Supplementary Material

Supplementary Data for the Paper: Signatures of Selection on Growth, Shape, Parr Marks, and SNPs among seven Canadian Atlantic Salmon (*Salmo salar*) Populations

M. Culling, H. Freamo, K. Patterson, P. Berg, S. Lien and E.G. Boulding

TABLE OF CONTENTS: APPENDIX A: DETAILED METHODS AND RESULTS FOR GEOMETRIC MORPHOMETRICS ANALYSIS USING ALIGNED CO-ORDINATES.

APPENDIX A: DETAILED METHODS FOR GEOMETRIC MORPHOMETRICS ANALYSES USING ALIGNED CO-ORDINATES

Morphological Data Acquisition

Landmark based geometric morphometrics were used to acquire shape data. Fifteen landmarks (see Fig. **2** in main paper) were digitized on each image using TPSDIG 1.37 [33]. The power of aligned landmark and geometric morphometric analyses increases proportionally to the number of landmarks for a given number of specimens [34], therefore 15 consistently repeatable landmarks were chosen to capture the overall body shape. Furthermore, an important factor in the selection of landmarks was that they should encapsulate morphological traits that have shown variation in earlier studies of Atlantic salmon [10].

We used the program TPSRELW 1.31 [35] which uses generalized orthogonal least squares Procrustes procedures [36], to align, scale and rotate the landmark configurations to superimpose the digitized landmarks. The shape variables used in the statistical analyses were aligned landmarks and partial warps acquired from the superimposed specimens using TPSRELW 1.31 [35]. The aligned TPS co-ordinates were used to estimate the centroid size for each fish using TPSREGR 1.31 [37]. Centroid size is deemed to be a reliable size measure theoretically [38], and empirically, based on earlier fish studies, which have found strong correlations between centroid size and standard length [39, 84, 85].

Discriminant Function Analyses of Aligned Landmarks

In the main paper we investigated subtle patterns of differentiation using five populations but here we use all seven populations to investigate patterns of differentiation by including the CHA and STC populations (Fig. **2S**). The CHA and STC populations were not included in the main paper because the former is so differentiated it masks subtle differences between the other populations and the latter population sample was largely hybrid in origin because 60% of the families the progeny were derived from were the result of crosses between STC males and TOB females. All 15 aligned landmarks in XY co-ordinate space were used as independent variables in a discriminant function analysis (DFA) of seven wild Atlantic salmon populations, in SPSS v. 15 that resulted in six canonical variates that summarized shape variation and allowed interpretation of the relative importance on each landmark co-ordinate to the observed shape variation. We also plotted canonical variates ordinations of five (Fig. **1S**), and seven (Fig. **2S**), Atlantic salmon populations from the DFA of aligned landmark coordinates using all combinations of functions in order to be able to examine the influence of each function on population differentiation.

In order to investigate the change in orientation of each landmark in each population for all 15 landmarks, relative to the overall mean landmark for all populations, we plotted the mean position of each landmark in each population (for all 15 landmarks separately), against the overall mean for every one of the 15 landmarks (Fig. **5S**). We were then able to plot the change in each landmark per population, relative to the overall mean landmark for all populations, as a vector on a schematic illustration of an Atlantic salmon (Fig. **6S**).

Detailed Results for Geometric Morphometrics Analyses using Aligned Co-ordinates

Aligned Coordinates Morphological Variation

Shape variation from the DFA of the aligned landmarks for the seven populations showed that canonical variate 1 (cv1), accounting for 76.9% of the total variation expressed an increase in body depth, and elongation of the pectoral fin and the posterior end of the fish (Table **2S**). We also carried out these analyses with the CHA and STC populations removed with broadly similar results (main paper Table **2**, Table **1S**).

Seven Populations

The finding of significant differences ($\Lambda = 0.011$, χ^2 ($df = 156$, $N = 242$) = 1013.7, $p < 0.001$) among populations of juvenile Atlantic salmon (242 individuals) from different natal rivers, in their overall body shape, as defined by aligned landmarks, suggests a relationship between water velocity and body shape in salmonids that may represent an adaptive response to water flow (Tables **2S**, **7S**). Furthermore, these individuals could be accurately classified into the seven populations ($\Lambda = 0.048$, χ^2 ($df = 125$, $N = 242$) = 683.8, $p < 0.001$) based upon their aligned coordinates, after removing the effects associated with the first discriminant function (Table **5S**). Six discriminant functions were interpreted (Table **2S**) because of the significance of both multivariate and univariate tests. Differences among the seven population groups accounted for 76.9% of the variability on the first discriminant function, and fell to 17.7% for the sixth and final discriminant function (Table **5S**). Correct group membership was determined from 91.7% to 51.6% of the time for all seven populations (Table **6S**). Overall, 65.3% cross-validated grouped cases were correctly classified.

Having potentially established a relationship between natal river and body shape in our study populations that could represent an adaptive response to water flow, we went on to investigate the location and distribution of these potentially adaptive changes in body shape. Shape variation from the DFA of the aligned landmarks was investigated for seven populations as well as five populations (see main paper) to extract further patterns of differentiation. The analyses revealed that morphological variation largely corresponded to the five population analyses with variation in the anterior and posterior regions (Fig. **2Sa**). Canonical variate 1 (cv1), accounted for 76.9% of the total variation, expressed an increase in body depth, and elongation of the pectoral fin and the posterior end of the fish (Fig. **2S**; Table **5S**). Significant pairwise differences between populations on cv1 are summarized in (Table **7S**). This morphological variation, significant in only 28.6% of the pairwise population comparisons (Table **7S**), was significant for 66.6% of pairwise comparisons involving the CHA population). Increasing length of the posterior end of the fish, a shortening and narrowing of the pectoral fin, and an increase in body depth and snout height are represented by canonical variate 2 (cv2), and accounted for 69.4% of variation across body form (Fig. **2S**). The morphological variation represented by this variate was significant in 85% of the pair-wise population comparisons. An increase in the length but a decrease in the height of the ventral trunk, elongation and broadening of the pectoral fin and a deepening of the general body are represented by canonical score 3, and accounted for 55.2% of the variation across body form. This morphological variation was significant in 85% of the pair-wise population comparisons (Table **7S**). Canonical variates 4 and 5 showed similar body form variation to cv2, while canonical variate 6 was similar to cv1.

Examination of the graphs for all the functions from the DFA of aligned landmark coordinates plotted against each other, for the seven populations, reveals that function 1 against function 3 allows differentiation of the seven populations on the y axis. Function 1 expresses an increase in body depth, and elongation of the pectoral fin and the posterior end of the fish while function 3 accounts for an increase in the length but a decrease in the height of the ventral trunk, elongation and broadening of the pectoral fin and a deepening of the general body form. On the x axis differentiation of the populations is revealed in a plot of function 1 against function 6 which encompasses very similar morphological variation to function 1 (Fig. **2S**). When plotting the all combinations of the five population DFA analyses, function 2 against function 3 allows differentiation of the populations on both axes while function 2 against function 4 separates the populations on the x axis only (Fig. **1S**). Function 2 expressed a decreasing body depth, shortening and a narrowing of the pectoral fin, while function 3 showed an increasing length of the posterior of the fish, increasing body depth, shortening and narrowing of the pectoral fin and a lowering of the snout. Function 4 encompassed a decreasing length of the posterior of the fish, increasing body depth, lengthening and broadening of the pectoral fin and lowering of the snout.

