Fisheries and Oceans Canada | Pêches et Océans Canada
Ecosystems and Oceans Science | Sciences des écosystèmes et des océans

Canadian Science Advisory Secretariat (CSAS)

Research Document 2018/044

Maritimes Region

Genetic Change in Inner Bay of Fundy Atlantic Salmon (*Salmo salar*) Across Three Generations of Captive Breeding and Rearing

Patrick T. O'Reilly[1], Carolyn Harvie[1], Sherisse McWilliam[2], Beth Lenentine[2], and Ross Jones[3]

[1]Fisheries and Oceans Canada
Bedford Institute of Oceanography
PO Box 1006, 1 Challenger Drive
Dartmouth, NS, B2Y 4A2

[2]Fisheries and Oceans Canada
Coldbrook Biodiversity Facility
1420 Fish Hatchery Lane
Coldbrook, NS, B4R 1B5

[3]Fisheries and Oceans Canada
Gulf Fisheries Centre
PO Box 5030, 343 University Avenue
Moncton, NB, E3C 9B6

November 2019

Canada

Foreword

This series documents the scientific basis for the evaluation of aquatic resources and ecosystems in Canada. As such, it addresses the issues of the day in the time frames required and the documents it contains are not intended as definitive statements on the subjects addressed but rather as progress reports on ongoing investigations.

Published by:

Fisheries and Oceans Canada
Canadian Science Advisory Secretariat
200 Kent Street
Ottawa ON K1A 0E6

http://www.dfo-mpo.gc.ca/csas-sccs/
csas-sccs@dfo-mpo.gc.ca

ISSN 1919-5044

Correct citation for this publication:

O'Reilly, P.T., Harvie, C., McWilliam, S., Lenentine, B., and Jones, R. 2019. Genetic Change in Inner Bay of Fundy Atlantic Salmon (*Salmo salar*) Across Three Generations of Captive Breeding and Rearing. DFO Can. Sci. Advis. Sec. Res. Doc. 2018/044. iv + 8 p.

Aussi disponible en français :

O'Reilly, P.T., Harvie, C., McWilliam, S., Lenentine, B., et Jones, R. 2019. Évaluer la réussite de la conservation des caractéristiques génétiques de la population de saumons atlantique (Salmo salar) de l'intérieur de la baie de Fundy sur trois générations de reproduction et d'élevage en captivité. Secr. can. de consult. sci. du MPO, Doc. de rech. 2018/044. iv + 10 p.

TABLE OF CONTENTS

ABSTRACT

The Live Gene Bank program for inner Bay of Fundy Atlantic Salmon (*Salmo salar*) has now been in operation for a little over 15 years, and across 3 generations. In this review, we assess expected rates of loss of genetic variation through to 2015 in the Stewiacke River reference population, using 3 different approaches. First, the program PMx and multi-generation pedigree information is used to monitor gene diversity, founder allele retention, and levels of inbreeding in spawner-year groups (sets of salmon spawned in a given year) across years; since kinship estimates of G0 (founder) salmon were included in the pedigree, metrics reported actually pertain to the unseen, unsampled parents (G-1 salmon) of the G0 founders. Second, we report levels of molecular genetic variation, including number of observed alleles, allele richness, effective heterozygosity, and observed heterozygosity, for spawner-year groups across years. Third, we use demographic information (including adult census population size, sex ratio information and variance in family size at maturity) to estimate the effective number of breeders and, in some instances, effective population size; where appropriate, this information is used to estimate expected rates of loss of gene diversity and accumulation of inbreeding through time. Overall, we expect some, but generally minimal, loss of genetic variation and accumulation of inbreeding between G0 and G2 or G3 generation salmon. Much of the genetic variation that is expected to be lost, and inbreeding accrued, can be directly related to very uneven G0 half-sib family size and, ultimately, the timing of the original founder collections.

