Enhancement of a North American Atlantic salmon genetic baseline for individual and stock identification and application of the baseline to historical scales collected at West Greenland

This draft final report was prepared in fulfilment of a contract to undertake genetic analysis of (1) Atlantic salmon from two rivers in the United States and (2) scales collected at West Greenland for the period prior 1968-1998. The project sponsors were Fisheries and Oceans Canada (DFO) and the International Atlantic Salmon Research Board (IASRB) of the North Atlantic Salmon Conservation Organisation (NASCO)

Contributors to the report were Ian Bradbury (DFO), Tim Sheehan (NOAA), Lorraine Hamilton (DFO)

Background. Multi-sea-winter Atlantic salmon from the western Atlantic migrate to West Greenland where they feed before returning to home waters two or more years later (Reddin 1988; Reddin et al. 2012). Fisheries targeting mixtures of populations have traditionally occurred either during this common feeding period off Greenland or during the return migratory phase of the life cycle (Gauthier-Ouellet et al. 2009; Reddin and Friedland 1999). Failure to identify the composition of these mixed harvests risks the over exploitation and extinction of small and vulnerable populations, the loss of which may threaten the inherent biodiversity of Atlantic salmon and ultimately the stability and persistence of populations and fisheries (Hilborn et al. 2003; Schindler et al. 2010). Recent work by Fisheries and Oceans Canada (DFO) to address this issue has resulted in the establishment of a North American genetic baseline for Atlantic salmon, which enables individuals or catches to be assigned back to their region of origin (Bradbury et al. 2014a; Moore et al. 2014a). However, several regions remain under-represented in the current baseline due to limited sampling and therefore the need for further work remains. In particular the regional group characterizing United States Atlantic salmon is represented by a single sample and likely not representative of the regional diversity present. Also the establishment of genetic baselines provides new opportunities for analysis of archived samples (e.g., Palstra et al. 2007) which can illuminate catch composition in past fisheries and reveal migratory routes for salmon at sea. However, again the utility of archived samples with this baseline has yet to be fully evaluated. Both further refinement of regional groups in the baseline, and testing of the potential use with archived samples (i.e. scales) are required to determine the full potential of this genetic baseline and ultimately meet the needs management and conservation efforts.

Principal goal and sub-goals: The purpose of this project was two-fold. First, to supplement the existing genetic baseline with analysis of additional samples from other rivers/populations currently listed as endangered in the United States. Second, to evaluate the appropriateness of the baseline for assigning historical Atlantic salmon harvested at Greenland from archived scales samples. A better characterization of baseline populations will improve accuracy and confidence in assignments of fishery proportions and the advice provided to managers on stock-specific exploitation rates. Validating the appropriateness of processing techniques and the ability of the contemporary baseline for assigning historical samples will inform researchers and managers of the utility of this approach for evaluating historical changes in stock status and characterization of the West Greenland stock. The over arching objective is to improve our ability to provide scientific advice to DFO-Fisheries Management, International Council for the Exploration of the Seas Working Group on North Atlantic Salmon (ICES WGNAS), and the North Atlantic Salmon Conservation Organization (NASCO), regarding the composition of mixed stock harvests of Atlantic Salmon in the Northwest Atlantic. This information will then be used to assist in the identification and implementation of appropriate management measures to ensure the long-term sustainability of Atlantic salmon stocks harvested in fisheries in Labrador, Saint Pierre et Miquelon, and West Greenland. Salmon exploitation and migration data will also be used to improve Atlantic salmon stock assessments, both regionally and internationally.

Methods

Baseline Samples

Baseline samples encompassed 12409 individuals spanning 194 individual river samples (See Table S1) ranging from Ungava Bay in the north to the Penobscot River in Maine to the south (Figure 1). Data included in the baseline represented a combination of previously analyzed datasets (see Bradbury et al. 2014j; Dionne et al. 2008 for regional analyses and further details) and new data (Table S1) and were collected using an ABI 3130xl (or standardized from ABI 3100 following Gauthier-Ouellet et al. 2009). See Bradbury et al. (2014a) and Moore et al. (2014a) for methods and database details. Two new samples were added to US reporting group and included the Sheepscot (n=119) and the Narraguagus Rivers (n=119). Samples for these two rivers spanned two collection years (2012, 2013).

DNA extraction and genotyping fishery samples

DNA extraction and microsatellite genotyping of all fishery samples were carried out at the Aquatic Biotechnology Laboratory (DFO) and DNA was extracted using the Qiagen DNeasy 96 Blood and Tissue extraction kit (Qiagen) following the guidelines of the manufacturer. DNA was quantified using QuantIT PicoGreen (Life Technologies) and diluted to a final concentration of 10 ng/μL in 10mM Tris (Buffer EB, Qiagen). Microsatellite polymorphisms were scored at 15 loci as follows: Ssa85, Ssa202, Ssa197 (O'Reilly et al. 1996), SSOSL417 (Slettan et al. 1995), SsaD85 (T. King, unpublished), SsaD58, SsaD71, SsaD144, SsaD486 (King et al. 2005), MST-3 (hereafter referred to as U3) (Presa and Guyomard 1996), SSsp2201, SSsp2210, SSsp2215, SSsp2216 and SSspG7 (Paterson et al. 2004). Genotyping of fishery samples follows the methods outlined in Bradbury et al. (2014a); and Bradbury et al. (2014j). In short, loci were multiplexed into three panels either by combining loci amplified individually prior to electrophoresis, or by multiplexing at the PCR stage. The PCR reactions for single locus amplification were set up in a 10 μL volume containing 20 ng DNA, 1X PCR buffer (KCl buffer or $(NH_4)_2SO_4$ (Fermentas) (Table S4)), 1.5-2.5 mM $MgCl_2$ (Fermentas)(Table S4), 0.2 mM dNTP's, 0.1μM of each primer and 0.5 U Taq (Fermentas). For multiplex amplification, the PCR reactions were set up in a 10 μL volume containing 10 ng DNA, 1X Type-it Multiplex PCR Master Mix (from Type-it Microsatellite PCR kit (Qiagen)) and primer mix (Table S4). PCR products were size separated on an AB 3130xl (Life Technologies) capillary electrophoresis system using Gene Scan 500 as the internal size standard (labelled in LIZ (Life Technologies)). The resulting electropherograms were analyzed using Gene Mapper 4.0 (Life Technologies). See Bradbury et al. (2014j) for further details.