Table 1S. Standardized Discriminant Function Coefficients from an Aligned Landmark Discriminant Analysis of Five Populations for 15 External Morphological Landmarks Characteristic for Wild Atlantic Salmon (*Salmo salar*)

Landmark		Function			
	Order	1	2	3	4
Tip of snout	1X	0.585	-3.207	-1.779	-1.764
Tip of snout	1Y	0.695	-1.111	-0.609	-1.300
Anterior beginning of lateral line	2X	0.823	0.804	0.424	-1.110
Anterior beginning of lateral line	2Y	0.276	-0.303	0.727	0.835
Anterior insertion of dorsal fin	3X	1.913	3.659	0.853	-2.808
Anterior insertion of dorsal fin	3Y	0.819	1.781	1.563	2.676
Anterior insertion of adipose fin	4X	3.974	9.525	2.731	-3.046
Anterior insertion of adipose fin	4Y	-1.124	0.978	3.384	5.261
Dorsal insertion of caudal fin	5X	4.064	9.911	3.483	-2.964
Dorsal insertion of caudal fin	5Y	-0.730	0.470	1.443	3.732
Posterior mid-point of caudal peduncle	6X	2.329	6.089	2.184	-1.173
Posterior mid-point of caudal peduncle	6Y	-0.624	0.428	2.227	4.144
Fork point in caudal peduncle	7X	4.114	11.447	3.571	-3.028
Fork point in caudal peduncle	7Y	-1.514	1.794	6.311	12.631
Ventral insertion of caudal fin	8X	3.396	9.903	3.025	-1.002
Ventral insertion of caudal fin	8Y	-0.767	0.653	2.013	4.549
Anterior insertion of anal fin	9X	3.056	7.853	3.184	-0.928
Anterior insertion of anal fin	9Y	-0.493	0.506	1.372	4.570
Anterior insertion of pelvic fin	10X	2.254	5.242	2.175	-0.464
Anterior insertion of pelvic fin	10Y	-0.825	-1.722	2.056	4.380
Posterior point of pectoral fin	11X	3.322	6.734	2.392	-1.079
Posterior point of pectoral fin	11Y	-1.147	-1.471	1.890	5.270
Ventral insertion of pectoral fin	12X	0.507	1.250	0.068	0.071
Ventral insertion of pectoral fin	12Y	-0.404	-0.785	-0.014	0.784
Dorsal insertion of pectoral fin	13X	0.193	0.230	0.542	-0.756
Operculum joins outline of body	14Y	-0.465	-1.339	0.212	0.533
Eigenvalue		3.826	1.704	0.597	0.414
% variation explained		66.60	48.40	26.50	13.98

Note: Eigenvalues from each discriminant function are listed below the column of coefficients; Variables 13Y, 14X, 15X and 15Y not used in the analyses because they failed the tolerance test.

Table 2S. Standardized Discriminant Function Coefficients From an Aligned Landmark Discriminant Analysis of Seven Populations for 15 External Morphological Landmarks Characteristic for Wild Atlantic Salmon (*Salmo salar*)

Landmarks	Order	1	2	3	4	5	6
Tip of snout	1X	-2.398	-1.981	-1.853	-1.751	-1.483	-2.444
Tip of snout	1Y	-0.916	1.125	-0.373	-1.480	-0.916	-1.362
Anterior beginning of lateral line	2X	-0.535	1.464	0.748	0.465	0.120	-1.029
Anterior beginning of lateral line	2Y	0.577	0.832	-0.815	0.059	0.573	0.672
Anterior insertion of dorsal fin	3X	-1.229	4.698	3.875	1.567	0.468	-1.741
Anterior insertion of dorsal fin	3Y	2.186	3.225	-0.030	0.984	1.097	2.611
Anterior insertion of adipose fin	4X	-0.390	10.08	7.247	4.300	1.282	-0.622
Anterior insertion of adipose fin	4Y	3.632	0.707	-1.564	2.161	2.919	4.008
Dorsal insertion of caudal fin	5X	-0.614	10.80	8.335	4.677	2.582	-0.361
Dorsal insertion of caudal fin	5Y	2.891	1.386	-0.978	0.847	1.093	3.065
Posterior mid-point of caudal peduncle	6X	-0.059	6.088	4.431	2.834	1.299	0.157
Posterior mid-point of caudal peduncle	6Y	3.297	0.901	-1.709	1.257	1.971	3.512
Fork point in caudal peduncle	7X	0.180	12.04	9.012	5.964	2.169	-0.364
Fork point in caudal peduncle	7Y	8.614	2.842	-4.067	3.642	5.780	10.443
Ventral insertion of caudal fin	8X	0.614	11.08	7.937	5.626	1.930	0.975
Ventral insertion of caudal fin	8Y	3.150	0.784	-1.425	1.279	1.735	3.617
Anterior insertion of anal fin	9X	0.468	8.611	5.806	3.975	2.174	0.649
Anterior insertion of anal fin	9Y	2.738	0.745	-1.405	0.738	1.574	4.061
Anterior insertion of pelvic fin	10X	0.262	5.862	3.644	2.594	1.369	0.545
Anterior insertion of pelvic fin	10Y	3.090	-0.390	-3.538	0.554	1.567	2.758
Posterior point of pectoral fin	11X	-0.375	8.033	4.939	3.683	1.498	0.253
Posterior point of pectoral fin	11Y	3.718	0.034	-3.312	0.143	1.599	3.835
Ventral insertion of pectoral fin	12X	-0.308	1.512	1.342	0.659	-0.184	-0.154
Ventral insertion of pectoral fin	12Y	0.562	-0.052	-0.654	-0.795	0.090	0.528
Dorsal insertion of pectoral fin	13X	0.096	0.675	0.319	-0.445	0.487	-0.486
Operculum joins outline of body	14Y	0.507	-0.813	-1.077	-0.066	-0.006	-0.009
Eigen Value		3.376	2.334	1.334	0.962	0.356	0.274
Percent of variation explained		39.1	27.0	15.4	11.2	4.1	3.2

Note: Eigenvalues from each discriminant function are listed below the column of coefficients; Variables 13Y, 14X, 15X and 15Y not used in the analyses because they failed the tolerance test.