We also investigate genetic change in inner Bay of Fundy Atlantic Salmon associated with possible introgression of genes from nearby wild and farm Atlantic Salmon sources. Evidence from several different analyses is consistent with historic gene flow from wild outer Bay sources into nearby inner Bay of Fundy river populations along the New Brunswick side of the inner Bay. Changing population dynamics post 1995 have likely resulted in ongoing and possibly increased rates of gene flow from outer Bay to Chignecto Bay populations in recent years; populations on the Minas Basin side of the inner Bay may have been more isolated. We also find evidence for the presence and successful spawning of European farm salmon in several rivers (across multiple years) of the inner Bay of Fundy, particularly those on the Chignecto Bay side.

EXECUTIVE SUMMARY

In the 1970s and 1980s, thousands of adult Atlantic Salmon (*Salmo salar*) returned yearly to the approximately 50 rivers of the inner Bay of Fundy (iBoF) to spawn. Numbers of returning adults declined precipitously starting in the latter half of the 1980s, continuing all through the 1990s, to an estimated 250 individuals in 1999. In 1998-2001, Fisheries and Oceans Canada (DFO) collected some of the last wild iBoF Atlantic salmon juveniles from the Stewiacke (STW) and Big Salmon (BSR) rivers, and began a Captive Breeding and Rearing or Live Gene Banking (LGB) program for this group of endangered Atlantic salmon, likely preventing its complete extirpation. In 2001, a LGB program was also initiated for the phenotypically distinct Gaspereau (GAK) population of iBoF Atlantic salmon. However, maintenance of small closed populations in captivity across multiple generations is not without risks, and the iBoF LGB program has now been in operation for 15-20 years, or for 3 to 4 salmon generations. The purpose of this document is to review genetic change (loss of genetic variation, including accumulation of inbreeding, and drift-induced changes in allele frequency distributions) associated with neutral processes over the course of the program (from 1998-2015). Some of the possible changes associated with non-neutral processes (e.g., selection for captive conditions or domestication) were assessed in a companion document (Harvie et al. 2018). Results were primarily based on the STW LGB population because

a) the G1-G4 portion of the STW pedigree is both very accurate and complete,

b) the BSR and GAK LGB populations were/are not closed, and levels of within-population variation in the former appear to be increasing over time, and

c) time constraints limited inclusion of information from additional LGB populations.

In this report, we also investigate genetic change associated with possible introgression of nearby wild and farm salmon genes into populations of iBoF Atlantic salmon, which might be expected to have increased after 1990 along with marked reductions in the sizes of iBoF populations relative to some potential source populations in the area.

Kinship analysis of the approximately 1000 original G0 generation STW salmon (wild parr collected in 1998-2001) indicates the presence of considerable family structuring, with over half of all individuals clustering into as few as 10 large half-sib (HS) groups, and over 10% into a single very large HS group of 117 individuals. In the year 2000, these estimates of first-order relatedness were used to minimize inbreeding in the next generation; in subsequent years, this information was also used to estimate mean kinship (MK) values and to prioritize salmon for spawning.

Consideration of the demographic characteristics of the STW population between 1967 and 2015, trends in family diversity observed across the four collection year groups (1998-2001) analyzed, the expected degree of HS family overlap between collection year groups produced before versus after 1995, and other information in this report, suggest that initial and later levels of founder genetic diversity would have been considerably higher, and present-day inbreeding lower, if collections were carried out even two years earlier (in 1996-1999). Slightly earlier collections would likely have

1) contained more HS and full-sib (FS) families,

2) exhibited lower variance in HS family size, and

3) exhibited less divergence between mean HS and FS (full-sib) family size, as is typically observed in juvenile collections obtained from larger nearby populations of Atlantic Salmon from the Southern Upland region of Nova Scotia.

Decisions regarding when to initiate Captive Breeding and Rearing (CBR) or LGB programs for salmonids are influenced by many factors, including the recognition that releasing animals produced and reared in captivity can harm wild populations, and that such activities likely pose a greater risk to small endangered populations than to large healthy populations. However, these results also highlight some of the costs of delaying CBR actions beyond a certain point (even by a just a few years), including lower levels of initial founder genetic variation, increased rates of loss of genetic variation into the future, and more rapid accumulation of inbreeding, all of which could impact subsequent efforts to restore wild self-sustaining populations in the future.