Reporting groups were previously identified (e.g., Bradbury et al. 2014a; Moore et al. 2014a) and largely approximate regional clusters identified in landscape analyses of population structure (e.g., Bradbury et al. 2014j; Dionne et al. 2008) and were evaluated for use in mixture analysis here (Figure S1, but see Bradbury et al. (2014a)). In total, 12 reporting groups were used for individual assignment and mixture analysis (Figure 1). Two general approaches for individual assignment and mixture analyses were utilized. First, we used a conditional maximum likelihood method to estimate mixture proportions and assign individuals (Millar 1987) as implemented in the program ONCOR (Kalinowski et al. 2007). Mixture proportions are estimated using the EM algorithm, and genotype probabilities are calculated using the

method of Rannala and Mountain (1997). Individuals in a fishery sample are assigned to populations associated with probabilities of membership >0.70. Second, we used a Bayesian mixture model from Pella and Masuda (2001) as implemented in cBAYES (Neaves et al. 2005) for mixture analysis. In this analysis eight 100 000 iteration Monte Carlo Markov chains were produced, each with starting values set at 0.90. Convergence was assessed using a shrink factor (< 1.2 indicating convergence) and the last 5,000 iterations of each chain were combined and used to calculate stock composition. Both approaches are used for examination of baseline performance, but cBAYES was selected for analysis of fishery samples (see Bradbury et al. 2014a).

Fishery samples

The West Greenland Atlantic salmon fishery was sampled in 2011-2014. Samples were collected by the North Atlantic Salmon Conservation Organization (NASCO) as part of an international sampling program. Fin clips and scale samples were collected from all individuals sampled. Fin clips were stored in 95% ethanol or RNAlater. For all samples, continent of origin was estimated independently using a microsatellite panel and assignment testing (King et al. 2001). Region specific mixture analysis was conducted on samples from 3 to 4 NAFO subdivisions each year with NAFO divisions 1B, 1C, and 1F consistently analyzed (no 1C sample available in 2011), in most years with samples analyzed from 1A and 1D in 2011. Analysis of these contemporary samples was outside of the scope of this fundings, but is reported here for comparison with the historical sample results. In addition to contemporary tissue samples, archived scale samples were available for analysis for 1968, 1978, 1988, 1998 and averaged ~100 scales per year. These scales were selected to test the application of genetic mixed stock analysis to archived samples over a range of years. For each individual sample, 10 scales were selected for DNA extraction as above.

Results and Discussion

Additional baseline samples

Samples from two additional US rivers (Sheepscot (n=119) and Narraguagus (n=119)) were obtained and analyzed. F_{ST} values among the three US samples ranged between 0.005 (PEN-NGR) to 0.011 (NGR-SHP). The revised US reporting group clustered with other southern regions (e.g., Nova Scotia, New Brunswick, Figure 2). Performance of the revised baseline was evaluated using 100% simulations to ensure that the addition of two new samples did not alter accuracy from levels previously reported (Bradbury et al. 2014a). Overall accuracy of mixture analysis was 97.6% across all reporting groups (Figure 3). Mixture analysis accuracy remained high for the US reporting group (~99%, Figure 3). Similarly correct individual assignment to the US reporting group also remained high (~94%, Figure 3). We simulated increasing contributions of US salmon to a fishery to explore detection thresholds for mixture contributions (Figure 4). Here contributions of US salmon were clearly detectable at levels of 0.5-1.0% (Figure 4). The revised baseline was applied to samples of Atlantic salmon from West Greenland from 2011-2014 (Figure 5). In the West Greenland harvest, we analysed 2336 individuals from 2011-2014 of which 342 were identified as European in origin (Figure 5). North American contributions were largely from Labrador, the Gulf of St. Lawrence, and the Gaspe Peninsula Again, this pattern was stable over the four years examined (Figure 6). US contributions to the fishery across all years were generally 1% or less of North American contributions. Estimates of stock composition here are consistent with previous genetic mixed stock analysis for Atlantic salmon from West Greenland using a partial North American baseline (Gauthier-Ouellet et al. 2009).

Accuracy of genetic mixed stock analysis is directly dependent on the representation of wild genetic diversity in the baseline. Unsampled or ghost populations or poorly resolved reporting groups could potentially influence the types of analyses reported here. In instances where the reporting groups represent regional clusters of rivers often supported by environmental associations (Bradbury et al. 2014a; Bradbury et al. 2014j; Dionne et al. 2008), unsampled populations should likely be represented by regional groupings. As the US reporting group has been previously represented by a single sample (Bradbury et al. 2014a; Moore et al. 2014d), it was unclear how accurately this sample reflected regional diversity. The addition of two new US samples ensures regional diversity is better reflected in the

baseline and minimizes the impact of "ghost populations" on mixture estimates and individual assignments. These three rivers together account for 75% of total US production (USASAC 2014) of Atlantic salmon and are thus likely representative of diversity in this reporting group.