Table 5S. Four Discriminant Functions Based on Either the Five and Seven Population Set and Using the Aligned Landmark Coordinates From Geometric Morphometrics. Percentage in Variability of the Scores that was Accounted for by Differences Among the Populations. The CHA and STC Populations were Removed from the Five Population Dataset Because the Former is so Differentiated it Masks Subtle Differences Between the Other Populations and the Latter Because the Crosses Measured here Were Largely Hybrid in Origin

	Aligned Landmark Coordinates	
Function	**5 Populations**	**7 populations**
1	66.60%	76.90%
2	48.40%	69.40%
3	26.50%	55.20%
4	13.98%	43.60%
5		25.40%
6		17.70%

Table 6S. Classification of Each Population based on Discriminant Function Analyses of Five and Seven Population sets from the Aligned Landmarks. The Percentage of Correct Classification for Each Population is Represented by Each Cell. The CHA and STC Populations were Removed Because the Former is so Differentiated it Masks Subtle Differences between the Other Populations and the Latter Because the Crosses Measured Here were Largely Hybrid in Origin

	Aligned Landmark Coordinates	
Population	**5 Populations**	**7 Populations**
HAM	60.60%	63.64%
SER	66.70%	71.79%
CHA		91.67%
STC		84%
USR	80.60%	77.42%
BSR	54.80%	51.61%
TOB	80.00%	85.00%

Table 7S. Significant Pair-wise Shape Differences (t-test, Bonferroni α=0.00047) between Seven Wild Atlantic Salmon Populations for Variation in Aligned Landmarks Expressed by Canonical Variates 1 to 6 (Below Diagonal)

Populations	HAM	SER	CHA	STC	USR	BSR
HAM						
SER						
CHA	1,2	1,2				
STC	2,3,4	2,3,4	1,2,3,4			
USR	2,3	2,3	1,2,3	2,3,4,6		
BSR	3	1,3	1,2,3,6	2,3,4	6	
TOB	3,4	4	1,2,3,4	2,3,4	2,3,4	3,4,6