The efficacy of the LGB program in minimizing rates of loss of genetic variation under current conditions and over the duration of the program (from 1998 to as late as 2015) was directly assessed using 3 approaches. First, the program PMx was used to assess levels of pedigree genetic variation in groups of salmon spawned in any one year (referred to here as spawner-year groups) in the production of the next generation of LGB salmon from 2000 through to 2012-2015, depending on the metric reported. Since kinship estimates of relatedness among G0 generation salmon were incorporated into the STW pedigree, estimates of founder contributions and rates of loss of founder genetic variation relate back to the G-1 generation, and specifically, to the unseen, unsampled parents of G0 generation salmon spawned in the production of G1 LGB salmon. Results, therefore, reflect not only the effects of program management of the STW LGB population from 2000-2015, but also the number of spawners and mating structure that produced the G0 generation of wild-collected parr. The G-1 HS female parents of the large G0 HS families mentioned above left a large number of G2 generation descendants, and contributed disproportionally to the genomes of G2 generation salmon (compared to both G-1 HS female parents of small G0 HS groups, and all G-1 FS male parents of G0 FS groups). Founder allele (*FA*) retention of G-1 HS female parents of even moderately large G0 HS groups approached 1.0, indicating that the multiple additional genetic contributions observed for the G-1 HS female parents of the very large G0 HS families were likely to add little to individual and overall mean *FA* retention. On the other hand, *FA* retention of G-1 HS female parents of small G0 HS groups, and all G-1 FS male parents of G0 FS groups, was markedly lower; additional founder contributions would likely have greatly increased *FA* retention of these G-1 generation salmon. Not unexpectedly, the total number of founder alleles retained in the STW LGB population declined at a fairly high rate over time, from a maximum of approximately 380 observed in the 2004 spawner-year group to a low of approximately 175 in the 2014 spawner-year group, though declines across spawner-year sets (consisting of multiple spawner-year groups approximately corresponding to individual iBoF salmon generations) were less pronounced for this period (approximately 600 to 400). Observed reductions likely reflect finite effective population size in the program in G1 to G2 generations, but also original G-1 founder parent contribution unevenness. Other measures of founder genetic variation, including gene diversity (*GD*) changed little over the duration of the program, ranging from 0.990 to 0.995 for different spawner-year groups, possibly increasing slightly in the early years (likely due to the mixing of the original 4 collection-year groups and associated sets of somewhat non-overlapping families), and potentially declining thereafter (from 2005 to 2014), but only slightly (to approximately 0.992).

The mean inbreeding of families or offspring (a measure of within-individual genetic variation) produced by male and female salmon spawned in a given year spanning 2000-2015 varied between $F \sim 0.001$ and 0.005, generally decreasing from 2003 through to 2008, to then increase from 2008 to 2012, before possibly decreasing again through to 2014. Analyses of the exact origins of inbreeding indicate that elevated early mean inbreeding was due to the occurrence of a very few (3-5) highly inbred crosses ($F = 0.125$), in large part a result of large G0 HS family size and associated high probabilities that unplanned or incorrectly prescribed crosses involved half siblings, and the shallow pedigree depth of ancestors common to early

pairs of spawners. Early declines reflected the transition from G0 to G1 spawners, associated marked reductions in family sizes, and resulting low probabilities that unplanned or incorrectly prescribed crosses involved siblings. Obligatory changes in pedigree assignment methodology from kinship- to parentage-based approaches across these two generations (see report for details), and increasing pedigree depth of G-1 HS ancestors common to pairs of G1 versus G0 generation salmon spawned, also contributed to observed reductions. Increases in mean family inbreeding observed from 2009 to 2012 parallel

1) the approximate transition of offspring from G2 to G3 generations,
2) a marked increase in the number of crosses where the male and female parent share at least one common ancestor, and
3) increases in the number ancestors common to male and female salmon spawned in a given year or, in other words, the beginning of unavoidable delayed inbreeding.

