Fisheries and Oceans Canada | Pêches et Océans Canada
Science | Sciences

C S A S | **S C C S**

Canadian Science Advisory Secretariat | **Secrétariat canadien de consultation scientifique**

Research Document 2009/095 | **Document de recherche 2009/095**

Conservation of genetic variation in the inner Bay of Fundy Atlantic salmon captive breeding and rearing program

Conservation de la variation génétique dans le programme de reproduction et d'alevinage en captivé du saumon atlantique de l'arrière-baie de Fundy

Patrick T. O'Reilly and Carolyn J. Harvie

Population Ecology Division, Science Branch
Department of Fisheries and Oceans
P.O. Box 1006, 1 Challenger Drive
Dartmouth, Nova Scotia
B2Y 4A2 Canada

ISSN 1499-3848 (Printed / Imprimé)
ISSN 1919-5044 (Online / En ligne)

Canada

TABLE OF CONTENTS

Correct citation for this publication:

O'Reilly, P.T., and C.J. Harvie. 2010. Conservation of genetic variation in the inner Bay of Fundy Atlantic salmon captive breeding and rearing program. DFO Can. Sci. Advis. Sec. Res. Doc. 2009/095. viii + 53 p.

ABSTRACT

In response to the precipitous decline of inner Bay of Fundy Atlantic (iBoF) salmon, the Department of Fisheries and Oceans (DFO) initiated a captive breeding and rearing program, intended to prevent the imminent extirpation of this *Species at Risk Act* (SARA)-listed species. The goal of the program was to minimize all genetic changes relative to the founding population, including the loss of genetic variation through drift, until conditions at sea improve, and wild self-sustaining populations can be established. Loss of genetic variation was minimized by maintaining large effective population sizes, and by assessing the genetic value, or Mean Kinship (*MK*), of candidate spawners. Mean kinship determinations were based on pedigree information, which was estimated by reconstructing kinship in the founding generation using molecular genetic marker data and recently developed algorithms for assessing first-order relatedness in the absence of parental information and, in subsequent generations, through molecular genetic marker data and parentage analyses. The objective of this report is to evaluate the ability of the above program to minimize the rate of loss of neutral molecular genetic variation.

Program efficacy was first evaluated by determining the percent of genetic variation present in the large (>1,000) wild juvenile parr collections, captured in the parents (founders, G0) selected for spawning. Loss of gene diversity (H_e) and number of observed alleles (#*A*), two commonly used measures of genetic variation, in the selection of founders was minimal, approximately 1%. We also evaluated the rate of loss of genetic variation in the production of the first generation (G1) of salmon from the original G0 founders. In this analysis, we assessed loss of genetic variation *by spawning year*, and *by year class*. Evaluations by *spawning year* involved comparisons of genetic variation in all parents spawned in a given year and all offspring later recovered and spawned, regardless of the year they matured and were genotyped. Gene diversity estimates were generally greater in the offspring relative to the parents and, when less, declined by only 0.24%. Parent-offspring reductions in #*A* were quite variable between spawning years, and could be quite high, 3.9 to 18.2 percent over one generation. However, because of a number of operational aspects of the current program, including monitoring of egg mortality and previous spawning history of individuals and families, this measure of loss of genetic variation is likely to over-represent the true loss of genetic variation.

In analyses of rates of loss *by year class,* genetic variation was compared between all parents obtained from the wild in a given year and selected for spawning, and all offspring of these same parents selected for spawning themselves, regardless of the year in which the parents were spawned, and regardless of the year the offspring were recovered and spawned in the production of the next (G2) generation. Gene diversity estimates were again generally greater in the offspring relative to the parents and, when less, declined by only 0.25% or 1.1%. Rates of loss of #*A* when assessed by year class were variable and sometimes very large (7.8% to 52.0%). However, much of the reported loss for any one year class was due to (1) spawnings having occurred between year classes or generations (though genetic variation will be recovered in these between-group classes, it could not be tabulated and incorporated into within-year class statistics), and (2) incomplete maturity and recovery of offspring, though further genetic variation is expected to be recovered when these offspring mature in subsequent

years. When comparing combined year class groups, thereby reducing the effects of factor (1) above, rates of loss of #*A* between parents and offspring were between 3.5 and 4%. Once offspring of parents collected in 2000 and 2001 mature, rates of loss of #*A* are expected to be lower, probably on the order of 2-3%. Overall, these results indicate that the rates of loss of genetic variation were indeed low during the period assessed, below 3-4% and probably between 0.25% and 3%, depending on the measure of genetic variation assessed. These analyses indicate that the captive breeding and rearing program being carried out on inner Bay of Fundy Atlantic salmon is indeed effective at maintaining neutral genetic diversity.

RÉSUMÉ

Face au déclin précipité du saumon atlantique de l'arrière-baie de Fundy (abF), le ministère des Pêches et des Océans (MPO) a lancé un programme de reproduction et d'alevinage en captivité dans le but d'empêcher une disparition imminente de ce poisson inscrit sur la liste d'espèces en péril établie en vertu de la *Loi sur les espèces en péril* (LEP). Ce programme visait à réduire le plus possible les changements génétiques par rapport à la population fondatrice, notamment la perte de variation génétique par dérive, jusqu'à ce que les conditions de vie en mer s'améliorent et que des populations sauvages autonomes puissent être établies. La perte de variation génétique a été réduite au minimum grâce à la création et au maintien de vastes effectifs de population ainsi qu'à l'évaluation de la valeur génétique, ou apparentement moyen, des frayeurs possibles. L'apparentement moyen a été établi selon l'information généalogique, estimée par reconstitution de la filiation dans la génération fondatrice à partir de données sur les marqueurs génétiques moléculaires et les algorithmes élaborés récemment pour évaluer la parenté au premier degré en l'absence d'information sur le lien parental; dans le cas des générations subséquentes, ce sont les données sur les marqueurs génétiques et des analyses de parenté qui ont été utilisées. Le présent rapport vise à évaluer la capacité du programme susmentionné à réduire le taux de perte de diversité génétique moléculaire neutre.

Pour évaluer l'efficacité du programme, nous avons d'abord déterminé quel était le pourcentage de variation génétique présent dans les vastes collections (>1 000) de tacons sauvages juvéniles provenant des parents (fondateurs, G0) sélectionnés pour la reproduction. La perte de diversité génétique (H_e) et le nombre d'allèles observés (#*A*), deux mesures courantes de la variation génétique, dans la sélection des fondateurs était minime, se chiffrant à environ 1%. Nous avons aussi évalué le taux de perte de variation génétique dans la production de la première génération de saumons (G1) issue des fondateurs (G0). Dans cette analyse, nous avons estimé la perte de variation génétique par *année de fraye* et *par classe d'âge*. L'évaluation par *année de fraye* faisait appel à des comparaisons de la variation génétique chez tous les parents qui ont été accouplés une année donnée et chez tous leurs descendants récupérés ensuite et accouplés, indépendamment de l'année de leur arrivée à maturité et de leur génotypage. Les estimations de la diversité génétique étaient en général plus élevées dans la descendance que chez les parents et dans les cas où elles étaient inférieures, elles n'avaient diminué que de 0,24 %. La réduction du nombre d'allèles (#*A*) entre les parents et leur descendance était relativement variable d'une année de fraye à une autre et pouvait être assez élevée, allant de 3,9 % à 18,2 % sur une génération. Toutefois, en raison de divers aspects opérationnels du programme actuel, comme la surveillance de la mortalité des œufs et les antécédents de fraye des individus et des familles, cette mesure de la perte de variation génétique constitue vraisemblablement une surestimation de la perte véritable.

Dans les analyses du taux de perte *par classe d'âge,* la variation génétique a fait l'objet d'une comparaison entre, d'une part, tous les parents prélevés dans le stock sauvage une année donnée et sélectionnés pour le fraye et, d'autre part, tous leurs descendants sélectionnés eux aussi pour le fraye, indépendamment de l'année où les parents ont été accouplés et indépendamment de l'année où les descendants ont été récupérés et accouplés pour produire la génération suivante (G2). Les estimations de la diversité génétique étaient là aussi généralement plus élevées chez les descendants que chez leurs parents et dans les cas où elles étaient inférieures, elles n'avaient diminué que de 0,25 % ou 1,1 %. Dans l'évaluation par classe d'âge, les taux de perte d'allèles (#*A*) étaient variables et parfois très élevés (de 7,8 % à 52,0 %). Toutefois, une bonne partie de la perte estimée pour une classe d'âge quelconque était due 1) au fait que des frayes étaient survenues entre les classes d'âge ou les générations (bien que dans ce cas la variation génétique sera recouvrée dans le total intergroupes, elle n'a pu être calculée et intégrée dans les statistiques par classe d'âge) et 2) à une maturité et une

récupération incomplètes des descendants, quoiqu'on devrait recouvrer une partie de la variation génétique quand ces descendants arriveront à maturité les années subséquentes. Il ressort de comparaisons entre groupes de classes d'âge combinées, permettant de réduire les effets du facteur 1) susmentionné, que les taux de perte d'allèles (#*A*) entre parents et descendants se situaient entre 3,5 % et 4 %. Une fois que les descendants des parents prélevés en 2000 et 2001 arriveront à maturité, les taux de perte d'allèles devraient diminuer et être probablement de l'ordre de 2 % à 3 %. Dans l'ensemble, ces résultats révèlent que les taux de perte de variation génétique ont bel et bien été faibles pendant la période considérée, se situant sous les 3 % ou 4 % et probablement entre 0,25 % et 3 %, selon la mesure de la variation génétique évaluée. Ces analyses montrent que le programme de reproduction et d'alevinage en captivité du saumon atlantique de l'arrière-baie de Fundy réussit effectivement à maintenir une diversité génétique neutre.

INTRODUCTION

The Bay of Fundy is situated between the provinces of Nova Scotia and New Brunswick, Canada, and is home to some of the highest tides in the world. These high tides and associated currents transfer nutrients from the bottom of the Bay into the water column, and result in incredibly high levels of productivity. Many birds and mammal species migrate great distances to take advantage of these nutrient-rich waters, feeding on small invertebrates and fishes in summer months before travelling to warmer southern areas to over-winter. Atlantic salmon populations occupying the approximately 35 most internal rivers of this ecologically unique Bay (Figure 1) have long been recognized as distinct from nearby populations (Huntsman, 1931). In addition to showing evidence of local marine migration (Jessop, 1976), inner Bay of Fundy (iBoF) salmon exhibit a very high rate of maturing as grilse (after one year at sea), and a high rate of repeat spawning (iteroparity) (Ducharme, 1969). Recent common-garden experiments have also demonstrated that phenotypic differences observed between iBoF salmon and nearby distant-migrating salmon from the Southern Upland region of Nova Scotia, including growth rates, levels of compensatory growth, and morphology, do indeed have a genetic basis, and that the direction of variation observed is consistent with predictions based on migratory differences (Fraser *et al*., 2007; Fraser *et al*., 2010). Inner Bay of Fundy salmon also exhibit a unique lineage of mitochondrial DNA (mtDNA) observed nowhere else in the species global distribution; the presence of this lineage, and the distribution of mtDNA variation in the area in general, may indicate a refugial origin of at least some salmon from this region (Verspoor *et al*., 2002).

In the mid-to-late 1980s, coincident with the arrival of aquaculture to the Gulf of Maine-Bay of Fundy area, the exponential increase in grey seal numbers off Sable Island, and the collapse of Atlantic cod along much of the east coast of North America, a potentially large-scale ecological change in itself, numbers of iBoF salmon began what was to be a sharp and unrelenting decline. Whereas average annual recreation catches of salmon from iBoF rivers, representing a portion of the adult returns, were approximately 2,000 in the period spanning 1970-1990 (DFO, 1997), the total adult population in recent years had declined to fewer than 250 individuals (DFO, 2009). Following listing of this unique group of salmon in 2001 as Endangered under Canada's *Species At Risk Act* (SARA), a recovery team consisting of representatives from the Federal Department of Fisheries and Oceans (DFO), Parks Canada, the provinces of Nova Scotia and New Brunswick, salmon conservation groups, and Aboriginal peoples was formed, with the goal of eventually restoring iBoF Atlantic salmon to these rivers (DFO, 2009). Although the exact causes of the decline (and continued suppression) of iBoF populations are not known, there is a general agreement that, like other populations along the Atlantic coast from Cape Breton to the southern end of their distribution, high marine mortality of post-smolts and adults is an important contributing factor (DFO, 2008). In fact, for salmon of the iBoF, smolt return rates are likely very low, less than 1% (DFO, 2008). For these populations, restoration of freshwater habitat or changes to fishing practices or quotas, often prescribed management actions for declining salmonid populations, are unlikely to have much of an impact on returning numbers of salmon or population persistence (DFO, 2008). The recovery team, therefore, identified the need "to harbour and protect what remains of the residual populations" (DFO, 2009) until causes of high marine mortality can be identified and, if possible, mitigated.

THE INNER BAY OF FUNDY ATLANTIC SALMON RECOVERY PROGRAM

The iBoF Atlantic salmon recovery program, as outlined in the document "Recovery Strategy for the Atlantic Salmon *(Salmo salar),* inner Bay of Fundy populations [Proposed]" (DFO, 2009) is multi-faceted and includes elements of habitat protection (river monitoring), public awareness and community outreach, and captive breeding and rearing (CBR). The CBR component serves

several functions, including the production of animals for conservation research (e.g., providing animals for at-sea tracking of smolts; see Lacroix and Knox, 2005; Lacroix *et al*., 2005) and studies of the effects of one or more pathogens on marine survival of post-smolts (O'Neil, pers. comm.). Captive breeding and rearing is also key to other conservation initiatives, which will be grouped into two main categories: 1) Live Gene Banking (LGB), defined as "the maintenance of genetic variation in living populations of individuals in captive or semi-captive environments for the future restoration of wild populations", and 2) supplementation, defined as "the use of artificial propagation, while conserving genetic resources, for the goal of restoring or augmenting self-sustaining populations" (Kapuscinski, 1991; Miller and Kapuscinski, 2003). Although these two programs are in some ways interrelated, they do have different guiding principles and objectives, and involve different activities. The Live Gene Banking program is primarily concerned with minimizing genetic change through time (discussed further below), and less with maximizing numbers of juveniles for release into iBoF rivers. Salmon considered surplus to the LGB program (e.g., additional representatives of particular families) are utilized in the supplementation program. Generally speaking, surplus salmon, either adults or juveniles, are released into secondary vacant iBoF rivers of non-native origin (Table 1). The primary objectives of the supplementation program are to 1) maintain salmon in natural river habitat, possibly serving important ecological functions, and also to provide mates for any returning female salmon, 2) monitor changes in freshwater survival of secondary iBoF rivers, 3) monitor changes in marine survival, and 4) begin the recovery of iBoF salmon when marine survival improves. While resident in these rivers, released juveniles also serve as additional LGBs, providing a further source of these populations in the event of a catastrophic loss of the captive and main semi-wild populations. Because the native populations in these rivers are extirpated, and since the return rate of salmon to the iBoF is so low, the primary consideration of this program is the maximum production of juveniles for release.

LIVE GENE BANKING OF INNER BAY OF FUNDY SALMON

The primary objective of the LGB program is to minimize all genetic changes in the lineage of iBoF salmon to be used to restore wild anadromous runs when ocean conditions and marine survival improve. These changes include loss of genetic variation and the accumulation of inbreeding, genetic adaptation to captivity, accumulation of deleterious alleles, and overall changes in allele frequencies that could represent shifts in populations from adaptive optima that may have existed prior to collapse. Preservation of genetic variation is an important goal of conservation programs because of its potential role specifically in the restoration of future wild populations (Frankel and Soulé, 1981). The number of different alleles observed (#*A*) in a population may set limits on the ability of released animals to adapt to future environmental challenges (Nevo, 1978), including natural long-term ecological fluctuations and changes associated with human impacts, such as global warming and the introduction of invasive species. Gene diversity (H_e), or the likelihood that two alleles sampled at random from a population are different, may be of particular importance for rapid re-adaptation to native wild conditions, as the rate of microevolution is expected to be determined by additive genetic variation (Moritz, 1999; Doyle *et al.*, 2001). Gene diversity, particularly at some loci, may also play a role in the success of restoration programs by affecting the survival of released animals directly. In Atlantic salmon, challenge experiments have demonstrated associations between particular MHC IIβ alleles and *Aeromonas salmonica* (Langefors *et al*., 2001), the causative agent of Furunculosis, and Hematopoietic Necrosis virus (IHNV) (Miller *et al*., 2004). Individuals heterozygous at MHC loci may, therefore, exhibit increased lifetime survival when they encounter multiple different pathogens sequentially. The importance of genetic variation at genes involved in pathogen resistance for restoration success has recently been demonstrated empirically. Captive-bred guppies exhibiting reduced immunogenetic variation showed lower

survival relative to more genetically variable guppies, also reared in captivity, when released into semi-natural environments containing native parasite fauna (Van Oosterhout *et al*., 2007).

Genetic change associated with loss of fitness in the wild is also an important and growing concern. An increasing body of evidence from many different taxa, recently reviewed by Frankham (2008), indicates that captive breeding and rearing may bring about a greater and more rapid loss of wild fitness than previously thought. In fact, studies from salmonids suggest that changes may come about in as few as one or two generations (Araki *et al*., 2007a, 2007b). Below, we outline the iBoF LGB program, the steps taken to minimize loss of genetic variation and domestication selection, and some of the constraints and limitations imposed by the biology of the species and fish culture technology presently available.

The iBoF LGB program began in 1998, with relatively large collections of juveniles from the Big Salmon and Stewiacke rivers (Tables 2 and 3). These rivers were selected for conservation efforts because they harboured the largest remaining runs, expected to exhibit the highest levels of genetic variation (Table 1). This choice of rivers also ensured that the ecologically distinct group of Chignecto Bay drainages, generally characterized by rocky substrates and steep gradients, and Minas Basin drainages, generally characterized by longer estuaries, shallow gradients and muddy substrates, were represented. Juveniles from the Stewiacke River are being housed at the Coldbrook Biodiversity Facility and juveniles from the Big Salmon River at the Mactaquac Biodiversity Facility. Prior to maturation, all salmon were tissue sampled and tagged to permit individual identification. DNA was extracted from tissue samples and genetic variation analyzed at typically nine microsatellite loci. Microsatellite genetic variation was then used to estimate first-order relationships using the kinship reconstruction programs *Pedigree* (Smith *et al*., 2001) and *Colony* (Wang, 2004). This information was used to help guide the selection of founders, so as to maximize recovery of founder genetic variation (discussed below). Soon after hatching, offspring were either released into native river habitat for development under natural conditions or reared in captivity (Figure 2). The following year, after some siblings from all families were exposed to wild river conditions and natural selection, juveniles from captive and river environments were sampled, tagged, tissue sampled, and reared in common captive environments through to maturity (Figure 2). Offspring were then assigned to actual crosses using parentage analyses, and mating strategies were developed so as to minimize loss of genetic variation due to drift (discussed below). This part of the LGB, the spawning of adults, release into the wild, and rearing in captivity, is to be repeated until restoration of wild self-sustaining populations becomes feasible.

ESTIMATING KINSHIP AMONG THE WILD STEWIACKE RIVER JUVENILE COLLECTIONS

The use of molecular genetic information in estimating first-order relatedness (kinship) in the absence of parental genotype information is a key component of this conservation program, and the authors are unaware of any other instances of its use in prioritizing or selecting founders in actual captive breeding and rearing programs, though Rudnick and Lacy (2008) explore the utility of doing so under some conditions using simulation analyses. However, its use in assessing relatedness among founder collections in conservation programs has been recommended (Blouin, 2003). The primary challenge of such analyses is the astronomically large number of alternate full- and half-sib groupings possible ($>10^{10}$) in even modest-sized groups (several hundred) of individuals. Moreover, a pair of siblings may share 0, 1, or 2 alleles in common at a given locus, and occasionally siblings will share few, if any, alleles across several loci, thereby appearing to be unrelated. *Colony* and *Pedigree* programs mitigate this latter problem by making use of information from growing full-sib and kin groups in the

reconstruction of relatedness in large groups of individuals. For example, two true full sibs sharing few alleles in common may be grouped through a third sibling that shares more alleles with each. Still, such kinship estimates based on even 10 or more very variable loci are not exact, and small full-sib (fewer than four) and half-sib groupings can easily occur by chance (Hansen and Jensen, 2005). However, when reasonable levels of family structuring exist in a population, as was observed in the 1998 collection of Stewiacke River juveniles (Figure 3), reliable estimates can be made for many kin groupings (Herbinger *et al*., 2006). The ability of these programs to accurately reconstruct larger kin groupings is also evident from observations of concordance observed between *Colony* and *Pedigree* programs in estimating relatedness in Stewiacke River and other iBoF populations (Figure 4; Herbinger *et al*., 2006). Another graphic example of the ability of these programs to estimate relatedness in the absence of parental genotype information is given in Figure 5. Here, relatedness among 292 Big Salmon River salmon from the Minto hatchery were assessed using *Colony*, and results contrasted with UPGMA clustering of pairwise distances based on allele-sharing measures (Bowcock *et al*., 1994). Numbers next to major branch groupings, and like-colouring of terminal branches, represent full-sib assignments made by *Colony*, whereas the branching pattern itself was produced by the UPGMA clustering method. Overall concordance between these two methods was greater than 97%.