Archived scale analysis

In total 420 individuals were analyzed from archived scales from the West Greenland Atlantic Salmon fishery (1968: 94; 1978: 110; 1988: 108; 1998: 108). Genetic analysis of the archived scales revealed successful genotypes in all years examined. 12 of 15 loci examined amplified successfully in the archived scale samples. These included the following loci: Ssa197, Ssa202, Ssa85, SsaD144, SsaD486, SsaD85, Ssosl417, Sssp2210, Sssp2215, Sssp2216, SsspG7, u3. For the 12 loci that were successfully typed, the amount of missing data per individual declined with age of sample, 1968 ~2.5 loci per individual were missing, 1978 ~2.5 loci per individual were missing, 1988 ~0.7 loci per individual were missing, 1998~0.4 loci per individual were missing. Our success with amplifying DNA from archived scale samples is consistent with other work reporting successful amplification from scales archived since the 1950's and 1960's (e.g., Palstra et al. 2007) and suggests retrospective analyses of stock composition and migration patterns are possible.

To demonstrate the utility of archived samples for retrospective genetic mixed stock analysis we analyzed samples from a single NAFO subdivision that was most prevalent across the full time period. Initially, individuals were screened for continent of origin using known origin European individuals (Figure 5) based on independent assignment (King et al. 2005). Assignments to European baseline ranged from 10-20% of the individuals analyzed for the period 1968-1998 (see below). These individuals had been pre-screened to remove European individuals based on scale characteristics, so their presence suggests they were missed by the scale based analysis. Interestingly, recent samples (2012-2014) were not pre-screened and estimates of European contributions were still ~20%, this is consistent with recently declining rates of occurrence of European origin salmon at West Greenland (ICES 2013; Reddin and Friedland 1999). For the North American contributions, regional specific contributions were consistent with contemporary samples (Figure 6) and dominated by Gaspe Peninsula, Southern Gulf of St. Lawrence, and Labrador (Figure 7). A single small sample showed a large contribution of Nova Scotia salmon, but given the sample size (n=28) this is difficult to interpret. The successful analysis of archived scales dating back to the 1960's strongly supports

the possibility of further retrospective analysis of long term changes in the stock composition of salmon harvested in West Greenland, and elsewhere such as Labrador and Newfoundland where large collections of archived material exist from past fisheries. The potential for large scale historical examinations of stock composition from marine samples has the potential to dramatically improve our understanding of marine spatial distributions in salmon, migratory routes, and environmental drivers of stock specific productivity.

Summary

This study has both improved our ability to identify contributions to mixed stock Atlantic salmon fisheries in the Northwest Atlantic and demonstrated the potential of this approach for retrospective analysis of archived scales in the West Greenland fishery. Our further development of the United States reporting group through the addition of new samples ensures that estimates of US contribution are accurate. Given the current endangered status of river stocks within this reporting group, accurate estimates of sources of marine mortality are vital to the persistence of salmon in the region. Moreover, the demonstration that archived scales from 40-50 years ago can provide DNA suitable for mixture analysis presents valuable opportunities for retrospective analysis of migration routes and spatial patterns. Together these advances represent a significant step forward in providing the science needed for the management and conservation of Atlantic Salmon in the Northwest Atlantic.

References

Bradbury, I.R., Hamilton, L.C., Rafferty, S., Meerburg, D., Poole, R., Dempson, J.B., Robertson, M.J., Reddin, D.G., Bourret, V., Dionne, M., Chaput, G., Sheehan, T.F., King, T.L., Candy, J.R., and Bernatchez, L. 2014a. Genetic evidence of local exploitation of Atlantic salmon in a coastal subsistence fishery in the Northwest Atlantic. Can. J. Fish. Aquat. Sci. **72**(1): 83-95.

Bradbury, I.R., Hamilton, L.C., Robertson, M.J., Bourgeois, C.E., Mansour, A., and Dempson, J.B. 2014j. Landscape structure and climatic variation determine Atlantic salmon genetic connectivity in the northwest Atlantic. Can. J. Fish. Aquat. Sci. **71**(2): 246-258.

Dionne, M., Caron, F., Dodson, J.J., and Bernatchez, L. 2008. Landscape genetics and hierarchical genetic structure in Atlantic salmon: the interaction of gene flow and local adaptation. Mol. Ecol. **17**(10): 2382-2396.

Gauthier-Ouellet, M., Dionne, M., Caron, F., King, T.L., and Bernatchez, L. 2009. Spatiotemporal dynamics of the Atlantic salmon (*Salmo salar*) Greenland fishery inferred from mixed-stock analysis. Can. J. Fish. Aquat. Sci. **66**(12): 2040-2051.

Hilborn, R., Quinn, T.P., Schindler, D.E., and Rogers, D.E. 2003. Biocomplexity and fisheries sustainability. Proc. Natl. Acad. Sci. U. S. A. **100**(11): 6564-6568.

ICES. 2013. Report of the Working Group on North Atlantic Salmon (WGNAS), 3–12 April 2012, Copenhagen, Denmark. ICES CM.

Kalinowski, S.T., Manlove, K.R., and Taper, M.L. 2007. ONCOR: software for genetic stock identification. Montana State University, Bozeman.