Detailed analyses indicate that nearly all ancestors common to male and female spawner pairs crossed in a given year were G-1 generation individuals, and nearly all of these were G-1 HS female parents of very large G0 HS families.

Recognizing the impacts of spawning multiple representatives of G0 HS families (in an attempt to capture genetic variation associated with the multiple FS families nested within) on rates of loss of *FA* and accumulation of inbreeding in the future, we carried out a PMx-based modelling analysis to assess the efficacy of alternate founder selection regimes in minimizing loss of genetic variation and accumulation of inbreeding through to the G10 generation. Neither the selection of one representative per G0 HS group (maximizing G-1 HS female contribution evenness at the expense of G-1 FS male representation) nor one representative per FS group nested within (maximizing G-1 FS male parent representation, at the expense of G-1 HS female evenness) resulted in the lowest rates of loss of genetic variation or accumulation of inbreeding over time (under conditions explored). An intermediate regime, where a maximum of five G0 salmon per HS group (each from a different FS family nested within) were selected as founders in the production of the G1 generation, resulted in slightly lower rates of loss of genetic variation and marked reductions in the accumulation of inbreeding through to the G10 generation. Ranked Mean Kinship (a recently development broodstock management regime previously demonstrated to perform well when family size is very uneven) achieved similar results, and is operationally easier to implement.

In the second approach used to assess program efficacy in retaining genetic diversity, we monitored changes in levels of molecular genetic variation across three generations of captive breeding and rearing. Mean observed heterozygosity (H_o) of STW spawner-year groups generally increased from approximately 0.840-0.850 early in the program to a maximum value of just over 0.880 in the year 2013, a likely result of ongoing mixing of the 4 early founder year collections (1998-2001), which were comprised of slightly overlapping but somewhat different sets of HS families. Observed heterozygosity then appeared to sharply decline across 2 years in a row to approximately 0.845 in the year 2015, a value similar to that observed in the beginning of the study. This possible decline beginning in the year 2014 is precisely 4 years (one generation) after observed increases in the occurrence and prevalence of G-1 HS female ancestors common to male and female pairs of parents spawned in a given year (beginning in 2010), and parallels increases in mean family inbreeding expected in this generation. Mean expected heterozygosity (H_e) of STW spawner-year groups varied much less across years, generally ranging from 0.840 to 0.850, with the first and last spawner-year groups exhibiting nearly identical values (just above 0.840). The observed number of alleles (#*A*) seen in exclusively STW spawner-year groups varied more over time, increasing initially from approximately 16.0 in the year 2000, to nearly 17.5 in 2004 (a likely result of early mixing of the

four founder year collections and associated families). Values then generally slowly declined through to the end of the study to just below 15.0. Estimates of allele richness (N_a) observed in STW spawner-year groups were similar to and paralleled those reported for *#A*. Detailed analyses indicate that the actual alleles that were lost were either short suspect European variants (actively selected against) or, generally speaking, very rare alleles (1 in approximately 2000 G0 alleles assayed) that were not expected to be retained.

Allele frequency distributions of early STW spawner-year groups (2000-2003) were significantly different from late spawner-year groups (2011-2015) in genic and genotypic tests, and levels of differentiation (F_{ST}) between some pairs of early and late spawner-year groups were moderate in magnitude (e.g., F_{ST} = 0.00394 for 2000 vs 2013 sample collection pairs). However, the largest differences observed in this study were between early neighbouring pairs of spawner-year groups from the period spanning 2000-2003 (e.g., F_{ST} = 0.00407 for 2001 vs 2003 and 0.00611 for 2000 vs 2003). These early spawner-year groups were comprised of varying proportions of G0 salmon collected in each of the years 1998-2001, which were in turn comprised of somewhat different sets of families; later generation salmon consisted of increasingly homogenized lineages from most original G0 families. Divergence observed between some early and late pairs of spawner-year groups likely reflects the complete absence of a subset of families in some early spawner-year groups, but presence of most families (or their descendants) in later spawner-year groups, as opposed to drift-induced change due to small population size. This interpretation is supported by the following observations:

