

ARTICLE

DOI: 10.1038/s42003-018-0112-9

OPEN

Extensive hybridization following a large escape of domesticated Atlantic salmon in the Northwest Atlantic

Brendan F. Wringer^{1,2}, Nicholas W. Jeffery^{1,3}, Ryan R.E. Stanley⁴, Lorraine C. Hamilton⁵, Eric C. Anderson⁶, Ian A. Fleming², Carole Grant¹, J. Brian Dempson¹, Geoff Veinott¹, Steven J. Duffy¹ & Ian R. Bradbury^{1,2,3} 

Domestication is rife with episodes of interbreeding between cultured and wild populations, potentially challenging adaptive variation in the wild. In Atlantic salmon, *Salmo salar*, the number of domesticated individuals far exceeds wild individuals, and escape events occur regularly, yet evidence of the magnitude and geographic scale of interbreeding resulting from individual escape events is lacking. We screened juvenile Atlantic salmon using 95 single nucleotide polymorphisms following a single, large aquaculture escape in the Northwest Atlantic and report the landscape-scale detection of hybrid and feral salmon (27.1%, 17/18 rivers). Hybrids were reproductively viable, and observed at higher frequency in smaller wild populations. Repeated annual sampling of this cohort revealed decreases in the presence of hybrid and feral offspring over time. These results link previous observations of escaped salmon in rivers with reports of population genetic change, and demonstrate the potential negative consequences of escapes from net-pen aquaculture on wild populations.

¹Science Branch, Fisheries and Oceans Canada, 80 East White Hills Road, St. John's, Newfoundland A1C 5X1, Canada. ²Department of Ocean Sciences, Memorial University of Newfoundland, St. John's, Newfoundland A1C 5S7, Canada. ³Faculty of Computer Science, Dalhousie University, Halifax, NS B3H 4R2, Canada. ⁴Science Branch, Fisheries and Oceans Canada, Bedford Institute of Oceanography, Dartmouth, NS B2Y 4A2, Canada. ⁵Aquatic Biotechnology Laboratory, Fisheries and Oceans Canada, Bedford Institute of Oceanography, Dartmouth, NS B2Y 4A2, Canada. ⁶Fisheries Ecology Division, Southwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Santa Cruz, CA 95060, USA. Correspondence and requests for materials should be addressed to I.R.B. (email: ian.bradbury@dfo-mpo.gc.ca)

The process of domestication results in genetically-based phenotypic divergence from wild populations, through both intentional and unintentional selection^{1–3}. Modern genomic data in both plant^{4–6} and animal systems^{3,7,8} have revealed that recurrent hybridization and gene flow between cultured and wild populations can occur, not only during the early stages of domestication⁹, but throughout the entire period of culture¹⁰. Repeated episodes of hybridization between cultured and wild populations can be detrimental for wild populations, resulting in the introduction of non-native alleles⁶, erosion of adaptive diversity in the wild^{4,11}, and ultimately a loss-of-wild population viability^{12,13}. The management and conservation of wild populations, confronted with domesticated conspecifics, requires the accurate quantification of potential genetic and ecological impacts to inform risk assessment and mitigation strategies.

The Atlantic salmon, *Salmo salar*, is of considerable socio-economic value both in culture and in the wild. Domestication of Atlantic salmon was initiated in 1969 in Norway¹⁴, and separately in 1979 in Eastern Canada¹⁵. Despite this short period, the process of domestication has resulted in genetic differences between cultured and wild Atlantic salmon^{16–18} which are likely maladaptive, and lead to lower relative survival of cultured salmon in the wild¹⁹. Domesticated Atlantic salmon exhibit lower relative fitness and spawning success compared to wild Atlantic salmon^{13,19–21}, and interbreeding can impart lasting, heritable, population-level reductions in fitness to wild populations¹². Escapes from Atlantic salmon net-pen aquaculture are a regular occurrence²², and the number of escapees can equate to an appreciable fraction of, or exceed, wild census size^{23,24}. As such, genetic changes in wild populations consistent with introgression from domesticated salmon have been detected in nearly all regions where salmon aquaculture and wild populations co-occur, including: Norway^{25,26}, Ireland^{27,28}, Northern Ireland^{29,30}, and Canada³¹. Furthermore, methodological and theoretical improvements^{32,33} have allowed the degree of hybridization within a single river^{29,30} or the cumulative impact of introgression at large spatial scales (i.e., >100 populations in Norway^{34,35}), to be resolved. Nonetheless, the unequivocal quantification of the magnitude and geographic scale of domestic-wild hybridization associated with single-escape events across a broad landscape of wild salmon populations has remained elusive.

Here we quantify the presence and magnitude of hybridization between wild and escaped domestic individuals following an escape of ~20,000 sexually mature, domestic Atlantic salmon from a single aquaculture net-pen in southern Newfoundland, Canada. This event occurred on September 18, 2013, just prior to the natural spawning period for salmon in this region (Fig. 1a). The southern Newfoundland region is analytically favorable for the detection of hybrids; because the domestic broodstock currently in use originates from a single-non-local source (Saint John River, New Brunswick, Canada), the magnitude of industry production in the region has been limited until recently, and finally estimates of the abundance of wild salmon throughout southern Newfoundland (~20,000 individuals) are approximately equal to the magnitude of the escape^{22,36}. Juvenile salmon were collected from the region and screened using 95 single nucleotide polymorphisms (SNPs) to identify hybrids, hybrid classes, and feral individuals present following this escape event. Next, we evaluated factors influencing the distribution of hybrids, and the magnitude of hybridization. Finally, using repeated temporal sampling, we examined and compared relative changes in the abundance of various hybrid classes over time. We report widespread evidence of hybridization (27.1% and hybrids detected in 17/18 rivers) following this escape event. Hybrids were observed in higher frequency in smaller rivers, and repeated annual

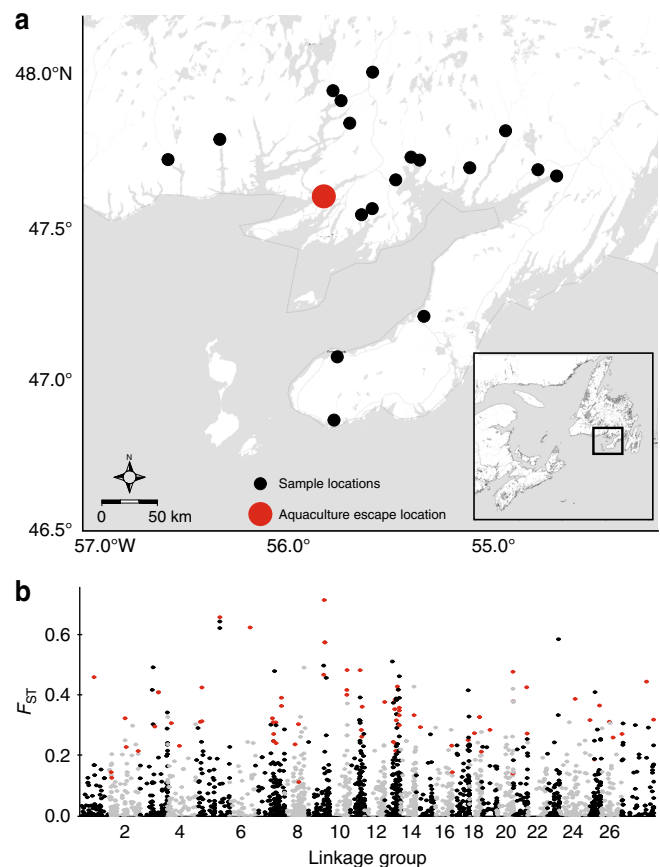


Fig. 1 Geographic distribution of sampling relative to aquaculture escape event and genome-wide comparison of wild and domestic salmon. **a** Map of southern Newfoundland, location relative to eastern Canada shown in inset. Black dots represent rivers surveyed; the red dot denotes the location of the 2013 aquaculture escape event. **b** Manhattan plot illustrating the genome-wide genetic differentiation (F_{ST}) between the wild and aquaculture baseline samples (Supplementary Table 7) used in the validation of the SNP panel accuracy. The red circles indicate the loci included in the 95 SNP collectively diagnostic panel. Linkage positions are from Brenna-Hansen et al.⁶⁴