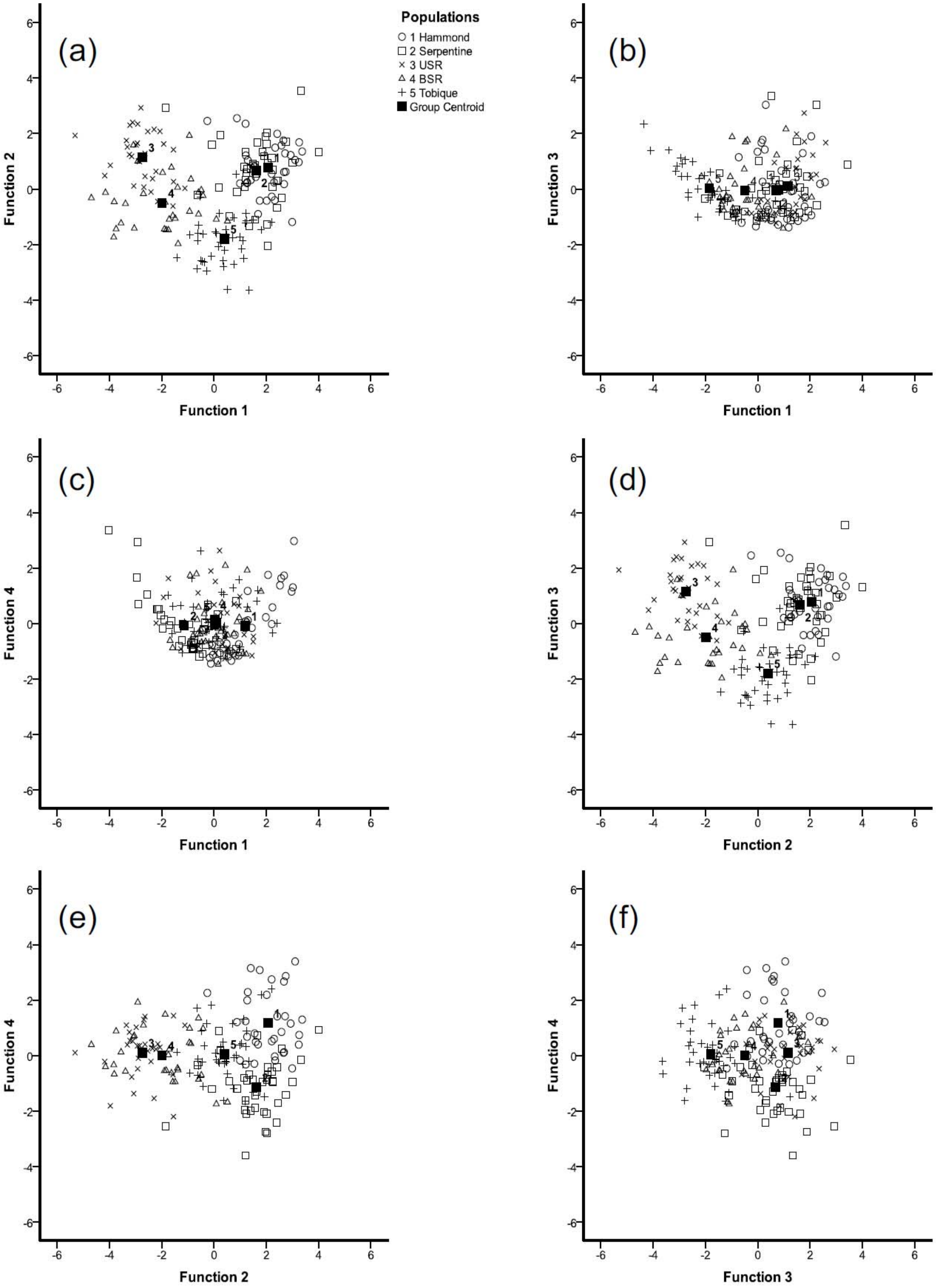
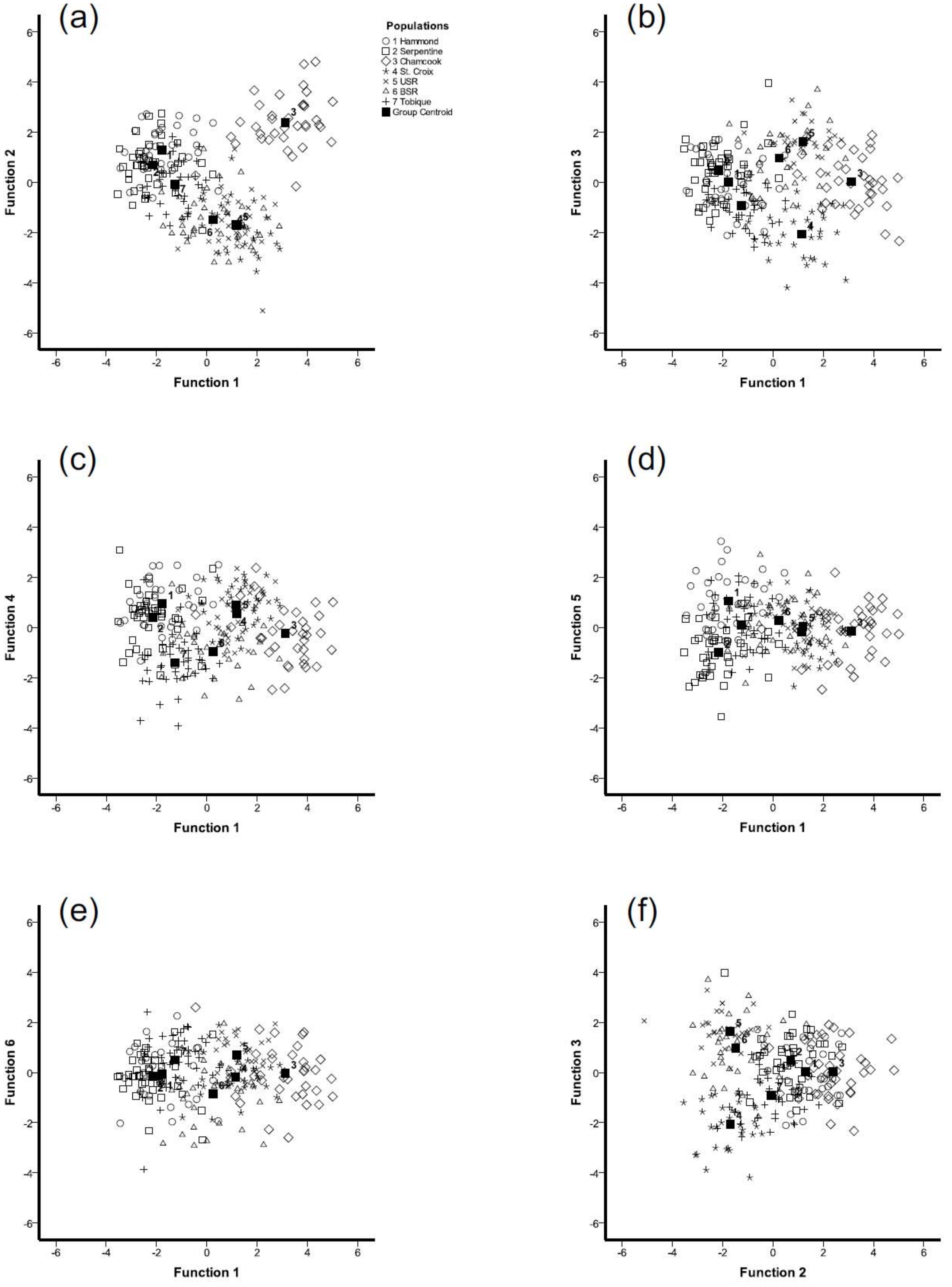
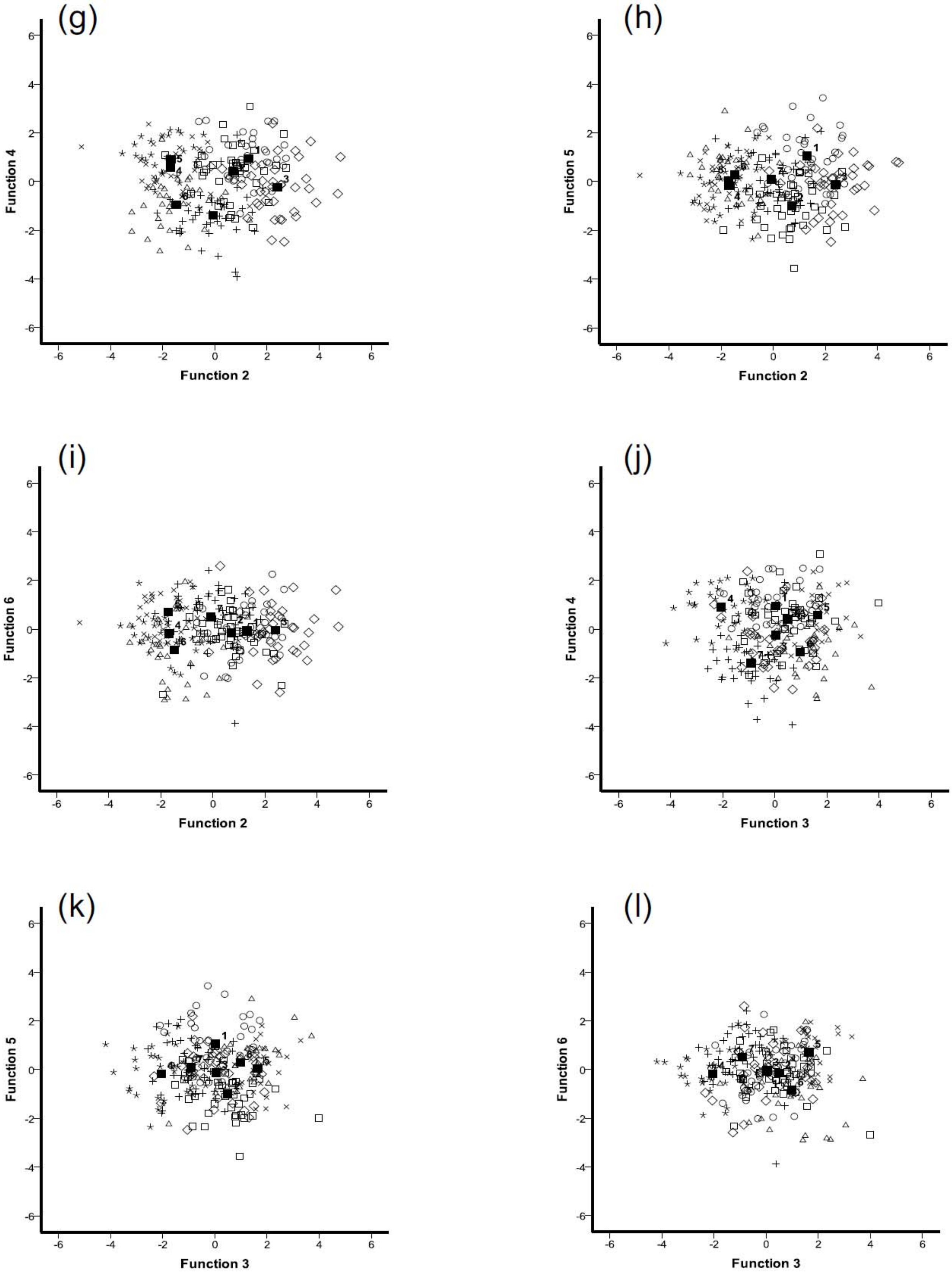


Fig. (1S). Canonical variates ordination of five Atlantic salmon populations from discriminant analysis of aligned landmark coordinates. **a** to **f.** Pairwise combinations of the six canonical variables.