1) divergence between either 2002 or 2003 spawner-year groups (comprised of representatives of most or all of the four 1998-2001 collection year groups) and each of the last (2011-2015) spawner-year groups was much reduced (F_{ST} = 0-0.00273, $\overline{X}$ = 0.0010),
2) the combined G0 founder group (all 1998-2001 collection year groups combined) was not significantly different from any late (2011-2015) spawner-year group, and
3) pairwise estimates of F_{ST} between the former and the latter were all very low (F_{ST} = 0-0.00047, $\overline{X}$ = 0.000152).

Finally, in the third approach used to assess expected program efficacy in minimizing rates of loss of genetic variation, we evaluated several population demographics parameters, including number of male and female spawners, sex ratio, variance in family size of offspring at maturity (at spawning), and the effective number of breeders (N_b) for the years 2000 to 2012, and this information was used to estimate expected rates of loss of genetic variation and accumulation of inbreeding over time. Approximately 100 males and 100 females (or more) were spawned in each year spanning 2000-2012, and sex ratios in any one year were within a few percentage points of 1.0, with the exception of the year 2000, when 149 males and 174 females were spawned, and the ratio of males to females was 0.856. Variance in family size (V_k) at maturation (spawning) ranged from 0.609 (in 2011) to 3.274 (in 2002), but was generally less than 1.0; variation in family size standardized to the mean (V_k/K) was less than 1.0 in all years (sometimes markedly so), except years 2000 (1.185) and 2002 (1.864). The effective number of breeders (taking into account both V_k and sex ratio departures from 1.0) ranged from a low of 153.4 in the year 2002 to a high of 473.5 in the year 2006, and was generally 250 or greater in 2003 and all subsequent years. Ratios of N_b/N_c ranged from 0.667 to 1.589, and were above 1.0 in 2003 and all subsequent years. We also estimated N_e based on individual lifetime reproductive success as reported by the program PMx for each of the years 2000-2010. Estimates of N_e ranged from a low of 104.3 in 2001 to a high of 271.1 in 2010, and were, in any one year, generally reduced relative to estimates of N_b. These slightly diverging results reflect the spawning of some low MK individuals in more than one year, and associated large family sizes, and higher V_k. However, from 2006 on, estimates of N_b and PMx-based estimates of N_e

were more similar, with the latter generally spanning 150 and 200. From 2013 on, the five STW LGB year class groups will exhibit non-overlapping population characteristics, and $N_b = N_e$. If a similar number of parents are spawned each year (100 males and 100 females), a similar sex ratio achieved (nearly 1.0), spawners are obtained exclusively from wild-exposed offspring groups managed similarly to those produced in 2010-2012 such that V_k is comparable to that observed for these three year class groups, and pedigree information from a similar number of genotyped offspring is available and used to reduce V_k, V_k/K is expected to range from approximately 0.4 to 0.7, and N_e from 250-300. Under these conditions, F is expected to accumulate at an approximate rate of $F = 0.002$ per generation, increasing to about 0.035 by the G20. Given these same demographic conditions, GD is expected to decline from 1.0 to approximately 0.998 by the G1 generation, and to 0.96 by the G20. Similar inbreeding and genetic variation retention goals can also be achieved using several other management regimes discussed in this report, each varying with respect to

a) the availability and use of genotype and pedigree information;

b) the use of

   i) exclusively captive-reared,

   ii) exclusively wild-exposed, or

   iii) captive-reared and wild-exposed offspring groups; and

c) the number of adults spawned in any given year (Nc).