sampling revealed decreases in the presence of hybrid and feral offspring over time. These results demonstrate the potential genetic consequences of a single-escape event from net-pen aquaculture on wild Atlantic salmon populations.

Results

Hybrid identification and genomic-based screening. In 2014, we collected 1704 young-of-the-year (YoY; i.e., fertilized the fall of the year of the escape, and hatched in the spring of the year of sampling) salmon from 18 rivers in the area adjacent to the escape event (Fig. 1; Table 1). Samples were again collected in 2015 ($n = 836$ of YoY and the 2014 cohort as 1+ juveniles; Table 1). All samples were screened using 95 genome-wide SNPs that were selected to maximize hybrid identification power and accuracy (Fig. 2; see Methods for further details as well as Supplementary Figures 1–3).

Our panel of 95 highly informative genome-wide SNPs identified 27.1% of the sampled YoY in 2014 as being of aquaculture ancestry based on a posterior probability assignment >0.80 (i.e., any of feral, F_1 , F_2 , and backcrosses, Fig. 3a). Hybrids were detected in 17 of the 18 rivers sampled (Fig. 3a, b), and feral

Table 1 Sample sizes of the juvenile Atlantic salmon screened for hybridization and introgression, the river from which they were collected, and the location of the river mouths

River name	Abbreviation	2104 YoY	2015 1+	2015 YoY	Lat (°N)	Long (°W)
Bottom Brook	BTB	32	33	0	47.765	56.322
Conne River	CNR	370	0	20	47.866	55.765
Dollard's Brook	DLR	25	24	22	47.708	56.555
Northwest Brook	FBN	41	0	0	47.720	55.393
Garnish River	GAR	199	50	56	47.239	55.353
Grand Bank Brook	GBB	42	26	15	47.104	55.754
Grand LaPierre	GLP	118	76	14	47.674	54.781
Long Harbour River	LHR	137	94	49	47.780	54.948
Salmonier Brook	LMS	40	22	89	46.865	55.775
Little River	LTR	130	0	0	47.809	55.743
Mal Bay Brook	MAL	17	70	36	47.669	55.131
Northeast Brook	NEB	115	19	0	47.723	55.367
Old Bay Brook	OBB	18	0	0	47.563	55.593
Southeast Brook	SEB	31	19	0	47.920	55.750
Simm's Brook	SMB	69	53	30	47.641	55.458
Taylor Bay Brook	TBB	120	0	0	47.543	55.637
Terrenceville Brook	TEB	120	0	0	47.671	54.711
Tailrace Brook	TRB	80	50	0	47.940	55.772

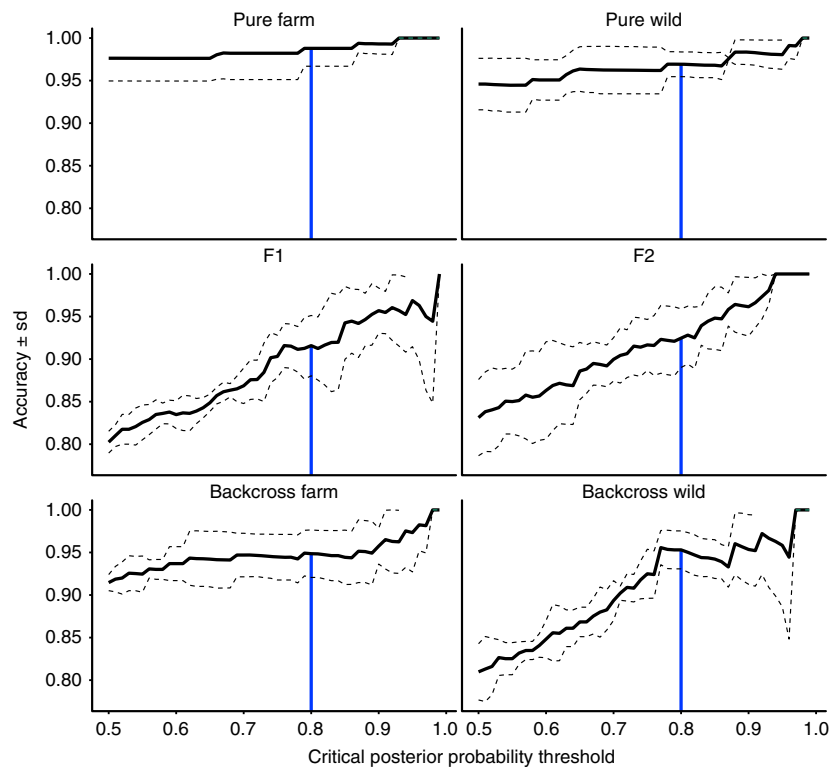


Fig. 2 Accuracy of detection of each of the genotype frequency classes across a range of critical posterior probability thresholds for the 95 SNP panel used in this study. The black line represents the mean of three replicate analyses of each of three independently simulated datasets and the dotted lines are the standard deviation. The vertical blue line is meant to highlight the critical posterior probability of assignment threshold (>0.8) used in this study

(i.e., offspring of two domestic salmon) offspring were detected in 13 rivers (Fig. 3a), revealing that the impacts of this escape event were substantial and region-wide. F₁ hybrids were the most common hybrid class detected in 2014, but F₂ and backcross individuals were also present (Fig. 3b). Observations of post-F₁ hybrids (i.e., F₂ and backcrosses) in 2014 YoY reveals that escape

events had occurred prior to 2013, and that genetic introgression was occurring in some rivers. Observations of feral offspring indicative of successful reproduction among escapees has not been previously reported to our knowledge within the natural range of Atlantic salmon¹⁸. However, the potential for the establishment of feral populations remains unclear. Sibship