SELECTION OF BROODSTOCK FROM WILD PARR COLLECTIONS AND THE RECOVERY OF FOUNDER DIVERSITY

Kinship analyses of Stewiacke River founders in years 1999 through to 2001 revealed a similar pattern of relatedness as that reported for the Stewiacke 1998 collection by Herbinger *et al*. (2006): a handful of large kin groups comprised of many full-sib families and a larger number of smaller kin groupings consisting of two or three individuals (Figure 3). A random sub-sampling of a small or modest number of broodstock from these groups would likely result in the selection of multiple founders from the larger kin groups and the exclusion of many smaller kin groups (see results). Starting in 2001, we employed a broodstock selection and spawning program intended to maximize recovery of founder genetic variation present in the large collections of juveniles, given the specific capabilities and limitations of the fish culture facilities in a given year (Tables 4 and 5). In this strategy, termed Mean Kinship Assist (*MKA*) here, Mean Kinship (*MK*) (Ballou and Lacy, 1995; see methods section) values were calculated once and all offspring sorted by *MK* values. Because all full siblings exhibit identical *MK* values, and as half siblings from similarly sized full-sib groups tend to exhibit similar *MK* values, family groupings and *MK* values were usually concordant, with siblings appearing immediately adjacent or nearby in lists of individuals sorted by *MK* values. Next, the lowest *MK* female from the lowest *MK* kin group was selected and designated for spawning with the lowest *MK* male from the lowest *MK* kin group, as long as the two did not belong to either the same full-sib or kin grouping. In instances where two individuals belonged to the same full-sib or kin grouping, the male exhibiting the next lowest *MK* value in any other kin grouping was chosen for spawning with the selected female (thus minimizing the chances of inbreeding in the next generation at the full- or half-sib level). This process was repeated a number of times equivalent to the total number of pairwise spawnings allowable in a given year for a given group of salmon; the number of pairwise spawnings allowable varied from year to year, and was a function of the capacity of the relevant fish culture facility at a given point in time.

Once the prescribed pairwise matings discussed above were designated, this same spawner selection process was used to identify additional salmon for spawning, though the pairing of individual males and females for specific crosses happened at the time of spawning, and was less controlled (explained below). Proceeding down the list of *MK*-sorted salmon, and picking up

from where the last paired spawners were selected, individuals were chosen from each kin grouping, following the same rules above, until the capacity of the fish culture facilities was met for this group of *non-paired* spawners. Individuals so chosen were also relegated to low or medium *MK* groups, and care was taken to approximately equalize the total number of males and females in each of the two groups. Later, at the time of spawning, a female was first obtained from a given group (e.g., the low-*MK* group) and a male sampled from that same group (exhibiting a relatively similar *MK* value). Crosses were only allowed between individuals from the same *MK* group that were not full or half siblings. This within-*MK* group spawning process was implemented so as to minimize genetic linkage between rare and common genes, as this would impede future attempts to increase the frequency of rare alleles (Ballou and Lacy, 1995). Offspring from paired and medium-*MK* spawnings were incorporated into the LGB program for their rivers of origin, whereas offspring from low-*MK* spawnings were released into vacant non-native rivers.

A disadvantage of this spawner selection process is that fewer individuals from some larger kin groups were often chosen for spawning than were selected from the smaller kin groupings; this occurred, for example, when the carrying capacity of the facilities was met somewhere in the middle of the *MK* distribution. However, in the present context, this may have several advantages. First, both *Colony* and *Pedigree* may artificially create small groups of two to four individuals (Smith *et al*., 2001; Hansen and Jensen, 2005). In such instances, spawning a second individual from these groups will increase the chances that individuals that are actually sole members of families (though incorrectly grouped with two or three other fish) may be chosen for spawning and, therefore, incorporated into the LGB program. Second, because of the large size of kin groupings from which a single brood fish was selected, there would be a greater likelihood that at least one representative would survive another year, compared to the smaller families of two or three, thus affording a second opportunity for spawning and incorporation into the LGB program. There is also a possible limitation of this approach. If the kin-group size of juveniles sampled in the wild is in part due to among-family differences in survival, then selection for reduced fitness in the wild may be an unintended consequence. However, variance in family size may also be a result of sampling error and spatial heterogeneity in the abundance of siblings from particular families, due to limited dispersal of offspring from redds, and the proximity of electrofishing-based sampling to these sites. Nevertheless, this is being assessed by comparisons of founder kin-group size in parr versus smolt samples obtained from the Point Wolfe River (unpublished data). Additionally, we are currently investigating whether *MK* values of parents are associated with survival of their offspring in wild freshwater habitat (Table 11).

Practical realities, such as high early juvenile mortality and specific procedures used in fish culture facilities to pool egg lots, resulted in the reduced likelihood of the recovery of offspring from some crosses in some years. Variance in survival of different families, due either to pre-existing genetic differences among founders, or variation in the amount of cumulative domestication selection among family lineages, may also have contributed to poor representation of some families. These, and other considerations, necessitated the tracking of past spawning histories of individuals and the inclusion of such information in the selection of broodstock in some years (Tables 4 and 5). Other requirements, such as the balancing of sex ratios and the production of offspring within groups (G1 versus G2, etc.) for later research and monitoring purposes, were also considered in the prioritization and spawning of salmon (Tables 4 and 5).

Inaccuracies of kinship assignments can result in reduced recovery of founder variation, increased inbreeding in the next generation, and increased loss of genetic variation over time. However, ignoring high levels of kinship in the selection of spawners by performing random

matings, or by practicing maximum avoidance of inbreeding, may be even more risky, possibly resulting in higher levels of loss of founder diversity (Doyle *et al*., 2001) and genetic variation over time (but see Rudnick and Lacy, 2008).

MINIMIZING LOSS OF GENETIC VARIATION IN G1 AND SUBSEQUENT GENERATIONS OF INNER BAY OF FUNDY SALMON

In addition to losing founder genetic variation in the selection of parents from the original collection of wild juveniles, genetic variation in LGB salmon could also be lost in the production of offspring (G1) iBoF salmon. Loss of genetic variation in production of G1 salmon was again minimized by genotyping G1 offspring, assigning parentage, and then using pedigree information to estimate mean kinship values for G1 and all other (remaining G0) available spawners. Mean kinship and other information, including previous spawning history, were then used in the selection of spawners and arrangement of matings.

We also mitigated the loss of genetic variation in iBoF salmon by cryopreserving milt from 90 G0- and G1-generation Stewiacke River males in 2006 and 2007 (Tables A1a and A1b) and a similar number of additional males each year after that (unpublished data). This material could be used to restore lost genetic variation in future generations (Stoss, 1983), but may also be useful in minimizing genetic changes brought about through adaptation to captive conditions (O'Reilly and Doyle, 2007). Although this may be counterproductive in terms of reducing adaptation to changing conditions in both freshwater and marine wild environments, it would seem that differences in the wild environments over a few generations would be much reduced relative to differences between captive and wild environments. Furthermore, decisions to use milt for this purpose can be made at some point in the future, perhaps based on experimental results of offspring survival using less-domesticated cryopreserved milt versus fresh milt from multigenerational LGB males.

The primary objective of this study is to report on the ability of the LGB program to minimize loss of genetic variation in iBoF Atlantic salmon. Genetic variation may be lost during the selection of founders from the large samples of wild-collected juveniles and through drift in the production of G1 and subsequent generations. We evaluate the efficacy of the program in recovering founder diversity by comparing levels of genetic variation (number of observed alleles #*A*, allele richness N_a, and gene diversity H_e) in the total collection of wild Stewiacke River juveniles, the selected parents or founders, and a similarly sized group of randomly chosen parents with identical ratios of males to females, selected at the adult stage. We also test the ability of four different broodstock selection programs, Minimization of Mean Kinship or *MMK*, Mean Kinship Assisted or *MKA*, Minimization of Pairwise Relatedness or *MPR*, and Random Selection and Mating or *RSM*, to recover founder variation, assuming a more limited number of founders (50 males and 50 females) and the family structuring observed in the wild Stewiacke River juveniles from 1998. Loss of genetic variation in the production of the first generation of LGB salmon is evaluated by comparing the above measures of genetic variation in the parent (G0) salmon spawned and their offspring (G1), using several different approaches (by spawning year and by year class). Expected loss of H_e is also assessed using demographic information from the G0 generation to estimate the effective number of breeders (N_b). This information, and known relationships between N_e and single-generation loss of H_e, is used to predict rates of loss expected, and this is compared with observed rates of loss in the iBoF LGB program reported here. Expected loss of #*A*, based on N_e and observed allele frequencies, is also estimated and compared to rates of loss observed for this statistic. This report does not attempt to assess rates of loss of quantitative genetic variation or genetic variation underlying fitness-related traits, though investigations are underway.

METHODS

Laboratory analyses: All tissue samples (typically fin clips from juveniles between 5 and 25 mg) were stored in 1 to 2 ml of 95 to 99% ethanol immediately after collection, in 1.5 ml screw-cap tubes. Tissue was transferred to Qiagen's 96-well DNeasy plates, and DNA extracted and purified following the manufacturer's specifications. Polymerase Chain Reaction (PCR) amplifications were carried out in 10 µl volumes, containing between 50 and 100 nanograms of template DNA, 2 mM each dNTP, 0.5 µM labelled and unlabelled primers, 50 mM KCl, 0.5 units of Taq DNA polymerase supplied by MBI Fermentis and 2.0 mM $MgCl_2$. Thermal cycling conditions were as follows: (94°C for 3 min.)X1, (94°C for 1 min., 58°C for 30 sec., 72°C for 30 sec.)X5, and (90°C for 30 sec., 58°C for 30 sec., and 72°C for 30 sec.)X 30, followed by a 15-minute extension step at 72°C. Primer sequences for loci Ssa 171, Ssa 197 and Ssa 202 are given in O'Reilly *et al*. (1996); 2210, 2215, 2216, 1G7 and 1605 are given in Paterson *et al*. (2004); and 144 and 486 are given in King *et al*. (2005); additional statistics on individual loci are given in Table A3.

PCR products were combined and salt, unincorporated dNTPs, and unincorporated labelled and non-labelled primers removed using Qiagen's 96-well PCR Purification plates, as specified by the manufacturer. Fragments were size-fractionated and detected using either a Hitachi MJ Research Basestation automated fragment analyser/sequencer or an Applied Biosystems 3130 XL. Genotype calls were cross-standardised between platforms and between data sets by including a different set of 2 of 10 individuals in each group of 96 samples; this group of 2 of 10 standards also internally labelled all groups of 96 samples, minimizing the chances of confusing or mixing batch identities. One sample from each strip of eight tubes was duplicated in wells 87 to 96 to identify sample placement errors, strip inversions, and plate inversions (Figure A1). Duplication of samples also permitted quantification of rates of genotyping errors.

Kinship analysis: Kinship was assessed in the samples of Stewiacke River wild founders collected in 1998 (G0-98), 1999 (G0-99), 2000 (G0-00), and 2001 (G0-01) (Table 2) using either the program *Pedigree* (Smith *et al*., 2001) or *Colony* (Wang, 2004). Both are likelihood-based approaches and utilize allele frequency information from growing groups of siblings and rules of Mendelian inheritance. The Class I (upper allele drop-out) and Class II (stochastic) error rates in all *Colony* analyses were set to 0.01, or approximately one single-locus genotyping error in every fifth individual analyzed at nine microsatellite loci. Additional details of both analyses as performed here are available in Herbinger *et al*. (2006).

Parentage analysis: Parentage of offspring obtained from captivity and from the wild was ascertained using microsatellite genotype information and, typically, both exclusion and likelihood methods. In all exclusion analyses, offspring were tested against known sets (prescribed crosses) of parents only, and a single-locus offspring-parental pair mismatch was allowed. Given the number of alleles involved in every offspring-parental pair group involving nine microsatellite loci and the proportion of errors that could exclude parentage assignments in such comparisons (approximately 60%), we expect 1 in every 10, and 1 in every 20, true parental pairs to be incorrectly excluded given error rates of 0.01 and 0.005, respectively, under these conditions. In many instances, mismatches at two loci were investigated further to identify additional possible sets of true parents.

The likelihood-based parentage analysis carried out here was performed using the program *Cervus* 3.0.3 (Marshall *et al*., 1998) with recent modifications to maximum likelihood equations (Kalinowski *et al*., 2007). The program *Cervus* uses allele frequency information to calculate LOD scores for all parent-offspring pairs; LOD values are the log of the likelihood of observing a given multilocus genotype in the candidate parent and given offspring assuming parent-offspring

relatedness *over* the likelihood of observing the two multilocus profiles assuming that the two are unrelated. The higher the score, the more likely the two are a parent-offspring set. In this analysis, offspring were tested against all candidate male and female parents, in the absence of cross information. Parentage results so obtained were then compared to the list of actual crosses. In 95% or greater of all such analyses, the pairs of parents implicated by *Cervus* at the 95% certainty level (the vast majority of assignments) were the actual crosses performed, despite the very large number (often >100) of hypothetical crosses involving one true parent and another non-true parent. Also, when *Cervus* and exclusion methods were both applied to a given set of parents and offspring, the overall concordance between the two approaches was greater than 96%. Where the two methods disagreed, one or two single-locus genotype errors were often involved (out of the nine loci used) and further analyses indicated that sometimes the parents identified by *Cervus* were actually the true parents, and sometimes the parents suggested by the exclusion analysis were the true parents.

Estimation of within-sample genetic variation: To permit comparisons of numbers of alleles observed across sample collections of varying size, the standardized number of alleles (N_a), or allele richness, was estimated using *FSTAT* version 2.9.3.2 (Goudet, 1995; Goudet, 2001), which is based on the rarefaction procedure of Hurlbert (1971). In this approach, estimates of the expected number of different alleles for each population are made by repeated sampling of 2N genes, where N is the smallest sample size of diploid genotypes present among the populations under study, for a given locus, and are given in Tables 6, 8, and 9. The observed number of alleles, *#A*, was simply the number of different alleles observed in a sample collection, with no attempts to control for sample-size effects. The observed heterozygosity was simply the proportion of genotypes exhibiting two different alleles. Gene diversity, H_e, also referred to as effective heterozygosity, was also estimated using *FSTAT*, and is the likelihood that two alleles randomly drawn from a sample are different. The extent of non-random mating, $F_{IS,}$ (*f* from Weir and Cockerham, 1984), approximately equal to (Hs-Ho/Hs), where Hs is the expected heterozygosity and Ho the observed heterozygosity within a population, and significance of departures from zero, were estimated using *FSTAT*.

Phylogenetic analyses of the G0 BSR Minto founders: Genetic distances between all pairs of 292 Minto Big Salmon River (BSR) founders (as discussed in Herbinger *et al.*, 2006) were estimated as the negative natural logarithm of the proportion of shared alleles (Bowcock *et al*., 1994) using the program *Microsat* (hpgl.stanford.edu/projects/microsat), and values were input into the *NEIGHBOUR* program of *PHYLIP* v 3.6 for Unweighted Pair Group Method with Arithmetic Mean (UPGMA) clustering (Felsenstien, 1995). Output was visualized using the program *TreeView* 1.6.6 (Roderic, 2001).

Estimation of number of eggs: The number of eggs was the total number of eggs produced by a given female, less the number of eggs estimated to have died during development or that were not successfully fertilized, through to post-shocking (the pouring of eggs from one container to another at the eyed-egg stage to cause unfertilized egg membranes to rupture and to permit their identification for removal). The total number of eggs produced per female was estimated from fork-length data for Stewiacke River LGB females, according to the equation:

$$FEC = 231.7 \times e^{(0.0513 \times FL)}$$

where *FL* is the fork length of the female and *FEC* is fecundity. Mortality was determined by tabulating daily egg loss per family through to egg shocking, after which some mortality still may have occurred, but could not be quantified for individual females, as eggs were combined for communal rearing.

Estimates of effective number of breeders and expected loss of genetic variation: Effective population size or, more correctly, variance effective population size, can be thought of as the number of individuals in an ideal population that would lose genetic variation at the same rate as the population under study. That ideal population would be one in which the number of males is equal to the number of females, where all individuals have an equal likelihood of contributing to the next generation, population size remains constant over time, and where generations are non-overlapping (Allendorf and Luikart, 2007). The actual effective population size, correcting for one or more of these factors, can be estimated using the equations to follow. It should be noted, however, that the presence of overlapping generations can complicate estimates of effective population size, and that in such instances we are really estimating the effective numbers of breeders. On the other hand, we are estimating the effective population size of a group of parents in the context of assessing loss of genetic variation in the production of their direct offspring. Given these objectives, it is not clear how important the distinction between the effective number of breeders and effective population size actually is in this study. Note, however, that these results should not be extrapolated to estimate rates of expected loss of genetic variation over multiple generations.

The effective number of breeders for a given cohort, accounting for departures from the idealized 1:1 sex ratio only (N_{b1}), was estimated using:

$N_{b1} = 4N_f N_m / (N_f + N_m)$ (approx.)

where N_f is the number of breeding females and N_m is the number of breeding males.

Single-generation effective population size for a given cohort, due to the influence of variance in family size (N_{b2}), was estimated using:

$N_{b2} = (Nk - 1) / [k - 1 + (V_k / k)]$

where N is the number of individuals, k is the mean family size, and V_k is the variance in family size.

Effective population size accounting for departures from the idealized sex ratio and variance in family size (N_{b3}) was estimated by first estimating N_{b2} for females and males separately using:

$N_{ef} = (N_f k - 1) / [k - 1 + (V_k / k)]$
$N_{em} = (N_m k - 1) / [k - 1 + (V_k / k)]$

where N_f is the number of females, N_m is the number of males, and k and V_k are the family size mean and variance for the respective sex group.

The estimates of N_{b2} are combined for the two sexes using:

$N_{b3} = 4N_{ef}N_{em} / (N_{ef} + N_{em})$

Variance in family size was based on the number of mature contributing spawners produced in the next generation from a given male or female. All formulas were from Chapter 10 of Frankham *et al*. (2002).

Expected per-generation reductions in gene diversity (effective heterozygosity) were estimated using $1/(2N)$, where N is N_{b3}, above.

Expected reductions in the number of different alleles for some comparisons were investigated by first estimating the expected total number of alleles remaining after a single generation at a given population size (*N*) using:

$$E(A') = A - \sum_{j=1}^{A} (1 - p_j)^{2N}$$

where *A* is the original number of alleles and p_j is the frequency of the *j*th allele (Allendorf and Luikart, 2007).

Analysis of the efficacy of the MKA founder selection strategy in recovering founder genetic variation: In order to evaluate the efficacy of the methods used here to select parents from the pool of available founders (Table 2) in terms of recovering genetic variation in the ancestral (G-1) generation, we compared demographic and genetic statistics between all salmon collected from the wild in a given year with both 1) the actual broodstock selected for spawning using the Mean Kinship Assisted (*MKA*) approach employed here and described above, and 2) the same number of adult salmon chosen at random, given the same sex ratio (Table 6). Results provided for the random group were obtained by randomly sampling from the pool of parents from a given year (or set of years) without replacement, five times, and averaging the five results.

Analysis of the efficacy of alternate founder-selection strategies on the recovery of genetic variation when population size is constrained: Given the objective of minimizing loss of genetic variation and the accumulation of inbreeding over time, and the initial large spawning and rearing capacity of the biodiversity facilities, a substantial percentage of the original founders were actually spawned and incorporated into the LGB program. Additionally, several hatchery conservation practices were implemented to maximize effective population size (balancing of sex ratios and selecting 5 or 10 offspring per family prior to combining egg lots). Therefore, in terms of recovering founder diversity, both the *MKA* method employed here and the random selection of spawners would be expected to recover greater than 98% of genetic variation, limiting evaluation of the *MKA* approach. In order to test the efficacy of the mate selection and spawning strategy employed here under more difficult conditions where fewer wild parr could be spawned, we tested the ability of *MKA* and three other methods to recover founder diversity using the large group of 402 wild parr collected from the Stewiacke River in 1998 (G0-98) and kinship estimates from Herbinger *et al*. (2006). Specifically, we compared levels of genetic variation in four groups of 100 founders, selected using different strategies (discussed below) from the total collection of wild founders. We chose this particular sample because 1) of its large size (N=402), 2) it may be fairly representative of the final generations of wild salmonid populations (few returning adults spawning with each other and multiple remaining male parr (Herbinger *et al*., 2006), and 3) details of the first-order relatedness of this group have been published and are available elsewhere (Herbinger *et al*., 2006).