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 analyses. Mol. Ecol. Notes **5**(1): 130-132.

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. Mol. Ecol. **10**(4): 807-821.

Millar, R.M. 1987 Maximum likelihood estimation of mixed stock fishery composition. Can. J. Fish. Aquat. Sci. **44**: 583–590.

Moore, J.-S., Bourret, V., Dionne, M., Bradbury, I., O'Reilly, P., Kent, M., Chaput, G., and Bernatchez, L. 2014a. Conservation genomics of anadromous Atlantic salmon across its North American range: outlier loci identify the same patterns of population structure as neutral loci. Mol. Ecol. **23**(23): 5680-5697.

Moore, J.-S., Bourret, V., Dionne, M., Bradbury, I., O'Reilly, P., Kent, M., Chaput, G., and Bernatchez, L. 2014d. Conservation genomics of anadromous Atlantic salmon across its North American range: outlier loci identify the same patterns of population structure as neutral loci. Mol. Ecol. **23**.

Neaves, P.I., Wallace, C.G., Candy, J.R., and Beacham, T.D. 2005. CBayes: computer program for mixed stock analysis of allelic data. pac.dfo-mpo.gc.ca/sci/mgl/Cbayes_e.htm.

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.

Palstra, F.P., O'Connell, M.F., and Ruzzante, D.E. 2007. Population structure and gene flow reversals in Atlantic salmon (*Salmo salar*) over contemporary and long-term temporal scales: effects of population size and life history. Mol. Ecol. **16**(21): 4504-4522.

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**(2): 160-162.

Pella, J., and Masuda, M. 2001. Bayesian methods for analysis of stock mixtures from genetic characters. Fish. Bull. (Seattle) **99**(1): 151-160.

Presa, P., and Guyomard, R. 1996. Conservation of microsatellites in three species of salmonids. J. Fish Biol. **49**(6): 1326-1329.

Rannala, B., and Mountain, J.L. 1997. Detecting immigration by using multilocus genotypes. Proc. Natl. Acad. Sci. U. S. A. **94** 9197–9201.

Reddin, D.G. 1988. Ocean life of Atlantic salmon (*Salmo salar* L.) in the Northwest Atlantic. *In* Atlantic salmon: Planning for the Future. *Edited by* D. Mills. Croom Helm Kent , UK. pp. 483-511.

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. Fish. Res. (Amst.) **43(1-3)**: 221-235.

Reddin, D.G., Hansen, L.P., Bakkestuen, V., Russell, I., White, J., Potter, E.C.E., Dempson, J.B., Sheehan, T.F., Maoiléidigh, Ó., Smith, G.W., Isaksson, A., Jacobsen, J.A., Fowler, M., Mork, K.A., and Amiro, P. 2012. Distribution and biological characteristics of Atlantic salmon (*Salmo salar*) at Greenland based on the analysis of historical tag recoveries. ICES J. Mar. Sci. **69**: 1589-1597.

Schindler, D.E., Hilborn, R., Chasco, B., Boatright, C.P., Quinn, T.P., Rogers, L.A., and Webster, M.S. 2010. Population diversity and the portfolio effect in an exploited species. Nature (Lond.) **465**(7298): 609-612.

Slettan, A., Olsaker, I., and Lie, Ø. 1995. Atlantic salmon, *Salmo salar*, microsatellites at the SSOSL25, SSOSL85, SSOSL311, SSOSL417 loci. Anim. Genet. **26**(4): 281-282.

USASAC (U.S. Atlantic Salmon Assessment Committee). 2014. Annual Report of the U.S. Atlantic Salmon Assessment Committee. Report No. 26 - 2013 Activities, Old Lyme, CT. 149 pp.

Table 1. Reporting groups in the North American Atlantic Salmon microsatellite baseline. See Figure 1 for locations and Bradbury et al. 2015 for sample details.

Region	Code	Individual Samples	Rivers
1. Ungava Bay and Northern Labrador	UNG	191	4
2. Central Labrador	LAB	1501	25
3. Quebec Lower North Shore and Southern Labrador	QLS	579	10
4. Newfoundland	NFL	3531	48
5. Avalon Peninsula, NL	AVA	1302	14
6. Quebec Higher North Shore and Quebec City	QUE	710	15
7. Gaspe Peninsula	GAS	1055	21
8. Anticosti Island	ANT	140	3
9. Southern Gulf of St. Lawrence	GUL	1580	30
10. Nova Scotia	NOS	734	13
11. Inner Bay of Fundy	FUN	406	8
12. United States of America	USA	338	3
13. Europe	EUR	342	NA
	Totals	12409	194