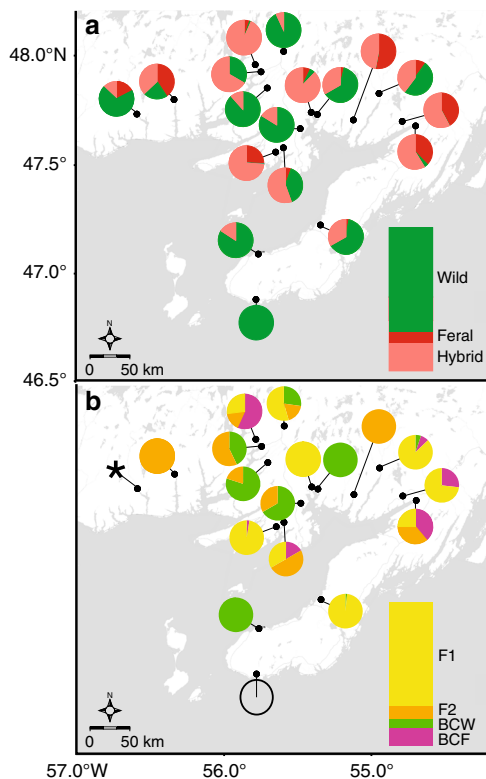


Fig. 3 Distribution and extent of hybridization following a large escape event of domestic Atlantic salmon. **a** Geographic distribution of wild, feral, or hybrid young-of-the-year Atlantic salmon across sample locations in 2014. **b** River-specific proportions of hybrid young-of-the-year salmon partitioned by hybrid genotype class (i.e., F₁, F₂, backcross wild (BCW), and backcross farm (BCF)). The open circle indicates a sample in which no hybrids were found, the asterisk signifies a location where accurate assignment to hybrid class was not possible. Bar graphs represent the overall proportions of each class in the entire sampling range, taking into account the varying sizes of the sampled populations (i.e., weighting by the axial distance, the distance along a straight line along the longest axis of the river), and colors therein are used as the legend

reconstruction revealed multiple unique parents for the hybrid and feral individuals in each river, suggesting that the over representation of a few families did not skew the river-specific estimation of hybrid proportion (Supplementary Tables 1-3).

Factors influencing hybridization. Levels of hybridization detected in 2014 were significantly associated with wild population size (Fig. 4; Supplementary Table 4). This was evident in significant associations between levels of hybridization and two proxies for salmon population size: river axial distance (i.e., the length of a straight line along a river's path), and average annual angling harvest (2010–2014), which correlate with salmon population size in this region (see Methods and Supplementary Figure 4). The proportion of hybrid YoY was negatively related to axial distance (Fig. 4b) and average annual angling (Fig. 4c); whereas, the opposite was true of the proportion of wild YoY (both $p < 0.001$, Supplementary Table 4). However, there was no statistical relationship between the proportion of feral YoY and either axial distance or average annual angling harvest (both $p > 0.10$; Supplementary Table 4). There was no evidence that distance between the location of the large escape event and river mouths influenced the proportion of wild, feral, or hybrid

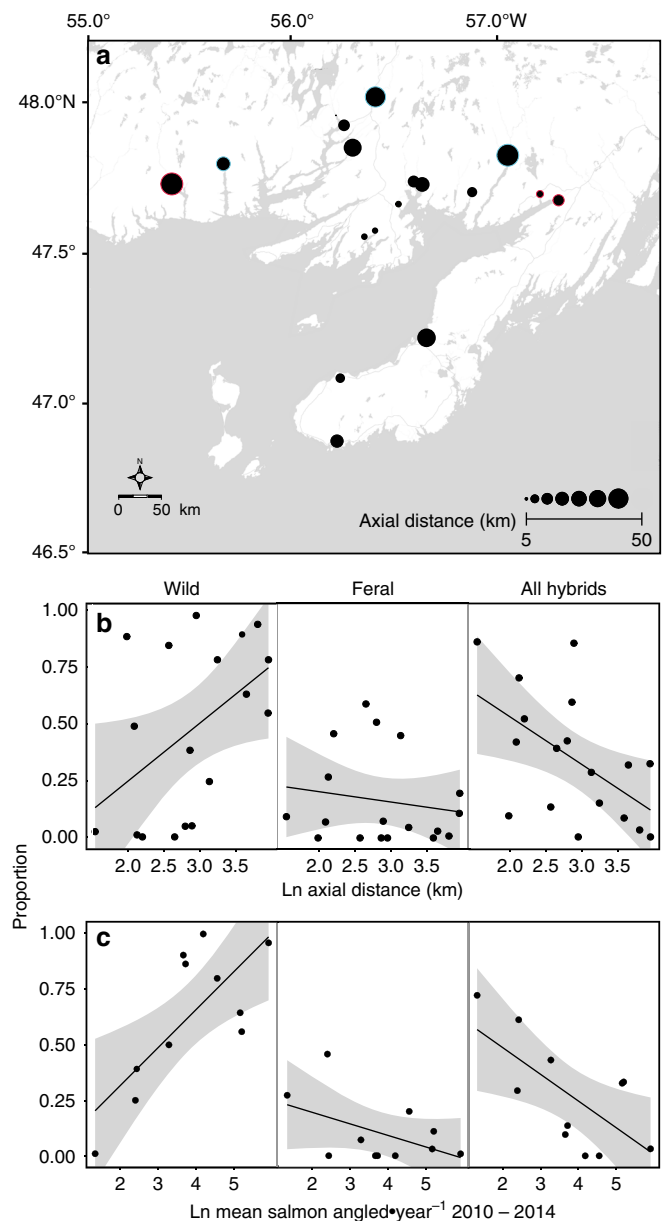


Fig. 4 Association between wild population size and levels of hybridization. **a** River axial distance (i.e., the distance along a straight line along the longest axis of the river). **b** Relationship between river axial distance and the proportions of wild, feral, and all hybrid (i.e., sum of proportions of F₁, F₂, BCW, and BC farm) young-of-the-year Atlantic salmon sampled in 2014. **c** Relationship between mean number of salmon angled (2010–2014) and the proportions of wild, feral, and hybrid young-of-the-year Atlantic salmon sampled in 2014. The gray shading is the 95% CI of the prediction of the linear models. See Supplementary Table 4 for model parameter estimates

offspring detected in the year following the escape event (all $p > 0.28$; Supplementary Table 5).

Temporal variation. To explore changes in the relative proportion of hybrids within the 2014 cohort over time, YoY and 1 + juvenile salmon were sampled and analyzed in 2015 from across the region. In comparison to the YoY sampled in 2014, these 2015 YoY samples revealed an almost complete absence of feral

