



# Migration effort and wild population size influence the prevalence of hybridization between escaped farmed and wild Atlantic salmon

Emma V. A. Sylvester<sup>1,\*</sup>, Brendan F. Wringe<sup>2</sup>, Steven J. Duffy<sup>1</sup>,  
Lorraine C. Hamilton<sup>3</sup>, Ian A. Fleming<sup>4</sup>, Ian R. Bradbury<sup>1</sup>

<sup>1</sup>Science Branch, Department of Fisheries and Oceans Canada, 80 East White Hills Road, St. John's, NL A1C 5X1, Canada

<sup>2</sup>Science Branch, Department of Fisheries and Oceans Canada, Bedford Institute of Oceanography, Dartmouth, NS B2Y 4A2, Canada

<sup>3</sup>Aquatic Biotechnology Laboratory, Fisheries and Oceans Canada, Bedford Institute of Oceanography, Dartmouth, NS B2Y 4A2, Canada

<sup>4</sup>Memorial University of Newfoundland, Department of Ocean Sciences, St. John's, NL A1C 5S7, Canada

**ABSTRACT:** Hybridization of escaped farmed Atlantic salmon *Salmo salar* with wild populations occurs throughout their native range and can threaten wild population stability and persistence. The extent of hybridization is often population-specific and can drive changes in phenotype and genotype. Current understanding of the forces that contribute to the spatial distribution of hybridization is insufficient despite its potential to inform conservation and management efforts. Using a panel of 95 single nucleotide polymorphisms previously validated for identifying parr of farmed, wild, or hybrid descent, we present a novel exploration of inter- and intra-river distribution of hybrids from 33 locations across 9 rivers in southern Newfoundland, Canada. The proportion of hybrids varied significantly across ( $p < 0.001$ ) and within rivers ( $p < 0.05$  in 4 rivers). Binomial mixed models and logistic regression showed increased proportions of hybrid and feral offspring within smaller rivers ( $p < 0.0001$ ). Within-river distribution of hybrid parr was strongly associated with the migration effort required to reach spawning sites; the hybrid proportion decreased significantly ( $p < 0.05$ ) with increased elevation, geographic distance and the presence of obstructions. These observations support previous hypotheses that the distribution of escaped farmed Atlantic salmon can be restricted by migratory challenges, which result in the reduction of hybrid individuals in upstream spawning sites relative to downstream locations. Our research demonstrates that levels of hybridization vary spatially and are associated with landscape features. We suggest that consideration of spatial variation in levels of hybridization will be essential for the evaluation of impacts that escaped farmed salmon impose on wild Atlantic salmon populations.

**KEY WORDS:** Fish farming · Genetic identification · Fisheries interactions · Migration · *Salmo salar* · Aquaculture impacts · Introgression

## INTRODUCTION

Farmed Atlantic salmon *Salmo salar* differ both genetically and phenotypically from their wild counterparts due to breeding practices, domestication selection, random drift during domestication, and

differences in ancestry (Karlsson et al. 2011, Glover et al. 2017). As a result, escaped farmed salmon are poorly adapted to the wild environment (Gausen & Moen 1991, Einum & Fleming 1997, Skaala et al. 2012), resulting in reduced survivorship of escapees and hybrids relative to wild individuals (Fleming et

\*Corresponding author: emmasylvester7@gmail.com

© I. A. Fleming and Fisheries and Oceans Canada. Open Access under Creative Commons by Attribution Licence. Use, distribution and reproduction are unrestricted. Authors and original publication must be credited.

Publisher: Inter-Research · www.int-res.com

al. 2000, McGinnity et al. 2003, Wringe et al. 2018a). Estimates suggest that the number of escaped individuals may outnumber that of wild salmon annually in the North Atlantic (Glover et al. 2017), with reports of interactions with native populations (Hindar et al. 2006, Glover et al. 2013, Karlsson et al. 2016). Possible consequences of these interactions include increased competition for resources (Skaala et al. 2012), exposure to introduced disease and pathogens, and outbreeding depression (McGinnity et al. 2003). Changes in genetic composition of wild populations (Skaala et al. 2006, Glover et al. 2012) and associated phenotypic changes (Bolstad et al. 2017) may compromise the fitness and reproductive capacity of wild populations (Fleming et al. 2000, McGinnity et al. 2003).

The degree of hybridization (production of  $F_1$  offspring between pure wild and pure farmed fish), introgression (further reproduction of  $F_1$  fish, resulting in  $F_2$  and backcross hybrid classes), and susceptibility of a population to its effects, however, is often river- or area-specific (Harvey et al. 2016). Hybrid proportion and the degree of introgression varies across populations and is influenced by the proximity of a spawning location to aquaculture facilities (Karlsson et al. 2016, Keyser et al. 2018), as well as the number of escapees from these sites (Hansen & Windsor 2006, Heino et al. 2015, Keyser et al. 2018). Recent Norwegian estimates suggest that 47% of rivers show evidence of significant genetic alteration, with the proportion of farmed genetic introgression reaching up to 42% within a river (Karlsson et al. 2016). Stream or river size has also been linked to the occurrence of hybridization with farmed fish (Heggerget et al. 1993, Økland et al. 1995, Wringe et al. 2018a), possibly due to density-dependent competition associated with the population size at a spawning site (Glover et al. 2013). Spatial distribution of escaped farmed and wild individuals can also vary within rivers; wild fish have been found in greater proportions in upstream sites (Webb et al. 1991, Carr et al. 1997), consistent with the hypothesis that reduced migratory capacity or reduced homing motivation of escapees will limit their maximum upstream migration distance; however, the opposite has also been reported (Thorstad et al. 1998, Moe et al. 2016). Conflicting findings may be due to specific river characteristics such as obstructions in downstream areas of certain rivers, restricting migration of farmed salmon to spawning areas below the obstruction. Alternatively, escapees may be relegated to the lower reaches of rivers through competition with wild conspecifics for up-stream spawning sites (Moe

et al. 2016). Research directed towards investigating the particular forces that influence the distribution of farmed escapees both within and across rivers has been minimal in North America, despite the potential to apply this knowledge to mitigation and management efforts.

Here, we explored the spatial distribution of impacts of escaped farmed salmon—specifically, the distribution of hybrid individuals across and within several watersheds in southern Newfoundland, Canada (see Fig. 1). Our primary objective was to quantify spatial variation in levels of hybridization within and between rivers (populations), as well as to identify key factors that influence this variation. Landscape-level studies on the number of affected rivers, or the extent and genetic consequences of introgression within rivers have been conducted in Europe (e.g. McGinnity et al. 2003, Glover et al. 2013, Karlsson et al. 2016, Moe et al. 2016), but remain relatively unexplored in North America (but see Bourret et al. 2011, Keyser et al. 2018, Wringe et al. 2018a). This study builds directly on work examining the consequences of a single escape event that occurred in 2013 of approximately 20 000 farmed Atlantic salmon, which approximately doubled the number of salmon in the region (Keyser et al. 2018, Wringe et al. 2018a). We extended this previous research with a novel investigation of potential fine-scale landscape influences on hybridization in the region following this escape event and subsequent trickle escapes.