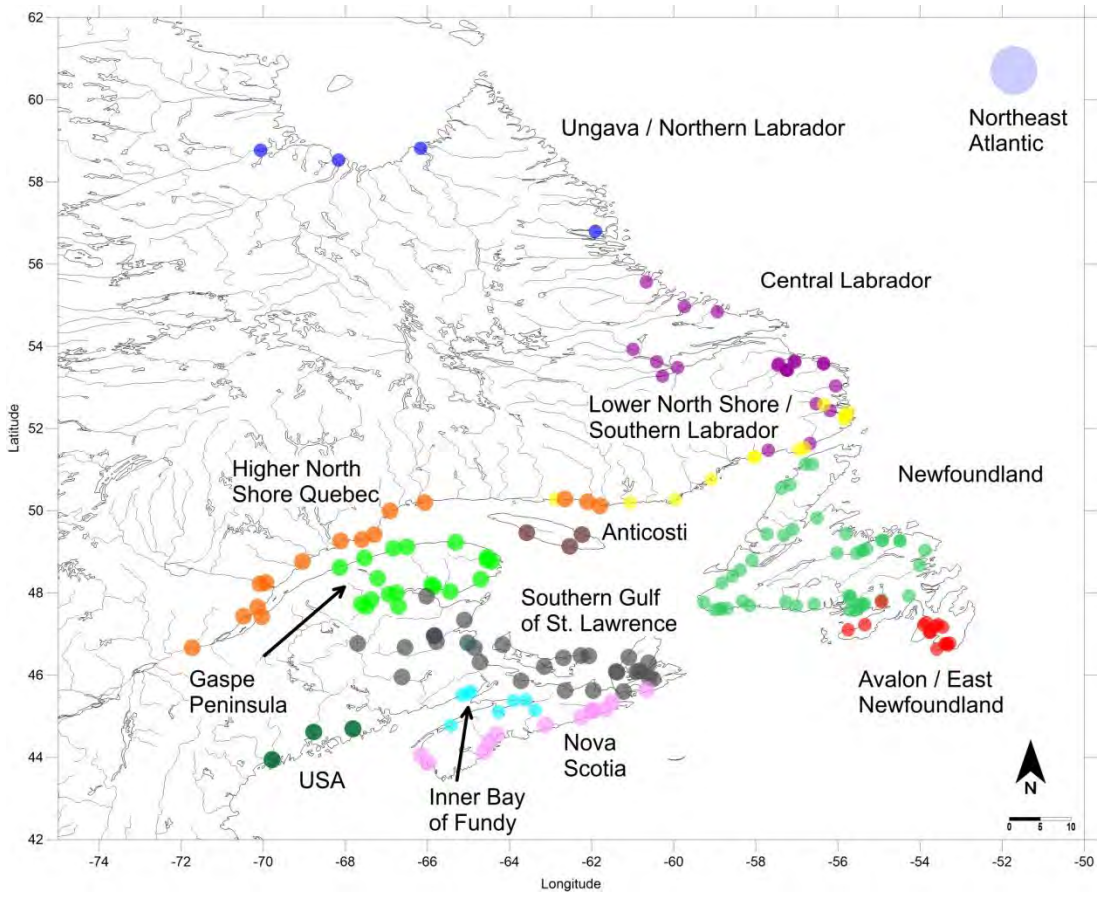

Figure 1. Map of sample locations used in microsatellite baseline for Atlantic salmon in North America. See Bradbury et al. 2015 for details and Table 1 for location abbreviations.