(a)
Populations
○ 1 Hammond
□ 2 Serpentine
◇ 3 Chamcook
✶ 4 St. Croix
× 5 USR
△ 6 BSR
+ 7 Tobique
■ Group Centroid
Function 2
Function 1
(b)
Function 3
Function 1
(c)
Function 4
Function 1
(d)
Function 5
Function 1
(e)
Function 6
Function 1
(f)
Function 3
Function 2
-6
-4
-2
0
2
4
6


(g)
Function 4
Function 2
(h)
Function 5
Function 2
(i)
Function 6
Function 2
(j)
Function 4
Function 3
(k)
Function 5
Function 3
(l)
Function 6
Function 3
-6
-4
-2
0
2
4
6
1
2
3
4
5
6
7

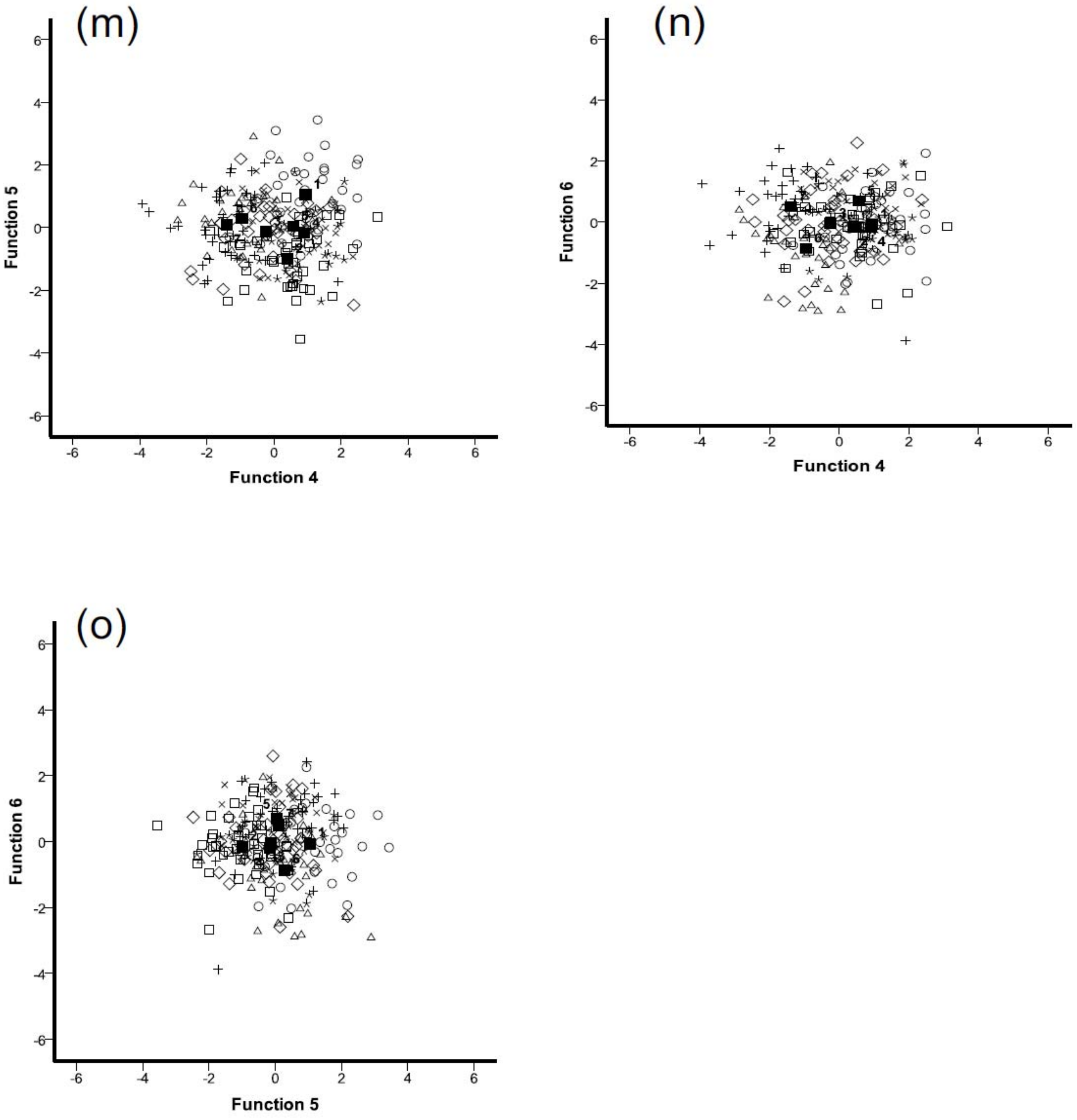
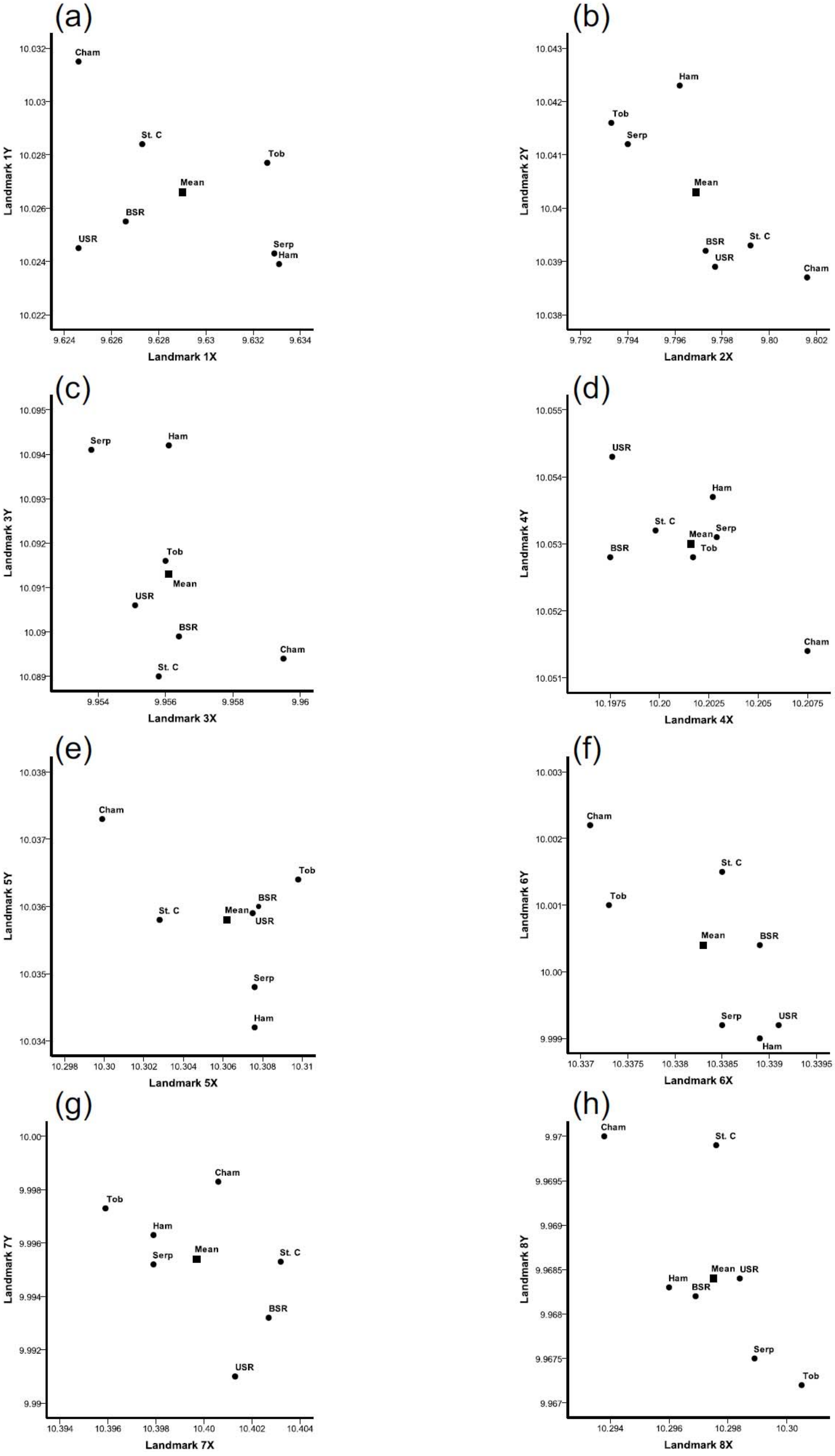