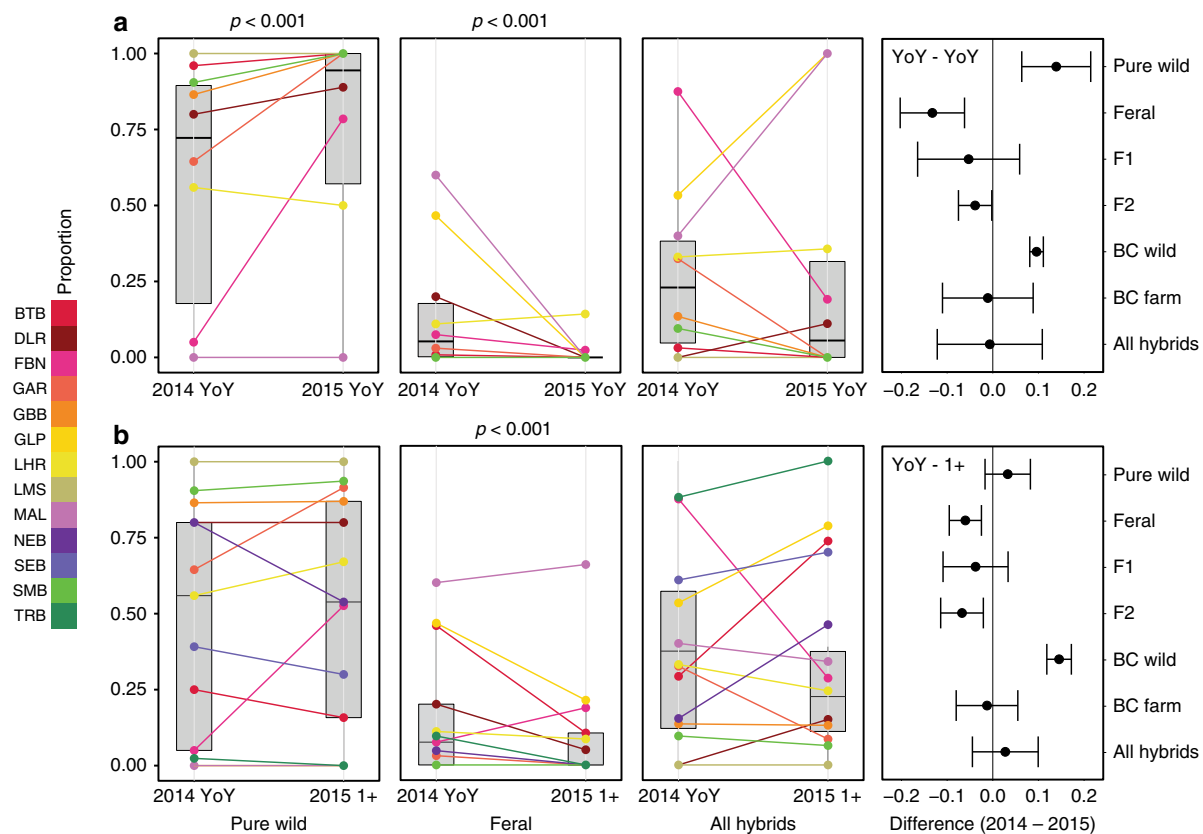


Fig. 5 Temporal variation (2014–2015) in levels of hybridization. **a** River specific and overall trends in the proportion of wild, feral, and hybrid young-of-the-year Atlantic salmon between 2014 and 2015. Gray shaded boxplots illustrate the overall proportions across all rivers, midline represents the medians, the upper and lower bounds the interquartile ranges, and the whiskers extend to 1.5 times the interquartile range. Black dots represent the mean difference (±SE) between 2014 and 2015 in the proportion of each pure and hybrid class present. All hybrids is the sum of proportions of F₁, F₂, BC wild, and BC farm. **b** River specific and overall trends in the proportion of wild, feral, and hybrid young-of-the-year and one year old (1+) Atlantic salmon sampled in 2014 and 2015, respectively. Gray shaded boxplots illustrate the overall proportions across all rivers. The midline represents the medians, the upper and lower bounds the interquartile ranges, and the whiskers extend to 1.5 times the interquartile range. Black dots represent the mean difference (±SE) across rivers between 2014 and 2015 in the proportion of each pure and hybrid class present of young-of-the-year and 1-year-old individuals. All hybrids is the sum of proportions of F₁, F₂, BC wild, and BC farm. See Fig. 1 and Table 1 for location information, and sample sizes

individuals and declines in the prevalence of most hybrid classes. This is likely reflective of overall lower numbers of escapees in 2014, a year in which no escape events were reported (Fig. 5a). The decline in feral individuals was significant ($p < 0.001$), as was the consequent increase in the proportion of wild ($p < 0.001$). However, whereas most hybrid classes were found to decrease, the change in the overall proportion of hybrids was offset by the increase in backcross wild individuals resulting in no significant difference between years ($p = 0.56$; Fig. 5a).

Potential offspring from the 2013 escape event (1+ individuals in 2015), showed that the proportion classified as feral declined significantly after a single year of selection in the wild ($p < 0.001$; Fig. 5b). There was no significant difference in the proportion of wild individuals ($p = 0.06$) and while there were decreases in most hybrid groups, the increase in backcross wild individuals muted any consistent statistical trend between years and among rivers and hybrid classes ($p = 0.20$; Fig. 5b). This decrease in the prevalence of offspring with part or full domestic ancestry is consistent with the reductions in relative hybrid survivorship observed in experimental studies^{13,21,37} and expected selection against these individuals in the wild. Nonetheless, the continued presence of F₂ and backcross individuals, as well as the observed increases in prevalence of wild backcross individuals indicates introgression is occurring.

Discussion

We report unambiguous landscape-scale evidence of interbreeding between wild and escapee Atlantic salmon resulting from a single-escape event, and of particular note, the first documented instance of which we are aware of feral offspring within the native range of Atlantic salmon¹⁸. The combination of a highly informative panel of genome-wide SNPs with a large escape event of non-local domestic individuals into largely pristine wild populations allowed unprecedented resolution of the magnitude and geographic scale of hybridization following a single-escape event. Hybrid and feral offspring were widespread geographically, occurring at distances of up to 100 km from the escape event, and accounted for ~27% of juvenile salmon surveyed. Moreover, the detection of F₂ and backcross individuals, presumably resultant from previous escape events, strongly supports the continued survival and reproductive viability of some hybrids, as well as the potential for significant demographic and genetic change as reported elsewhere¹⁸.

Our results demonstrate a clear association between the size of wild populations and the degree of hybridization (Fig. 4) suggesting that smaller salmon populations are at greater risk of hybridization and introgression with escaped domestic individuals as noted in Norway^{34,38}. This relationship is consistent with the dilution of domestic individuals in larger wild populations, as

well the consequences of increased competition between wild and domestic individuals both on the spawning grounds and at juvenile stages^{26,34}. Although, we lack actual estimates of wild population census size for many of the rivers included, the two correlates used here (river axial distance and annual angling harvest) are highly associated with population size on monitored rivers within the region (Supplementary Figure 5) and likely reflective of spatial trends in population census size.

Our results provide evidence consistent with declines in the proportion of offspring with domestic ancestry (e.g., hybrid, and feral) over time following the escape event. Comparison of the hybrid class composition of 1-year-old individuals sampled in 2015 relative to young-of-the-year sampled in 2014, revealed decreases in most hybrid classes, with only wild and wild backcrosses increasing in prevalence (Fig. 5b). Reduced wild-domestic hybrid survivorship for Atlantic salmon has previously been reported in experimental studies^{13,21,37}, but, we believe this is the first documentation following a single-escape event in the wild (Fig. 5b). The observed loss of feral and hybrid individuals over time is consistent with expected selection against these individuals in the wild. Interestingly, hybrid class composition of young-of-the-year sampled in 2015 revealed an almost complete absence of feral individuals and declines in the prevalence of most hybrid classes. This is consistent with an absence of reported escape events in 2015 and a reduced influence of the 2013 escape event. Despite evidence of declines in the proportion of domestic offspring or hybrids over time, the continued presence of F_2 s and backcrosses is clear evidence of introgression and that significant genetic change is occurring in these wild populations³⁹.