## MATERIALS AND METHODS

### Sampling and genotyping

Individual samples used as farmed references ( $n = 156$ ) were provided from 3 cage sites within Newfoundland, Canada, and used as baseline samples. As salmon cages in Atlantic Canada are presently stocked only with individuals from a single, non-local lineage from the Saint John River (Wringe et al. 2018b), these samples are representative of escapees in the region. Previously identified (with high certainty) pure wild individuals ( $n = 301$ ), collected from 9 rivers between 2008 and 2010 (see Wringe et al. 2018a), were included as the wild baseline. A total of 3198 juvenile Atlantic salmon *Salmo salar*, aged young-of-the-year (YoY), 1+, and 2+, were collected by electrofishing in 2016 from 33 sites across 9 rivers in southern Newfoundland (see Table 1, Fig. 1). Individuals were stored in 95% ethanol for later DNA extraction.

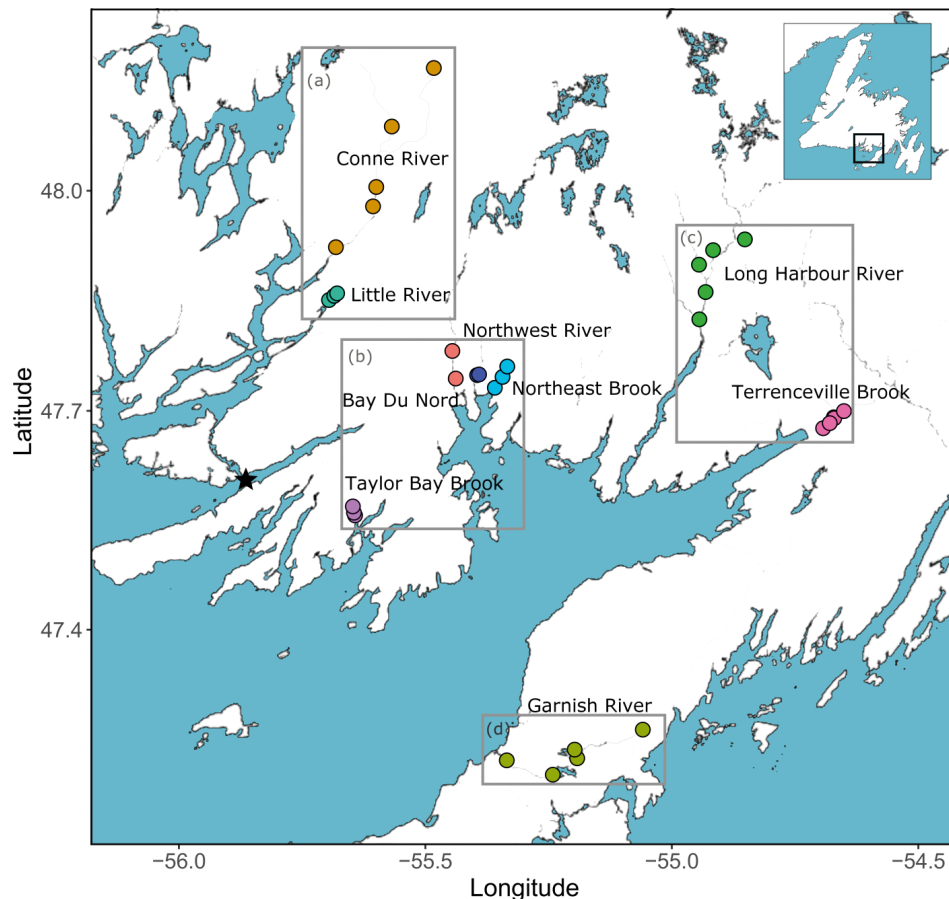


Fig. 1. Sampling sites of Atlantic salmon across 9 rivers in southern Newfoundland, Canada. Sampling sites within a given river are indicated by the same colour. Black star: location of the 2013 aquaculture escape event. Labelled panels correspond to panels of group proportions in Fig. 3

DNA was extracted using QIAamp 96 DNA QIAcube HT Kit (Qiagen) on a QIAcube HT (Qiagen) following the manufacturer's protocol. Tissue samples were disrupted using a Tissue-Lyser II (Qiagen) mixing  $2 \times 10$  s at  $20 \text{ s}^{-1}$ . DNA was eluted twice in  $100 \mu\text{l}$  Buffer AE (Qiagen) pre-heated to  $70^\circ\text{C}$ . DNA extracts were quantified using QuantiT PicoGreen ds-DNA Assay Kit (Thermo Fisher Scientific) and read on a FLUOStar OPTIMA fluorescence plate reader (BMG Labtech). Individuals were genotyped using single nucleotide polymorphism (SNP) Type assays (Fluidigm) following the manufacturer's protocols, targeting 95 SNPs previously established for classification of farmed and wild salmon in Newfoundland (Wringe et al. 2018b). Each plate extraction included 10 redundant samples to detect processing errors. To calculate the genotype error rate, 214 samples (representing 12.2% of the processed samples) were genotyped a second time; 130 of those were newly extracted, and for the remaining 84 samples with no tissue remaining, samples were genotyped from the original extraction.

### Statistical analyses

All R analyses were conducted in R v.3.4.1 (R Core Development Team 2008). Figures were created using the R package 'ggplot2' (Wickham 2009). Classification of individuals to pure wild, feral, or 1 of 4 'hybrid' classes ( $F_1$ ,  $F_2$ , backcross wild [BCW], backcross feral [BCF]) at a given posterior probability was conducted using NEWHYBRIDS (Anderson & Thompson 2002), which implements a Markov chain Monte Carlo Bayesian approach for assignment, producing a posterior probability per class, per individual, based on provided baselines. We applied centred baseline simulations from the R package 'hybridetective' (Wringe et al. 2017a). NEWHYBRIDS was implemented for each site independently, in the R package 'parallelnewhybrid' (Wringe et al. 2017b) with a burn-in of 50 000 and 100 000 sweeps. Baseline individuals were labelled following NEWHYBRIDS documentation, using the 's' parameter, indicating that baseline samples were not obtained from the sam-

pling site, to avoid biasing the expected sample proportions. To confirm our hybrid class assignment, we also applied STRUCTURE for hybrid classification as proposed by Karlsson et al. (2014). Each individual was run individually with 100 simulated, centred baselines for both wild and farmed classes. We ran STRUCTURE with 50 000 burn-in and 100 000 repetitions for 3 iterations for all 3198 samples. *Q*-values were then plotted by class assigned in NEWHYBRIDS to assess consistency of assignment (see Fig. S1 in the Supplement at [www.int-res.com/articles/suppl/q010p401\\_supp.pdf](http://www.int-res.com/articles/suppl/q010p401_supp.pdf)). For downstream analysis, we grouped individuals assigned to 1 of the 4 hybrid classes in NEWHYBRIDS into a single hybrid class and filtered individuals at a posterior probability threshold of 0.8 for assignment to a given group (pure farmed [feral], pure wild, or hybrid). Chi-squared fitness tests were conducted on count data for pure wild, feral, and hybrid groups (class proportion  $\times$  sample size at a given site), scaled by sample size. Chi-squared tests were conducted on each river independently to test for within-river variation, and across rivers, using overall river group counts to assess inter-river variation. Population structure was assessed for association with hybrid classification using principal components analysis (PCA). Analyses were conducted on all data (baseline and field samples), as well as field-sampled data alone, using the R package 'adegenet' (Jombart 2008) with 2 principal components retained.