The four broodstock selection strategies being tested here are 1) a Minimization of Mean Kinship (*MMK*; Ballou and Lacy, 1995), 2) the general Mean Kinship Assisted (*MKA*) program employed here, 3) a strategy based on Minimizing Pairwise Relatedness (*MPR*), and 4) Random Selection and Mating (*RSM*). Under *MMK*, *MK* estimates were made for all founders, as described in Ballou and Lacy (1995) using the program *PM2000* (www.vortex9.org/pm2000.html). Estimates of *MK* were based on first-order kinship estimates using *Pedigree* and *Colony*, as in Herbinger *et al*. (2006). The female and male with the lowest *MK* values were first identified. If they were unrelated based on available information (neither full nor half siblings), the pair were spawned to produce one offspring. *MK* values were then

recalculated with parents and the new single offspring included in the analysis, resulting, in some instances, in greater than one offspring being produced in successive iterations by a given pair of spawners. With the addition of each offspring, the *MK* values of the parents increased until a new female and male exhibited the lowest *MK* values. This pair was then spawned, as discussed above, and the process repeated until 50 males and 50 females were selected for comparison with the other approaches of sampling founders from the original group of wild parr collected in 1998.

The *MKA* approach, used here, was explained earlier in this report. This approach utilized *MK* values and other information to prioritize salmon for spawning.

In the *MPR* method, Ritland's (1996) pairwise genetic relatedness was first estimated between each G0-98 wild parr and each of the remaining 401 wild parr collected in 1998, using genotype information from nine microsatellite loci (see Table A3) and the program *GenAlEx 6* (Peakall and Smouse, 2006). Average pairwise distances between each wild parr and all other wild parr were calculated and the 50 males and 50 females were selected with the lowest average pairwise genetic relatedness. A limitation of this approach is that it may select multiple siblings from small families that are unrelated to other individuals in the collection, and may not represent the optimal use of pairwise distances in capturing founder genetic variation from a larger group of individuals collected from the wild.

In the *RSM* method, 50 male and 50 female salmon were chosen at random, without any information on pairwise genetic distance or kinship (first-order relatedness), from the group of 402 G0-98 salmon, as discussed above. This random sampling of 50 males and 50 females was repeated 100 times, and results averaged across all replicates.

Several statistics including observed number of alleles (#*A*), allele richness (N_a), effective heterozygosity (H_e), and observed heterozygosity (H_o) were estimated for all groups of 100 parents selected using these four different approaches, and results compared to amounts of genetic variation present in the original group of 402 G0-98 wild parr.

Tabulation of demographic and genetic information of parents and offspring: In the Stewiacke River LGB program, the original parents, or wild founders (G0), were collected from the wild in 1998 (G0-98), 1999 (G0-99), 2000 (G0-00), and 2001 (G0-01) (Table 2). These fish began to mature in 2000, 2001, 2002, and 2003, respectively (Table 3 and Figure A2). In 2000, only Stewiacke River salmon from 1998 were mature, and so only G0-98 x G0-98 crosses were carried out to produce the 2000 G1 (G1-00) group (all offspring are identified in accordance with the LGB generation that they represent and the year in which they were produced; 2001 G1 are identified as G1-01, 2002 G1 as G1-02, and 2003 G1 as G1-03). In subsequent years, salmon from one or more collection years were spawned together. For example, in 2001 previously spawned salmon from 1998, non-previously spawned salmon from 1998, and maturing salmon collected in 1999 were spawned (the crosses that year were G0-98xG0-98, G0-99xG0-99, and G0-98xG0-99). Moreover, in 2004 spawnings were carried out between wild founders and G1 produced in 2000 (e.g., G0-01xG1-00). In other words, there are no sets of parents from a given single year class and their offspring with which to estimate single-generation rates of loss of genetic variation, with the exception of the 1998 parental group (G0-98xG0-98) spawned in 2000, and their offspring (G1-00). Even here, comparisons involving this year class are complicated by the fact that G0-98 salmon were also spawned in 2001, 2002, etc., with other year classes of Stewiacke River Atlantic salmon.

In order to facilitate the assessment of rates of loss of genetic variation given the above constraints, parents and offspring were contrasted in two different ways. In the first,

comparisons were made *by spawning year* and were carried out between all salmon spawned in a given year (regardless of year class) and the offspring produced that same year (Table 8). Demographic and genetic information are contrasted between the parents spawned in a given year and their offspring as smolts and as spawning adults. We also compared genetic variation between the parents from a given year and offspring that would have been produced through a process involving the random selection of adults for spawning (assuming identical sex ratios, numbers of spawnings, and numbers of recovered offspring). Because of time constraints, only five random samples were drawn and information averaged across samples. These additional comparisons were made to demonstrate the potential loss of genetic variation resulting from mortality between the smolt and adult stage and the selection process used to identify individuals for spawning. In these analyses, several genetic statistics, such as allele richness, have also been tabulated in such a way as to compare changes in levels of genetic variation between parents and offspring (e.g., N_a across) and between groups of parents or between groups of offspring over time (e.g., N_a down). This latter comparison is confounded by 1) changing levels of genetic variation in the founders from 1998 through to 2001 (Table 2), 2) variable numbers of parents across years, and 3) variable proportions of G0-98 to G0-01 founder groups in the different spawning years. Nonetheless, this comparison may provide a general indication of the amount of variation present in the LGB through the seven years of operation.

The second main analysis carried out here, for the purpose of assessing rates of loss of genetic variation in production of G1 salmon from the G0 founders, involved the tabulation of genetic and demographic information for parents and offspring *by year class***.** Here, genetic variation was compared between parents collected in a given year, and offspring produced by those parents regardless of the year the parents were spawned. For example, for the parental group collected in 1998 (G0-98), genetic variation was compared between the parents involved in G0-98xG0-98 crosses and their offspring, regardless of the year in which the offspring were produced. It should be noted that year class of founders, as used here, refers to the year the offspring were collected from the wild and not the year the offspring were produced; offspring collected as parr in 1998 may have been produced via spawning in 1998, 1997, or 1996. Although initially this may appear to be a preferable approach to assessing rates of loss of genetic variation, it is complicated by the frequent spawning of salmon across year classes and, occasionally, between G0 and G1 generations. In other words, parents from a given year class may also contribute offspring via spawning with parents from another year class (or a parent from another generation), but such offspring cannot be included in assessments of recovery of variation from the year class of either of their two parents. Therefore, estimates of genetic variation in the offspring of a given year class should be considered minimal estimates of "recovery" of genetic variation from a respective parental group. The degree to which such estimates under-represent recovery of genetic variation relative to the parents can be inferred from the *Px* statistics, which is the proportion of offspring excluded because only one parent was from a given year class or generation being considered. This problem was further addressed by combining sets of parents (e.g., G0-98 and G0-99) and assessing variation in this larger group of parental crosses (e.g., G0-98xG0-98, G0-99xG0-99 and G0-98xG0-99) and their offspring (the benefit coming from the impact on *Px*).

Comparisons of demographic information for Stewiacke River juveniles recovered from captive and wild river environments: In order to minimize the selection for captive conditions and loss of fitness in the wild, siblings of offspring reared in captivity were released into wild river habitat as unfed or six-week feeding fry. These individuals were then captured as late-stage parr, after having experienced their early juvenile life in the wild, exposed to natural selection. Following their capture, fish were tagged, fin-clipped and genotyped as described above. This information was then used to assess parentage and to tabulate the number of

males and females from each family, number of families recovered, average family size, minimum family size, maximum family size, and variance in family size. The number of exclusive families represents the number of families recovered from one or the other environment, but not both. The number of individuals maturing and number of individuals spawned were also tabulated using parentage information.

Cryopreservation of salmon milt: Starting in 2006, and continuing each year thereafter, milt from approximately 40-50 males from different low-*MK* Stewiacke River families was cryopreserved for future use in restoring lost genetic variation (Tables A1a and A1b). First, extender or cyroprotectant solution was prepared as described in Jodun *et al*. (2006). Salmon milt from a ripe Stewiacke River male was then expressed through a small bore tube into a plastic tray. Approximately 1 ml of milt was mixed with 3 ml of extender. The above solutions, air temperature, and storage/transfer vessels were maintained at approximately 5 degrees Celsius. Milt was then drawn into 0.25 ml straws supplied by IMV (www.imv-technologies.com) and sealed using plugs supplied by the manufacturer. Straws were positioned above a tray of liquid nitrogen at a variable distance of 1 to 10 cm so as to control the rate of cooling to approximately 20-30 degrees Celsius per minute (Stoss, 1983), until the temperature was below -100 degrees Celsius, at which time they were submersed in liquid nitrogen (-196 degrees Celsius) for long-term storage. Mobility of fresh milt, and thawed previously frozen milt, was checked under a microscope and the fertilization success of thawed milt was compared to that of fresh milt from each of several males (data not shown).

RESULTS

In evaluating the amount of genetic variation lost in the selection of founders, we compared several measures of genetic diversity in wild parr collections (*Wild Parr*) from a given year class with a) broodstock chosen to produce the next generation of salmon (*Parents Selected*), and b) an equivalent number of randomly chosen parents (*Parents Random*). All measures of genetic variation were based on sets of microsatellite loci that varied somewhat from year to year (see methods section for details), so comparisons in this table are meaningful only within year classes or year class groups (e.g., across *Wild Parr*, *Parents Selected*, and *Parents Random* groups from G0-98, etc.). The number of loci on which #*A*, N_a, H_e, H_o, and F_{IS} statistics were based for each year class are given in Table 6 (# loci). For estimates of allele richness (N_a), comparisons were standardized so as to account for the differences in sample sizes across groups, with the standardized number of individuals given in Table 6 (# indiv.).

In almost all within-year class comparisons of *Wild Parr* collections and parental groups (*Parents Selected* and *Parents Random*), levels of #*A* were lower in the parent groups, though only the G0-01 *Parents Random* group was significantly lower than its respective *Wild Parr* group (p =0.016, Wilcoxon signed-rank test without correction for multiple tests, Table 6). The percent decline in the *Parents Selected* (the actual founders) relative to the *Wild Parr* collections was very small, between 0 and 3.3%, averaging 1.6% for single-year class comparisons, and as low as 1.2% for the year class set G0-98-01. In the randomly selected group of parents, however, rates of loss were as high as 6.4%, averaging 2.9% for single-year class comparisons, and 1.6% for the year class set G0-98-01. In the G0-98 year class comparison, estimates of N_a were also lower in the *Parents Selected* relative to the *Wild Parr* group, though this difference was not significant (p>0.10, Wilcoxon signed-rank test). In all other single-year class comparisons, N_a estimates were actually significantly higher (p<0.02, Wilcoxon signed-rank test without corrections for multiple tests, Table 6) in the *Parents Selected* relative to both the *Wild Parr* and *Parents Random* groups, except the *Parents Selected-Parents Random* pair for the G0-01 year class. Levels of gene diversity (H_e) were very similar across all three groups in all

year class comparisons, but were either equal or slightly larger in the *Parents Selected* group (Table 6). Observed heterozygosity (H_o) and F_{IS} values were very similar across all three groups for all year class comparisons, and no obvious trend was observed. Mean Kinship (*MK*) estimates were very similar between *Wild Parr* collections and *Parents Random* groups in all year class comparisons, but values were nearly always smaller in the *Parents Selected* samples, markedly so in the G0-01 comparison (0.0212 *Parents Selected*, 0.0373 *Wild Parr*, 0.0350 *Parents Random*).

Since the numbers of selected parents, or founders, was relatively large, particularly compared to the size of their respective *Wild Parr* collections (Table 6), it is not too surprising that a modest amount of genetic variation was lost in the selection of founders. To further test the efficacy of the Mean Kinship Assisted (*MKA*) program in minimizing loss of founder genetic variation, we compared it to three other methods, Minimization of Mean Kinship (*MMK*), Minimizing Pairwise Distances (*MPR*), and Random Selection and Mating (*RSM*) under more rigorous conditions, reducing the number of parents or founders sampled from the larger group of wild parr collected to 100 (Table 7). This exercise was carried out only on the 402 wild parr collected from the Stewiacke River in 1998.

When reducing the number of broodstock selected to 50 males and 50 females, the observed number of alleles (*#A*) recovered in the *Parents Selected* or founder groups relative to the original *Wild Parr* collection declined significantly for all broodstock selection methods tested (p<0.03, Wilcoxon signed-rank test without correction for multiple tests), except *MPR*, which declined only slightly (0.7%). Relative to the *Wild Parr* collection, *#A* declined by 8.3% in the *Parents Selected* under *MMK*, 10.3% under *MKA*, and 16.2% under *RSM*. Minimization of Mean Kinship (*MMK*), *MKA*, and *MPR* performed significantly better than *RSM* in minimizing loss of *#A* (p<0.04), and *MPR* performed significantly better than *MKA* (p=0.016, Wilcoxon signed-rank test without correction for multiple tests). No other pairwise comparisons involving *#A* were significantly different at p=0.05.

Levels of N_a in *MMK*, *MKA*, and *MPR parent* groups were significantly higher than in the 1998 *Wild Parr* collections (p<0.03, Wilcoxon signed-rank test without correction for multiple tests). Estimates of N_a in the *Wild Parr* and *RSM groups w*ere very similar.

Observed heterozygosity (H_o) values were slightly less than H_e (expected heterozygosity) for all groups, leading to small and positive F_{IS} values. Observed heterozygosity and H_e were very similar between the *Wild Parr* and *RSM Parents Selected* groups. The highest levels of H_e were observed in the *MPR Parents Selected* group, followed by the *MMK* and *MKA Parents Selected* groups. The number of full-sib families (*#FSF*) and half-sib families (*#HSF*) present in the *Wild Parr* collection was higher than all four parental founder groups, particularly so for *#FSF*. The largest numbers of half-sib and full-sib families were observed in the *MKA Parent Selected* group, and lowest in the *RSM Parent Selected* group.

We first assessed loss of genetic variation in the offspring (G1) relative to their parents (G0) *by spawning year*, where levels of genetic variation present in all parents or founders (G0 salmon) spawned in a given year were compared with levels of genetic variation present in all offspring produced that same year by those same parents, regardless of when the offspring matured and were genotyped (Table 8). So as to provide insight into where the genetic variation was lost, we compared parents (*Parents Selected*, G0 salmon) to offspring (G1) salmon as 1) smolts three years after their spawning year and before the selection of broodstock was performed (*Offspring as Smolts*), 2) adults selected as broodstock (*Offspring as Selected Adults*), and 3) adults randomly sampled (*Offspring Randomly Sampled*). The number of *Offspring as Smolts* greatly exceeded the number of offspring eventually selected as broodstock at the adult stage, and the

number of *Offspring as Selected Adults* was equal to the number of *Offspring Randomly Sampled* (Table 8). The number of observed alleles (*#A*) is very sensitive to sample size, so comparisons of this statistic between groups will reflect both the number of individuals sampled and levels of genetic diversity within the groups. On the other hand, N_a (allele richness) accounts for differences in sample size between groups, and so reflects only levels of genetic variation within a given sample, though this statistic too suffers from limitations when sample sizes vary markedly (discussed below). Because the statistic N_a across only adjusts for sample size differences across the table (between parents and offspring groups from a given year), this statistic is not useful for comparisons down the table, among years. Estimates of N_a down involve adjustments for sample size variation between years and allow comparisons of levels of genetic variation between years within a given parent or offspring group. The number of loci (# loci acr.) and number of individuals (# ind. N_a acr.) on which N_a across estimates for a given spawning year are based are given, as well as the number of loci (# loci down) and number of individuals (# ind. N_a down) on which N_a down estimates are based. Comparisons of *#A* should not be made down the table between years for any parent or offspring group because no adjustment for differences in sample size is possible and because estimates in different years are based on slightly different sets of loci that differ in average levels of variability.

Estimates of gene diversity (H_e) are given for comparisons between parent and offspring groups within a given year (H_e across) and for comparisons between spawning years within a parent or offspring group (H_e down). Although H_e is much less affected by sample size than *#A*, comparisons should only be made within the appropriate groups (across parent-offspring groups or down spawning years) because statistics for the two sets of comparisons were estimated using two different sets of microsatellite loci (# loci acr. and # loci down; Table 8).

Statistics on levels of genetic variation were only provided for offspring groups for the spawning years 2000 to 2003 because offspring for later years (2004-2007) had not yet matured or been genotyped at the time this report was prepared. Because information from these offspring groups was not available, estimates of N_a across and associated statistics for the parental groups from 2004 on could not be determined. Estimates for N_a down and H_e down are given for parents from 2004-2007 so that levels of genetic variation of parents (broodstock) could be compared across a longer (8-year) time span.

The observed number of alleles (*#A* across) in the offspring groups were lower than in their respective parental groups in 9 of 12 comparisons (Table 8). The largest declines in numbers of observed alleles were in the 2000 spawning year, where single-generation rates of loss were as high as 18% in the *Offspring as Selected Adults* group. Rates of loss in *#A* in the two groups O*ffspring as Selected Adults* and *Offspring Randomly Sampled* were lower relative to their parental groups in 2001 (7.1% and 8.1%, respectively) and 2002 (4.0% and 3.3%, respectively), but were greater in 2003 (11.0% and 15.3%, respectively). In 2000 and 2002, rates of loss were slightly greater in the *Offspring as Selected Adults* group relative to the *Offspring Randomly Sampled* group, but in 2001 and 2003, rates of loss were larger in the *Offspring Randomly Sampled* group, particularly in the 2003 comparison (11.0% versus 15.3%). In all spawning year comparisons, the smallest declines in *#A* relative to parental groups were in the *Offspring as Smolts* groups.

Allele richness values, standardizing for sample size differences across parental and offspring groups (N_a across), were less different between the parental and respective offspring groups compared to *#A*, and values were more often higher in the parental groups, and when not they were very similar. Values of N_a across were more similar still between the very large *Offspring as Smolts* groups and the smaller groups of offspring sampled at the adult stage, demonstrating again the importance of sample size (or number of individuals) on estimates of *#A* reported

above. In fact, N_a values were very similar between *Offspring as Smolts* and *Offspring Randomly Sampled* (typically differing by less than 1%), and between *Offspring as Selected Adults* and *Offspring Randomly Sampled*, though quite a bit higher in the former in 2003 (13.58 versus 12.94, respectively). Estimates of gene diversity (H_e across) were much more similar between parental and respective offspring groups, though values were often slightly higher in offspring groups, particularly in the *Offspring as Selected Adults* group (Table 8).

Estimates of allele richness, standardizing across years (N_a down), appeared to increase in both parental and offspring groups from 2000 to 2001, then decreased in 2002 and 2003, though this decline was much reduced in the *Offspring as Selected Adults* group. Gene diversity values in the parents (H_e down) were very similar across the years 2000 to 2003 (0.832) and overall tended to increase in later years (Table 8). Gene diversity values in all three offspring groups (H_e down) were more variable, and may have declined slightly over time.

Loss of genetic variation in the offspring (G1) relative to their parents (G0) was also evaluated *by year class*. Here, levels of genetic variation in parents from a given year class (e.g., 1998), spawned in any year, were compared with levels of genetic variation in their offspring, recovered as adults, in any subsequent year (Table 9). Some matings occurred between salmon from different year classes (e.g., G0-98 x G0-99) or generations (e.g., G0-00 x G1-01); such offspring were not from any single year class and so could not be directly included in tabulations of recovery of genetic variation from any particular year class. To address this, we included the statistic P_x, which is the proportion of offspring with one parent from a particular year class under consideration (e.g., 1998) and the second parent from another year class (e.g., 1999) or generation. This statistic represents the potential for additional recovery of genetic variation that could not be directly accounted for here in any single-year class (or single-generation) parent-offspring comparison. Only half of the genes in the offspring of these inter-year class (or inter-generation) crosses are from parents from the single year class under consideration, but collectively they still represent a substantial potential for recovery of genetic variation, particularly in some years. To further address this limitation, we also compared levels of genetic variation between parents from combined year class groups (e.g., 1998-2001) and their respective offspring, thus minimizing *Px*. However, this approach was not ideal because 1) the LGB program did change over time, in a direction that is expected to minimize loss of genetic variation, and 2) offspring from later year classes have not yet matured and thus could not be considered in tabulations of recovery of genetic variation in the offspring (discussed further below). We also compared genetic variation in smaller combined year class groups (1998+1999 and 2000+2001) to address expected changes in recovery of genetic variation resulting from operational changes in the program over time (Table 5) and reduced maturity of offspring of parents from later year classes.

The number of spawners (*Parents Selected*) in every year class that may have contributed offspring is indicated in the top row (*N*) and declined through time from 363 in 1998 to 108 in 2001. The number of offspring recovered from a given year class also declined through time, but much more precipitously, to the extent that all estimates of genetic variation for the 2001 group should be viewed with much caution, as they are based on too few individuals (9) for meaningful estimates of any genetic statistics presented here. This latter reduction is a result of both the offspring of these later year class parents having not yet matured or been genotyped and the number of inter-year class and inter-generation crosses having increased from 1998 though 2001, leaving fewer offspring from within-year class spawnings. This increase in crosses is due to salmon from earlier year classes maturing early in the program when only one or two year classes were present, limiting the likelihood of inter-year class and inter-generation crosses. However, in later years when salmon from the 2000 and 2001 year classes were maturing, many year classes were mature and available (1998 to 2001), increasing the

probability of inter-year class crosses. Additionally, many 2001 G0 salmon and first-generation (G1) offspring from the 2000 spawning were mature at the same time, resulting in many more inter-generation crosses.