The identification and quantification of introgression and hybridization between domestic and wild Atlantic salmon is a critical first step toward understanding, predicting, and managing the genetic impacts of net-pen salmon aquaculture on wild populations. Our clear resolution of hybridization and introgression between escapee and wild Atlantic salmon in the Northwest Atlantic is the first to our knowledge, and is consistent with observations of genetic perturbation from aquaculture escapees³¹ both in the Canadian Maritimes and in Europe^{14,34,35}. Our results link previous observations of escapes of domesticated Atlantic salmon with reports of population-level genetic changes^{31,35} and regional declines of Atlantic salmon populations³⁶. Moreover, these results further demonstrate the potential consequences of escapes from net-pen aquaculture on wild Atlantic salmon populations.

Methods

Development of collectively diagnostic SNP panel. The collection of wild samples used for the development of our single nucleotide polymorphism (SNP) panel has been previously detailed in Bradbury et al.⁴⁰. Briefly, juvenile Atlantic salmon ($n = 260$, 0+ to 3+ years of age), were collected via electrofishing during the summers of 2008–2010 (sample sizes are found in Supplementary Table 6; genetic differentiation between populations are described in Supplementary Table 7). All wild collections were conducted under the auspices of Fisheries and Oceans collection permits. Aquaculture samples ($n = 156$) were obtained from two cage sites located within the region shown in Fig. 1. No effort was made to screen for or remove potential sibs from these baseline groups⁴¹. These baseline individuals were first screened using a 5568 SNP-locus panel developed by the Centre for Integrative Genomics (CIGENE, Norway^{42,43}) as per Bradbury et al.⁴⁴. Locus calls were visually confirmed and loci were retained if call rates were >0.85 and with overall minor allele frequencies >0.01 or a minor allele frequency >0.05 in either population⁴⁴. The loci retained after quality control filtering were ranked by Weir and Cockerham's F_{ST} between the two pooled reference groups (wild and domestic salmon), and the 95 most informative loci for which suitable assays could be developed were incorporated into the custom Fluidigm EPI array (see below). Linkage disequilibrium was not considered explicitly, however, the final panel provided genome-wide coverage (Fig. 1).

For each candidate locus, sequences from identified targets were downloaded from GenBank (SNP database, www.ncbi.nlm.nih.gov) and submitted to D3 Assay Design application (www.d3.fluidigm.com) for SNP Type assay design (Fluidigm,

San Francisco, CA, USA). Assays were tested on samples with known genotype and the selection criteria for inclusion in the final panel included: correct genotypes for known samples and positive controls (see below); genotypes being reproducible across multiple chip runs; the ranking of the target SNP in the prioritized list; and assays not requiring the STA (Specific Target Amplification) step. Positive controls consisted of normalized solutions of synthesized double stranded DNA (gBlocks (Integrated DNA Technologies, Coralville, IA, USA))⁴⁶. SNP genotyping was performed using SNP type assays (Fluidigm) per the manufacturer's protocols, without the STA (Specific Target Amplification) step, using 96.96 genotyping Integrated Fluidic Circuits (IFC) and read on an EPI (Fluidigm) and analyzed using SNP Genotyping Analysis software (Fluidigm). Each 96-well plate extraction included 10 samples that were repeated on the plate (redundants) to detect processing errors (row or plate reversal) and ensure consistent clustering interpretation. The setup for each IFC also included positive controls (see above for details). To calculate the genotype error rate, 11.3% of the samples were reanalyzed from the original tissue where tissue samples were permitted. Based on Pompanon et al.⁴⁷, the genotype error rate was calculated to be 0.01%.

Hybrids. We used the R⁴⁸ package *hybriddetective*⁴⁹ to simulate pure wild, farmed, and hybrid populations to evaluate the power of this panel to identify hybrids and hybrid classes. Using *hybriddetective* we simulated multigenerational (viz. pure wild, pure farm, F_1 , F_2 , and backcrosses to wild and farm) hybrid datasets based on the genotypes of our wild and farmed baselines at the 95 SNPs in our panel. A random subset of 90% of the individuals from the wild and farmed baselines was first taken. A centered wild baseline was created by randomly sampling two alleles per locus from those of the randomly sampled subset without replacement. The same was done to create a centered farmed baseline. Centering was done following Karlsson et al.³³ and has the effect of removing linkage and deviances from Hardy-Weinberg equilibrium that may have been present in a pooled sample of populations. Next, using the centered baselines, individuals in generation $t+1$ were created by randomly sampling without replacement one allele per locus from each of the parental populations (i.e., wild baseline subsample and farmed baseline subsample) at time t ⁴⁹. Three independently simulated datasets were each in turn analyzed three times in parallel using *NewHybrids*³² and the R package *parallel-newhybrid*⁵⁰, with a burn-in of 50,000 followed by 100,000 sweeps. *NEWHYBRIDS* calculates the posterior probability that an individual belongs to each of, in our case, six hybrid classes³². The results of the analyses of these simulated datasets were used to determine the efficiency and accuracy⁵¹ of our 95 SNP panel.

To evaluate the efficacy of our panel, two metrics were considered: the panel's accuracy and its efficiency. For both these measures, we use the definitions provided by Vähä and Primmer⁵¹. First, accuracy is the proportion of all individuals that were assigned to a hybrid class that truly belong in that hybrid class (i.e., number of individuals correctly assigned to a hybrid class divided by the total number of individuals assigned to that class), and is calculated independently for each hybrid class. Efficiency, is also calculated independently for each hybrid class, and measures the proportion of individuals that are known a priori to belong to a hybrid class that were assigned to that class (i.e., number of individuals correctly assigned to a hybrid class divided by total number of individuals known a priori to belong to a class). The accuracies and efficiencies calculated from the analyses of these simulated datasets across a range of posterior probability of assignment thresholds are shown in Fig. 2 and Supplementary Figures 1–3. From Fig. 2 (and also Supplementary Figure 3) it can be seen that the proportion of simulated individuals correctly assigned as either pure wild or feral are the highest across all posterior probability of assignment thresholds, while F_1 , F_2 , and backcross wild were comparatively lower. However, at all posterior probabilities of assignment shown, the accuracy for all hybrid classes was $>80\%$, suggesting the potential impact of miss-assignments is low. Similarly, efficiencies (Supplementary Figures 1 and 2) were above 90% for posterior probability of assignments thresholds between 0.5 and 0.8 (used in this analysis), suggesting the majority of individuals were assigned. Taken together, the high accuracy indicates that of those individuals assigned to a given class the majority were assigned correctly (i.e., little false assignment bias), while the high efficiency suggests that most individuals were assigned. A posterior probability of assignment threshold of 0.8 for individual classification was chosen based on the simulations and calculation of efficiency and accuracy (Fig. 2, Supplementary Figures 1–3). Individuals which did not meet the 0.8 posterior probability threshold for any hybrid class were considered only for the assignment as wild, farmed, or hybrid, and excluded from analyses focusing on specific hybrid classes. Convergence of the MCMC chains in *NewHybrids* was also confirmed using *hybriddetective*⁴⁹.

We evaluated both assignment to each of the six genotype frequency classes (Fig. 2 and Supplementary Figure 1), and pooled hybrid class identification (Supplementary Figures 2 and 3) separately, and accepted individual assignments to a class if their posterior probability of assignment to that class met, or exceeded a threshold of 0.8. We chose the threshold of 0.8, which is more conservative than what is typically used (e.g., 0.5^{51,52}), because, we wanted to maximize the accuracy of assignments (Fig. 2 and Supplementary Figures 3).