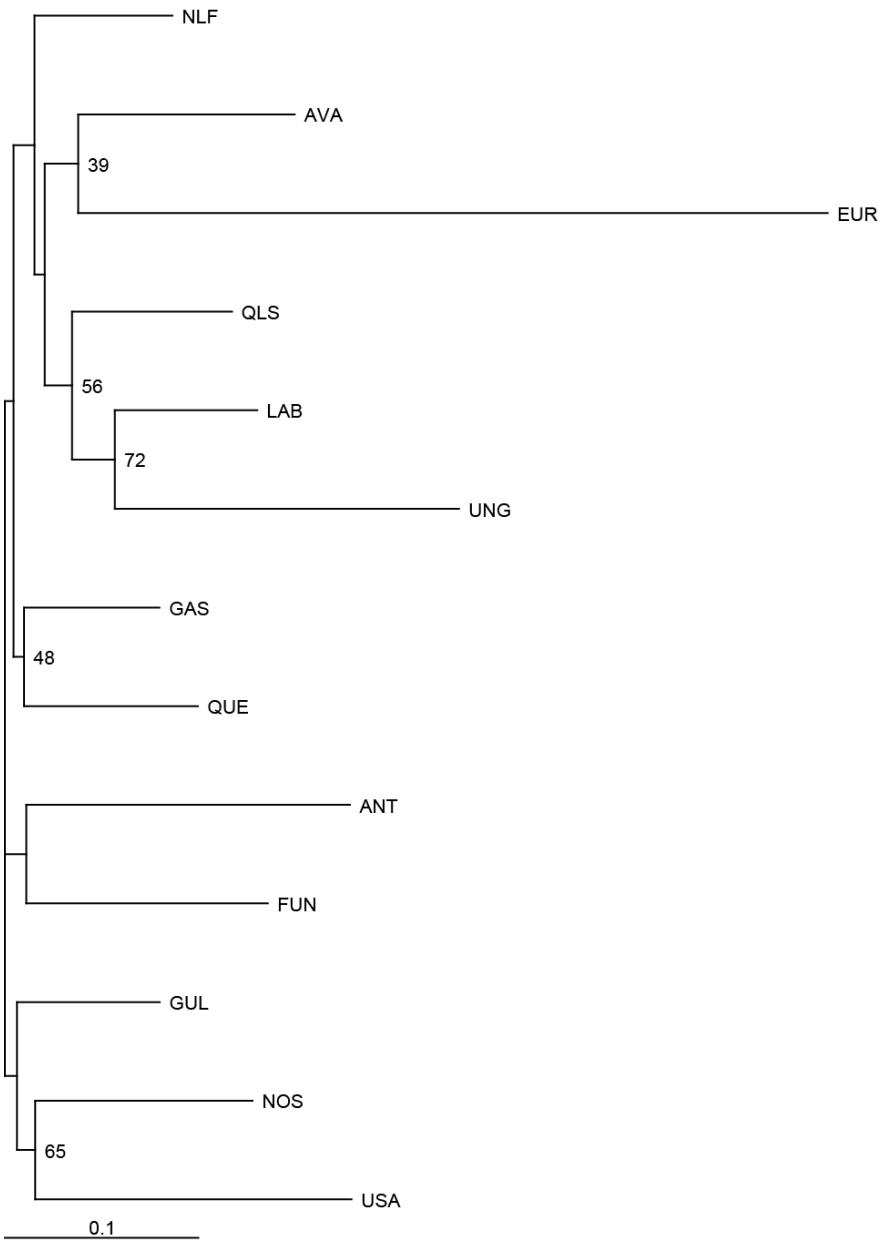

Figure 2. Neighbor-joining tree of Cavalli-Sforza and Edward's genetic distance among baseline groups in North American Atlantic salmon baseline. See Figure 1 for locations and Table 1 for location abbreviations.

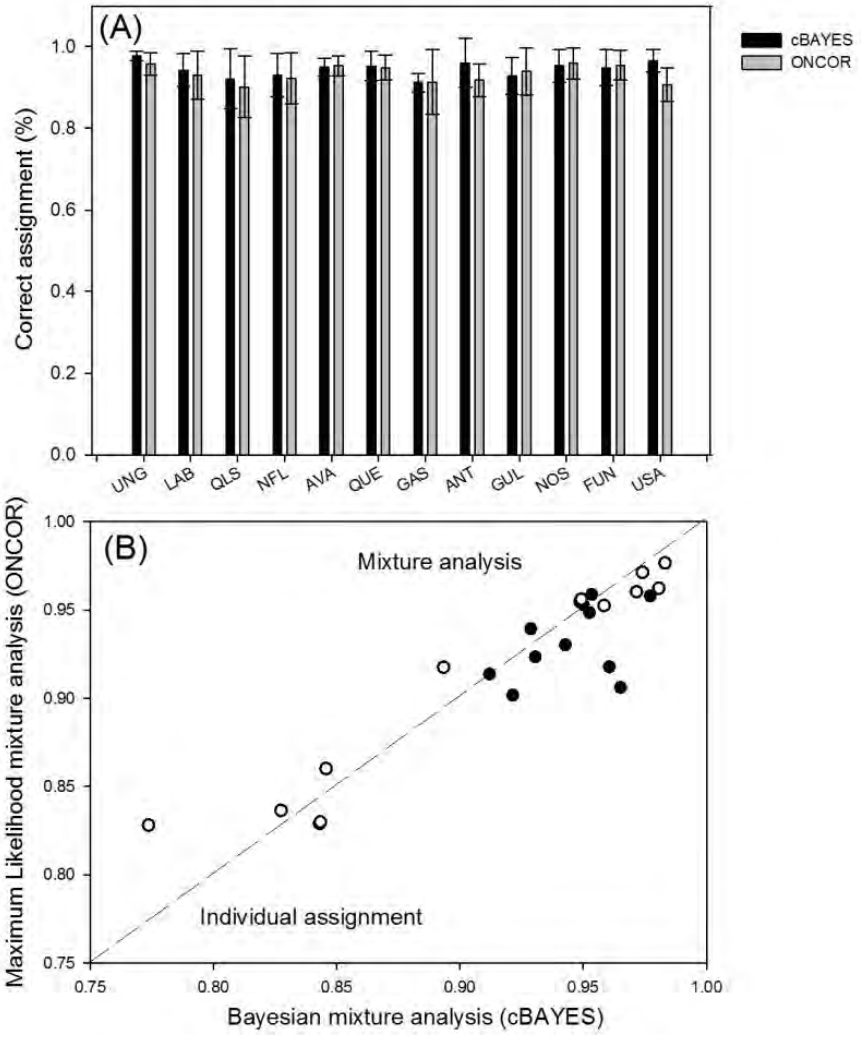

Figure 3. Comparison of Bayesian (cBAYES) and maximum likelihood mixture estimates and individual assignment accuracy using the North American Atlantic salmon baseline. (A) Estimates of mixture composition of 100% simulated mixtures of each baseline group. (B) Comparison of mixture analysis and assignment estimates for both approaches. See Figure 1 for locations and Table 1 for location abbreviations. Figure modified from Bradbury et al 2015.

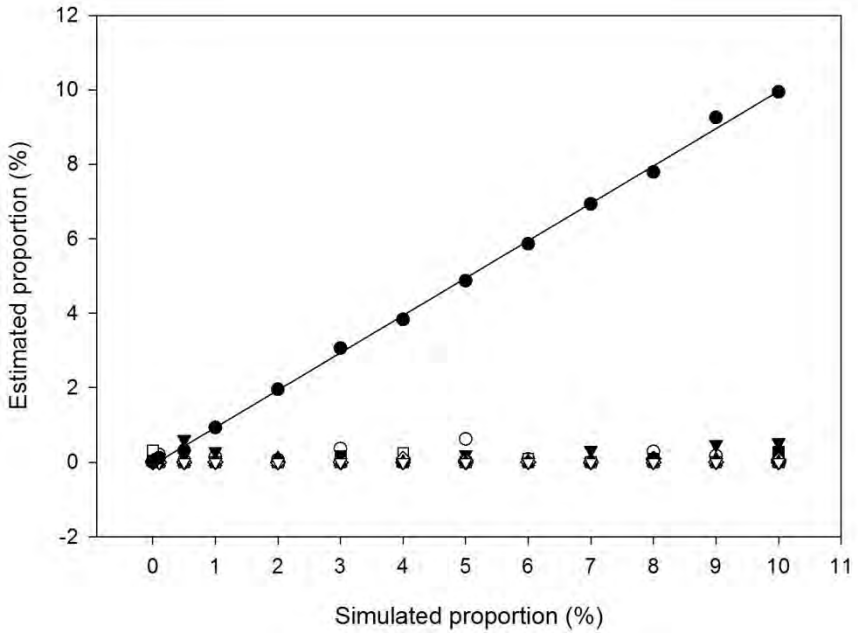

Figure 4. Detection threshold analysis for rare stock contributions. Simulated various contributions (0-10%) of USA reporting group into an otherwise single stock Labrador fishery and associated estimates of contribution. USA contributions represented by sold black dots.