Due to platform changes and funding availability, the suite of microsatellite loci analyzed changed somewhat over time, resulting in a variable number of loci common between parents and offspring analyzed in a given year with which to estimate various statistics (*#A*, N_a across, H_e, etc.) (Table 9). Hence, for several statistics, two estimates were provided, one maximizing the number of loci at the expense of the number of individuals (or sample size), and a second maximizing the number of individuals at the expense of the number of loci. Estimates based on more individuals but fewer loci are given inside the parentheses. For example, for the 1998 parental group, the *#A* estimate of 16.25 is based on 43 individuals and 8 loci, whereas the *#A* estimate of 13.71 is based on 92 individuals and 7 loci.

The observed number of alleles (*#A*) decreased in the offspring (*Offspring Selected*) relative to their respective parental groups (*Parents Selected*) in all comparisons (Table 9). The largest single-year class decline in *#A* between the parents and offspring was in the 2001 comparison (51.7%), but the decline in 2000 was also considerable (22.4%). Loss of genetic variation was much more modest in the combined 1998+1999 group (5.5%, 8 loci, 226 individuals; 3.7%, 7 loci, 308 individuals) and in the combined 1998-2001 group (4.1%, 8 loci, 320 individuals; 3.6% 7 loci, 400 individuals). Estimates of N_a across were very similar between parents and their respective offspring groups, sometimes slightly larger and sometimes slightly smaller in the latter (Table 9). Gene diversity (H_e) levels were also very similar between the parent and respective offspring groups (within 1.2%), but were usually larger in *Offspring Selected* groups (Table 9). Levels of H_o were also quite similar between the parental and respective offspring groups, though more often slightly larger in the latter (Table 9). Mean Kinship (*MK*) values generally tended to increase in the parental groups from 1998 through 2001, reflecting the comparatively few large kin groups, perhaps resulting from reduced numbers of returning adult spawners in latter years. The low *MK* values in parental groups from combined years (e.g., 1998+1999) likely reflect the largely different parental origins of salmon from different collection years and, therefore, reduced average relatedness among the combined group of individuals. The increase in *MK* values of the offspring over time across single-year class groupings is in part a function of the increased relatedness of the parents discussed above, but may also be an artifact of the very small sample sizes (e.g., $N = 9$) in latter years.

Estimates of the effective number of breeders, taking into account sex ratio (N_{b1}), variance in family size (N_{b2}), and both sex ratio and family size (N_{b3}) for any single parental year class, were all quite large, ranging from a minimum of 93.19 to a maximum of 362.38. For some year classes, N_{b1} estimates were smaller than N_{b2} estimates, but in other year classes, the reverse was observed, suggesting that departures from ideal sex ratios and completely even family sizes slightly impacted effective population sizes. Estimates of N_{b3} for any given parental year class group were sometimes smaller than N_{b1} or N_{b2}, as would be expected at first, but were sometimes larger. Estimates of N_b generally reflected N, except for the 1999 parental group where the estimates of effective number of breeders which considered variance in family size (N_{b2} and N_{b3}) were much reduced relative to N.

Expected parent-offspring loss of effective heterozygosity was generally small, less than half a percent, and ranged from 0.173% (1998) to 0.411% (1999) (Table 9). Rates of loss for the combined year class groups were consistently lower than for single year class groups, with less than 0.1% expected loss for the 1998-2001 combined year class group. Expected percent loss of the observed number of alleles was much higher, ranging from 3.01% (2000) to 8.42% (1999) (Table 9). The expected rate of loss of observed number of alleles for any year class or set of

year classes did not reflect numbers of parents associated with that group or the effective number of breeders estimated using any of the three approaches. Expected loss of observed number of alleles also did not seem to parallel losses in *#A* reported; in four instances, the observed loss was greater than expected, and in three instances observed loss was less than expected. Moreover, the greatest loss of alleles was observed when the second-lowest loss was expected (51.7% versus 3.84%, 2001, Table 9) and the smallest loss for any single-year class group was observed when the largest loss was expected (7.8% versus 8.42%, 1999).

In 2002, 35 crosses were carried out and offspring from each family were reared in captivity and released into wild river habitat as unfed or six-week feeding fry. Demographic information on offspring recovered from both environments at the time of collection, the late parr stage, is provided in Table 10, and estimates of family-specific egg mortality and numbers of eggs present, prior to the combining of egg lots, are given in Figure 6. Based solely on a visual inspection of the results, family size did not appear to co-vary between the captive and wild-exposed groups. Additionally, neither egg mortality nor egg number seemed to correspond with the number of offspring recovered from wild or captive environments. Given the slightly larger number of wild-exposed individuals sampled (collected and genotyped) relative to the captive-reared group, the number of families recovered per unit sampling effort was similar (30 versus 26, respectively). The average family size, maximum family size, and variance in family size were larger in the wild-exposed group compared to the captive-reared group. Twice the number of families were observed exclusively in the wild-exposed group compared to the captive group. Whereas all salmon produced in 2002 that were exposed to wild river conditions matured in 2006, slightly more than half of those reared in captivity matured that same year. Approximately 40% more wild-exposed offspring were spawned in 2006 compared to the captive-reared group.

Milt from 49 and 41 Stewiacke River males was cryopreserved in 2006 and 2007, respectively, as part of annual ongoing efforts to preserve milt from G1 males from as many families as possible. The identification of males, date of cryopreservation, and location in the storage facilities are given in Tables A1a and A1b. We used simulation analyses to estimate the proportion of genetic variation present in the original 1,029 juveniles collected between 1998 and 2001, potentially preserved and realistically recoverable, in the milt preserved to date. Specifically, we randomly sampled one of the two alleles present at each locus, in all 90 individuals, and summed the number of different alleles recovered across all loci in the sample, and expressed this as a percent of the total number of alleles observed in the 1,029 juveniles sampled from the wild. This resampling exercise was repeated 100 times. It should be noted that because we are preserving millions of gametes from each male, essentially all of the genetic variation present in the males was preserved in the frozen milt. However, given fertilization success achieved to date, and using present-day reproductive technology, the milt can only be used to fertilize the eggs from one or a few females. Furthermore, the number of adult broodstock produced using cryopreserved milt from the now many dozen, and later many hundreds, of males preserved will also be restricted. Therefore, the simulation analyses were designed to represent the recovery of genetic variation from one haploid nucleus per cryopreserved male, reflecting variation recovered in the production of one adult broodstock per cryopreserved male. The majority of genetic variation present in the 1,029 wild juveniles was represented in the samples of a single allele (representing a single gamete) from each of the 90 cryopreserved males (average 87.93%, minimum 82.89%, maximum of 93.15%), using the above criteria.

DISCUSSION

Relative to the original group of Stewicke River parr collected from the wild in 1998-2001, very little neutral molecular genetic variation has been lost in production of the latest generation of iBoF salmon to have matured and spawned. Not only was most of the genetic variation present in the original large group of wild juveniles (> 1,000 individuals) maintained in the group of founders chosen and spawned, but very little genetic variation was lost in the production of the first generation of LGB salmon from the original parents. These conclusions were based on several different approaches to assessing single-generation loss of genetic variation, and several different measures of genetic variation.

Recovery of genetic variation in the Stewiacke River founders selected for spawning: The reduction in the observed number of alleles (*#A*) in the *Parents Selected* relative to the *Wild Parr* groups in nearly all year classes and sets of year classes indicates that some genetic variation was lost when selecting founders or parents from the larger collections of wild parr. In the G0-98 year class comparison, this reduction (2%) was actually larger than the *Parents Random* group (0.8%) relative to the *Wild Parr* collections, though is still overall very small. The primary spawning year for the G0-98 year class was 2000, when kinship information was used only to avoid matings between hypothesized full and half sibs (Table 5); in this year only, the process of spawner selection, though not the pairing of males and females for spawning, was very similar to random (discussed further below). The spawners from the G0-99 and subsequent year class groups were all selected using Mean Kinship and family information (Table 5)**.** Percent declines in *#A* from the *Wild Parr* to the *Parents Selected* groups were much reduced for G0-99 salmon, but increased for G0-00 and G0-01 comparisons. This greater reduction in *#A* of the *Parents Selected* relative to the *Wild Parr* in these two latter comparisons is likely due to 1) the lower percentage of wild parr chosen for spawning as selected parents (65% for G0-00 and G0-01 versus 90% for G0-98 and G0-99), and 2) the fact that representatives of families in later year classes would have been encountered and spawned in earlier year classes. Collections from any one year consisted of fry and parr, and juveniles can remain resident for two or even three years before exiting the river as smolt. Indeed, kinship analyses indicated that families did span collections made across two and sometimes three years (data not shown). Prior spawning history of family groups was used in deciding whether to spawn an individual in any given year (Tables 4 and 5). Therefore, the best overall estimate of loss of founder genetic variation between the *Wild Parr* and *Parents Selected* groups is likely the combined year class group G0-98-01 (1.2%), which is close to the average for the four single-year class groups (1.6%). Declines for the *Parents Random* group ranged from 0.6% to 6.4%, averaging 2.9%.
Allele richness (N_a) values were very similar across the three groups, in all year class comparisons, though nearly always slightly higher in the *Parents Selected* group, indicating slightly higher levels of genetic variability, when accounting for differences in sample size. Gene diversity estimates were nearly always as high or higher in the *Parents Selected* group relative to the *Wild Parr* and *Parents Random* groups. By these measures, very little genetic variation was lost in the selection of founders from the larger wild parr collections.

The rationale behind *MMK* (Ballou and Lacy, 1995) and many other breeding strategies designed to minimize loss of genetic variation (Fernandez *et al*., 2001) is that rare alleles are the first to be lost from small isolated populations, that such alleles are likely to occur in low-*MK* individuals, and that by giving spawning preference to low-*MK* individuals the overall rate of loss of genetic variation in the population will be minimized. Although we could not follow a true *MMK* breeding program here, the prioritization of individuals for spawning based on *MK* values did seem to minimize loss of genetic variation. Levels of genetic variation (*#A*, N_a, H_e) were similar or even higher in the *Parents Random* group in the G0-98 comparison relative to the *Parents Selected* group. Values of *MK* were also very similar in these two groups, and the primary

spawning year for this year class was 2000, when spawners were essentially chosen at random. In all later years, individuals were prioritized by *MK* values, reflected in lower average *MK* values in the *Parents Selected* groups relative to the *Parents Random* groups, and all measures of genetic variation were higher in the *Parents Selected* groups relative to their respective *Parents Random* groups, except H_e for G0-00, where reported values were identical. Where levels of family structuring were highest (the G0-01 year class group, data not shown) and the potential benefits of *MK* prioritization expected to be greatest (and indeed, where the difference in average *MK* values between *Parents Selected* and *Parents Random* was the most extreme), the magnitude of the difference in performance of the two methods of selecting spawners was quite large for all three measures of genetic variation, and the greatest for any single-year class comparison. Differences in *MK* values between *Parents Selected* and *Parents Random* for the G0-99 and G0-00 single-year class groups were moderate, as were increases in performance of the *MKA* spawner selection process compared to random selection of spawners, in capturing genetic variation present in the larger wild parr collection.

Upon reducing the number of founders or spawners to only 50 males and 50 females from the original 1998 collection of 402 wild parr in the simulation study, the overall rate of loss of genetic variation increased, as did the degree of difference in performance of spawner selection methods employed (Table 7), relative to that reported above for the larger set of 363 spawners in the *Parents Selected* and *Parents Random* groups (see Table 6). All methods employing information from molecular genetic markers and the prioritization of spawners appeared to perform better than the approach involving the random selection of an equivalent number of founders from the 402 wild parr. The *MPR* method of selecting founders, not tested in the larger true founder collections discussed above, appeared to perform better than all other methods (Table 7). However, since this approach was based directly on the same nine markers that were used to estimate genetic variation in the wild parr groups, it may reflect superior performance at conserving variation at these loci specifically, and not represent levels of conservation of genetic variation across the genome. Since *MMK* and *MKA* methods are based on kinship estimates, values given here should more accurately reflect conservation across the genome, to the extent that the kinship estimates are indeed accurate. The *MMK* method performed better than *MKA* in terms of #*A*, *Na*, and *He*, but only slightly, and no differences in #*A* and *Na* were significant (Wilconxon signed-rank tests without correction for multiple tests). In terms of recovery of full- and half-sib families from the original collection of wild parr, *MKA* performed best, followed by *MMK* for half-sib families and *MPR* for full-sib families; random selection of founders performed the most poorly. The reduced performance of the *MMK* method relative to the *MKA* approach in selecting representatives from all families was surprising, but may be due to the low ranking of individuals from very large kin groups and the failure of this approach to select any representatives from some large families. Overall, although the *MMK* method may have conserved more genetic variation than *MKA*, given the modest possible increase in performance of *MMK*, the apparent increased loss of some families that the *MKA* approach conserved, and the marked increase in the degree of difficulty in carrying out all pairwise matings, this approach may not be warranted in this conservation program.

Loss of genetic variation in production of G1 salmon (evaluation by spawning year): In the first approach to evaluating the rate of loss of genetic variation in the production of G1 salmon from the G0 founders, we compared levels of genetic variation in the parents and offspring, *by spawning year*. In other words, we compared the G0 parents (*Parents Selected*) spawned in a given year (regardless of their year class) with their offspring (G1) produced that same year but selected later, at the adult stage (*Offspring as Selected Adults*), for spawning in the production of the next (G2) generation. In addition to contrasting parents (G0) and their offspring (G1) actually selected for spawning, we also evaluated levels of variation in 1) offspring (G1) sampled as smolts, prior to selection of broodstock (*Offspring as Smolts*), and

2) a hypothetical group of offspring (G1) consisting of the same number of individuals as the group of adults actually selected for spawning, by randomly sampling from the G1 generation, at the adult stage, just prior to spawning (*Offspring Randomly Sampled*). These multiple comparisons allowed evaluation of the effects of mortality from the smolt to the adult stage, independent of the sample selection process used here, and evaluation of the efficacy of the *MKA* sample selection process (whether it performed better than randomly sampling spawners).

In 2000, there was clearly a very large reduction in *#A* in the offspring groups relative to the parents spawned (14.8% to 18.0 %). In fact, even in the very large group of 429 *Offspring as Smolts*, nearly 15% of the alleles observed in the parents were lost. This major reduction in genetic variation in 2000 may reflect, at least in part, the effects of 1) the very large number of parents spawned that year (322) and, therefore, large levels of *#A* available to lose in this group (Tables 2 and 8), 2) the small number of loci (four) with which this statistic was estimated compared to other years and the increased possible role of chance in estimates of variability, and 3) the way offspring of families were sampled from all 174 families produced that year. In 2000 and 2001, 10 eggs were sampled from each family *after* egg lots were combined, instead of before the pooling of eggs. In addition to increasing variance in family size, this method of sampling also resulted in multiple families not being recovered in the G1 generation at all, despite the large number of offspring analyzed in this group (429 as smolts). The further reduction in levels of genetic variation in the *Offspring as Selected Adults* group mostly reflects the markedly reduced number of individuals (133) over which to recover genetic variation from the parental group relative to the *Offspring as Smolt* group. The decline in *#A* was higher in the *Offspring as Selected Adults* group relative to the random offspring group sampled at the same stage (as adults), though the difference was small (18.0% versus 17.4%, respectively). Gene diversity estimates (H_e across) were very similar across all four groups, though slightly higher in the *Offspring as Selected Adults*.

In 2001, *#A* in the *Offspring as Smolts* group did not appear to decline relative to the *Parents Selected* group, and the two groups of individuals exhibited similar levels of N_a. This high level of retention of genetic variation may reflect the large number of offspring sampled and genotyped (407) relative to the number of parents (226). Similarly, the reduction in *#A* in the *Offspring as Selected Adults* and the *Offspring Randomly Sampled* groups probably reflects the smaller sizes of these groups relative to the parents (145). The loss in *#A* was greater in the *Offspring Randomly Sampled* compared to the *Offspring as Selected Adults*, and allele richness was highest in the *Offspring as Selected Adults*. The similar levels of N_a in the *Offspring Randomly Sampled* and *Offspring as Smolts* groups again suggests a minimum role of mortality between the smolt and adult stage. As in 2000, 10 individuals were sampled from each family after egg lots were combined. Again, gene diversity estimates (H_e across) were very similar across all four groups, highest in the *Parents Selected* and second highest in *Offspring as Selected Adults*, declining by only 0.24%.

In 2002, there appeared to be a modest increase in *#A* in the *Offspring as Smolts* group relative to the parental group (3.9%). Although the number of offspring sampled as smolt (444) was very high relative to the number of parents crossed (200) and little loss of genetic variation was expected based on the number of breeders alone, an actual increase in number of observed alleles was not biologically possible, except via mutation. These results may also be due to missing genotypes present in the parental group that involved rare alleles; if such alleles were passed on to offspring that were genotyped successfully at these loci, a perceived increase in *#A* would have been observed. Values of N_a across were less different. In the *Offspring as Selected Adults* and *Offspring Randomly Sampled* groups, levels of genetic variation (*#A*) were observed to decline relative to the parents, but by very little (3.9% and 3.3%, respectively).

Gene diversity estimates were higher in the *Offspring as Selected Adults* compared to the *Parents Selected* group.

In 2003, a slight increase in *#A* was observed in the *Offspring as Smolts* relative to the *Parents Selected* (1.3%), but a drop was noted in the *Offspring as Spawning Adults* (11.0%) and *Offspring Randomly Sampled* (15.3%) groups. The large loss of alleles in the adult groups relative to the smolt group most likely reflects the markedly reduced number of offspring (59) in these two groups relative to the *Offspring as Smolts* group (413). This reduction, relative to earlier years, is a result of incomplete maturation of offspring of 2003 spawnings at the time this report was compiled; although more offspring did mature, they were not genotyped in time for these analyses, but were incorporated into spawnings carried out in 2008 (and will likely be represented in subsequent years as well). The much higher levels of genetic variation in the *Offspring as Selected Adults* group relative to the *Randomly Sampled Offspring* may reflect the fact that 2003 was the primary spawning year of the 2001 year class, where the highest level of family structuring was observed (data not shown), as reflected by the markedly higher values of *MK* in this year class relative to other year classes (see Tables 6 and 9). Because of the higher levels of structuring, a spawner selection process utilizing *MK* values would have the greatest opportunity to demonstrate increased performance relative to random selection of spawners. The increased performance of the *MKA* approach for the group of salmon spawned in 2003 was also reflected in the N_a estimates of the four groups (Table 8). Gene diversity values (H_e across) were very similar in all offspring groups, and were slightly higher compared to the *Parents Selected* group.

Looking across spawning years in Table 8, levels of genetic variation did not appear to decline in the pool of parents spawned, or in their resulting offspring. Estimates of *#A* varied slightly over the years but no trend was apparent. Levels of allele richness (N_a down) actually appeared to increase over time in the parental group, with the exception of 2007**,** where N_a down was slightly lower. With the exception of the 2006 spawning year, levels of gene diversity (H_e down) were generally higher in the parental group from the second half of the period monitored relative to the first half. However, this may be a result of the presence of an increasing number of different year classes, each consisting of slightly different family compositions (see Table 8, Parental Group types column). For example, the group of salmon spawned in 2000 consisted of only 1998 spawners, those spawned in 2001 of founders from 1998 and 1999, and in 2001 founders from 1998, 1999 and 2000, etc.). No obvious trend in N_a down or H_e down in the offspring groups was apparent.

Overall, when viewed by spawning year, levels of *#A* appeared to have declined in the offspring groups relative to their parents by 3.9% to 18.0%. However, these comparisons do not provide complete accounting of the genetic variation retained in the production of G1 salmon for several reasons. First, alleles present in the 1998 broodstock spawned in 2000 and not recovered in the offspring produced that same year may have been recovered through the spawning of G0-98 salmon in subsequent years. This may be particularly important here because the spawning history of individuals during the prioritization of spawners was considered in 2001 onward. Second, alleles present in 1998 broodstock spawned in 2000 and not recovered in the offspring produced that same year may have been present in other year class collections, and may have been recovered in their offspring spawned in subsequent years. Third, the high estimate of loss of genetic variation in 2000 was based on four loci, and is, therefore, susceptible to various stochastic effects, as discussed above. Fourth, this same estimate was from a spawning year employing sub-optimal egg sampling regimes not used in 2002 onward, and is, therefore, not representative of the program in all but the earliest years. Fifth, the reported loss in 2003 is due to the fact that less than half (59) of the offspring produced that year could be genotyped and spawned; the majority of offspring were genotyped and spawned the following year. Incomplete

maturity and spawning also impacted results from 2002, but to a lesser extent. Finally, it is important to note that in addition to monitoring the spawning history of all genetically important individuals, as noted above, we also monitored previous spawning history of families, egg survival to hatching, and recovery of offspring from individual spawners and families. This information was incorporated into prioritizing individuals in later spawning years, allowing recovery of additional genetic variation. Nonetheless, these analyses do indicate the magnitude of potential loss of genetic variation possible, despite large numbers of breeders, in offspring produced by parents spawned in a given year, without genetic monitoring and adaptive program management.