Sample collection and analysis. On 18 September 2013, 20,000 sexually mature Atlantic salmon weighing between 4.5 and 7 kg (10–15 lbs) escaped from an open

cage culture facility in southern Newfoundland, Canada. A number of these escapees were subsequently detected and captured in nearby rivers by technicians working for Fisheries and Oceans Canada (DFO). Gross morphological examination, in addition to necropsies conducted by DFO employees, showed that the recovered salmon were sexually mature, and in spawning condition. In 2014, the year following the large escape event, young-of-the-year (YoY) salmon were collected by electrofishing stream and river habitats in the 18 rivers shown in Fig. 1. Sampling included both rivers with historical records of established salmon populations (Conne River, Little River, Garnish River) and smaller streams lacking prior information on the presence or status of Atlantic salmon populations. With the exception of a few monitored rivers, information on the status of the wild populations in these rivers is largely lacking³⁶; what information does exist suggests recent declines in abundance.

Individuals were approximately age-binned based on an expected size–age distribution from 200 K aged Newfoundland parr with a 97.5% accuracy in YoY identification. All YoYs captured were euthanized and stored whole in 95% ethanol for later DNA extraction. Sample sizes by year for each river are listed in Table 1. Sampling was repeated in 2015 using the same methodology, with the exception that both YoY and 1+ individuals were retained. The 1+ individuals collected in 2015 belong to the same cohort of fish that were spawned following the escape event in 2013, and collected as YoY in 2014. Conversely, the YoY collected in the 2015 sampling were spawned in 2014, a year in which no escape events were reported in Newfoundland, and are thus expected to be reflective of the background rates of hybridization and introgression.

DNA was isolated from tissue samples using QIAamp 96 DNA QIAcube HT Kit (Qiagen, Toronto, ON, Canada) on a QIAcube HT (Qiagen) per the manufacturer's protocol with some modifications. Tissue samples were manually disrupted using a Tissue Lyser II (Qiagen) mixing 2×10 s at 20 s^{-1} . DNA was eluted twice in 100 μL buffer AE (Qiagen) pre-heated to 70°C . DNA extracts were quantified using Quant-iT PicoGreen dsDNA Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA) and read on a FLUOStar OPTIMA fluorescence plate reader (BMG Labtech, Ortenberg, Germany). All individuals were screened using the custom Fluidigm SNP panel and NEWHYBRIDS was used to quantify the proportion of individuals from different genotype frequency classes present in a river sample³². Samples from each river, and each year, were run independently. Prior information on allele frequencies of baseline farm and wild salmon were also provided to NEWHYBRIDS during analyses by including simulated pure farm and pure wild individuals (i.e., the same individuals used in the testing of the accuracy and efficiency of the panel described above). The known class (i.e., pure wild and pure farm) were indicated to NEWHYBRIDS, as well as the fact that they were not to be included as part of the mixture⁵³. Like in the determination of the efficacy of our panel described above, NEWHYBRIDS was run with a burn-in of 50,000 followed by 100,000 sweeps, which was found to be sufficient to ensure convergence during the panel testing. Proportions assigned to the various hybrid classes are shown in Supplementary Tables 8–10.

COLONY⁵⁴ was used to simultaneously infer the parentage and sibships of the YoY sampled in 2014, and the YoY and 1+ individuals sampled in 2015. Each river, sampling year, and year class was analyzed separately in COLONY and parents were assigned an ancestry (wild, farm, or F_1) based on the hybrid class in question (i.e., if an individual was feral, both parents must be farmed, if an individual was an F_1 , one parent must be farmed and the other wild, if an individual is an F_2 , both parents must themselves have been F_1 s, etc.). For each river, sampling year, and year class, locus-specific allelic dropout rates were estimated using the “missing” function in PLINK^{55,56}, and these were provided to COLONY. Allele frequencies were estimated by COLONY from the data provided. In running COLONY, because all samples were wild caught, no information about numbers of candidate males or females provided. Both sexes were assumed to be polygynous, and “long” runs with “VeryHigh” precision were used. Because we were not attempting to assign parentage, merely estimate the number of families present in each sample, and show that the proportions of hybrid classes detected was not the result of over representation of one, or a few families, the full-sib grouping for each individual with the highest probability was accepted. It should also be noted that because no parental genotypes were provided to COLONY, it was unable to meaningfully assign sexes to parents. Therefore, the total number of parents are presented.

Statistical analyses. All statistical analyses were conducted in R version 3.4⁴⁸. The proportion of wild, feral, and hybrid at each location was explored for associations with wild population size; in this case two proxies were used (axial river distance and average annual harvest). For the Newfoundland region, wild population size⁵⁷ is associated with river axial distance⁵⁸ (the distance along a straight line along the longest axis of the river; linear model, $R^2 = 0.6944$, $F_{1,8} = 18.18$, $p < 0.01$; Supplementary Figure 4) and as such, axial distance was used as a proxy for population size. We also used average annual harvest (2010–2014) as a proxy for population size; because, the two were related (linear model, $F_{1,8} = 40.47$, $R^2 = 0.835$, $p < 0.001$). Harvest statistics are collected annually by Fisheries and Oceans Canada⁵⁹, and counts of population size and estimates of annual harvest were available for 10 rivers (Supplementary Table 11).

Exponential models for effect of distance from the escape event were used because straying of farmed salmon generally follows a negative exponential distribution⁶⁰. The relationship between the proportion feral, wild, and hybrids

detected in each river and the distance between the river mouths and the site of the escape were tested and fit using the R function *nls*. No significant relationships were found for distance from the escape event (all $p > 0.28$; Supplementary Table 5), so this factor was not considered further. The impact of the relative size of the native salmon populations in respective rivers on proportions was tested using linear models with the R function *lm*. The proportion wild, the proportion feral, and the proportion hybrid were tested separately as a function of axial distance, and then average annual angling harvest between 2010 and 2014.

We tested for differences in proportion of wild, feral, or hybrid individuals between years within the same cohort (i.e., the YoY collected in 2014 and the 1+ collected in 2015), and between years with and without reported large escape events (i.e., YoY collected in 2014 and YoY collected in 2015) using binomial mixed-effects models with river as the random effect using the R function *glmer*⁶¹. The proportions of wild, feral, and hybrid were tested with separate models, and p -values were adjusted using the false discovery rate⁶².