Fig. (2S). Canonical variates ordination of seven Atlantic salmon populations from discriminant analysis of aligned landmark coordinates. **a** to **o.** Pairwise combinations of the six canonical variables.


(a)
Landmark 1Y
Landmark 1X
10.032
10.03
10.028
10.026
10.024
10.022
9.624
9.626
9.628
9.63
9.632
9.634
Cham
St. C
Tob
Mean
BSR
USR
Serp
Ham
(b)
Landmark 2Y
Landmark 2X
10.043
10.042
10.041
10.04
10.039
10.038
9.792
9.794
9.796
9.798
9.80
9.802
Ham
Tob
Serp
Mean
BSR
St. C
USR
Cham
(c)
Landmark 3Y
Landmark 3X
10.095
10.094
10.093
10.092
10.091
10.09
10.089
9.954
9.956
9.958
9.96
Serp
Ham
Tob
Mean
USR
BSR
Cham
St. C
(d)
Landmark 4Y
Landmark 4X
10.055
10.054
10.053
10.052
10.051
10.1975
10.20
10.2025
10.205
10.2075
USR
Ham
St. C
Serp
Mean
BSR
Tob
Cham
(e)
Landmark 5Y
Landmark 5X
10.038
10.037
10.036
10.035
10.034
10.298
10.30
10.302
10.304
10.306
10.308
10.31
Cham
Tob
BSR
St. C
Mean
USR
Serp
Ham
(f)
Landmark 6Y
Landmark 6X
10.003
10.002
10.001
10.00
9.999
10.337
10.3375
10.338
10.3385
10.339
10.3395
Cham
St. C
Tob
Mean
BSR
Serp
USR
Ham
(g)
Landmark 7Y
Landmark 7X
10.00
9.998
9.996
9.994
9.992
9.99
10.394
10.396
10.398
10.40
10.402
10.404
Cham
Tob
Ham
Serp
Mean
St. C
BSR
USR
(h)
Landmark 8Y
Landmark 8X
9.97
9.9695
9.969
9.9685
9.968
9.9675
9.967
10.294
10.296
10.298
10.30
Cham
St. C
Mean
USR
Ham
BSR
Serp
Tob

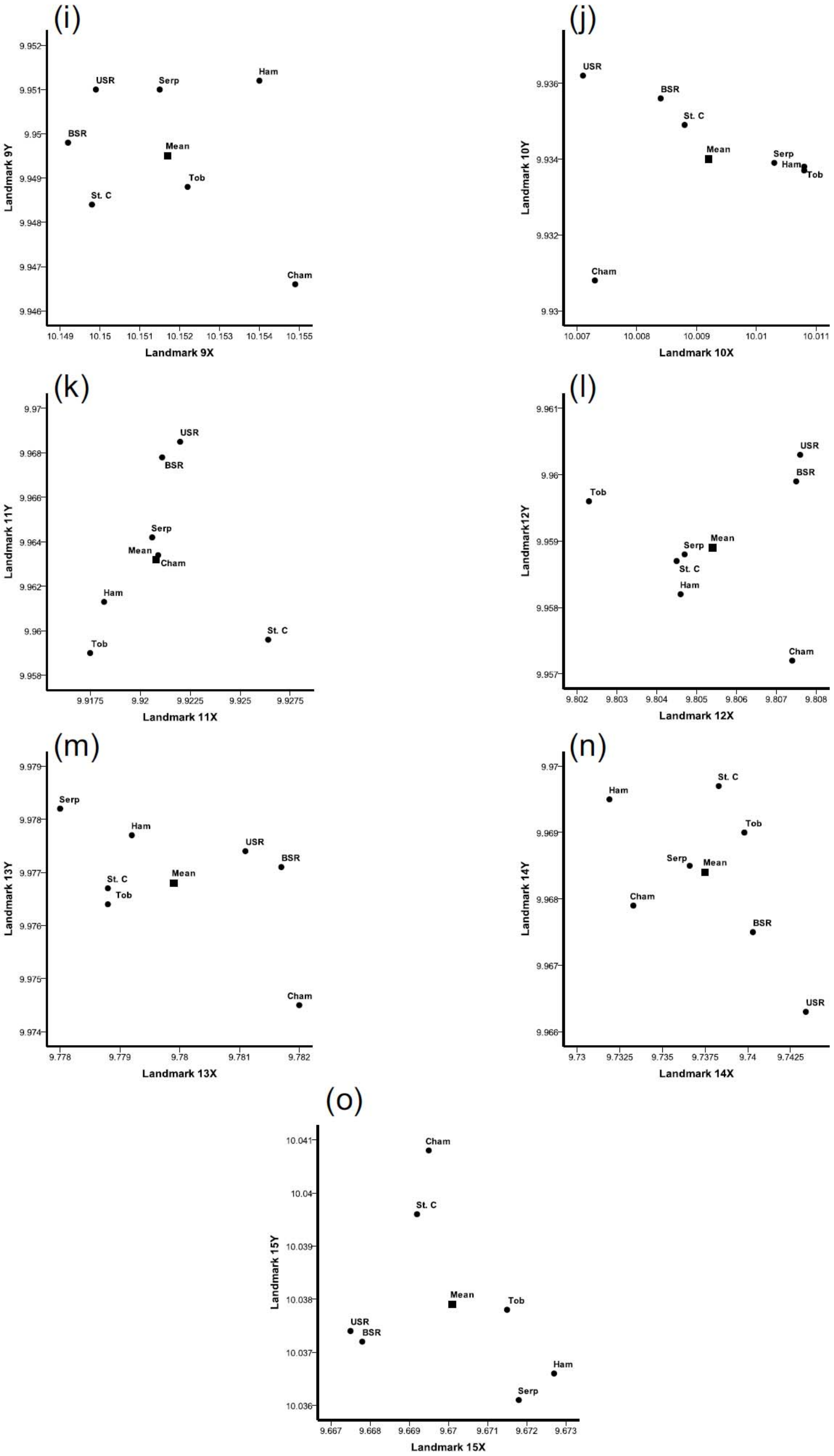
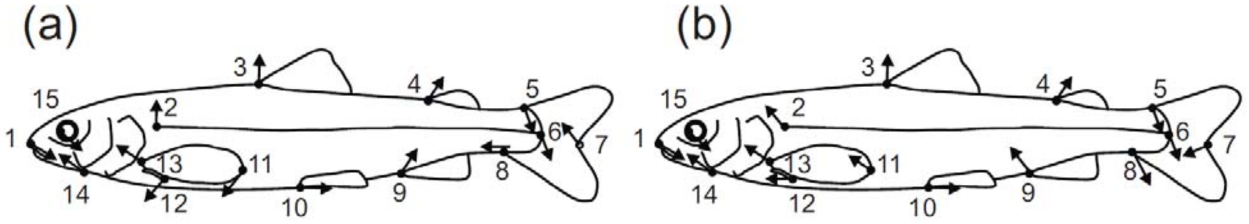
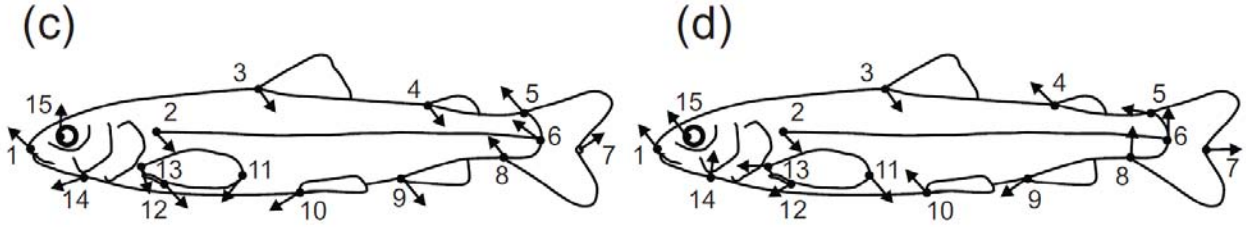