Loss of genetic variation in production of G1 salmon (evaluation by year class): When loss of genetic variation between G0 and G1 salmon is evaluated *by year class*, a more complete assessment of program efficacy is possible. In the single-year class comparisons, the reported loss of *#A* between *Parents Selected* and respective *Offspring Selected* groups was variable and could be quite high, particularly for the 2001 year class comparison (51.7%). Here, although spawner selection criteria and family or egg sampling methods were optimized, very few offspring were produced by G0-01 x G0-01 spawnings, and even fewer of these within-year class offspring were mature and genotyped at the time this report was produced (N = 9). Clearly, the very high rate of loss mostly reflects the fact that the vast majority of the offspring of these parents were not accounted for. The second-highest parent-offspring loss of *#A* was for the 2000 comparison (22.4%). Factors associated with the reduction in *#A* include those described for the G0-01 comparison. Because of a one-time decision made in 2002 to delay spawning of the primary year class group of salmon (2000) by one year, this year class group is as affected by incomplete maturity and genotyping information as the 2001 year class group discussed above.

In the 1999 year class, loss of observed number of alleles was only 7.8%. Here, most of the offspring had matured and were genotyped for inclusion in estimates of loss of genetic variation, and the only obvious factors contributing to the moderately high levels of loss were 1) still high values of *Px*, and 2) the sampling of offspring after pooling families. In the 1998 comparison, the estimate of decline in *#A* based on 7 loci (9.3%) is probably more accurate than that estimated from 8 loci (16.1%), as the latter is based on relatively few individuals (43). Factors associated with the decline observed in this single-year class parent-offspring comparison include those reported for the 1999 comparison, but the large number of parents spawned, and hence large amounts of variation to recover, likely contributed.

Combining year class groups reduced *Px*, yet very different results were observed for the 1998+1999 and 2000+2001 combined groups. In the 1998+1999 comparison, where *Px* was fairly low (0.064), the parent-offspring percent decline in *#A* was between 3.7% and 5.5%. Although a marked improvement relative to any single-year class comparison was observed, parents from this set of years were spawned early in the program, prior to the optimization of procedures designed to minimize loss of genetic variation (e.g., *MK* information was not used in ranking salmon in 2000 and families were sampled from combined egg lots, greatly increasing family variance and resulting in some families not being recovered). This and the small, but not negligible, value of *Px* indicate that these values do not represent overall program performance in minimizing loss of genetic variation to date. For the 2000+2001 comparison, the reported decline in #A was much higher, 13.7%, despite improvement in methods used in prioritizing and selecting salmon for spawning, and in the sampling of families. Here, the large reduction in #A from parent to offspring groups is in part a function of the very high *Px* value (0.373) even after combining adjacent years. The higher *Px* value in this combined group is a result of an increasing number of year classes maturing and spawning in later years of the program, and the occurrence of *among-generation* crosses beginning in 2004, involving G0-01 and G1-00

salmon. Another very important factor contributing to the reported decline for this group is the incomplete maturity of offspring of these latter year classes, reflected by the very small number of mature genotyped offspring (64). In the 1998-2001 comparison, however, where *Px* dropped to its lowest level, 0.016, the parent-to-offspring loss in #A was 3.6% and 4.1%, again varying slightly depending upon the set of loci (and individuals) on which this statistic is based. Even here, the reported recovery of #A (~96%) is an underestimate of actual program performance because *Px* is not zero and genetic variation in some G0-01 parents spawned with G1-00 adults may be recovered in the resulting offspring but, more importantly, as mentioned above for the 2000+2001 comparison, many offspring of the G0-01 parents have not yet matured or been genotyped.

In comparison, the percent expected loss of *#A*, based on the number of breeders and observed allele frequency distributions following Allendorf and Luikart (2007), varied from 1.90% to 8.42% for various year class comparisons. Most importantly, the expected loss of *#A* for the combined year class group 1998-2001 was fairly high, 7.92%. The magnitude of loss expected is surprising given the number of parents (821) and effective number of breeders (467.58 to 815.85), as loss of genetic variation due to drift is expected to be inversely proportional to effective population size. In this analysis, the substantial reduction expected appears to reflect large numbers of very rare alleles observed at the loci surveyed in this combined collection. The better than expected performance may reflect the use of *MK* values to rank and spawn salmon. As discussed above, salmon with low *MK* values are likely to exhibit rare alleles (Ballou and Lacy, 1995). By giving spawning preference to these salmon, the chances of losing rare alleles may be reduced, below that expected given allele frequencies and effective population size.

Allele richness values in the parents and offspring were much more similar (generally within 5%) and were sometimes greater in the parents and sometimes greater in the respective offspring group. Not surprising, allele richness estimates were slightly higher in the offspring relative to their parent groups in the 1998+1999 and 1998-2001 comparisons, where reductions in *#A* were lowest. Gene diversity estimates were higher in the offspring groups relative to their parental groups in nearly all comparisons, indicating that this measure of genetic variation did not decline in the offspring relative to the parents. In all comparisons, average *MK* values were quite a bit higher in offspring relative to their parent groups. Although *MK* is expected to truly increase through time in small closed populations, the marked increases observed here are also an artifact of the fact that *MK* values of offspring are based on an additional generation of pedigree information compared to their parents.

The effective number of breeders based on sex ratio (N_{b1}), variance in family size (N_{b2}) or both sex ratio and family size (N_{b3}) were all relatively large, approximately 100 or greater. In some instances, estimates of N_b were larger than the number of spawners, and in others smaller, but never less than half, and usually greater than three-quarters of the census population size. That no consistent trend in the effects of different variables on effective population size were observed (sometimes N_{b1} estimates were larger than N_{b2} and vice versa) indicates that different factors impacted effective population size to varying degrees in different years. Effective population size did not appear to be a good predictor of percent loss of *#A* (Table 9). The second-smallest loss of *#A* for any single-year class comparison (1998, 16.1%) was associated with the largest estimates of effective population size (362.38, 345.73, 289.63), even though loss of genetic variation should be minimal given this many effective breeders (see Fraser, 2008), and the smallest reduction (1999, 7.8%) with the lowest estimates of number of breeders (170.99, 107.79, 121.67). This observation is consistent with the assertion that other factors, such as *Px* and incomplete maturity or recovery of offspring from later year classes, are primarily responsible for reports of lower levels of genetic variation recovered in single-year

class analyses, though the chance number of rare alleles in different year class collections is also likely very important.

The expected loss of H_e, based on 1/(2N) where N is the effective number of breeders (Frankham *et al*. 2002, p. 190), was relatively low, less than 0.5% in all year class comparisons and less than 0.25% in all but two of seven year class comparisons.

Genetic variation in fitness-related traits may be lost through drift as discussed above for neutral loci, but also through unintentional selection for captive conditions. To minimize the effects of selection on reducing genetic variation for such traits, offspring are exposed to native river habitat as early juveniles, and recovered as late-stage parr or smolt for later genetic analysis, pedigree placement, and possible spawning to produce the next generation (Figure 2). The extent to which this strategy can be effective depends in part on the number of families recovered from the wild, the number of wild-exposed families spawned, and the relative contribution of wild-exposed and captive-reared siblings to the next generation.

Analyses of a group of 35 families from controlled captive and wild native river environments yielded a similar number of families per unit of sampling effort (Table 10). This high rate of recovery of families from the wild river environment was somewhat surprising given that 1) family size was equalized in the captive but not the wild environment, and 2) family size was small (approximately 10) in the captive group, but was unrestricted (and in excess of 1,000) in the wild group. Both situations should have contributed markedly to variance in family size in the sample from the wild, thereby reducing recovery of families. Although variance in family size was indeed higher in the similarly sized wild collection (13.58 versus 7.83), the effect on recovery of families was minimal (Table 10). In fact, eight families were found only in the collection of wild-exposed offspring, whereas only four were exclusive to the captive-reared group.

A greater number of wild-exposed salmon were spawned from this group in 2006, relative to their captive-reared siblings (61 versus 44). Additional bias in favour of wild-exposed siblings was constrained by 1) the very high variance in family size and the very large size of many wild-exposed families (Figure 6), 2) the presence of some families in the captive-exposed group only, 3) the need for additional females in 2006 to balance sex ratios in the remaining broodstock, and 4) research requirements associated with an evaluation of the effects of early wild exposure as juveniles on subsequent offspring survival (Table 11). Wild-exposed siblings were more heavily favoured over captive siblings in the 2007 selection of spawners, including the incorporation of information on the family size of wild-exposed groups (Table 5).

A considerable body of research is currently being carried out on Stewiacke River and other iBoF populations that has direct relevance for the management of iBoF salmon and the ability of the present program to conserve genetic variation and fitness (Table 11). In many instances, studies are hoped to be prescriptive, and may serve to guide management decisions regarding issues for which there are potentially important tradeoffs. Examples include whether to carry out spawnings within populations at the risk of increased inbreeding and possible inbreeding depression, or whether to carry out crosses among salmon from different iBoF populations and risk possible outbreeding depression and the loss of local adaptation. Another study aims to investigate the possible fitness costs of preferential spawning of fish with low *MK* values relative to possible benefits in terms of maintaining genetic variation, with the concern being that rare fish in the wild have a survival disadvantage relative to more common fish.

CONCLUSION

Overall, these results indicate that relative to amounts of genetic variation present in the very large collection of wild juvenile parr from 1998 to 2001 (>1000 individuals), little neutral molecular genetic variation has been lost to date through genetic drift. Approximately 99% of the original founder genetic variation has been recovered in the selection of broodstock, regardless of the measure of genetic variation considered. Rates of loss of genetic variation in the production of offspring (G1) from the parents or founders (G0) were also minimal. Gene diversity estimates were usually higher in the selected offspring relative to their parents, and lower in only two of seven year class comparisons by 0.25% and 1.2%. The reductions that were observed were generally similar to that expected based on N_b, and toward the lower end of the distribution of that reported for other captive populations of Atlantic salmon (0.1% to 2.5%, with most greater than 0.5%; Fraser, 2008). Even for those few populations for which some loss was reported, reductions were in line with or less than the maximum rate of 1% suggested by Frankel and Soulé (1981) for captive breeding programs. Clearly, loss of gene diversity, a less sensitive indicator of population bottlenecks than number of alleles, is likely minimal in the iBoF LGB program. The more common observation of slightly higher levels of gene diversity in the offspring relative to their respective parental groups may reflect the combined effects of minimizing loss of genetic variation (a result of the large number of breeders and the *MKA* method used to rank and select parents) and the slight outbreeding resulting from the intentional avoidance of matings between full and half siblings. Analyses of the other key measure of genetic variation assessed, number of observed alleles (*#A*), a much more sensitive measure of population bottlenecks, indicate that a maximum of 3-4 % of the alleles present in the G0 parents was lost in production of G1 offspring. However, even these rates of loss were based on offspring groups that under-represent the amount of genetic variation recovered in the program principally because 1) many crosses were between year classes and generations, and 2) many offspring of G0-00, and especially G0-01, had not matured and were not genotyped at the time this report was produced. Therefore, rates of loss of observed alleles may be quite a bit lower than the 3-4% reported, possibly as low as 2-3%.

Additional genetic variation, approximately 90% of that observed in the 1,029 juveniles collected from the Stewiacke River in 1998-2001, is also being maintained in banks of cryopreserved milt housed at the Bedford Institute of Oceanography. This limited gene pool is likely to be comprised of mostly common or moderately common alleles and, to a lesser extent, rare alleles, whereas most of the alleles lost from the Live Gene Bank salmon are likely to be the more rare alleles. However, because somewhat different rare alleles are expected to be maintained in the two collections, this material will help mitigate loss of genetic variation in the living lineages of salmon maintained in captivity.

ACKNOWLEDGMENTS

A large number of individuals from several different organizations have contributed either samples, analyses, operational support, or advice, all of which were essential in the design and implementation of the inner Bay of Fundy Live Gene Banking program analyzed here. Peter Amiro, Shane O'Neil, Ross Jones, and Jason Flanagan participated in the original collection of founders and the ongoing collection of wild-exposed juveniles. Beth Lenentine and Bev Davison and crew at Coldbrook and Mersey Biodiversity facilities, and Trevor Goff and Danielle MacDonald and colleagues from the Mactaquac Biodiversity Facility, performed all captive breeding and rearing and helped with release of juveniles into the wild. Roger Doyle, Eric Verspoor, Ellen Kenchington, and Christophe Herbinger assisted with the design of program elements associated with the maintenance of genetic variation and minimization of adaptation to

captive conditions. Staff from the Diadromous Fish Division and members of the inner Bay of Fundy Planning and Recovery teams, including representatives of the Atlantic Salmon Federation and Nova Scotia Salmon Association, Fort Folly First Nations, the Province of Nova Scotia, and the Province of New Brunswick, helped develop and guide the overall program direction and implementation. Manon Cassista-Da Ros, Kris Gordos, Andrea Cox, Danielle Mathews, Daniela Denti, Julie Porter, and Eric Verspoor and colleagues from the FRS Freshwater Laboratory, Pitlochry, Scotland, UK, assisted with laboratory analyses. Thanks to participants of the CSAS Biodiversity Facility Review for reviewing earlier drafts of this manuscript, and to Dylan Fraser for comments on the near final version.

REFERENCES

Allendorf, F.W., and G. Luikart. 2007. Conservation and the genetics of populations. Blackwell Publishing, USA.

Araki, H., B. Cooper, and M.S. Blouin. 2007a. Genetic effects of captive breeding cause a rapid, cumulative fitness decline in the wild. Science 318: 100-103.

Araki, H., W.R. Ardren, E. Olsen, B. Cooper, and M.S. Blouin. 2007b. Reproductive success of captive-bred steelhead trout in the wild: Evaluation of three stocking programs in the Hood River. Conservation Biology 21: 181-190.

Ballou, J.D., and R.C. Lacy. 1995. Identifying genetically important individuals for management of genetic variation in pedigreed populations; pp. 76-111. In J. Ballou, M. Gilpin, and T.J. Foose (Eds.). Population management for survival and recovery: Analytical methods and strategies in small populations. Columbia University Press, New York, New York, USA.

Blouin, M.S. 2003. DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. Trends in Ecology and Evolution 18: 503-511.

Bowcock, A.M., A. Ruíz-Linares, T. Tomfohrde, E. Minch, J.R. Kidd, and L.L. Cavalli-Sforza. 1994. High resolution human evolutionary trees with polymorphic microsatellites. Nature 368: 455-457.

DFO. 1997. Atlantic salmon inner Bay of Fundy SFA 22 and part of SFA 23. DFO Science Stock Status Report D3-12.

DFO. 2008. Recovery Potential Assessment for inner Bay of Fundy Atlantic salmon. DFO Canadian Science Advisory Secretariat Science Advisory Report 2008/050.

DFO. 2009. Recovery Strategy for the Atlantic salmon (*Salmo salar*), inner Bay of Fundy populations [Proposed]. In *Species at Risk Act* Recovery Strategy Series. Ottawa: Fisheries and Oceans Canada.

Doyle, R.W., R. Perez-Enriquez, M. Takagi, and N. Taniguchi. 2001. Selective recovery of founder genetic diversity in aquacultural broodstocks and captive, endangered fish populations. Genetica 111: 291-304.

Ducharme, L.J.A. 1969. Atlantic salmon returning for their fifth and sixth consecutive spawning trips. Journal of the Fisheries Research Board of Canada, 26: 1661–1664.

Felsenstien, J. 1995. PHYLIP. Phylogeny inference package, version 3.5. Department of Genetics, University of Washington, Seattle, Washington, USA.

Fernandez, J., M.A. Toro, and A. Caballero. 2001. Practical implementation of optimal management strategies in conservation programmes: A mate selection method. Animal Biodiversity and Conservation 24: 17–24.

Frankel, O.H., and M.E. Soulé. 1981. Conservation and evolution. Cambridge University Press, Cambridge, UK.

Frankham, R. 2008. Genetic adaptation to captivity. Molecular Ecology 17(1): 325-333.

Frankham, R., J.D. Ballou, and D.A. Briscoe. 2002. Introduction to conservation genetics. Cambridge University Press, Cambridge, UK.

Fraser, D.J. 2008. How well can captive breeding programs conserve biodiversity: A review of salmonids. Evolutionary Applications 1(4): 535-586.

Fraser, D.J, A.S. Houde, P.V. Debes, P.T. O'Reilly, J. Eddington, and J.A. Hutchings. 2010. Consequences of farmed-wild hybridization across divergent wild populations and multiple traits in salmon. Ecological Applications 20: 309-327

Fraser, D.J., L.K. Weir, T.L. Darwish, J. Eddington, and J. Hutchings. 2007. Divergent compensatory growth responses within species: Linked to contrasting migrations in salmon? Oecologia 153: 543-553.

Goudet, J. 1995. A computer program to calculate F-statistics (Version 1.2). Journal of Heredity 86(6): 485-486.

Goudet, J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from www.unil.ch/popgen/softwares/fstat.htm

Hansen, M.M., and L.F Jensen. 2005. Sibship within samples of brown trout (*Salmo trutta*) and implications for supportive breeding. Conservation Genetics 6: 297-305.

Herbinger, C.M., P.T. O'Reilly, and E. Verspoor. 2006. Unravelling first-generation pedigrees in wild endangered salmon populations using molecular genetic markers. Molecular Ecology 15: 2261-2275.

Huntsman, A.G. 1931. The Maritime salmon of Canada. Biological Board of Canada, Bulletin XXI.

Hurlbert, S.H. 1971. The nonconcept of species diversity: A critique and alternative parameters. Ecology 52: 577-586.

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. Journal of the Fisheries Research Board of Canada 33: 829–833.

Jodun, W.A., K. King, P. Farrell, and W. Wayman. 2006. Methanol and egg yolk as cryoprotectants for Atlantic salmon spermatozoa. North American Journal of Aquaculture 69: 36–40.

Kalinowski, S.T., M.L. Taper, and T.C. Marshall. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Molecular Ecology 16: 1099-1106.

Kapuscinski, A.R. 1991. Genetic analysis of policies and guidelines for salmon and steelhead hatchery production in the Columbia River Basin. Prepared for the Northwest Power Planning Council, Agreement 90-037, Portland, Oregon, USA.

King, T.L., M.S. Eackles, and B.H. Letcher. 2005. Microsatellite DNA markers for the study of Atlantic salmon (*Salmo salar*) kinship, population structure, and mixed-fishery analyses. Molecular Ecology Notes 5: 130-132.

Lacroix, G.L., and D. Knox. 2005. Distribution of Atlantic salmon (*Salmo salar*) postsmolts of different origins in the Bay of Fundy and Gulf of Maine and evaluation of factors affecting migration, growth, and survival. Canadian Journal of Fisheries and Aquatic Sciences 62: 1363-1376.

Lacroix, G.L., D. Knox, and J.W. Stokesbury. 2005. Survival and behaviour of post-smolt Atlantic salmon in coastal habitat with extreme tides. Journal of Fish Biology 66: 485-498.

Langefors, A., J. Lohm, M. Grahn, O. Anderson, and T. von Schantz. 2001. Association between major histocompatibility complex class IIB alleles and resistance to *Aeromonas salmonica* in Atlantic salmon. Proceedings of the Royal Society of London, Series B, 268: 479-485.

Marshall, T.C., J. Slate, L.E.B. Kruuk, and J.M. Pemberton. 1998. Statistical confidence for likelihood-based paternity inference in natural populations. Molecular Ecology 7: 639-655.

Miller, K.M., J.R. Winton, A.D. Schulze, M.K. Purcell, and T.J. Ming. 2004. Major histocompatibility complex loci are associated with susceptibility of Atlantic salmon to infectious hematopoietic necrosis virus. Environmental Biology of Fishes 69: 307-316.

Miller, L.M., and A.R. Kapuscinski. 2003. Genetic guidelines for hatchery supplementation programmes; pp. 329-355. In E.M. Hallerman (Ed.). Population genetics: Principles and applications for fisheries scientists. American Fisheries Society, Bethesda, Maryland, USA.

Moritz, C. 1999. Conservation units and translocations: Strategies for conserving evolutionary processes. Hereditas 130: 217-228.

Nevo, E. 1978. Genetic variation in natural populations: Patterns and theory. Theoretical Population Biology 13: 121-177.

O'Reilly, P., and R. Doyle. 2007. Live Gene Banking of Endangered Populations of Atlantic Salmon; pp. 436-482. In E. Verspoor, L. Stradmeyer, and J.L. Nielsen (Eds.). The Atlantic salmon: Genetics, conservation and management. Blackwell Publishing.

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

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

Peakall, R., and P.E. Smouse. 2006. GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6: 288-295.

Ritland, K. 1996. Estimators for pairwise relatedness and individual inbreeding coefficients. Genetic Research 67:175–185.

Roderic, D.M. 2001. TreeView (http://taxonomy.zoology.gla.ac.uk/rod/treeview.html).

Rudnick, J.A., and R.C. Lacy. 2008. The impact of assumptions about founder relationships on the effectiveness of captive breeding strategies. Conservation Genetics 9: 1439-1450.