Rivers of the iBoF, particularly the most northwest drainages along the Chignecto Bay side of the iBoF, are a few tens of kilometres from the Saint John River (SJR) which, from 1990-2005, could have been a potential source of large numbers of hatchery and wild strays. These same iBoF rivers were also just a few more tens of kilometres removed from Passamaquoddy/ Cobscook Bay areas where, all through the 1990s, large and increasing numbers of farm salmon post-smolt and adults were reared for marine grow-out. At the same time, river populations of the iBoF were in decline and, by the end of this period, likely consisted of a few dozen adult returns. Given the population dynamics of these potential nearby sources (wild SJR strays and farm salmon escapees) and recipient iBoF populations, there existed the potential for high rates of gene flow from both wild outer Bay of Fundy (oBoF) and farm salmon genes into populations of the iBoF.

Reduced pairwise F_{ST} estimates between Tobique (TOB) or Nashwaak (NSH) sample collections obtained in approximately the year 2001 and contemporaneous BSR sample collections relative to that observed between these two SJR collections and those obtained from the STW and GAK rivers at or around the same time suggest that, even before the latest dramatic iBoF population declines, gene flow from the oBoF into nearby rivers on the Chignecto Bay side of the iBoF was elevated relative to more distant Minas Basin populations. Elevated early gene flow from SJR to Chignecto Bay relative to Minas Basin populations is further suggested by the following:

1) *structure* results of these same early samples showing reduced partitioning of genomes of individual Atlantic salmon obtained from the Minas Basin into SJR baseline sample collections compared to those obtained from the BSR;

2) frequency-based individual assignment test results showing

   i) high self-assignment (using leave-one-out procedures) of STW and GAK baseline samples, and

ii) low cross-assignment of these same individuals to SJR baseline sample collections compared to that observed for BSR baseline samples;

3) elevated linkage disequilibrium observed for BSR baseline sample collection compared to baseline sample collections obtained from smaller and less genetically variable STW and GAK populations analyzed; and

4) markedly reduced prevalence (in most instances, complete absence) of clade 1-3 mitochondrial DNA (mtDNA) genotypes in rivers draining into the Chignecto Bay side of the iBoF compared to those draining into the Minas Basin.

Gene flow from large wild SJR populations and both captive and LGB BSR populations appears to be ongoing, as indicated by a) declining pairwise F_{ST} estimates between early NSH (year 2000) and increasingly later (2000-2015) yearly collections of both wild-exposed (LGB-origin) and wild-produced smolt obtained from the BSR and b) phylogenetic analyses involving this same NSH sample collection and multiple different sample collection types obtained from the BSR, with BSR adult sample collections clustering closest to the NSH sample collection, followed by (generally speaking) wild-produced smolt collections, and then wild-exposed (LGB-origin) smolt collections. Ongoing gene flow from external sources into the BSR is also consistent with observed increases in *#A* and N_a across 2000-2015 spawner-year groups of BSR salmon; these metrics appear to slowly decline in STW spawner-year groups across this same period, as expected for closed populations of finite effective population size. The number of observed alleles (*#A)* and N_a also appear to increase across sets of a) LGB-origin BSR smolt, b) wild-origin BSR smolt, and c) wild-origin BSR adults, produced in each of the years 2000-2012.

Investigations into the presence and spawning success of Atlantic Salmon from nearby farm salmon sources in the iBoF (and inferences regarding hybridization and introgression) are hindered by

a) the non-availability of temporally relevant baseline sample collections from all industry sources in the area;

b) the purported local origin (SJR) of some broodstock used in the area and low *(and potentially declining)* differentiation between this and iBoF salmon, particularly those of the BSR (F_{ST} ~ 0.02); and

c) the small N_e of some iBoF populations (in some years) and expected rapid drift-induced changes in allele frequency distributions expected from generation to generation.

Additionally, even if SJR-origin salmon could be accurately identified in iBoF populations using traditional assignment tests, it might not be possible to determine whether suspect SJR salmon were either a) SJR-origin farm salmon escapees, or b) wild SJR strays.