Distance from each sampling site to the respective river mouth and elevation at each site was approximated using Google Earth Pro (<https://google.com/earth>). Basin area, mean river width, basin perimeter, axial length (length of river along the down-valley axis), and maximum basin relief was obtained for each river from published data (Porter et al. 1974). Average propagule pressure was calculated from aquaculture data (available from 2005 to 2012) in concordance with Keyser et al. (2018). The number of obstructions along the path to each site and whether or not a site was within a tributary or the main stem was determined from the river descriptions of Porter et al. (1974) and Google Earth Pro. Environmental variables were standardized prior to variable selection (see Fig. S2). Mixed models were run in the R package 'lme4' (Bates et al. 2015) with each group set as a proportional binomial response (i.e. wild vs. non-wild, feral vs. non-feral, and hybrid vs. pure). Binomial rather than multinomial models were run to allow for the interpretation of the relative contribution of tested variables to group proportion for each group. River was set as a grouping variable (random factor) for each model, with a forward selection

approach for each environmental variable (fixed factor) considered. The interaction of distance to river mouth and elevation was considered as a function of the migration effort (Moore et al. 2017) required to reach a given spawning site (hereafter referred to as minimum migration effort, MME). Model fit was assessed by comparing Akaike's information criterion (AIC) (Akaike 1974) and the significance of an additional term across models tested. Models were first run with each environmental variable and the elevation  $\times$  distance interaction term tested independently with the grouping variable included in the model (see Table 2, Models 2 to 10). A variable was considered to be important and included in subsequent model tests if (1) its inclusion in the model was found to be statistically significant at  $\alpha = 0.05$  and (2) it reduced the AIC from Model 1 (see Table 2) in at least 2 of the 3 analyses tested. The model with the best fit was used as a base model for subsequent models to test significant variables in combination. A variable was not considered in subsequent models if the criteria for importance were not met. As AICs allow only for relative model comparison, we also assessed the overall model fit using the marginal and conditional  $R^2$  value, or the proportion of the variance explained by only fixed effects and fixed plus random effects, respectively (Nakagawa & Schielzeth 2013), using the R package 'MuMIn' v.1.40.4 (Barto 2013). Increases in marginal and conditional  $R^2$  values suggest improved model fit, though these were not applied as selection criteria. To visualize and evaluate individual variable importance, we assessed the linear fit of each group proportion regressed against the log-transformed parameters determined to be important through the above modelling approach. To investigate possible cohort effects, we applied our forward selection modelling approach on all data combined, as well as on each age group (YoY, 1+, and 2+ parr) separately.

## RESULTS

Sample sizes ranged from 28 to 235 ind. (mean = 96.9) site<sup>-1</sup> (Table 1). The genotype completion rate (the percentage of loci across all samples that yielded a genotype, excluding samples with >9 genotypes that were not called) was 99.8%. The genotype error rate was calculated to be 0.13% (Pompanon et al. 2005). Across all 6 genotype classes, 90% of individuals were assigned above the posterior probability threshold of 0.8, with 96% of individuals reaching this threshold when hybrid classes ( $F_1$ ,  $F_2$ , BCW and

Table 1. Site locations (in decimal degrees), sample sizes (N) and  $\chi^2$  statistics of intra-river group composition of juvenile Atlantic salmon from 9 rivers in southern Newfoundland, Canada. Axial length: length of river along the down-valley axis.  
\* $p < 0.05$ ; \*\* $p < 0.001$

River	Code	Site	Latitude (° N)	Longitude (° W)	N	$\chi^2$	Axial length (km)	Elevation (m)	Distance (km)	Obstacles
Bay Du Nord	BDN	1	47.744	55.438	45	<1	67.57	1	1.7	1
		2	47.781	55.445	123			11	6.18	1
Conne River	CNR	1	47.923	55.681	110	<1	45.21	5	1.7	1
		2	47.978	55.606	108			65	11	1
		3	48.005	55.599	137			75	14.3	2
		4	48.087	55.568	102			139	24.4	2
		6	48.166	55.483	71			176	43	3
Garnish River	GAR	1	47.22	55.335	235	27.4**	38.61	5	1.27	1
		3	47.223	55.192	109			23	14.9	1
		4	47.235	55.197	82			26	18.7	1
		5	47.2	55.241	92			19	9.44	1
		6	47.262	55.059	149			104	32.4	1
Little River	LTR	1	47.851	55.695	102	2.4	36.36	3	0.9	1
		2	47.857	55.685	119			10	2.1	1
		3	47.86	55.679	120			14	2.69	1
Long Harbour River	LHR	1	47.825	54.944	122	116.97**	52.29	15	1.9	1
		2	47.862	54.931	49			34	6.22	1
		4	47.919	54.916	61			96	13.4	2
		5	47.934	54.852	76			139	24.4	3
		6	47.899	54.945	122			80	10.9	3
Northeast River	NEB	1	47.732	55.359	106	2.7	25.74	4	0.27	2
		2	47.746	55.343	116			12	2.55	2
		3	47.761	55.334	94			33	4.48	2
Northwest River	NWR	1	47.749	55.395	88	<1	18.02	21	1.48	5
		2	47.75	55.391	42			20	1.78	5
Taylor Bay Brook	TBB	1	47.557	55.643	28	66.24**	8.36	2	0.24	1
		2	47.561	55.645	153			23	0.71	1
		3	47.569	55.647	51			40	1.82	2
Terrenceville River	TEB	1	47.676	54.693	89	40.15**	16.41	1	0.4	1
		2A	47.691	54.671	38			22	3.1	1
		2B	47.69	54.67	90			10	2.88	1
		3	47.683	54.68	53			6	1.78	1
		4	47.7	54.651	116			23	4.72	1
Total					3198					

BCF) were combined into a single group (Fig. 2). Hybrids (at least 1 of the 4 hybrid classes represented) were detected in all 9 rivers, in 69.7% of all sites. As well, feral parr (i.e. the offspring of escaped farmed parents) were found in 5 rivers and 39.4% of all sites (Fig. 3). The proportion of feral individuals within a given site ranged from 0 to 0.38 (mean = 0.06), with hybrid proportions ranging from 0 to 0.76 (mean = 0.15). The proportion of pure wild individuals ranged from 0 to 1 (mean = 0.79). Chi-squared tests indicated significant differences in group composition across rivers ( $\chi^2 = 443.06$ ,  $p < 0.001$ ). Four of the 9 rivers tested showed significant evidence of intra-river (inter-site) variability in group composition (Table 1). We opted to include all rivers in subse-

quent analyses to avoid biasing trends in intra-river variation.