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

Stoss, J. 1983. Fish gamete preservation and spermatozoan physiology; pp. 305-350. In W.S. Hoar, D.J. Randall, and E.M. Donaldson (Eds.). Fish Physiology, Vol. 9, Part B. Academic Press, New York, New York, USA.

Van Oosterhout, C., A.M. Smith, B. Hanfling, I.W. Ramnarine, R.S. Mohammed, and J. Cable. 2007. The guppy as a conservation model: implications of parasitism and inbreeding for reintroduction success. Conservation Biology 21(6): 1573-1583.

Verspoor, E., M. O'Sullivan, A.L. Arnold, D. Knox, and P.G. Amiro. 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.

Wang, J. 2004. Sibship reconstruction from genetic data with typing errors. Genetics 166: 1963-1979.

Weir, B.S., and C.C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38: 1358-1370.

TABLES

Table 1. Information considered in the prioritization of inner Bay of Fundy rivers for conservation efforts.

Sample Collection	Basin	MtDNA Clade 1-3^{+}	Primary Life-history Characteristics	Primary Responsible Agency	Allele Richness (Variance)	Gene Diversity
Primary iBoF Live Gene Bank populations						
Big Salmon R.	C	no	L	DFO	11.35 (197.04)	0.848
Stewiacke R.	M	yes	L	DFO	10.61 (187.86)	0.832
Secondary iBoF Live Gene Bank populations						
Gaspereau R.	M	yes	D	DFO	8.48 (188.38)	0.756
Great Village R.	M	yes	L	DFO	8.68 (202.35)	0.807
Economy R.	M	yes	L	DFO	6.81 (135.72)	0.738
Harrington R.	M	n/a	L	DFO	n/a	n/a
Black R.	C	no	L	DFO	n/a	n/a
Irish R.	C	no	L	DFO	n/a	n/a
Upper Salmon R.	C	n/a	L	EC	9.27 (140.45)	0.791
Non inner Bay reference population						
Saint John R.	n/a	no	D	DFO	11.67	0.850

Modified from O'Reilly and Doyle (2007). Estimates of gene diversity and allele richness, discussed within, are based on sample collections from the year 2000, except for Gaspereau, Saint John, and Upper Salmon collections, which were obtained in 1999, 2000, and 2001, respectively.

C	Chignecto Basin.
M	Minas Basin.
MtDNA clade 1-3^{+}	Two related mtDNA haplotypes found at high frequency in multiple Minas Basin rivers that have not been observed outside the inner Bay of Fundy (Verspoor *et al*., 2002).
L	Local migration, high one-sea-winter component, high incidence of multiple repeat spawning.
D	Distant migration, one- and multi-sea-winter components.
DFO	Department of Fisheries and Oceans, Canada.
EC	Environment Canada.
Allele richness	Standardized number of alleles observed, estimated by standardising to the smallest sample size (N=42) using resampling procedures.

Table 2. Number and stage of collection of wild salmon (founders) recruited into the primary Live Gene Bank (LGB) programs.

River of Origin	Year	Number of Juveniles Collected	Abbreviation Used in Text	Gene Diversity (H_e)	Allele Richness (N_a)	Observed Number of Alleles (#A)	Number of Juveniles recruited INTO Primary IBoF LGBs
Stewiacke	1998	402	G0-98	0.8310	12.420	14.57	363
Stewiacke	1999	188	G0-99	0.8255	13.168	13.71	171
Stewiacke	2000	273	G0-00	0.8315	12.369	12.86	179
Stewiacke	2001	166	G0-01	0.7947	12.386	12.42	108
Big Salmon	1998	268	G0-98	0.8398	16.003	NA	NA
Big Salmon	1999	216	G0-99	0.8406	14.416	NA	NA
Big Salmon	2000	313	G0-00	0.8479	13.217	NA	NA
Big Salmon	2001	304	G0-01	0.8379	12.755	NA	NA
Big Salmon	1998	292	G0-98	0.8182	11.017	NA	NA

Allele richness based on 155 (Stewiacke) and 213 (Big Salmon) individuals and 7 loci, and calculated as described within.

Table 3. Schedule of inner Bay of Fundy Live Gene Bank operations.

1	2		3	4	4	1	2			3	4	4
Collection of founder (G0) broodstock	Captive rearing through to maturity	Tissue sampling, DNA fingerprinting, pedigreeing G0 broodstock	Spawning of G0 founder broodstock*	Retention/ river release of G1 fry, G1 parr	River release of G1 smolt	Capture of wild-exposed G1 fry as late parr/ smolt **	Captive rearing of wild-exposed G1 parr/ smolt through to maturity	Possible adult return and wild spawning of G1 released as smolts	Tissue sampling, finger-printing, Pedigreeing captive-reared G1	Spawning of G1 salmon; priority given to wild-exposed individuals***	Retention/ river release of G2 fry, G2 parr	River release of G2 smolt
1998	1998-2000	2000	2000	2001	2002	2003	2003-2005	2003/2004	2004	2004	2005	2006
1999	1999-2001	2001	2001	2002	2003	2004	2004-2006	2004/2005	2005	2005	2006	2007
2000	2000-2002	2002	2002	2003	2004	2005	2005-2007	2005/2006	2006	2006	2007	2008
2001	2001-2003	2003	2003	2004	2005	2006	2006-2008	2006/2007	2007	2007	2008	2009
2002	2002-2004	2004	2004	2005	2006	2007	2007-2009	2007/2008	2008	2008	2009	2010
2003	2003-2005	2005	2005	2006	2007	2008	2008-2010	2008/2009	2009	2009	2010	2011

Modified from O'Reilly and Doyle (2007).

* Some individuals will not mature and spawn for an additional one to two years, and all subsequent steps will be delayed accordingly; less common broodstock that do spawn in year one may be spawned a second or third time in subsequent years.

****** 2+ parr and 3+ smolt from a given year class may be captured the following year.

*** Captive and wild-exposed salmon from a given year class will also mature for the first time in one or two following years.

Table 4. Factors considered in the ranking and spawning of iBoF Atlantic salmon.

Factor	Impact
MK value	Retention of genetic variation/accumulation of inbreeding
Gender	Spawning success
Relatedness (pairwise)	Maximum avoidance of inbreeding
Previous spawning history	Retention of genetic variation/minimization of inbreeding in next generation
Full-sib grouping	Retention of genetic variation/minimization of inbreeding in next generation
Kin grouping	Retention of genetic variation/minimization of inbreeding in next generation
Previous spawning history of sibs	Retention of genetic variation/minimization of inbreeding in next generation
Wild versus captive environment as juveniles	Early exposure to natural selection/maintenance of wild fitness
Relative size of family in the wild	Early exposure to natural selection/maintenance of wild fitness
Present research needs	Minimization of unintended consequences of domestication and maximization of success for iBoF and other salmonids
Future research potential	Minimization of unintended consequences of domestication and maximization of success for iBoF and other salmonids

Table 5. Changes to spawning, rearing and release strategies for the Stewiacke River LGB through time.

	Year							
	2007	**2006**	**2005**	**2004**	**2003**	**2002**	**2001**	**2000**
Spawning (criteria for selection of spawners)								
Use of kinship information to minimize spawnings between full and half sibs	x	x	x	x	x	x	x	x
Use of mean kinship information in the ranking and prioritization of spawners	x	x	x	x	x	x	x	
Attempts to spawn at least one fish from each full-sib group	x	x	x	x	x	x		
Number of F1 recovered from spawners			x	x		N/A		
Number of years of potential F1 recovery (previous spawnings) from spawners	x	x	x	x	x	x		
Number of F1 recovered from full-sib groups				x		N/A		
Number of years of potential F1 recovery from full-sib groups (summed over sibs)	x	x		x				
Number of years of potential F1 recovery from parents for F1 spawner				x	N/A	N/A		
Number of years of potential F1 recovery from full-sib group of parents for F1 spawner				x	N/A	N/A		
Is the individual a newly matured virgin spawner						x		
Maintenance of pure wild-exposed or pure captive-reared families for future research	x	x	N/A	N/A	N/A	N/A		
Spawning preference given to wild-exposed fish over captive-reared	x	x	N/A	N/A	N/A	N/A		
Additional spawning of large wild-exposed families	x							
Duration of wild exposure (if applicable)	x		N/A	N/A	N/A	N/A		
Rearing								
Selection of family representatives prior to pooling	x	x	x	x	x	x		
Two groups of 10 offspring taken per family						x		
Two groups of 5 eggs or offspring taken per family	x	x	x	x	x			
Release of offspring								
High and medium priority fish released throughout the river system, no attempts to equalize family size (offspring from low priority/redundant families were released in other rivers)			x	x	x	x	x	x
Equal number of individuals from each family released into one site, the rest distributed throughout the river system	x	x						

Table 6. Genetic and demographic characteristics of all wild parr (Wild Parr) collected from the Stewiacke River in 1998-2001, those selected as parents (G0, Parents Selected) for the production of the next generation of LGB salmon, and an equivalent number of juveniles selected at random (Parents Random), ***by year class.***

Year Class or Year Class Set	G0-98			G0-99			G0-00			G0-01			G0-98+99			G0-00+01			G0-98-01		
Group Type	Wild Parr	Parents Selected	Parents Random	Wild Parr	Parents Selected	Parents Random	Wild Parr	Parents Selected	Parents Random	Wild Parr	Parents Selected	Parents Random	Wild Parr	Parents Selected	Parents Random	Wild Parr	Parents Selected	Parents Random	Wild Parr	Parents Selected	Parents Random
N	402	363	363	188	171	171	273	179	179	166	108	108	590	534	534	439	287	287	1029	821	821
#A (% change)	16.11	15.78 (-2.0)	15.98 (-0.8)	15.56	15.56 (0)	15.47 (-0.6)	12.86	12.71 (-1.2)	12.34 (-4.0)	13.56*	13.11 (-3.3)	12.69* (-6.4)	17.78	17.56 (-1.2)	17.56 (-1.2)	13.43	13.14 (-2.2)	12.86 (-4.2)	18.00	17.78 (-1.2)	17.71 (-1.6)
N_a	15.85	15.71	15.89	15.36*	15.52*+	15.43+	12.38*	12.62*+	12.25+	12.48*	13.00*	12.57	17.49	17.48	17.49	12.89	13.08	12.80	16.92	17.07	17.04
H_e	0.844	0.844	0.845	0.839	0.841	0.839	0.831	0.831	0.831	0.792	0.812	0.796	0.846	0.846	0.846	0.822	0.828	0.823	0.846	0.846	0.846
H_o	0.830	0.833	0.831	0.804	0.801	0.802	0.843	0.841	0.844	0.838	0.839	0.837	0.822	0.823	0.822	0.837	0.836	0.838	0.827	0.826	0.826
F_{IS}	0.017	0.014	0.016	0.043	0.048	0.043	-0.014	-0.011	-0.015	-0.063	-0.038	-0.057	0.029	0.028	0.028	-0.019	-0.011	-0.019	0.022	0.024	0.023
# loci (*#A*, N_a, H_e, H_o, *FIS*)		9			9			7			9			9			7			9	
# indiv. (N_a)		333			165			156			98			504			259			526	
MK	0.0091	0.0089	0.0090	0.0143	0.0136	0.0150	0.0104	0.0108	0.0114	0.0373	0.0212	0.0350	0.0067	0.0066	0.0068	0.0155	0.0116	0.0154	0.0060	0.0051	0.0056

N = Number of individuals.
#A = Observed number of alleles in the wild juvenile and parental (G0 or founder) groups, uncorrected for differences in sample size and percent change from the wild parr group.
N_a = Allele richness or standardized numbers of alleles (Note: this statistic is standardized to the lowest numbers observed at a given locus for any single population or sample collection).
H_e = Effective heterozygosity or gene diversity.
H_o = Observed heterozygosity.
F_{IS} = Wright's inbreeding coefficient.
loci = Number of loci used in the analyses.
indiv. (N_a) = Number of individuals sub-sampled in allele richness estimate, in comparisons of N_a across wild juvenile and parental groups.
MK = Mean kinship.
* and + designate statistically significant comparisons (p<0.05).

Table 7. Performance of different founder selection and mating strategies in recovering founder genetic variation when the number of broodstock is reduced to 50 males and 50 females.

Group Type	**Wild Parr Collection (1998)**	**Parents Selected (Founders)**			
Parent Selection Method		**Minimization of Mean Kinship (*MMK*)**	**Mean Kinship Assisted (*MKA*)**	**Minimizing Pairwise Relatedness (*MPR*)**	**Random Selection and Mating (*RSM*; Avg. 100 Samples)**
#A	16.111*@	14.778* (-8.3%)	14.444+@ (-10.3%)	16.000+ (-0.7%)	13.500*+ (-16.2%)
N_a	13.208@$	14.533*$	14.234+@	15.794*+@	13.224*+
H_o	0.830	0.844	0.836	0.847	0.831
H_e	0.845	0.853	0.851	0.866	0.845
F_{IS}	0.017	0.010	0.016	0.022	0.018
# FSF	218	86	93	89	79
# HSF	55	47	54	41	36

#A = Actual number of alleles observed.
N_a = Allele richness or standardized numbers of alleles, based on 85 single-locus genotypes.
H_e = Effective heterozygosity or gene diversity.
H_o = Observed heterozygosity.
F_{IS} = Wright's inbreeding coefficient.
#FSF = Number of full-sib families.
#HSF = Number of half-sib families.
All estimates were based on 9 microsatellite loci.
*, +, @, and $ designate statistically significant comparisons (p<0.05).

Table 8. Genetic and demographic characteristics of Parents Selected (G0) and 1) all Offspring as Smolts in spawning year + 3 (G1), 2) Offspring as Selected Adults in spawning year + 4 (G1) and 3) an equivalent number of Offspring Randomly Sampled at the adult stage (G1), ***by spawning year****.*

	Parents Selected (G0 or G0+G1)									Offspring as Smolts in Spawning Year + 3 (G1)							Offspring as Selected Adults in Spawning Year + 4 and Subsequent Years (G1)							Offspring Randomly Sampled as Adults in Spawning Year + 4 and Subsequent Years (G1)						
Spawn Year	Parental group type(s)	Number male/ female parents (crosses)	H_e down	H_e across	#*A* across	N_a down	N_a across	# ind. N_a acr.	# loci acr.	# indiv.	H_e down	H_e across	#*A* across	N_a down	N_a across	% change #*A* across	# indiv.	H_e down	H_e across	#*A* across	N_a down	N_a across	% change #*A* across	# indiv.	H_e down	H_e across	#*A* across	N_a down	N_a across	% change #*A* across
# loci down			7			7					4			4				4			4				4			4		
# ind. N_a down						179								290							58							56		
2000	G0-98	149/173 (174)	0.832	0.864	**15.25**	12.63	13.68	124	4	429	0.861	0.861	**13.00**	12.89	12.35	**-14.8**	133	0.865	0.865	**12.50**	11.84	12.47	**-18.0**	133	0.863	0.863	**12.60**	11.69	12.55	**-17.4**
2001	G0-98-99	113/113 (113)	0.832	0.832	**14.14**	13.68	12.44	99	7	407	0.858	0.828	**14.29**	16.33	12.68	**1.0**	145	0.862	0.830	**13.14**	13.82	12.76	**-7.1**	145	0.862	0.826	**13.00**	13.62	12.62	**-8.1**
2002	G0-98-00	100/100 (100)	0.832	0.844	**15.88**	13.28	15.56	157	8	444	0.853	0.842	**16.50**	15.42	15.13	**3.9**	194	0.857	0.846	**15.25**	13.13	15.06	**-4.0**	194	0.853	0.842	**15.35**	12.95	15.12	**-3.3**
2003	G0-98-01	133/132 (133)	0.832	0.834	**15.40**	13.26	13.37	56	10	413	0.857	0.840	**15.60**	14.37	12.89	**1.3**	59	0.857	0.839	**13.70**	12.73	13.58	**-11.0**	59	0.858	0.840	**13.04**	12.48	12.94	**-15.3**
2004	G0-98-01, G1-00	173/173 (173)	0.837			13.58																							
2005	G0-99-01, G1-00-01	145/145 (145)	0.835			13.68																							
2006	G0-00-01, G1-00-02	148/148 (148)	0.831			13.93																							
2007	G1-00-03	117/122 (128)	0.838			13.37																							

Spawn year = Year in which a particular group of parents was spawned and their respective groups of offspring produced.
Parental group types = Parental type by year class and generation (see methods section for more details).
Number of male/female parents = Number of male parents followed by number of female parents, and number of crosses in parentheses.
H_e = Effective heterozygosity or gene diversity.
#*A* = Number of alleles observed.
N_a = Allele richness or standardized numbers of alleles.
indiv. N_a = Number of individuals sub-sampled in allele richness estimate.
loci = Number of loci involved in estimates.
% change #*A* = Percent change in observed number of alleles in offspring groups relative to their respective parental groups.
down = Comparing values down through the years.
across = Comparing values across parent/offspring groups within the same year.

Table 9. Genetic and demographic characteristics of Stewiacke River LGB parents (G0) and their respective offspring (G1) chosen for spawning as adults, ***by year class.***

	1998		1999		2000		2001		1998+1999		2000+2001		1998-2001	
	Parents Selected (G0)	Offspring Selected (G1)	Parents Selected (G0)	Offspring Selected (G1)	Parents Selected (G0)	Offspring Selected (G1)	Parents Selected (G0)	Offspring Selected (G1)	Parents Selected (G0)	Offspring Selected (G1)	Parents Selected (G0)	Offspring Selected (G1)	Parents Selected (G0)	Offspring Selected (G1)
N	363	222	171	125	179	34	108	9	534	451	287	64	821	545
P_x		0.323		0.516		0.600		0.763		0.064		0.373		0.016
#A	16.25 (13.71)	13.63 (12.43)	16.00	14.75	12.71	9.86	13.11	6.33	18.13 (15.57)	17.13 (15.00)	13.78	11.89	18.38 (15.86)	17.63 (15.29)
% change *#A*		-16.1 (-9.3)		-7.8		-22.4		-51.7		-5.5 (-3.7)		-13.7		-4.1 (-3.6)
N_a across	11.75 (11.55)	11.91 (11.75)	15.12	14.64	10.24	9.81	6.88	6.33	16.29 (14.41)	16.55 (14.72)	12.01	11.87	16.67 (14.44)	17.27 (15.12)
H_e	0.843 (0.831)	0.843 (0.832)	0.841	0.831	0.831	0.829	0.812	0.816	0.845 (0.833)	0.847 (0.835)	0.821	0.829	0.846 (0.833)	0.850 (0.838)
% change H_e		0.012		-1.173		-0.251		0.543		0.225		1.029		0.485
H_o	0.843 (0.834)	0.850 (0.848)	0.815	0.861	0.841	0.815	0.839	0.802	0.834 (0.824)	0.852 (0.841)	0.844	0.819	0.838 (0.828)	0.852 (0.839)
F_{IS}	0.0001 (-0.0017)	-0.0068 (-0.0166)	0.0314	-0.0340	-0.0107	0.016	-0.0303	0.0187	0.0134 (0.0123)	-0.0046 (-0.0056)	-0.0324	0.0130	0.0090 (0.0073)	-0.0020 (-0.0010)
# loci	8 (7)	8 (7)	8	8	7	7	9	9	8 (7)	8 (7)	9	9	8 (7)	8 (7)
# indiv. (N_a)	43 (92)	43 (92)	107	107	33	33	9	9	226 (308)	226 (308)	63	63	320 (400)	320 (400)
MK	0.0089	0.0116	0.0136	0.0209	0.0108	0.0327	0.0212	0.0772	0.0066	0.0067	0.0116	0.0151	0.0051	0.0058
N_{b1}	362.38		170.99		169.61		93.19		533.52		264.14		815.85	
N_{b2}	345.73		107.79		251.22		193.15		318.55		442.06		467.58	
N_{b3}	289.63		121.67		205.22		162.60		369.03		358.93		510.83	
% exp. loss H_e		0.173		0.411		0.244		0.308		0.135		0.139		0.098
% exp. loss *#A*		6.65		8.42		3.01		3.84		8.21		1.90		7.92

N = Number of individuals.
P_x = Proportion of the total number of offspring from a given year class or year class group exhibiting one parent from the relevant group and another from another group (see methods section for details).
#A = Observed number of alleles, not standardized for sample size.
% change *#A* = % change in number of alleles observed between parents and offspring.
N_a across = Allele richness or number of alleles standardized for *N* across parents and offspring.
H_e = Effective heterozygosity or gene diversity.
% change H_e = % change in allele richness between parents and offspring.
H_o = Observed heterozygosity
F_{IS} = Wright's inbreeding coefficient.
loci = Number of loci used in the analyses.
indiv. (N_a) = Number of individuals sub-sampled in N_a estimate.
MK = Mean kinship.
N_{b1} = Effective number of breeders (uneven sex ratio).
N_{b2} = Effective number of breeders (variance in family size).
N_{b3} = Effective number of breeders (uneven sex ratio and variance in family size).
% exp. loss H_e = Expected loss in H_e in offspring relative to parents.
% exp. loss *#A* = Expected loss in *#A* in offspring relative to parents.