Between 1989 and 1995, European (Landcatch) salmon were imported into Maine, and shortly thereafter permeated 30-50% of farm salmon broodstock used by US growers (Baum 1998). Genetic markers for European (EU) ancestry has since been found in Canadian-origin farm salmon escapees (direct hatchery-origin juvenile salmon collected from the Magaguadavic River), indicating the presence of EU farm salmon ancestry (and likely wholly EU salmon) in New Brunswick farm salmon strains. Because EU and North American (NA) salmon have been (generally speaking) reproductively isolated for approximately one million years (Hurst et al. 1999; Nilsson et al. 2001) and are, therefore, more genetically divergent, EU farm salmon escapees likely represent a greater risk to endangered iBoF salmon than local-origin (SJR) escapees. However, they (and their early descendants) should also be easier to detect in sample collections of Atlantic salmon obtained from rivers of the Bay of Fundy (BoF).

Nuclear Ss1 microsatellite alleles and mtDNA haplotypes commonly observed in Atlantic salmon from the EU continent, but rarely seen in populations from NA, have been observed in juveniles obtained from the iBoF Upper Salmon River (USR). Individuals exhibiting these variants also exhibit short EU-type Ssa202 alleles, previously reported to be common in wild EU salmon and EU farm escapees observed in the oBoF. These same USR juveniles also exhibited short alleles at three additional loci (SSsp1605, SSsp2215 and SSsp1G7) that are also observed in

a) EU farm salmon escapees collected from the Magaguadavic River,

b) suspect wholly EU farm salmon escapees obtained elsewhere in the Maritimes, and

c) wild salmon from the European continent, but were rare or absent in early sample collections obtained from geographically more isolated iBoF rivers of the Minas Basin.

These results represent some of the strongest evidence to date for the presence and spawning of farm salmon (and, in particular, EU farm salmon) in the iBoF. Short EU-type alleles at these same four loci (SSsp1605, SSsp1G7, SSsp2215 and Ssa202) for which genotype information in the BoF area is widely available have also been observed in several Atlantic Salmon collected from the STW in 1999. The extreme clustering of alleles across these 4 loci in just 3 of over 1000 STW salmon analyzed, the expected very low frequency of such genotypes if these short variants were indeed rare NA alleles, and other evidence reported (including uniparental inheritance of multiple short EU-type alleles across loci) indicate that these Atlantic salmon were likely early descendants (probably immediate offspring) of EU farm salmon escapees spawning in the STW River. These same short EU-type alleles involving these four nuclear microsatellite loci were also observed in early (2001-2004) sample collections obtained from the GAK River, were similarly clustered within just a few individuals, and also indicate the likely presence of EU farm salmon ancestry in this population. Individuals exhibiting short EU-type alleles at 2 of these 4 loci have also been observed in several other iBoF sample collections obtained from the Mispec, Black, and Point Wolfe iBoF rivers.

In early collections of juvenile and adult Atlantic Salmon obtained from the BSR (1998-2001), however, short EU-type alleles at these 4 loci were either absent or rare, and when observed did not exhibit any notable within-individual clustering like that reported for USR, STW, or GAK suspect EU farm salmon descendants, described above. Beginning in 2003 and continuing through to at least 2015 (the last year for which information was available), however, 1 or more individuals exhibiting short EU-type Ssa202 alleles were observed in nearly every large or small (e.g., N = 7) yearly sample collections of 'in-river' produced (non-LGB-origin) adults or smolt obtained on the BSR, comprising 5.0-33.3%, ($\overline{X}$ = 11.6) and 1.2-10.2 %, ($\overline{X}$ = 5.11) of their respective collections. Very high levels of clustering of short EU-type alleles at 2 to as many as four loci (including the locus Ssa202) in individual parr or smolt collected in multiple later years (between 2009 and 2015), and additional supporting evidence described, indicate a very likely local EU farm source of at least some of these short EU-type Ssa202 alleles. However, given the prevalence of post-2003 BSR smolt and adult salmon exhibiting short Ssa202 alleles at this one locus only, we explored possible alternative origins of short Ssa202 variants in the iBoF. Short Ssa202 alleles (<255 bp) were not observed in 43 of 45 sample collections of juveniles (representing over 2000 individuals obtained from 2 rivers of the Minas Basin and 4 from the Chignecto Bay side of the iBoF) acquired largely before 1996 and analyzed elsewhere; where short alleles were observed, they were not abundant (approximately 1 in 30 or more alleles assayed from a given sample collection). Furthermore, neither of the short Ssa202 variants (Ssa202-239 and Ssa202-247) commonly observed in iBoF samples collected recently were observed at all in these early sample collections. These and other results described in this report suggest that short Ssa202 alleles (common in EU wild and local EU farm escapees) were