Individual classification scores aligned near-perfectly with PCA plots (Fig. 4), suggesting that most detectable structure in these populations is dominated by farmed introgression. The first and second PC explained 11 and 4.6% of the variance, respectively, with sites aligning along this axis by hybrid class. Clustering of sampled individuals identified as either pure wild or feral with their respective baseline groups shows support for classification accuracy and demonstrates congruency between true genetic signals of hybridization and detectable structure in sampled individuals. Individuals that were not assigned above the posterior probability threshold generally



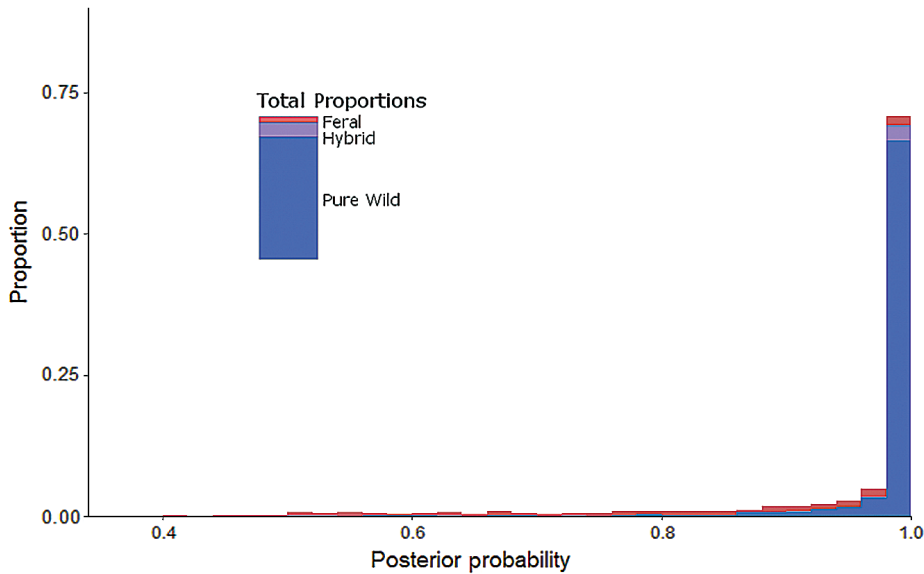


Fig. 2. Frequency distribution of posterior probability (PP) from NEWHYBRIDS classification. Colours indicate assigned group with all Atlantic salmon hybrid classes ( $F_1$ ,  $F_2$ , backcross wild [BCW], and backcross farmed [BCF]) combined into a single group. Embedded panel shows overall class proportions of individuals over the PP threshold, 0.8

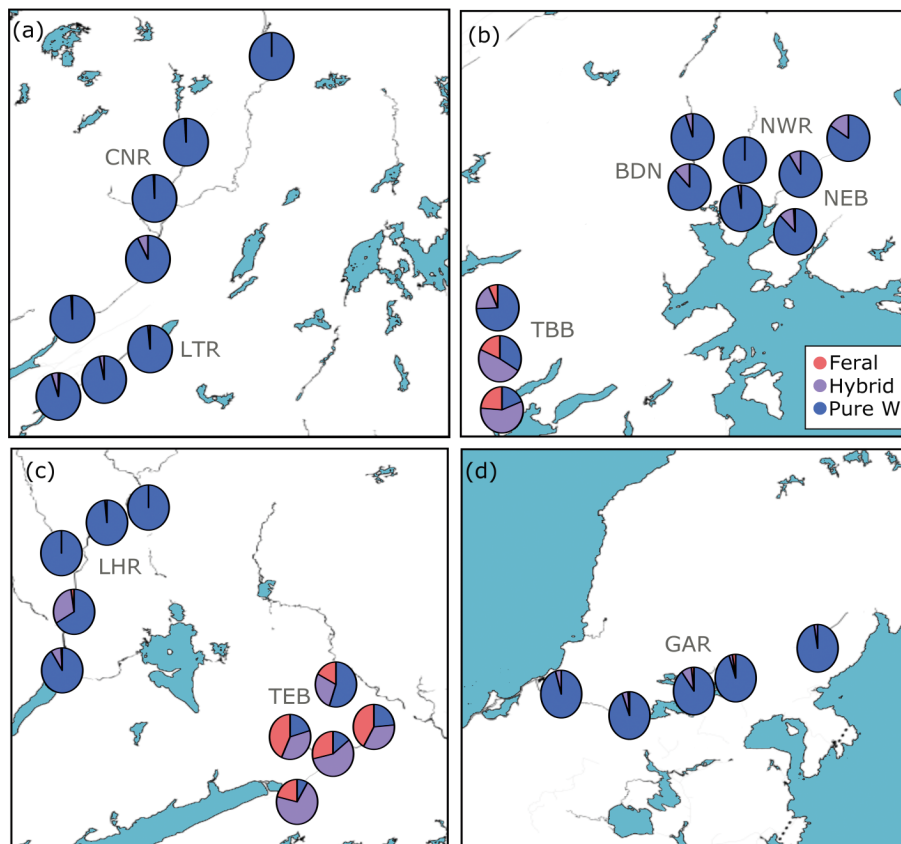


Fig. 3. Population composition of Atlantic salmon parr as feral, hybrid, or pure wild at all 35 sites across 9 rivers in southern Newfoundland, Canada. Panels correspond to geographic areas designated in Fig. 1. Abbreviations are listed in Table 1. Points have been slightly shifted for visibility of pie graphs; refer to Fig. 1

clustered between discrete classes, indicating areas of difficulty for the classification algorithm.