Table 10. Comparison of demographic characteristics of the captive-reared and wild-exposed groups of offspring produced via captive breeding in 2002.

	Rearing Environment (Juveniles)		
	Captive	**Wild**	**Total**
Number of individuals	106	128	234
Number of males	53	68	121
Number of females	53	60	113
Number of families	26	30	34 (out of 35)
Average family size	4.08	4.23	6.88
Variance in family size	7.83	13.58	N/A
Minimum family size	1	1	1
Maximum family size	11	14	20
Number of exclusive families*	4	8	N/A
Number maturing in 2006	59	128	187
Number of individuals spawned (2006)	44	61	105

* Number of families with all wild-exposed or all captive-reared individuals.

Table 11. Examples of conservation research currently underway involving inner Bay of Fundy Atlantic salmon.

Project	Affiliated Group	Anticipated Results
Relative effects of inbreeding and outbreeding on Chignecto Bay salmon in captivity and in the wild	Danielle McDonald, DFO	2009
Effects of wild exposure of parents on performance and survival of offspring in the wild	Nate Wilke, Ian Fleming, Memorial University	2010
Effects of generations of domestication selection on performance and survival in the wild	Nate Wilke, Ian Fleming, Memorial University	2010
Potential effects of mate choice/breeding competition on growth and performance in the wild	Nate Wilke, Ian Fleming, Memorial University	2009
Effects of mean kinship (family size of founders) on survival and performance of offspring	Nate Wilke, Ian Fleming, Memorial University	2010
Genomic scans of SNP variation to identify informative nuclear loci for identifying introgression of non-local salmon into LGB populations	Heather Freamo, Elizabeth Boulding, Guelph University	2009
Microarray-based studies of gene regulation in siblings reared in captivity and in the wild	Wendy Tymchuck and Trish Schulte, University of British Columbia, and Nate Wilke, Ian Fleming, Memorial University	2009
Effects of generations of domestication selection on dominance behaviour of inner Bay of Fundy salmon	Nate Wilke, Ian Fleming, Memorial University	2009
Effects of generations of domestication selection on predator avoidance behaviour in inner Bay of Fundy salmon	Louise de Mestral Bezanson, Christophe Herbinger, Dalhousie University	2009
Relatedness of iBoF salmon in the context of smolt run timing	Louise de Mestral Bezanson, Christophe Herbinger, Dalhousie University	2009
Relative effects of inbreeding and outbreeding on Minas Basin salmon in captivity and in the wild	Dylan Fraser and Jeff Hutchings, DFO, Dalhouise University	2010
Common garden experiments for investigating genetic basis of differences between farmed Saint John, wild inner Bay of Fundy and Southern Upland Atlantic salmon	Aimee Houde, Dylan Fraser and Jeff Hutchings, Dalhousie University	2010
Investigations into possible local adaptation in Bay of Fundy salmon to turbidity	Aimee Houde, Dylan Fraser and Jeff Hutchings, Dalhousie University	2011

FIGURES

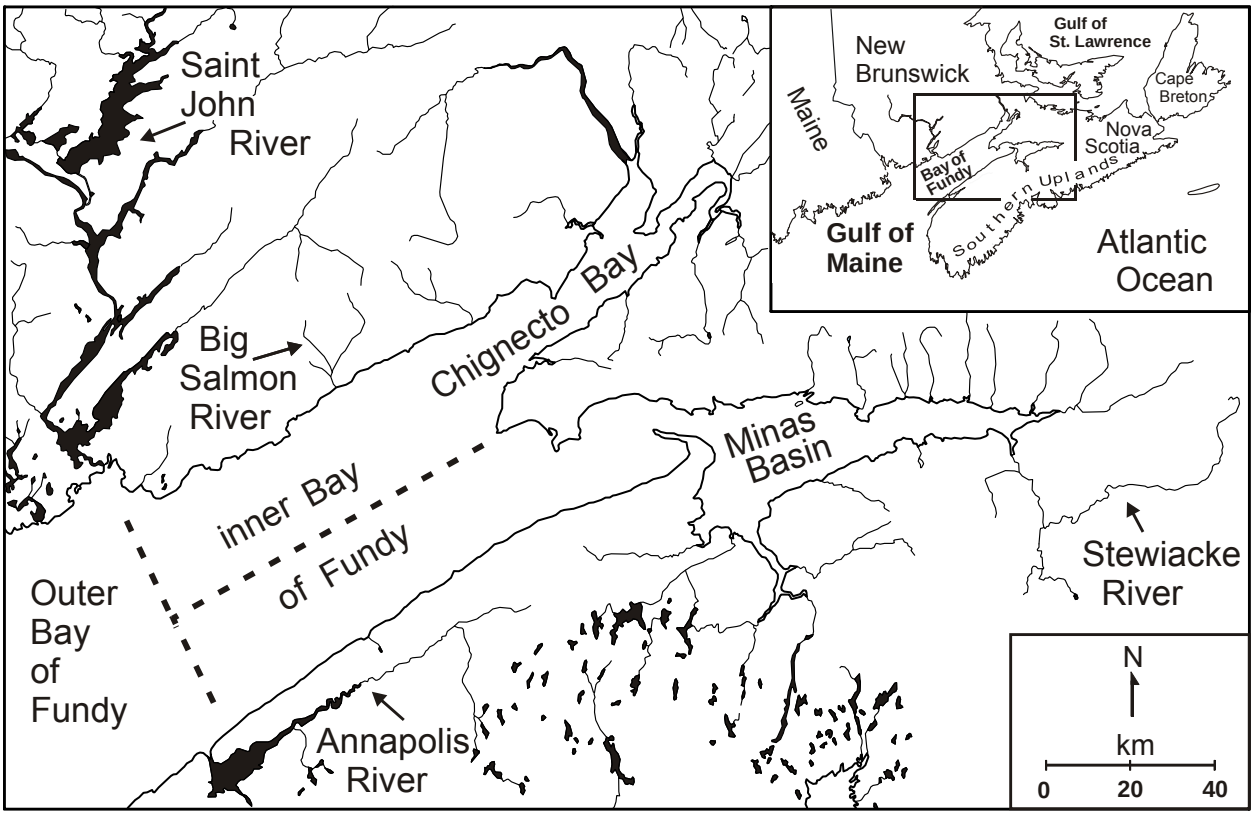

Figure 1. Geographic locations of inner Bay of Fundy rivers.

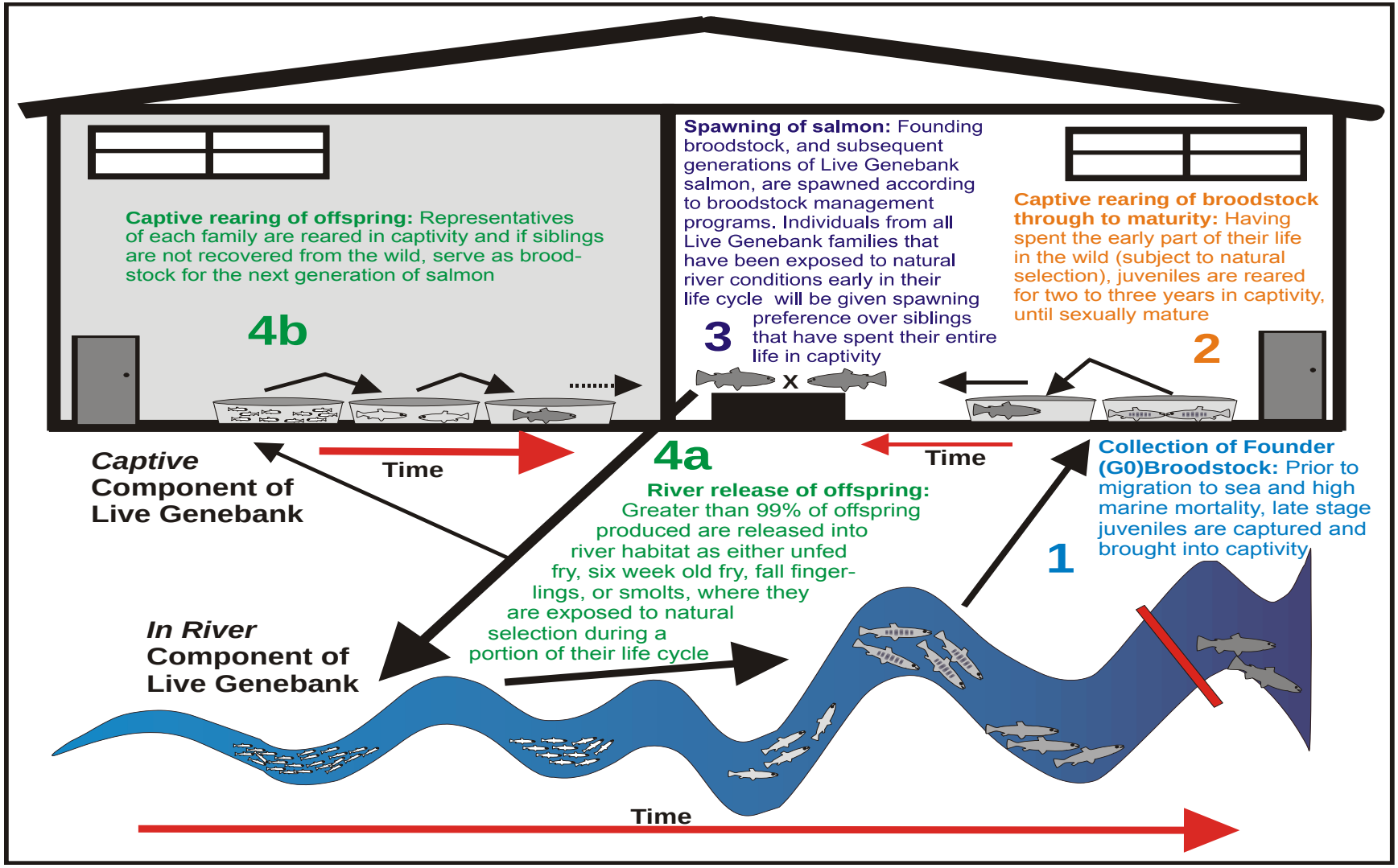

Figure 2. Schematic depicting the inner Bay of Fundy Live Gene Bank program, including captive and in-river components.

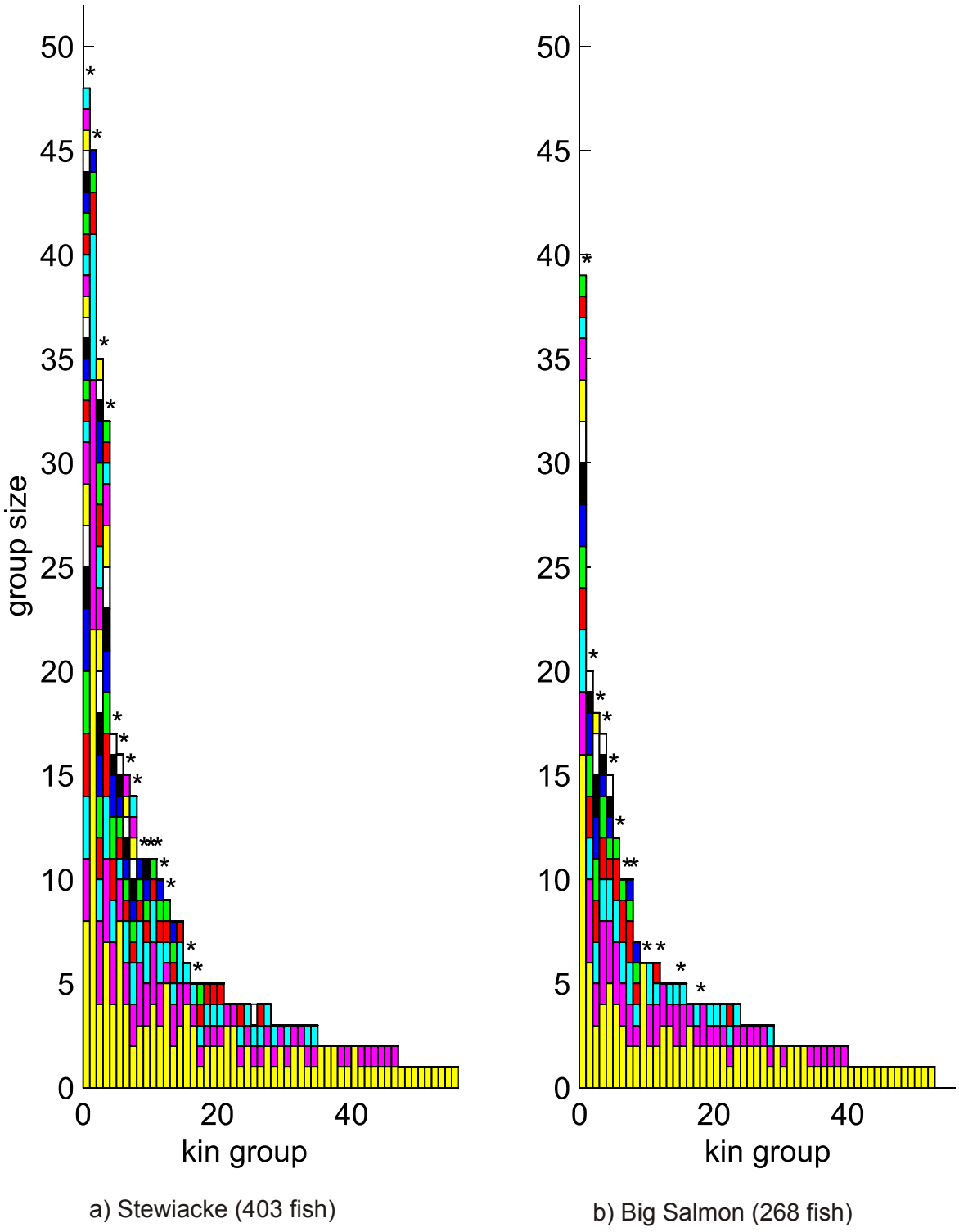

Figure 3. Distribution of Big Salmon River and Stewiacke River salmon into full-sib and kin groupings from Herbinger et al. (2006). Multi-coloured vertical bars represent individual kin groupings (56 and 53 in the Stewiacke and Big Salmon collections, respectively) and solid, coloured rectangles comprising each vertical bar represent full-sib groupings nested within. Asterisks indicate kin groupings that are unlikely to have occurred by chance alone.

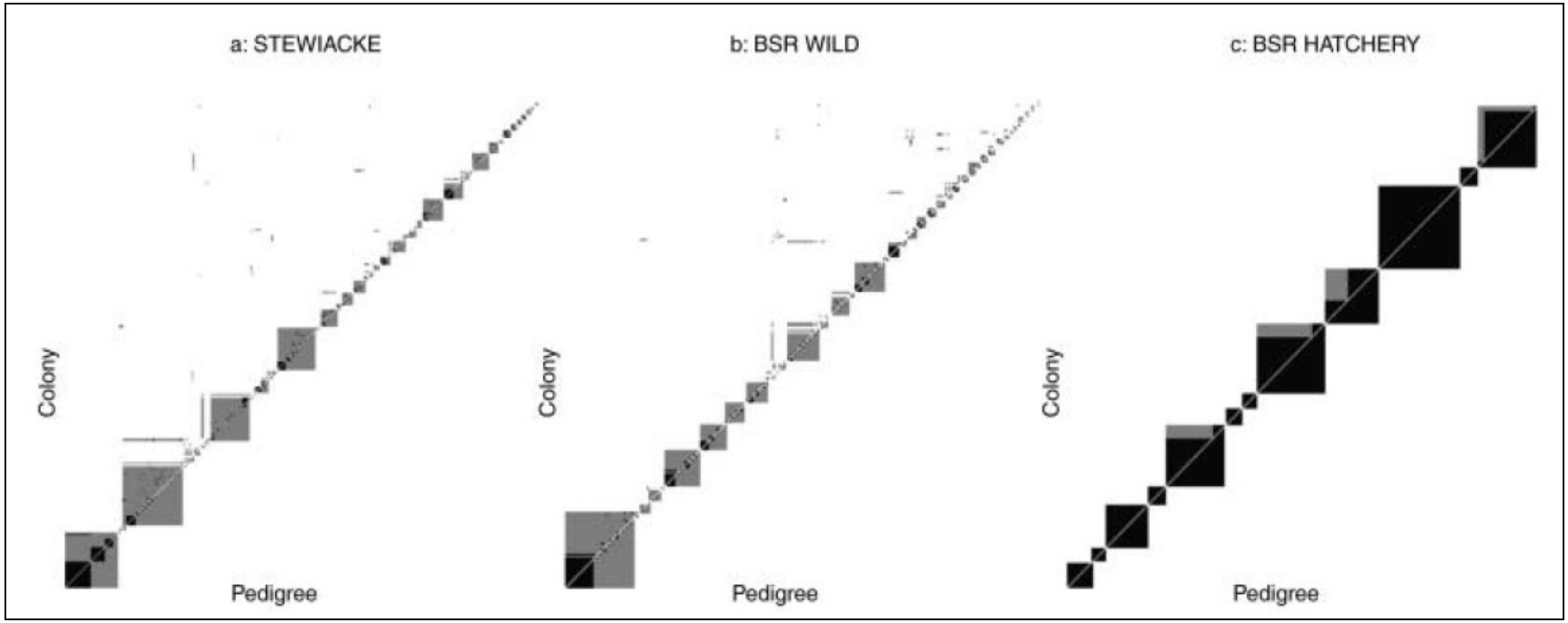

Figure 4. Graphic representation of concordance between two different methods and associated software, Colony and Pedigree, in assessing first-order relatedness for wild and hatchery populations from the inner Bay of Fundy (Herbinger et al., 2006). Symmetry about the diagonal reflects overall concordance in kinship estimates. Group size is indicated by the size of the triangle on either side of the diagonal, and shading denotes the nature of the kin grouping, black corresponding to full-sib groupings and grey to half-sib groupings.

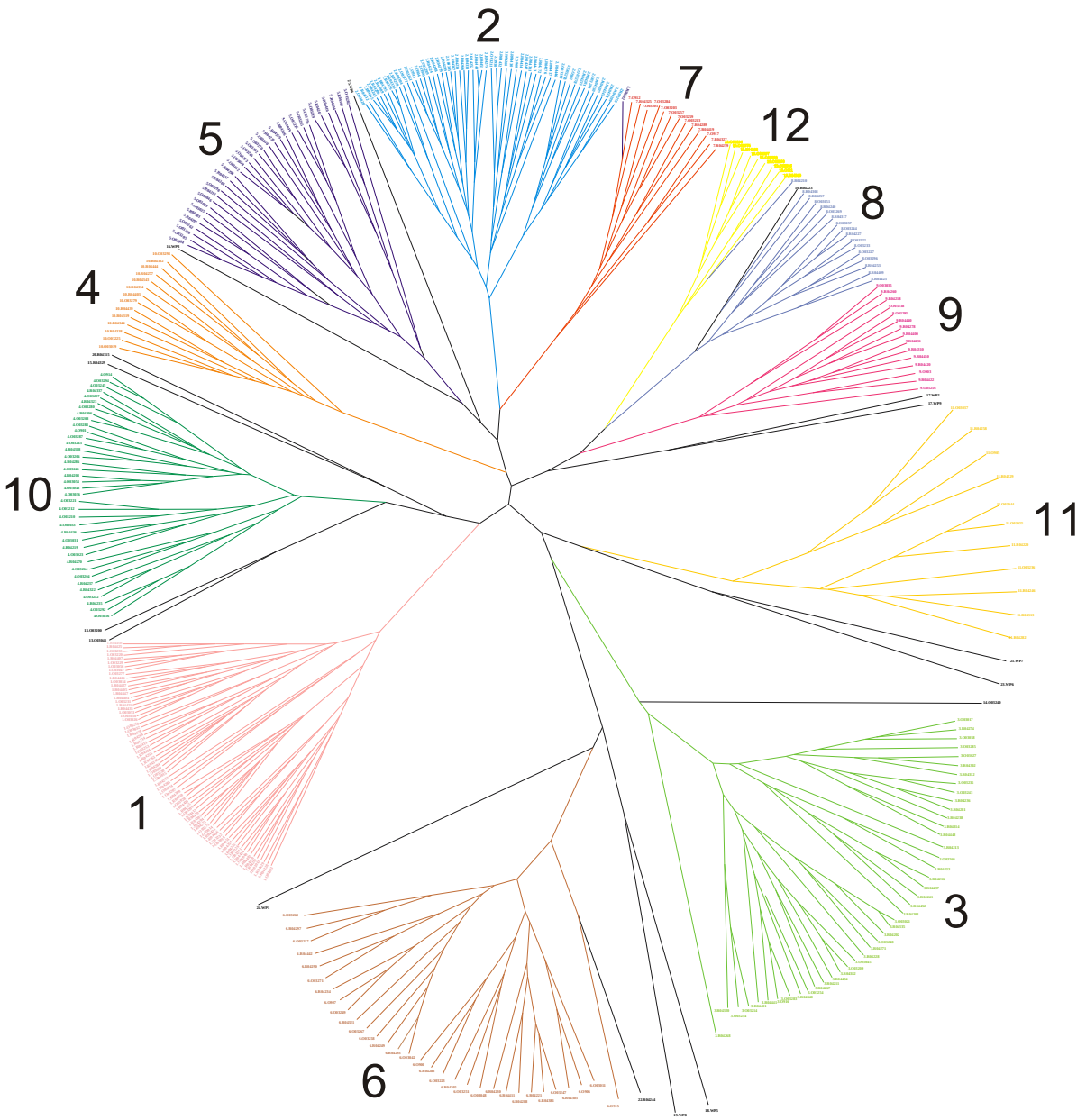

Figure 5. Concordance between first-order relatedness estimated using Colony (Wang, 2004) and UPGMA clustering of pairwise distances based on the allele-sharing metric (DPs) of Bowcock et al. (1994). Terminal branches and fish IDs of full siblings identified by Colony are like-coloured. The 12 major full-sib families identified by Colony also clustered into as many major branches identified by the clustering approach.