not present at moderate or high frequencies, and were not widely distributed in the iBoF, before 1996. The possible role of other potential sources of short Ssa202 alleles in contributing to patterns now observed in the iBoF were also explored, including recent emigration from the SJR, where a handful of short alleles at all 4 Continent of Origin (COO) loci have been observed. Given the low frequency of short Ssa202 allele-bearing salmon seen across multiple SJR sample collections analyzed, along with parentage assignment results indicating that a substantial proportion of BSR adults captured in any one year (approximately 30-70%) are at least second-generation BSR salmon, oBoF salmon appear to be a very inefficient source of short Ssa202 alleles, and are unlikely to account for the presence and prevalence of these variants in sample collections of adults or smolt obtained from this river in recent years (2003-2015). None of the remaining sources explored appear to be plausible major contributors to short Ssa202 alleles observed in the iBoF today.

If short Ssa202 alleles alone in recent collections of iBoF salmon do reflect the presence of EU farm ancestry, as much as approximately 10-25% of 'in-river' produced BSR smolt may exhibit EU farm genes, though the overall percentage of EU farm genes in the population is likely relatively low (less than 3%). The current practice of collecting smolt at the mouth of the BSR river, and the high prioritization of 'in-river' produced smolt for spawning (done for multiple reasons, and through several mechanisms) may be accelerating the introgression of putative EU farm genes into the BSR LGB population as well. Even if introgression of new EU genes is minimized going forward, EU farm genes will continue to spread throughout the BSR LGB population, though the percent of EU ancestry in any given individual should, on average, decline. Some limited EU farm ancestry associated with the original suspect G0 EU farm/NA hybrid founders still persists in the STW LGB population, but is being removed. European farm ancestry is somewhat more extensive in the GAK LGB population, and the benefits of using pedigree information to remove putative EU genes should be considered in the context of potential costs in terms of loss of native GAK genes.

REFERENCES CITED

Baum, E.T. 1998. History and Description of the Atlantic Salmon Aquaculture Industry of Maine. DFO. Can. Stock Assess. Sec. Res. Doc. 98/152.

Harvie, C., McWilliams, S., and O'Reilly, P. 2018. Effects of Captive Breeding and Rearing on Survival, Growth and Other Phenotypic Traits in Inner Bay of Fundy Atlantic Salmon (*Salmo salar*) Across Three Generations. DFO Can. Sci. Advis. Sec. Res. Doc. 2018/042.

Hurst, C.D., Barlett, S.E., Davidson, W.S., and Bruce, I.J. 1999. The Complete Mitochondrial DNA Sequence of Atlantic Salmon, *Salmo salar*. Gene 239: 237-242.

Nilsson, J., Gross, R., Asplund, T., Dove, O., Jansson, H., Kelloniemi, J., Kohlmann, K., Loytynoja, A., Nielsen, E.E., Paaver, T., Primmer, C.R., Titov, S., Vasemagi, A., Veselov, A., Ost, T., and Lumme, J. 2001. Matrilineal Phylogeography of Atlantic Salmon (*Salmo salar* L.) in Europe and Postglacial Colonisation of the Baltic Sea area. Mol. Ecol. 1: 89-102.