In the mixed modelling with all data combined, the grouping variable alone (river) resulted in  $R^2$  values of 0.39, 0.26, and 0.08 for pure wild, hybrid, and feral models, respectively. As fixed effects explained more variance in the data in subsequent models, the river contribution (difference between the conditional and marginal  $R^2$ ) was reduced in models that included fixed effects (Table 2). The most significant single fixed term in all 3 models was MME, the interaction of elevation and upstream migration distance. The inclusion of axial length improved model fit across all models, both when included as the only fixed term and in combination with other factors. The number of obstructions (also related to migration effort) was important in all models when included as the sole fixed factor and was significant in both wild and hybrid models in combination with other significant variables. Whether or not the site was within a tributary or along the main river stem was important for wild and hybrid models when included as the only fixed factor, but not in combination with other variables. When assessed independently using a  $t$ -test, there was no significant difference in class proportion between sites within a tributary or along the main stem in any of the classes (Fig. S3 in the Supplement). For all second level models (models 11 to 15; Table 2), the reduction in AIC and contribution to the fit of the model was minimal relative to the inclusion of the MME term. The final best model for the wild and hybrid analyses included MME, the number of obstructions along a migratory route, and axial length. The best model for the feral analysis included MME and axial length (Table 2). Applying modelling approaches to each age

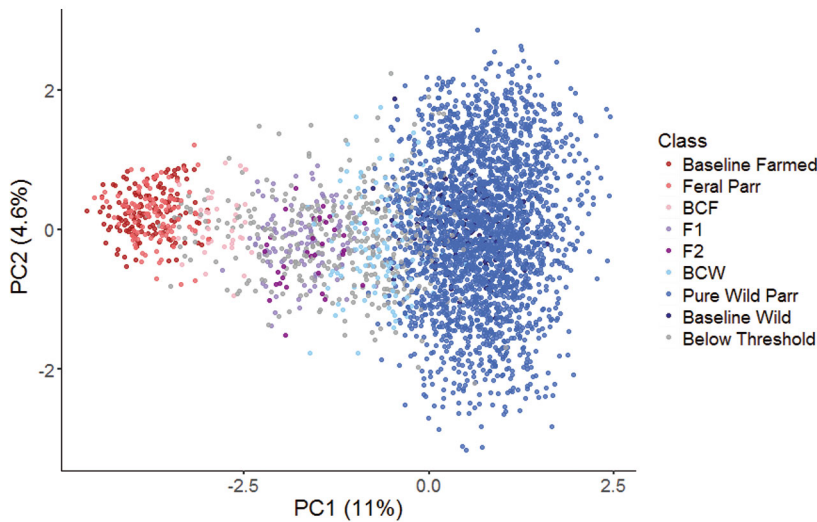


Fig. 4. Principal components analysis (PCA) of all sampled and baseline Atlantic salmon individuals, with axes indicating proportion of variance explained by each PC axis. Individuals are coloured according to classification in NEWHYBRIDS (Anderson & Thompson 2002), or, in the case of baseline individuals, known baseline class (see Fig. S5 in the Supplement for clustering by river of origin)

class independently revealed similar results for 1+ and 2+-only models, in which the single most significant term was MME (Table S1 & Fig. S4). The best models for all 1+ analyses included MME, number of obstructions, axial length, and the tributary term. Axial length was not a significant term for 2+-only models. The best YoY models included the number of obstructions and axial length; however, only elevation (not distance) was a significant factor in these models. Additionally, due to the very low proportion of feral

individuals assigned above the posterior probability threshold in YoY samples, the feral model could not be run on this age class alone due to insufficient variance.

Inspection of linear regressions of class proportion against log-transformed variables (Fig. 5) revealed decreasing frequency of fish of hybrid and pure farmed origin (and conversely, increasing frequency of wild fish) with increasing axial length ( $r^2 = 0.161$  to  $0.405$ ,  $p < 0.00001$ ). Similar trends were found for MME and the number of obstructions (MME:  $r^2 = 0.08$  to  $0.26$ ,  $p \leq 0.01$ ; obstructions:  $r^2 = 0.052$  to  $0.104$ ,  $p < 0.05$ , respectively). In all cases, the feral model was the least significant. For all environmental variables analyzed using linear regression, the magnitude of the slope and  $R^2$  was greatest in pure wild analyses and lowest in feral class proportions (Fig. 5).

## DISCUSSION

Despite increasing production of farmed Atlantic salmon *Salmo salar* in North America (ICES 2017) and extensive reports of escapees (Keyser et al. 2018), factors influencing the distribution of hybrid-

Table 2. Binomial mixed models (model number plus additional terms) and associated change in Akaike's information criteria ( $\Delta AIC$ ) from base (numbered) model, with marginal  $R^2$  ( $mR^2$ ) and conditional  $R^2$  ( $cR^2$ ). Asterisks indicate models in which additional parameters were statistically significant. **Bold** AICs indicate models with best fit (lowest AIC with significant additional term). Age-specific model results can be found in Table S1 in the Supplement at [www.int-res.com/articles/suppl/q010p401\\_suppl.pdf](http://www.int-res.com/articles/suppl/q010p401_suppl.pdf). \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$

Number	Model (No. + term)	$\Delta AIC$ ( $mR^2$ , $cR^2$ )		
		Wild	Hybrid	Feral
1	Group (river)	*** (–, 0.39)	*** (–, 0.26)	*** (–, 0.08)
2	1 + Propagule pressure	–1.3 (0.01, 0.39)	–1.7 (0.0, 0.26)	0 (0.10, 0.34)
3	1 + Basin area	–0.7 (0.05, 0.38)	–0.9 (0.03, 0.25)	1.2 (0.12, 0.30)
4	1 + Mean width	–0.9 (0.06, 0.39)	–0.7 (0.06, 0.26)	0.2 (0.07, 0.30)
5	1 + Perimeter	–0.2 (0.05, 0.38)	–0.2 (0.04, 0.25)	1.4 (0.14, 0.27)
6	1 + Axial length	1.8* (0.10, 0.37)	1.7* (0.07, 0.24)	2.4* (0.19, 0.28)
7	1 + Tributary	24.8*** (0.01, 0.40)	27.6*** (0.01, 0.26)	–1.4 (0.0, 0.08)
8	1 + Obstructions	237.1*** (0.24, 0.57)	225.7*** (0.21, 0.46)	3.8* (0.10, 0.34)
9	1 + Elevation	256.7*** (0.18, 0.43)	211.6*** (0.20, 0.32)	17.5*** (0.12, 0.39)
10	1 + Distance	107*** (0.14, 0.41)	118.9*** (0.17, 0.30)	–1.6 (0.06, 0.28)
11	1 + Elevation $\times$ distance	287*** (0.18, 0.43)	215.9*** (0.20, 0.31)	48.3*** (0.76, 0.85)
12	11 + Obstructions	17.2*** (0.24, 0.46)	38.3*** (0.25, 0.34)	–0.2 (0.79, 0.86)
13	11 + Axial length	2.4* (0.30, 0.46)	1.8* (0.29, 0.34)	<b>3.6** (0.69, 0.75)</b>
14	11 + Tributary	–1.4 (0.17, 0.43)	–1.2 (0.20, 0.31)	1.2 (0.70, 0.79)
15	12 + Axial length	<b>4.9** (0.42, 0.50)</b>	<b>2.7* (0.38, 0.39)</b>	2.1 (0.87, 0.87)