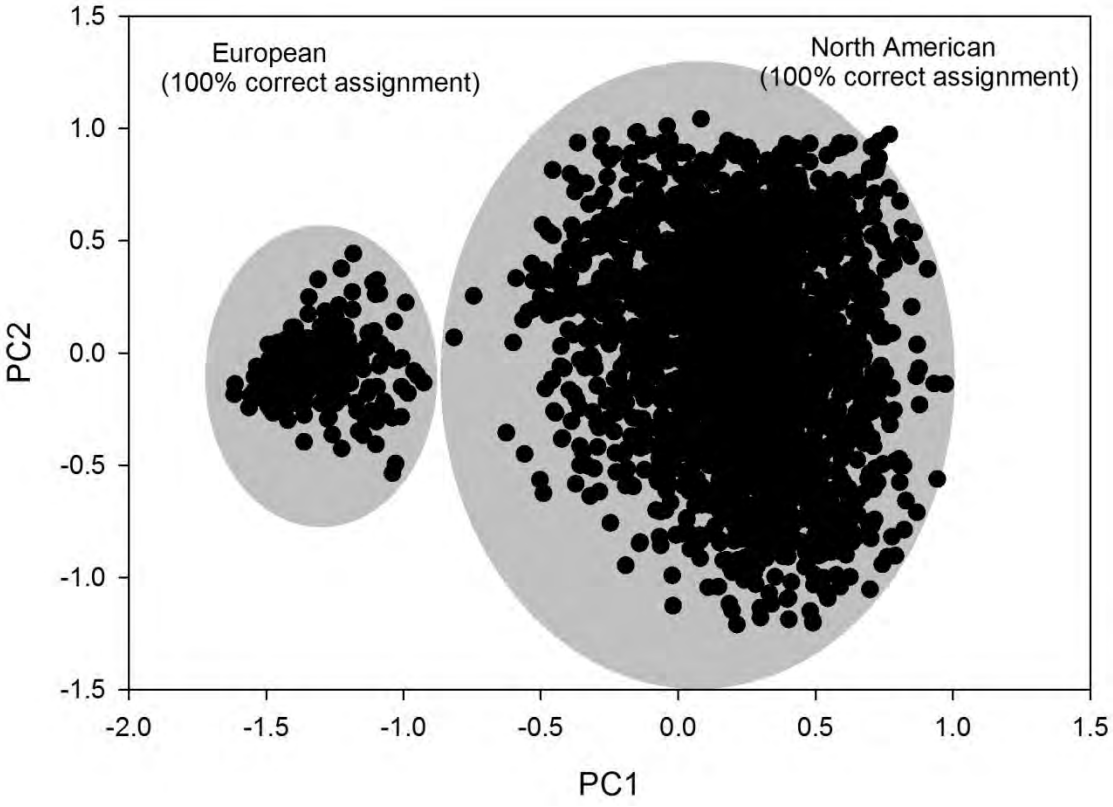

Figure 5. Principle component analysis of microsatellite data (N=15 loci) from Atlantic salmon sampled in West Greenland fishery 2011-2014. Clusters are compared to accepted COO assignment.