Data availability. Genotype, river characteristic, salmon and angling count data for this study are available in the Dryad Digital repository⁶³ at: <https://doi.org/10.5061/dryad.3k888n7>

Received: 6 December 2017 Accepted: 16 July 2018

Published online: 09 August 2018

References

- Meyer, R. S. & Purugganan, M. D. Evolution of crop species: genetics of domestication and diversification. *Nat. Rev. Genet.* **14**, 840–852 (2013).
- Sweeney, M. T. et al. Global dissemination of a single mutation conferring white pericarp in rice. *PLoS Genet.* **3**, e133 (2007).
- Frantz, L. A. F. et al. Evidence of long-term gene flow and selection during domestication from analyses of Eurasian wild and domestic pig genomes. *Nat. Genet.* **47**, 1141–1148 (2015).
- Feurtey, A., Cornille, A., Shykoff, J. A., Snirc, A. & Giraud, T. Crop-to-wild gene flow and its fitness consequences for a wild fruit tree: towards a comprehensive conservation strategy of the wild apple in Europe. *Evol. Appl.* **10**, 180–188 (2017).
- Hufford, M. B. et al. Comparative population genomics of maize domestication and improvement. *Nat. Genet.* **44**, 808–811 (2012).
- Oka, H. I. & Chang, W. T. Hybrid swarms between wild and cultivated rice species, *Oryza perennis* and *O. Sativa*. *Evolution* **15**, 418–430 (1961).
- Eriksson, J. et al. Identification of the yellow skin gene reveals a hybrid origin of the domestic chicken. *PLoS Genet.* **4**, e1000010 (2008).
- Kadwell, M. et al. Genetic analysis reveals the wild ancestors of the llama and the alpaca. *Proc. R. Soc. B Biol. Sci.* **268**, 2575–2584 (2001).
- Brown, T. A., Jones, M. K., Powell, W. & Allaby, R. G. The complex origins of domesticated crops in the Fertile Crescent. *Trends Ecol. Evol.* **24**, 103–109 (2009).
- Dempewolf, H., Hodgins, K. A., Rummell, S. E., Ellstrand, N. C. & Rieseberg, L. H. Reproductive isolation during domestication. *Plant Cell* **24**, 2710–2717 (2012).
- Cornille, A., Giraud, T., Smulders, M. J. M., Roldan-Ruiz, I. & Gladieux, P. The domestication and evolutionary ecology of apples. *Trends Genet.* **30**, 57–65 (2014).
- McGinnity, P. et al. Fitness reduction and potential extinction of wild populations of Atlantic salmon, *Salmo salar*, as a result of interactions with escaped farm salmon. *Proc. R. Soc. Lond., Ser. B: Biol. Sci.* **270**, 2443–2450 (2003).
- Fleming, I. A. et al. Lifetime success and interactions of farm salmon invading a native population. *Proc. R. Soc. Lond., Ser. B: Biol. Sci.* **267**, 1517–1523 (2000).
- Ferguson, A. et al. in *The Atlantic Salmon Genetics, Conservation and Management* (eds E. Verspoor, L. Stradmeyer, & J. L. Nielsen) pp. 357–398 (Blackwell Publishing Ltd, Oxford, 2007).
- Carr, J. W., Anderson, J. M., Whoriskey, F. G. & Dilworth, T. The occurrence and spawning of cultured Atlantic salmon (*Salmo salar*) in a Canadian river. *ICES J. Mar. Sci.* **54**, 1064–1073 (1997).
- Cross, T. F. & King, J. Genetic effects of hatchery rearing in Atlantic salmon. *Aquaculture* **33**, 33–40 (1983).
- Karlsson, S., Moen, T., Lien, S., Glover, K. A. & Hindar, K. Generic genetic differences between farmed and wild Atlantic salmon identified from a 7K SNP-chip. *Mol. Ecol. Res.* **11**(Suppl 1), 247–253 (2011).
- Glover, K. A. et al. Half a century of genetic interaction between farmed and wild Atlantic salmon: Status of knowledge and unanswered questions. *Fish. Fish.* **18**, 890–927 (2017).
- Fleming, I. A. & Einum, S. Experimental tests of genetic divergence of farmed from wild Atlantic salmon due to domestication. *ICES J. Mar. Sci.* **54**, 1051–1063 (1997).