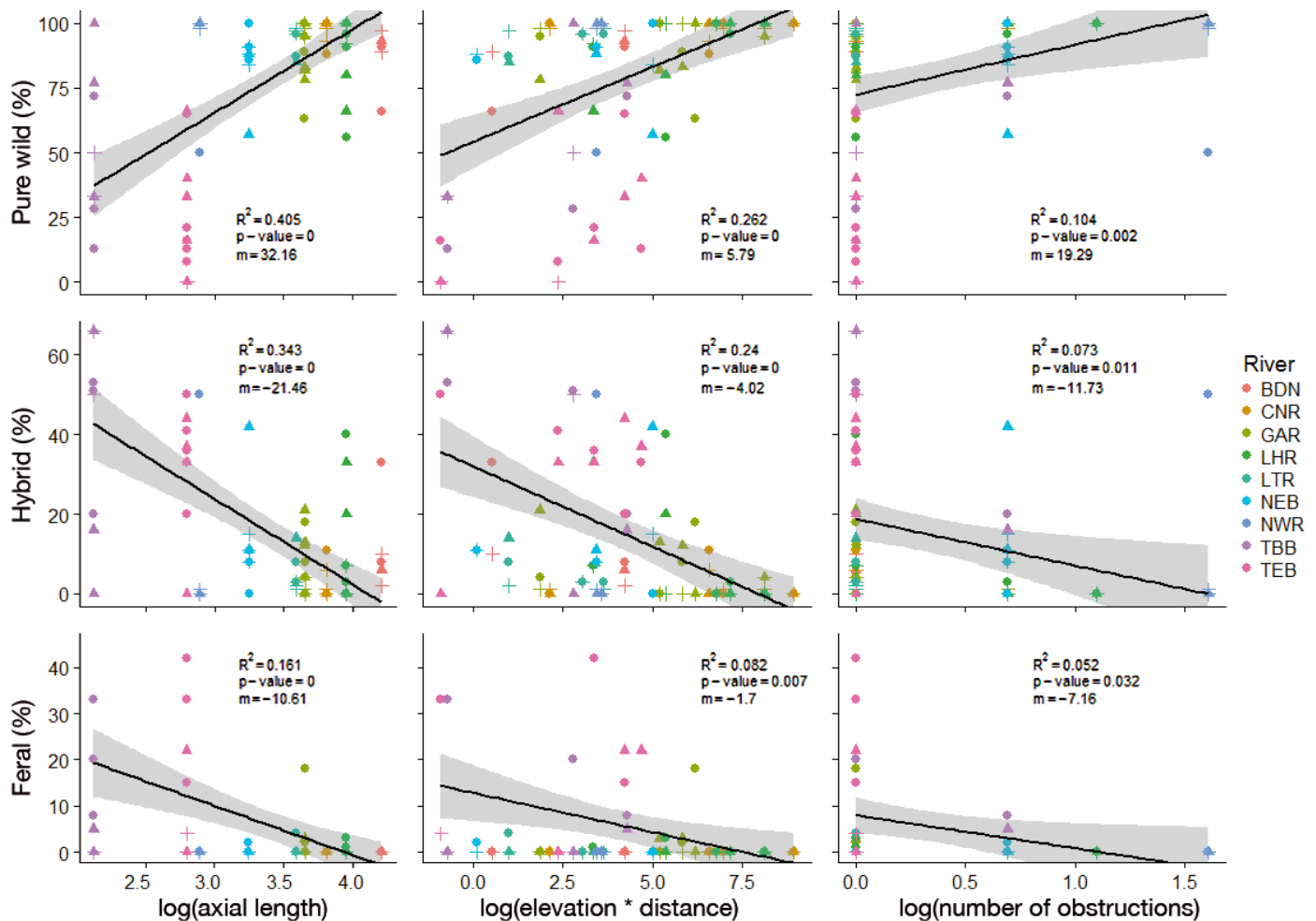


Fig. 5. Linear regressions of Atlantic salmon group composition against log-transformed variables: axial length, minimum migration effort (as estimated by the product of elevation and upstream migration distance), and the number of obstructions along a migratory route to a given site. Each point represents an age class per sampling site within a river. Crosses: young-of-year; circles: 1+ parr; triangles: 2+ parr. Sampling sites are indicated by colour. Shaded grey areas indicate standard error of the regression, with R<sup>2</sup>, p-values, and slope (m) shown for each regression analysis. River abbreviations are as in Table 1, corresponding to colours in Fig. 1

ization between wild and escaped farmed salmon remain poorly resolved (Wringe et al. 2018b). Here, we assessed spatial variation in hybridization and landscape variables influencing fine-scale distribution of hybridization across and within 9 rivers in southern Newfoundland, Canada. Our analysis indicated that inter-river distribution was largely driven by river or population size (axial length), while intra-river distribution was determined by factors influencing migration effort required to reach a given spawning site. This work builds on previous evidence of hybridization occurring in North America (Bourret et al. 2011, Keyser et al. 2018, Wringe et al. 2018a), and extends research conducted elsewhere exploring factors influencing the distribution

of escaped farmed salmon into wild populations (Webb et al. 1991, Carr et al. 1997, Moe et al. 2016). We provide novel insights into fine-scale, within-river spatial variation in levels of hybridization following escape events, demonstrating that population composition may be associated with landscape features influencing individual performance or migratory motivation. The resolution of this variability will ultimately be essential to evaluating impacts from escaped farmed salmon on wild Atlantic salmon populations.

Overall, we observed evidence of hybridization or introgression occurring within at least 1 site in all 9 rivers, with chi-squared results suggesting that variation among rivers was stronger than variation within



rivers. It is worth noting that our estimates of hybridization rates are also likely underestimated, as individuals that did not meet the posterior probability threshold in NEWHYBRIDS (and thus were not included in downstream analysis) often clustered near other hybrids or in between clusters in both principal component and STRUCTURE analyses. Also, rivers included in this study were selected to observe hybrid distribution following a nearby escape event, based on previous evidence of hybridization occurring in the area (Wringe et al. 2018a). As such, our estimates do not reflect rates of introgression across the broader region but may provide accurate estimates of introgression within rivers likely to be affected by aquaculture. Relative class proportions identified in the present study reflect that of previous estimates (Wringe et al. 2018a) and are congruent with evidence of overall farmed invasion in North American rivers (Morris et al. 2008). That feral parr and hybrid parr were present in the YoY age class in somewhat lower proportions than 1+ and 2+ individuals may suggest a decreased presence of escapees spawning in 2016 relative to previous years reflected in the 1+ and 2+ age class or reduced survivorship of feral and hybrid parr relative to their wild counterparts (McGinnity et al. 2003). As farmed escapee presence in rivers is expected to decrease following the initial surge from a large escape event, this is not surprising; however, the presence of second generation hybrids in all age classes indicates that trickle escapes are ongoing and influence the genetic composition of wild populations.