Fig. (5S). Plots of the position in XY space of the seven population means for each of the 15 aligned landmarks and the overall mean for each of the 15 aligned landmarks.



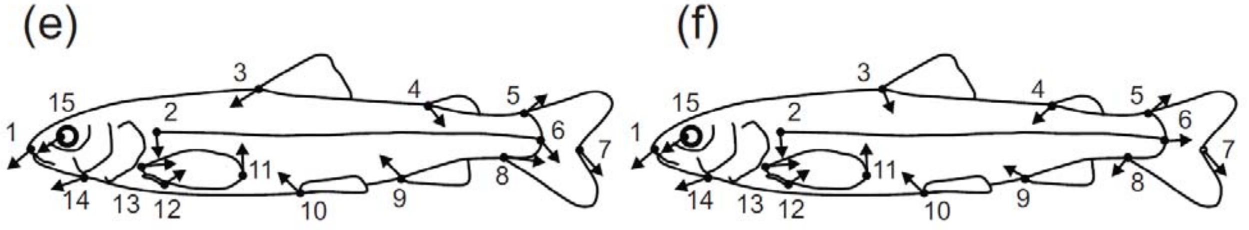
Hammond

Serpentine



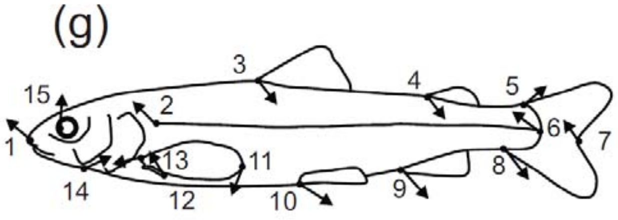
Chamcook

St. Croix



Upper Salmon River

Big Salmon River



Tobique

Fig. (6S). Schematic illustration of an Atlantic salmon with 15 landmarks (Fig. **2**) and the changes in X Y space of each landmark from plots of the mean aligned landmark coordinates for each of the seven populations.

Appendix B: Supplementary Data for the Paper: Signatures of Selection on Growth, Shape, Parr marks, and SNPs among seven Canadian Atlantic salmon (*Salmo salar*) Populations

M. Culling, H. Freamo, K. Patterson, P. Berg, S. Lien and E.G. Boulding

TABLE OF CONTENTS: APPENDIX B: DETAILED METHODS AND RESULTS FOR GEOMETRIC MORPHOMETRICS ANALYSIS USING PARTIAL WARPS.

Table 3S. Standardized Discriminant Function Coefficients from a partial warp landmark discriminant analysis of five populations for 15 external morphological landmarks.

Table 4S. Standardized Discriminant Function Coefficients from a partial warp landmark discriminant analysis of seven populations for 15 external morphological landmarks

Table 8S. Six and four discriminant functions for the seven populations based on partial warps from geometric morphometrics.

Table 9S. Classification of each population based on discriminant function analyses of seven populations partial warps from geometric morphometrics.

Fig. (3S). Canonical variates ordination of five Atlantic salmon populations from discriminant analysis of partial warp scores.

Fig. (4S). Canonical variates ordination of seven Atlantic salmon populations from discriminant analysis of partial warp scores.

Fig. (7S). Geometric morphometric deformation grids depicting shape differences from the extremes of the first two functions of the canonical variates ordination of seven wild Atlantic salmon populations.

Detailed Methods: Discriminant Function Analyses of Partial Warps

General patterns of geometric shape differentiation were investigated by performing a relative warp analysis on the partial warp scores. Each of the twenty-six partial warp scores were then regressed against the centroid size value saved for each fish in the SPSS program and the residual values were saved. To attempt to discriminate between the five Atlantic salmon populations, these size-free residual values (including those from Unix and UniY), were then used as the independent variables in a discriminant function analysis (DFA), in SPSS which resulted in four canonical variates that summarized shape variation (Fig. **3S**). We performed the DFA again after including the CHA and STC populations that were excluded for the reasons summarized in the main manuscript (Fig. **4S**).

Potential shape differences among groups were visualized by creating deformation grid plots. Partial warps were regressed against the first canonical variate scores from the DFA using the software TPSREGR 1.31 [37]. This procedure creates deformation grid plots that summarize how one form can be stretched, pulled, and shrunk in order to deform it into another related form. The deformation grids reflect the degree and type of shape change required to deform the overall consensus form into the most extreme form for each function (Fig. **7S**).