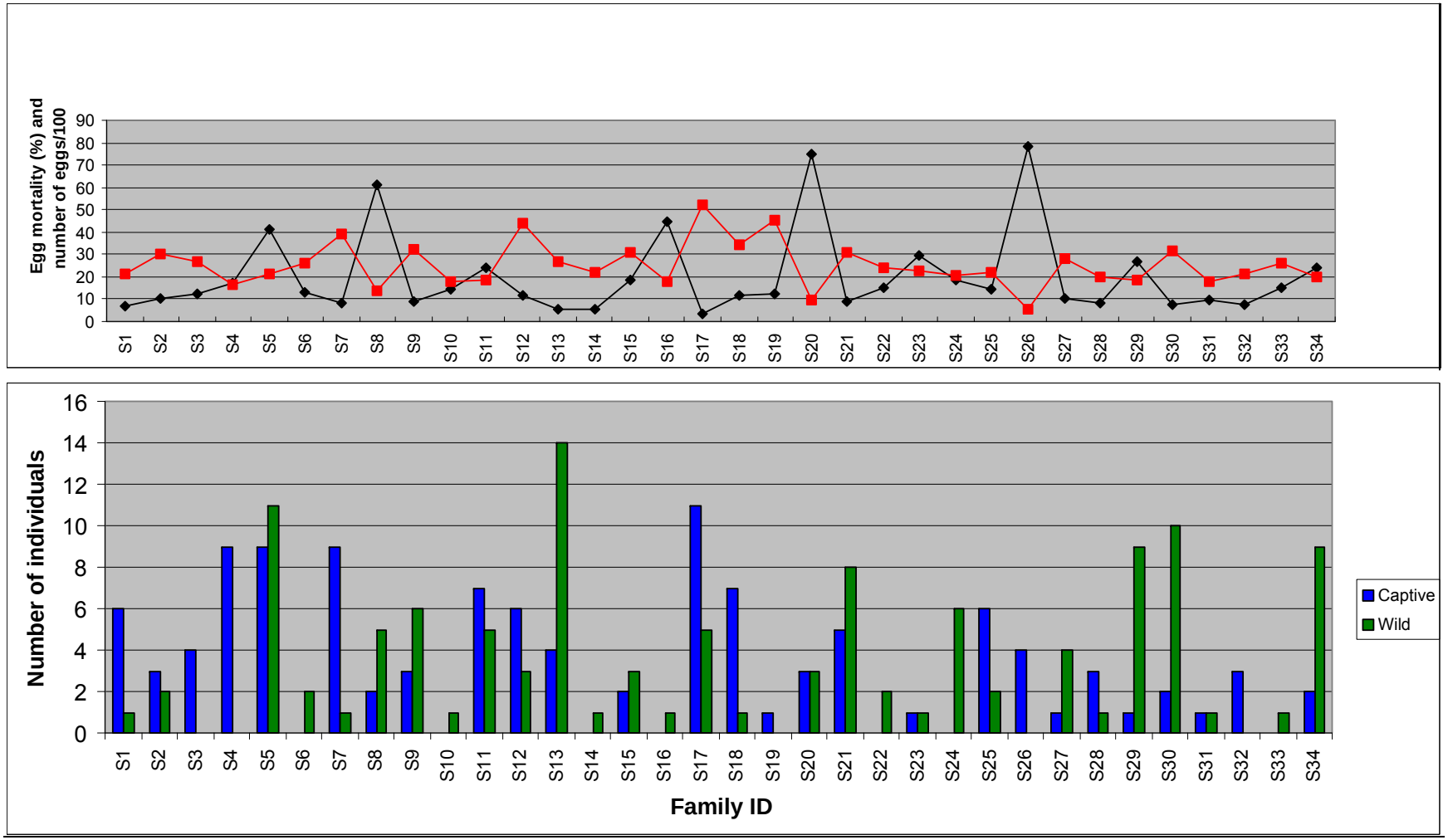

Figure 6. Number of siblings exposed to captive and wild environments as juveniles (bottom graph), percent mortality (top graph, black diamonds), and estimated number of surviving eggs (top graph, red squares) for 34 Stewiacke River families created in 2002.

APPENDIX 1

Table A1a. Information on the cryopreservation of Atlantic salmon milt at Coldbrook Biodiversity Facility in 2006.

Date Frozen (2006)	Hatchery of Origin	Stewiacke River Fish Tag	Cane #	Position in Cane	Canister*	Dewar*
Nov-16	Coldbrook	A45177	6	Bottom	1	1
Nov-16	Coldbrook	454827092	6	Top	1	1
Nov-16	Coldbrook	454B3C182B	2	Bottom	1	1
Nov-16	Coldbrook	454A240923	2	Top	1	1
Nov-16	Coldbrook	454B440D21	3	Top	1	1
Nov-16	Coldbrook	A45246	3	Bottom	1	1
Nov-16	Coldbrook	A45248	4	Bottom	1	1
Nov-16	Coldbrook	454B486772	4	Top	1	1
Nov-16	Coldbrook	454B555D56	?	?	1	1
Nov-16	Coldbrook	454B486772	5	Top	1	1
Nov-16	Coldbrook	454B555D5E	5	Bottom	1	1
Nov-16	Coldbrook	43600A3B17	1	ALONE	1	1
Nov-16	Coldbrook	421D7A2B2D	10	Top	2	1
Nov-16	Coldbrook	430E632C5A	10	Bottom	2	1
Nov-16	Coldbrook	4225707D68	9	Bottom	2	1
Nov-16	Coldbrook	A45183	9	Top	2	1
Nov-16	Coldbrook	A45166	8	Bottom	2	1
Nov-16	Coldbrook	431544610B	8	Top	2	1
Nov-16	Coldbrook	4237782970	7	Top	2	1
Nov-16	Coldbrook	42327A3117	7	Bottom	2	1
Nov-16	Coldbrook	430E493A2D	11	Bottom	2	1
Nov-16	Coldbrook	4225417E13	11	Top	2	1
Nov-16	Coldbrook	430E77064B	12	Bottom	2	1
Nov-16	Coldbrook	421D541C76	12	Top	2	1
Nov-16	Coldbrook	4233012A62	13	Bottom	2	1
Nov-16	Coldbrook	43155E3F38	13	Top	2	1
Nov-16	Coldbrook	435F7A7738	14	Bottom	3	1
Nov-16	Coldbrook	431042424C	14	Top	3	1
Nov-16	Coldbrook	4233007314	15	Bottom	3	1
Nov-16	Coldbrook	42353A3A68	15	Top	3	1
Nov-16	Coldbrook	430E651133	16	Bottom	3	1
Nov-16	Coldbrook	4232790614	16	Top	3	1
Nov-16	Coldbrook	4229120D28	17	ALONE	3	1
Nov-22	Mersey	4311331A67	M1	Bottom	4	1
Nov-22	Mersey	421E21187D	M1	Top	4	1
Nov-22	Mersey	4310664825	M2	Top	4	1
Nov-22	Mersey	430E671154	M2	Bottom	4	1
Nov-22	Mersey	435F71606F	M3	Bottom	4	1
Nov-22	Mersey	430E691C17	M3	Top	4	1
Nov-22	Mersey	431146471D	M4	Bottom	4	1
Nov-22	Mersey	4315574F7F	M4	Top	4	1
Nov-22	Mersey	42257E2F64	M5	Bottom	4	1
Nov-22	Mersey	421E0C1825	M5	Top	4	1

Date Frozen (2006)	Hatchery of Origin	Stewiacke River Fish Tag	Cane #	Position in Cane	Canister*	Dewar*
Nov-22	Mersey	430E445E02	M6	Bottom	4	1
Nov-22	Mersey	4225704A1B	M6	Top	4	1
Nov-22	Mersey	435F775463	M7	Bottom	4	2?
Nov-22	Mersey	435F587F0E	M7	Top	4	2?
Nov-22	Mersey	435F66053F	M8	Bottom	4	2?
Nov-22	Mersey	430E553068	M8	Top	4	2?

* Canister and Dewar numbers may have changed; however, cane number remains the same.

Table A1b. Information on the cryopreservation of Atlantic salmon milt at Coldbrook Biodiversity Facility in 2007.

Number	Date Preserved	Stewiacke River Fish Tag	Cane	Position in Cane	Canister	Dewar
1	Nov 16-2007	454B57620F	18	Bottom	2	B
2	Nov 16-2007	A45164	18	Top	2	B
3	Nov 16-2007	43110C1B17	19	Bottom	2	B
4	Nov 16-2007	435F74500A	19	Top	2	B
5	Nov 16-2007	486E173D20	20	Bottom	2	B
6	Nov 16-2007	43132C226F	20	Top	2	B
7	Nov 16-2007	4360134643	21	Bottom	2	B
8	Nov 16-2007	42290A5E03	21	Top	2	B
9	Nov 16-2007	4311083570	22	Bottom	2	B
10	Nov 16-2007	4315550727	22	Top	2	B
11	Nov 16-2007	4311440A7C	23	Bottom	2	B
12	Nov 16-2007	4549337745	24	Top	2	B
13	Nov 16-2007	D600	24	Bottom	2	B
14	Nov 16-2007	430E553551	25	Bottom	2	B
15	Nov 16-2007	454A270C63	25	Top	2	B
16	Nov 16-2007	435F795166	26	Top	2	B
17	Nov 16-2007	454957073B	26	Bottom	2	B
18	Nov 16-2007	421C477A59	27	Top	2	B
19	Nov 16-2007	435F5D172A	28	Bottom	2	B
20	Nov 16-2007	421E043203	28	Top	2	B
21	Nov 16-2007	431537666F	29	Top	2	B
22	Nov 16-2007	486C103C0B	29	Bottom	2	B
23	Nov 16-2007	431165084B	30	Top	3	B
24	Nov 16-2007	45494B5A74	30	Bottom	3	B
25	Nov 16-2007	4549363B78	31	Top	3	B
26	Nov 16-2007	4549086473	31	Bottom	3	B
27	Nov 16-2007	4548740618	32	Top	3	B
28	Nov 16-2007	454964232A	32	Bottom	3	B
29	Nov 16-2007	430E74563F	33	Top	3	B
30	Nov 16-2007	D849	33	Bottom	3	B
31	Nov 16-2007	454A1D1533	34	Top	3	B
32	Nov 16-2007	454B3F2906	34	Bottom	3	B
33	Nov 16-2007	486A35596C	35	Top	3	B
34	Nov 16-2007	4315570 E25	35	Bottom	3	B
35	Nov 16-2007	436010413D	36	Top	3	B
36	Nov 16-2007	454B617176	36	Bottom	3	B
37	Nov 16-2007	D534	37	Top	3	B
38	Nov 16-2007	43115B0747	37	Bottom	3	B
39	Nov 16-2007	D663	38	Top	3	B
40	Nov 16-2007	486A3F095B	38	Bottom	3	B
41	Nov 16-2007	D882	39	Top	3	B

Table A2. Single-locus allele richness statistics for sample comparisons for Tables 6b and 7.

		Locus										
		1605	2201	2210	2215	2216	1G7	197	202	486	144	Average
Table 6b	Spawning Year											
Parents selected	2000				11.155	15.920	12.739	14.901				13.679
	2001	10.822		6.847	12.528	16.107	14.091	14.258	12.444			12.442
	2002	12.348	33.000	6.801	16.277	15.717	12.808	14.740	12.816			15.563
	2003	10.321	26.359	6.287	12.861	13.642	12.862	11.961	11.521	6.584	21.341	13.374
Offspring selected	2000				10.932	13.939	11.000	13.991				12.466
	2001	11.616		5.998	11.964	15.538	16.941	14.641	12.631			12.761
	2002	9.980	31.321	6.980	16.759	14.585	12.877	13.993	13.958			15.057
	2003	10.947	27.525	5.998	12.940	13.894	11.931	11.892	10.000	7.847	22.858	13.583
Table 7	Founder Yearclass											
Parents selected	1998	9.399	21.768	5.897	9.645	13.436	10.987	12.755	10.151			11.755
	1999	11.215	31.678	7.266	14.782	14.536	15.015	13.675	12.824			15.124
	2000	9.631		5.930	11.353	12.301	11.080	11.115	10.248			10.237
	2001	7.350		3.853	7.247	7.231	7.959	6.590	6.692	4.625	10.406	6.884
	1998-99	11.967	33.828	7.524	14.674	16.700	16.145	15.707	13.791			16.292
	2000-01	10.208		6.129	12.456	14.115	12.332	12.511	11.324	6.785	22.234	12.010
	1998-01	12.423	34.955	7.587	15.095	17.186	16.544	15.727	13.855			16.672
Offspring selected	1998	8.756	22.000	5.368	10.341	12.961	12.404	12.856	10.558			11.906
	1999	11.781	27.000	6.877	14.848	14.839	15.787	12.995	12.971			14.637
	2000	10.911		6.000	9.941	11.940	11.000	8.970	9.941			9.815
	2001	9.000		5.000	7.000	6.000	6.000	5.000	6.000	4.000	9.000	6.333
	1998-99	12.636	32.000	7.663	16.271	16.473	17.656	14.872	14.863			16.554
	2000-01	10.984		6.000	12.969	13.000	13.000	10.969	10.000	6.984	22.937	11.871
	1998-01	12.954	34.000	7.760	16.704	17.830	19.074	14.930	14.929			17.273

Table A3. Single-locus statistics and reference information for the microsatellite loci used in this study.

Locus	# Alleles	Obs. Het.	Exp. Het.	PIC	Reference
SSsp 1605	14	0.7954	0.8097	0.7880	Paterson *et al.* (2004)
SSsp 2201*	36	0.9107	0.9318	0.9270	Paterson *et al.* (2004)
SSsp 2210	9	0.7321	0.7280	0.6786	Paterson *et al.* (2004)
SSsp 2213*	13	0.7313	0.8550	0.8396	Paterson *et al.* (2004)
SSsp 2215	18	0.8824	0.8726	0.8588	Paterson *et al.* (2004)
SSsp 2216	19	0.8240	0.8495	0.8346	Paterson *et al.* (2004)
SSsp 1G7	20	0.8814	0.8723	0.8586	Paterson *et al.* (2004)
Ssa 197	18	0.8291	0.8310	0.8125	O'Reilly *et al*. (1996)
Ssa 202	15	0.8551	0.8597	0.8443	O'Reilly *et al*. (1996)

* Statistics based on individuals from G0-98 and G0-99.
# Alleles = Number of different alleles observed.
Obs. Het. = Observed number of alleles exhibiting two different alleles.
Exp. Het. = Expected heterozygosity or gene diversity, described within.
PIC = Polymorphic Information Content.

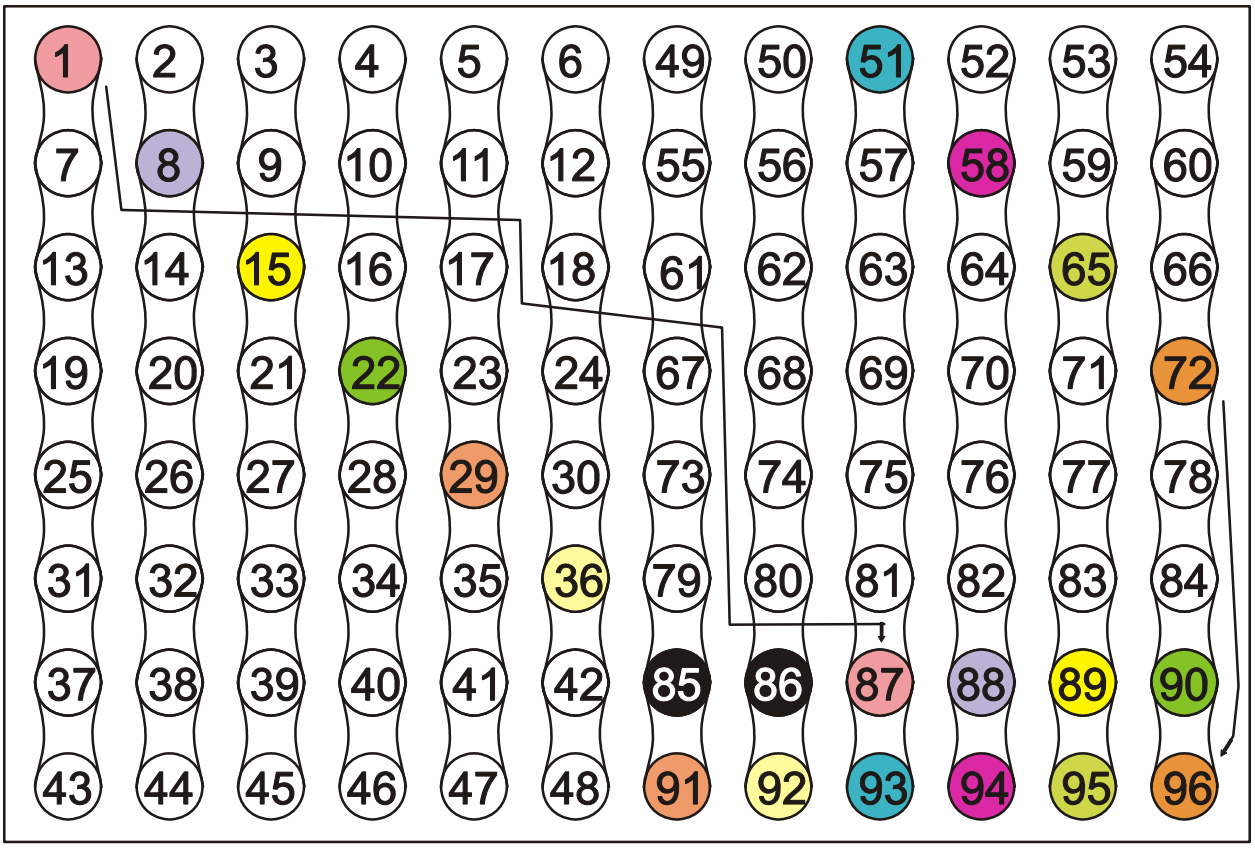

Figure A1. Schematic demonstrating the placement of samples for DNA extraction in a plate containing 96 salmon samples arrayed in 12 strips of 8 tubes. Wells similarly coloured indicate duplicated samples. Black wells indicate the positioning of 2 of 10 cross-gel standards, which are varied across batches so that their genotypes individually identify all batches independent of the physical labelling that also identifies all batches of samples.

Once microsatellite genotype information has been obtained, results are analyzed so as to test for the occurrence of most types of laboratory errors, including inverted strips, single-well sample displacement errors, and plate inversions. Once the accuracy of the information has been verified, data are uploaded into a secure Oracle database.

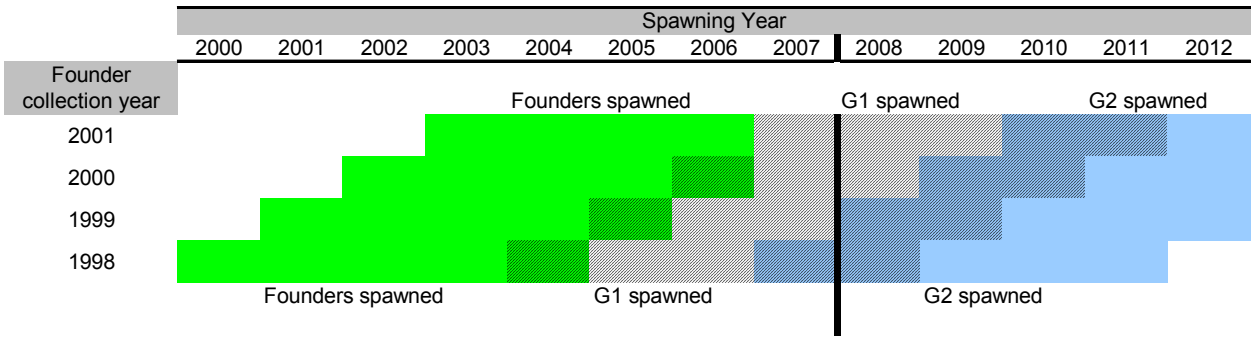

Figure A2. Schematic representation of the spawning timeline for founders, G1 salmon and G2 salmon through years 2000 to 2012.