Of the variables tested that were invariant within rivers (i.e. variables that were identical for each site within a river and varied only between rivers), and therefore potentially influencing inter-river distribution, only axial length was found to be significant. As an estimate of habitat size, axial length is reflective of population size in the region, which has been demonstrated to be indicative of the degree of farmed invasion in Newfoundland (Wringe et al. 2018a) and elsewhere (Heggberget et al. 1993, Økland et al. 1995). Although log-transformed linear trends were significant for all 3 classes against axial length, this relationship was most extreme (larger magnitude of slope) for hybrids compared to feral parr, resulting in higher relative proportions of hybrids in smaller rivers. The differential impact of river or population size on the proportion of feral or hybrid parr present at a given site is likely due to a dilution effect, reflecting a higher proportion of hybrids in the watershed overall. Alternatively, hybrids may demonstrate a competitive advantage over feral parr (Fleming et al.

2000), resulting in elevated presence in small, dense populations, but this inference is outside the scope of the current study.

Across all models, MME resulted in the greatest model improvements except when analyzing YoY parr alone. The number of obstacles between the river mouth and a site also improved model fit, suggesting further influence of difficulty during migration as a predictor of feral and hybrid presence at a site. Alternatively, obstacles may affect hybrid or feral presence by influencing dispersal. As escaped farmed fish often move throughout river systems more than their wild counterparts (Økland et al. 1995, Thorstad et al. 1998), the proportion of hybrid fish in spawning sites blocked by fewer passable obstacles may be underestimated at a given site relative to more obstructed areas. Linear regression analyses provide additional evidence that increased difficulty along a migration route contributes to the reduced hybrid proportions in upstream areas (Webb et al. 1991, Økland et al. 1995, Moe et al. 2016). Linear regression of feral proportions was the least significant with lowest  $R^2$  values across tests, likely because feral parr were present in low proportions overall, thereby providing little variance across sites with which to detect trends. While there are other factors that affect the effort required for an individual to reach a given spawning site (e.g. flow velocity, flow volume), our findings that the combination of distance and elevation (i.e. MME) is predictive of intra-river proportions of hybrid and wild parr strongly implicates the role of migration effort. Further, the interaction of elevation and distance along a migration route has been demonstrated to affect habitat choice in Arctic char (Moore et al. 2017), supporting its implementation in the present study. It is possible that spatial variation in fitness or survival, rather than landscape effects on migration performance, may contribute to variance in observed class proportion across sites. Regardless of the mechanism involved however, within-river variation observed in the present study demonstrates that sampling from multiple sites within rivers is likely to improve estimates in future work of the frequency of hybridization and introgression between wild Atlantic salmon and farmed escapees.

Propagule pressure was not found to be a significant factor in mixed models, despite existing evidence demonstrating elevated farmed escapees in proportion to propagule pressure in North American Atlantic salmon (Keyser et al. 2018), as well as other systems of interbreeding between farmed and wild individuals (e.g. mussels: Crego-Prieto et al. 2015,

and mink: Beauclerc et al. 2013, Bowman et al. 2017). The lack of correlation in this study is likely due to the relatively small geographic area observed and the targeting of rivers that were previously known to contain farmed escapees (Wringe et al. 2018a), such that propagule pressure did not sufficiently vary to detect an influence on hybrid distribution. Additionally, as propagule pressure influences actual numbers of farmed escapees entering river systems rather than proportions of resultant offspring, analyzing class proportions may obscure results due to differing wild population size across rivers.

Efforts to preserve genetic integrity and conserve wild Atlantic salmon populations are best directed towards reducing aquaculture escape events (Hindar et al. 2006); however, the quantity of escapees that continue to enter wild systems every year (ICES 2017) necessitates attention to reducing the impacts of invading farmed salmon. We have demonstrated that smaller wild populations are most at risk of detrimental genetic effects from farmed introgression due to increased presence of hybrid and feral parr in smaller rivers. Increased MME due to elevation and upstream migration distance and the number of obstructions to a spawning site reduces the presence of hybrid offspring, suggesting that wild salmon spawning in downstream sites are likely faced with higher competition from farmed invaders and elevated outbreeding depression compared to those in upstream spawning sites. Knowledge of the distribution of introgression can inform management practices to help maintain genetic integrity of highly vulnerable populations. Our work extends previous research that focused on the presence of farmed salmon mating with wild populations in Atlantic Canada (Bourret et al. 2011, Keyser et al. 2018, Wringe et al. 2018a) and provides novel insight into the forces that influence introgression, applicable to affected systems globally.

**Acknowledgements.** The authors thank the staff of the Newfoundland DFO Salmonids section for juvenile tissue sampling, as well as Freya Keyser and Ryan Stanley for assistance in analysis. 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.

## LITERATURE CITED

- ✦ Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Automat Contr* 19:716–723
- ✦ Anderson EC, Thompson EA (2002) A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* 160:1217–1229
- ✦ Barto K (2013) MuMIn: model selection and model averaging based on information criteria. <http://cran.r-project.org/web/packages/MuMIn/index.html> (accessed 13 Dec 2017)
- ✦ Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. *J Stat Softw* 67:1–48
- ✦ Beauclerc KB, Bowman J, Schulte-Hostedde AI (2013) Assessing the cryptic invasion of a domestic conspecific: American mink in their native range. *Ecol Evol* 3: 2296–2309
- ✦ Bolstad GH, Hindar K, Robertsen G, Jonsson B and others (2017) Gene flow from domesticated escapes alters the life history of wild Atlantic salmon. *Nat Ecol Evol* 1:0124
- ✦ Bourret V, O'Reilly PT, Carr JW, Berg PR, Bernatchez L (2011) 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
- ✦ Bowman J, Beauclerc K, Farid AH, Fenton H, Klütsch CFC, Schulte-Hostedde AI (2017) Hybridization of domestic mink with wild American mink (*Neovison vison*) in eastern Canada. *Can J Zool* 95:443–451
- ✦ Carr JW, Lacroix GL, Anderson JM, Dilworth T (1997) Movements of non-maturing cultured Atlantic salmon (*Salmo salar*) in a Canadian river. *ICES J Mar Sci* 54: 1082–1085
- ✦ Crego-Prieto V, Ardura A, Juanes F, Roca A, Taylor JS, Garcia-Vazquez E (2015) Aquaculture and the spread of introduced mussel genes in British Columbia. *Biol Invasions* 17:2011–2026
- ✦ Einum S, Fleming IA (1997) Genetic divergence and interactions in the wild among native, farmed and hybrid Atlantic salmon. *J Fish Biol* 50:634–651
- ✦ Fleming IA, Hindar K, Mjølnerød IB, Jonsson B, Balstad T, Lamberg A (2000) Lifetime success and interactions of farm salmon invading a native population. *Proc R Soc Lond B* 267:1517–1523
- ✦ Gausen D, Moen V (1991) Large-scale escapes of farmed Atlantic salmon (*Salmo salar*) into Norwegian rivers threaten natural populations. *Can J Fish Aquat Sci* 48:426–428
- ✦ Glover KA, Quintela M, Wennevik V, Besnier F, Sørvik AGE, Skaala Ø (2012) Three decades of farmed escapees in the wild: a spatio-temporal analysis of Atlantic salmon population genetic structure throughout Norway. *PLOS ONE* 7:e43129
- ✦ Glover KA, Pertoldi C, Besnier F, Wennevik V, Kent M, Skaala Ø (2013) Atlantic salmon populations invaded by farmed escapees: quantifying genetic introgression with a Bayesian approach and SNPs. *BMC Genet* 14:74
- ✦ Glover KA, Solberg MF, McGinnity P, Hindar K and others (2017) Half a century of genetic interaction between farmed and wild Atlantic salmon: status of knowledge and unanswered questions. *Fish Fish* 18:890–927
- ✦ Hansen LP, Windsor ML (2006) Interactions between aquaculture and wild stocks of Atlantic salmon and other diadromous fish species: science and management, challenges and solutions: an introduction by the Conveners. *ICES J Mar Sci* 63:1159–1161