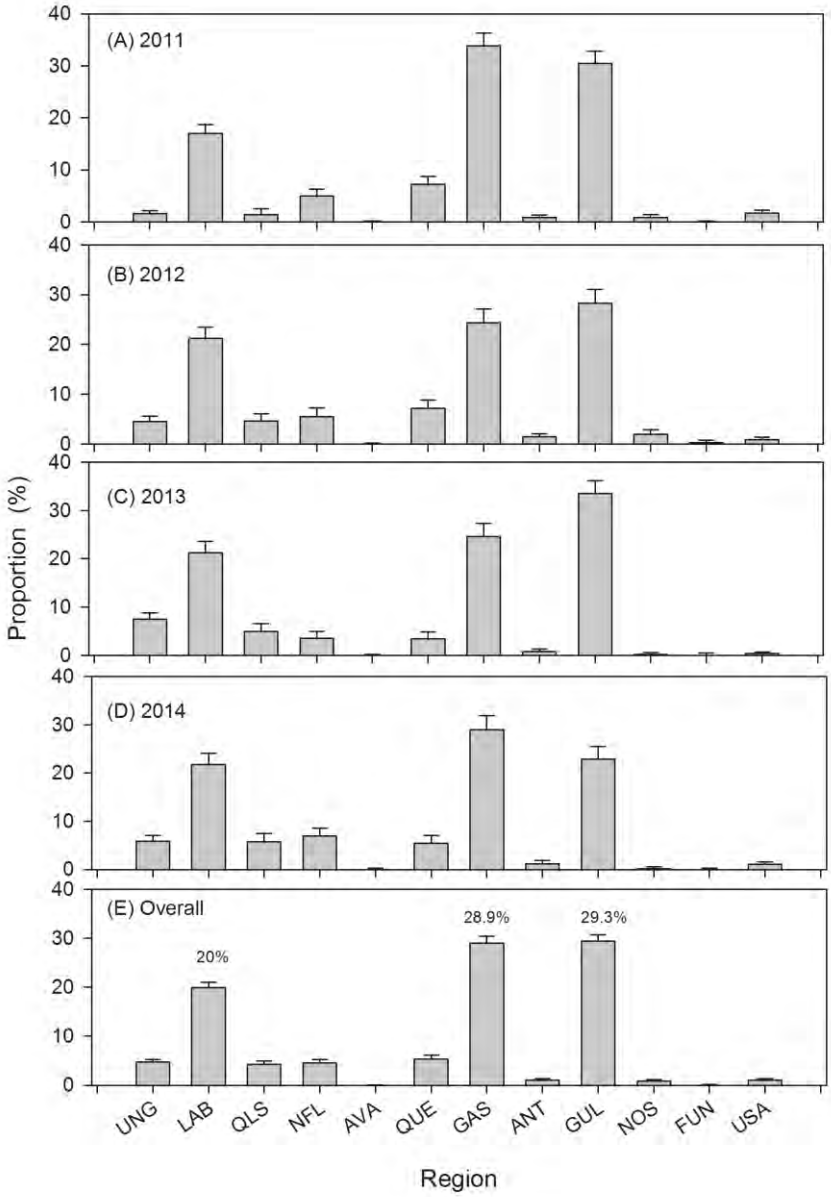

Figure 6. Bayesian estimates of mixture composition of samples from the West Greenland Atlantic salmon fishery 2011-2014 and overall. Baseline locations refer to regional reporting groups identified in Figure 1 and Table 1

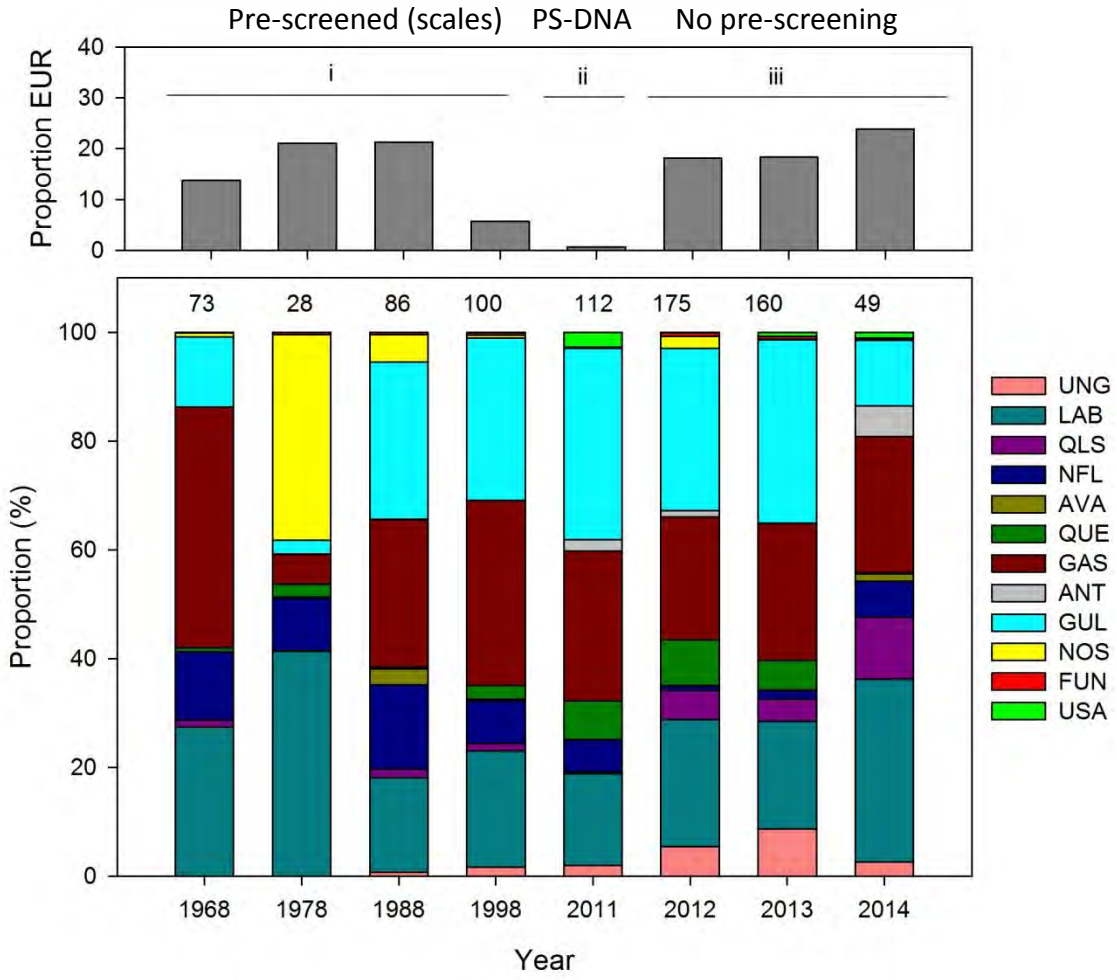

Figure 7. Estimates of European contribution and Bayesian estimates of mixture composition of samples from the West Greenland Atlantic salmon fishery 1968-2014. Baseline locations refer to regional reporting groups identified in Figure 1 and Table 1