20. Fleming, I. A., Jonsson, B., Gross, M. R. & Lamberg, A. An experimental study of the reproductive behaviour and success of farmed and wild Atlantic salmon (*Salmo salar*). *J. Appl. Ecol.* **33**, 893–905 (1996).
21. McGinnity, P. et al. Genetic impact of escaped farmed Atlantic salmon (*Salmo salar* L.) on native populations: use of DNA profiling to assess freshwater performance of wild, farmed, and hybrid progeny in a natural river environment. *ICES J. Mar. Sci.* **54**, 998–1008 (1997).
22. Keyser, F. et al. Predicting the impacts of escaped farmed Atlantic salmon on wild salmon populations. *Can. J. Fish. Aquat. Sci.* **75**, 506–512 (2018).
23. Morris, M. R. J. et al. Prevalence and recurrence of escaped farmed Atlantic salmon (*Salmo salar*) in eastern North American rivers. *Can. J. Fish. Aquat. Sci.* **65**, 2807–2826 (2008).
24. Skilbrei, O. T., Heino, M. & Svåsand, T. Using simulated escape events to assess the annual numbers and destinies of escaped farmed Atlantic salmon of different life stages from farm sites in Norway. *ICES J. Mar. Sci.* **72**, 670–685 (2015).
25. Skaala, O., Wennevik, V. & Glover, K. A. Evidence of temporal genetic change in wild Atlantic salmon, *Salmo salar* L., populations affected by farm escapees. *ICES J. Mar. Sci.* **63**, 1224–1233 (2006).
26. Glover, K. A. et al. Three decades of farmed escapees in the wild: a spatio-temporal analysis of Atlantic salmon population genetic structure throughout Norway. *PLoS ONE* **7**, e43129 (2012).
27. Clifford, S. L., McGinnity, P. & Ferguson, A. Genetic changes in Atlantic salmon (*Salmo salar*) populations of northwest Irish rivers resulting from escapes of adult farm salmon. *Can. J. Fish. Aquat. Sci.* **55**, 358–363 (1998).
28. Clifford, S. L., McGinnity, P. & Ferguson, A. Genetic changes in an Atlantic salmon population resulting from escaped juvenile farm salmon. *J. Fish. Biol.* **52**, 118–127 (1998).
29. Crozier, W. W. Evidence of genetic interaction between escaped farmed Salmon and Wild Atlantic Salmon (*Salmo salar* L.) in a Northern Irish River. *Aquaculture* **113**, 19–29 (1993).
30. Crozier, W. W. Escaped farmed salmon, *Salmo salar* L., in the Glenarm River, Northern Ireland: genetic status of the wild population 7 years on. *Fish. Manag. Ecol.* **7**, 437–446 (2000).
31. Bourret, V., O'Reilly, P. T., Carr, J. W., Berg, P. R. & Bernatchez, L. Temporal change in genetic integrity suggests loss of local adaptation in a wild Atlantic salmon (*Salmo salar*) population following introgression by farmed escapees. *Heredity* **106**, 500–510 (2011).
32. Anderson, E. C. & Thompson, E. A. A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* **160**, 1217–1229 (2002).
33. Karlsson, S., Diserud, O. H., Moen, T. & Hindar, K. A standardized method for quantifying unidirectional genetic introgression. *Ecol. Evol.* **4**, 3256–3263 (2014).
34. Glover, K. A. et al. Atlantic salmon populations invaded by farmed escapees: quantifying genetic introgression with a Bayesian approach and SNPs. *BMC Genet.* **14**, 74–93 (2013).
35. Karlsson, S., Diserud, O. H., Fiske, P. & Hindar, K. Widespread genetic introgression of escaped farmed Atlantic salmon in wild salmon populations. *ICES J. Mar. Sci.* **73**, 2488–2498 (2016).
36. COSEWIC. COSEWIC assessment and status report on the Atlantic Salmon *Salmo salar* (Nunavik population, Labrador population, Northeast Newfoundland population, South Newfoundland population, Southwest Newfoundland population, Northwest Newfoundland population, Quebec Eastern North Shore population, Quebec Western North Shore population, Anticosti Island population, Inner St. Lawrence population, Lake Ontario population, Gaspé-Southern Gulf of St. Lawrence population, Eastern Cape Breton population, Nova Scotia Southern Upland population, Inner Bay of Fundy population, Outer Bay of Fundy population) in Canada. xlvii+136 (Ottawa, 2010).
37. Skaala, O. et al. Performance of farmed, hybrid, and wild Atlantic salmon (*Salmo salar*) families in a natural river environment. *Can. J. Fish. Aquat. Sci.* **69**, 1994–2006 (2012).
38. Heino, M., Svåsand, T., Wennevik, V. & Glover, K. A. Genetic introgression of farmed salmon in native populations: quantifying the relative influence of population size and frequency of escapees. *Aquacult. Environ. Inter.* **6**, 185–190 (2015).
39. Bolstad, G. H. et al. Gene flow from domesticated escapes alters the life history of wild Atlantic salmon. *Nat. Ecol. Evol.* **1**, 0124 (2017).
40. Bradbury, I. R. et al. Landscape structure and climatic variation determine Atlantic salmon genetic connectivity in the Northwest Atlantic. *Can. J. Fish. Aquat. Sci.* **71**, 246–258 (2014).
41. Waples, R. S. & Anderson, E. C. Purging putative siblings from population genetic data sets: a cautionary view. *Mol. Ecol.* **26**, 1211–1224 (2017).
42. Bourret, V. et al. SNP-array reveals genome-wide patterns of geographical and potential adaptive divergence across the natural range of Atlantic salmon (*Salmo salar*). *Mol. Ecol.* **22**, 532–551 (2013).
43. Bourret, V., Dionne, M., Kent, M. P., Lien, S. & Bernatchez, L. Landscape genomics in Atlantic salmon (*Salmo salar*): searching for gene-environment interactions driving local adaptation. *Evolution* **67**, 3469–3487 (2013).
44. Bradbury, I. R. et al. Transatlantic secondary contact in Atlantic Salmon, comparing microsatellites, a single nucleotide polymorphism array and restriction-site associated DNA sequencing for the resolution of complex spatial structure. *Mol. Ecol.* **24**, 5130–5144 (2015).
45. Weir, B. S. & Cockerham, C. C. Estimating F-statistics for the analysis of population-structure. *Evolution* **38**, 1358–1370 (1984).
46. Richards-Hrdlicka, K. L. A genomic approach to finding and understanding how the amphibian chytrid fungus is genetically structured in Northeastern America, PhD thesis, Yale University, New Haven, Connecticut (2014).
47. Pompanon, F., Bonin, A., Bellemain, E. & Taberlet, P. Genotyping errors: causes, consequences and solutions. *Nat. Rev. Genet.* **6**, 847–859 (2005).
48. R Development Core Team. *R: A language and environment for statistical computing*. (R Foundation for Statistical Computing, 2016).
49. Wringe, B. F., Stanley, R. R. E., Jeffery, N. W., Anderson, E. C. & Bradbury, I. R. hybriddetective: a workflow and package to facilitate the detection of hybridization using genomic data in R. *Mol. Ecol. Resour.* **17**, e275–e284 (2017).
50. Wringe, B. F., Stanley, R. R., Jeffery, N. W., Anderson, E. C. & Bradbury, I. R. parallelnewhybrid: an R package for the parallelization of hybrid detection using newhybrids. *Mol. Ecol. Resour.* **17**, 91–95 (2017).
51. Vähä, J. P. & Primmer, C. R. Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Mol. Ecol.* **15**, 63–72 (2006).
52. Pritchard, V. L. et al. Single nucleotide polymorphisms to discriminate different classes of hybrid between wild Atlantic salmon and aquaculture escapees. *Evol. Appl.* **9**, 1017–1031 (2016).
53. Anderson, E. C. User's guide to the program NewHybrids Version 1.1 beta. (2003).
54. Jones, O. R. & Wang, J. L. COLONY: a program for parentage and sibship inference from multilocus genotype data. *Mol. Ecol. Res.* **10**, 551–555 (2010).
55. Purcell, S. PLINK 1.07 v. 1.07 (2007), <http://zzz.bwh.harvard.edu/plink/contact.shtml#cite>.
56. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
57. DFO. Atlantic Salmon (*Salmo salar*) Stock Status Update in Newfoundland and Labrador for 2014. *DFO Can. Sci. Advis. Sec. Sci. Resp.* 2015/023. (2015). http://publications.gc.ca/collections/collection_2015/mpo-dfo/Fs70-7-2015-023-eng.pdf
58. Porter, T. R., Riche, L. G. & Traverse, G. R. Catalogue of Rivers in Insular Newfoundland. Volume B. 287 (Environment Canada, Fisheries and Marine Service, Resource Development Branch, Newfoundland Region, 1974).
59. DFO. Stock Assessment of Newfoundland and Labrador Atlantic Salmon – 2016. *DFO Can. Sci. Advis. Sec. Sci. Advis. Rep.* 2017/035. (2017).
60. Jonsson, B., Jonsson, N. & Hansen, L. P. Atlantic salmon straying from the River Imsa. *J. Fish. Biol.* **62**, 641–657 (2003).
61. Bates, D., Machler, M., Bolker, B. M. & Walker, S. C. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
62. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate - a practical and powerful approach to multiple testing. *J. Roy. Stat. Soc. Ser. B.* **57**, 289–300 (1995).
63. Wringe, B. et al. Data from: extensive hybridization following a large escape of domesticated Atlantic salmon in the Northwest Atlantic. Dryad Digital Repository. <https://doi.org/10.5061/dryad.3k888n7>
64. Brenna-Hansen, S. et al. Chromosomal differences between European and North American Atlantic salmon discovered by linkage mapping and supported by fluorescence in situ hybridization analysis. *BMC Genom.* **13**, 432 (2012).

Acknowledgements

The authors wish to thank Laura K. Weir, Paul Bentzen, and Sarah J. Lehnert for helpful and insightful comments on previous versions of this manuscript. Staff of DFO Salmonids section Newfoundland Region were responsible for the collection of juvenile tissue samples. Acquisition of aquaculture salmon baseline samples was facilitated by G. Perry, C. Hendry, DFO Aquaculture section Newfoundland Region, and by industry partners Cooke Aquaculture and Northern Harvest Sea Farms. We also thank the staff of the Aquatic Biotechnology Laboratory of the Bedford Institute of Oceanography for their work in the SNP genotyping. This study was funded through the Program for Aquaculture Regulatory Research of Fisheries and Oceans Canada, the Genomics Research and Development Initiative of Canada, as well as the National Sciences and Engineering Research Council of Canada, and support from the Atlantic Salmon Federation.

Author contributions

B.F.W., I.R.B., C.G., and I.A.F. conceived of and designed the study. Statistical and genomic analyses were conducted by B.F.W., E.C.A., and L.C.H. Figures were produced by B.F.W., N.W.J., R.R.E.S., and I.R.B. All authors wrote and approved the final draft of the manuscript.

Additional information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s42003-018-0112-9>.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018