- Harvey AC, Glover KA, Taylor MI, Creer S, Carvalho GR (2016) A common garden design reveals population-specific variability in potential impacts of hybridization between populations of farmed and wild Atlantic salmon, *Salmo salar* L. *Evol Appl* 9:435–449
- Heggberget TG, Økland F, Ugedal O (1993) Distribution and migratory behaviour of adult wild and farmed Atlantic salmon (*Salmo salar*) during return migration. *Aquaculture* 118:73–83
- Heino M, Svåsand T, Wennevik V, Glover K (2015) Genetic introgression of farmed salmon in native populations: quantifying the relative influence of population size and frequency of escapees. *Aquacult Environ Interact* 6:185–190
- Hindar K, Fleming IA, McGinnity P, Diserud O (2006) Genetic and ecological effects of salmon farming on wild salmon: modelling from experimental results. *ICES J Mar Sci* 63:1234–1247
- ICES (2017) Report of the Working Group on North Atlantic Salmon (WGNAS). ICES CM 2017/ACOM:20. ICES, Copenhagen
- Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403–1405
- Karlsson S, Moen T, Lien S, Glover KA, Hindar K (2011) Generic genetic differences between farmed and wild Atlantic salmon identified from a 7K SNP-chip. *Mol Ecol Resour* 11:247–253
- Karlsson S, Diserud OH, Moen T, Kindar K (2014) A standardized method of quantifying unidirectional genetic introgression. *Ecol Evol* 4:3256–3263
- Karlsson S, Diserud OH, Fiske P, Hindar K (2016) Widespread genetic introgression of escaped farmed Atlantic salmon in wild salmon populations. *ICES J Mar Sci* 73: 2488–2498
- Keyser F, Wringe BF, Jeffery NW, Dempson JB, Duffy S, Bradbury IR (2018) Predicting the impacts of escaped farmed Atlantic salmon on wild salmon populations. *Can J Fish Aquat Sci* 75:506–512
- McGinnity P, Prodöhl P, Ferguson A, Hynes R and others (2003) 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 B* 270:2443–2450
- Moe K, Næsje TF, Haugen TO, Ulvan EM, Aronsen T, Sandnes T, Thorstad EB (2016) Area use and movement patterns of wild and escaped farmed Atlantic salmon before and during spawning in a large Norwegian river. *Aquacult Environ Interact* 8:77–88
- Moore JS, Harris LN, Le Luyer J, Sutherland BJG and others (2017) Genomics and telemetry suggest a role for migration harshness in determining overwintering habitat choice, but not gene flow, in anadromous Arctic char. *Mol Ecol* 26:6784–6800
- Morris MRJ, Fraser DJ, Heggelin AJ, Whoriskey FG, Carr JW, O'Neil SF, Hutchings JA (2008) Prevalence and recurrence of escaped farmed Atlantic salmon (*Salmo salar*) in eastern North American rivers. *Can J Fish Aquat Sci* 65:2807–2826
- Nakagawa S, Schielzeth H (2013) A general and simple method for obtaining  $R^2$  from generalized linear mixed-effects models. *Methods Ecol Evol* 4:133–142
- Økland F, Heggberget TG, Jonsson B (1995) Migratory behaviour of wild and farmed Atlantic salmon (*Salmo salar*) during spawning. *J Fish Biol* 46:1–7
- Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: causes, consequences and solutions. *Nat Rev Genet* 6:847–859
- Porter T, Riche L, Traverse G (1974) Catalogue of rivers in insular Newfoundland, Vol B. Resource Development Branch, Newfoundland Region. Fisheries and Marine Science, Department of the Environment, St. John's
- R Core Development Team (2008) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Skaala Ø, Wennevik V, Glover KA (2006) Evidence of temporal genetic change in wild Atlantic salmon, *Salmo salar* L., populations affected by farm escapees. *ICES J Mar Sci* 63:1224–1233
- Skaala Ø, Glover KA, Barlaup BBT, Svåsand T, Besnier F, Hansen MM, Borgstrøm R (2012) Performance of farmed, hybrid, and wild Atlantic salmon (*Salmo salar*) families in a natural river environment. *Can J Fish Aquat Sci* 69: 1994–2006
- Thorstad E, Heggberget T, Økland F (1998) Migratory behaviour of adult wild and escaped farmed Atlantic salmon, *Salmo salar* L., before, during and after spawning in a Norwegian river. *Aquacult Res* 29:419–428
- Webb JH, Hay DW, Cunningham PD, Youngson AF (1991) The spawning behaviour of escaped farmed and wild adult Atlantic salmon (*Salmo salar* L.) in a northern Scottish river. *Aquaculture* 98:97–110
- Wickham H (2009) ggplot2: elegant graphics for data analysis. Springer-Verlag, New York, NY
- Wringe BF, Stanley RRE, Jeffery NW, Anderson EC, Bradbury IR (2017a) hybriddetective: a workflow and package to facilitate the detection of hybridization using genomic data in R. *Mol Ecol Resour* 17:e275–e284
- Wringe BF, Stanley RRE, Jeffery NW, Anderson EC, Bradbury IR (2017b) parallelnewhybrid: an R package for the parallelization of hybrid detection using NEWHYBRIDS. *Mol Ecol Resour* 17:91–95
- Wringe BF, Jeffery NW, Stanley RRE, Hamilton LC and others (2018a) Extensive hybridization following a large escape of domesticated Atlantic salmon in the Northwest Atlantic. *Commun Biol* 1:108
- Wringe BF, Anderson EC, Jeffery NR, Stanley RRE, Bradbury IR (2018b) Development and evaluation of SNP panels for the detection of hybridization between wild and escaped Atlantic salmon (*Salmo salar*) in the West Atlantic. *Can J Fish Aquat Sci*, doi:10.1139/cjfas-2017-0394