Partial Warp Morphological Variation

Having analysed aligned landmark coordinates to established a relationship between natal river and body shape in our study populations, we also went on to investigate the distribution of these potentially adaptive changes in body shape using geometric morphometrics. Discriminant analysis was conducted on the residuals of the partial warp analysis to determine whether the 174 Atlantic salmon individuals could be accurately discriminated into five, and 242 individuals into seven, population groups based upon their body shape. The overall Wilks' lambda was significant, $\Lambda = 0.067$, χ^2 ($df = 84$, $N = 174$) = 431.49, $p < 0.001$), for five and seven ($\Lambda = 0.69$, χ^2 ($df = 102$, $N = 242$) = 46.2, $p < 0.001$), populations indicating there were significant differences among the five and among the seven population sets based on body shape (Tables **3S**, **4S**, **8S**). In addition, the Wilks' lambda was also significant for the residual five, $\Lambda = .255$, χ^2 ($df = 60$, $N = 174$) = 218.46, $p < 0.001$, and residual seven, $\Lambda = .89$, χ^2 ($df = 80$, $N = 242$) = 26.375, $p < 0.009$, population sets. Due to the significance of both sets of tests, we chose to interpret four and six discriminant functions for each set separately (Tables **3S** and **4S**). Among the five population set, the amount of variability of the scores that was accounted for by differences among the populations ranged from 66.6% for the first discriminant function down to 13.9% for the fourth discriminant function, and for the seven population analyses, 66.9% down to 10.9% (Table **8S**). This classification allowed us to determine how well we can predict group membership using a classification function. After using a prior based on actual group sample sizes, the mean number of cross-validated grouped cases correctly classified was 67.2% for the five population analysis and 65.3% for the seven population analysis. The cross-validated grouped cases correctly classified ranged from 87.5% to 67.74% for the five populations and 83.3% to 64.5% for all seven populations (Table **9S**).

Shape variation from the DFA canonical scores showed that the majority of the morphological variation was in the anterior region (Figs. **3S**, **4S**). Canonical variate 1, accounting for 66.9% of the total variation, in the seven population analyses, expressed the elongation/truncation of the head region with reference to the location of the tip of snout and the beginning of the lateral line. (Table **4S**). Variation across populations in dorsoventral flattening and deepening, and orientation of the pupil of each fish was also expressed in canonical variate 1. The position of the pectoral fin insertion relative to the head region represented 65.6% of the total variation found among individual fish. Variation of the position of pectoral fin insertion point in relation to the gill insertion (ventral), and the distance between the pectoral fin insertions is represented by canonical variate 2. The pectoral fin length and general jaw length are represented by canonical score 3, accounting for 46.1% of variation across the body form. Jaw length variation is shown in changes to the relative positions of the tip of the snout and the ventral gill insertion point (Table **4S** and **8S**). Canonical variate 4 shows similar body form variation in the relative position of the pectoral fin insertion and the gill insertions and accounts for only 24.7% of the total morphological variation. Significant differences were found in a majority of population comparisons across a canonical variate. Canonical variate 6, which accounts for 10.9% of total variation, represents further body form variation with more evidence for variation in pectoral fin length (Tables **4S** and **8S**). Similar results were obtained for the five population analyses (Tables **3S** and **8S**).

Interpretation of the changes in orientation of each of the 15 landmarks for each population, relative to the overall mean landmark for all populations, revealed that the CHA population landmarks were almost always on the opposite side of the overall population mean (for 1 or both axis), to all other populations 87% of the time (Fig. **5S**). Furthermore, the CHA population landmarks were orientated directly opposite those of the HAM and SER populations 40% of the time (Fig. **5S**). The HAM and SER populations' landmarks were mainly located in the same orientation (53% of the time) although they did differentiate in the location of anterior lateral line, anterior insertion of the dorsal, adipose and caudal fins (Fig. **5S**). The USR and BSR populations' landmarks were also mainly located in the same orientation (67%) with some differentiation in the location of anterior insertion of the adipose fin, the caudal peduncle (Fig. **5S**). The differentiation highlighted here is supported by examination of vectors plotted on a schematic illustration of an Atlantic salmon of the change in each landmark per population, relative to the overall mean landmark for all populations (Fig. **6S**). The orientation of the CHA population landmark vectors differ from the HAM and SER populations landmarks in all 15 landmarks. While the orientation of the USR and BSR populations landmark vectors differ from all other populations landmark vectors 47% of the time with tip of the snout, insertion of the caudal peduncle and pectoral fin particularly differentiated (Fig. **6S**).

Shape differences identified by geometric morphometric TPS analyses were interpreted through the variation being represented by deformation grids of the entire body shape (Fig. **7S a-d**). Through applying subtle magnification (x3) the positive extremes of canonical variates appear generally more compressed and elongate at the anterior of the fish while the posterior is shorter and deeper at the negative extremes.

Table 3S. Standardized Discriminant Function Coefficients from a Partial Warp Landmark Discriminant Analysis of Five Populations for 15 External Morphological Landmarks Characteristic for Wild Atlantic Salmon (*Salmo Salar*)

Landmarks	Order	1	2	3	4
Tip of snout	1X	0.41	-1.64	2.54	0.24
Tip of snout	1Y	0.47	-2.02	-0.92	0.80
Anterior beginning of lateral line	2X	-0.38	1.67	-3.04	-12.86
Anterior beginning of lateral line	2Y	-0.34	0.95	-1.37	1.57
Anterior insertion of dorsal fin	3X	1.03	1.81	-0.88	2.27
Anterior insertion of dorsal fin	3Y	-0.66	-0.93	0.35	-0.85
Anterior insertion of adipose fin	4X	-0.83	0.22	0.95	-1.38
Anterior insertion of adipose fin	4Y	-2.37	2.38	8.46	4.37
Dorsal insertion of caudal fin	5X	0.43	-0.03	-2.18	-1.17
Posterior mid-point of caudal peduncle	6X	1.58	-0.99	0.52	1.88
Posterior mid-point of caudal peduncle	6Y	1.65	-8.14	-0.20	-2.52
Fork point in caudal peduncle	7Y	0.29	0.85	-0.37	-0.01
Ventral insertion of caudal fin	8X	-0.70	2.85	-4.03	6.79

Table 3S. Contd…..

Landmarks	**Order**	**1**	**2**	**3**	**4**
Ventral insertion of caudal fin	8Y	-0.98	-0.34	0.73	0.45
Anterior insertion of anal fin	9Y	-1.59	3.63	-0.76	0.69
Anterior insertion of pelvic fin	10X	0.96	0.87	0.20	1.68
Posterior point of pectoral fin	11X	-0.36	0.08	-0.28	0.14
Ventral insertion of pectoral fin	12X	0.43	0.20	-0.09	-2.23
Ventral insertion of pectoral fin	12Y	1.25	-1.16	0.12	-0.06
Dorsal insertion of pectoral fin	13X	0.11	-0.01	0.46	0.11
Dorsal insertion of pectoral fin	13Y	0.00	0.08	0.22	0.14
Eigen Value		2.79	1.41	0.40	0.16
Percent of variation explained		58.59	29.55	8.48	3.38

Note: Eigenvalues from each discriminant function are listed below the column of coefficients. Variables 5Y, 7X, 9X, 10Y, 11Y, 14X, 14Y, 15X, 15Y, UniX and Uniy not used in the analyses because they failed the tolerance test.

Table 4S. Standardized Discriminant Function Coefficients from a Partial Warp Landmark Discriminant Analysis of Seven Populations for 15 External Morphological Landmarks Characteristic for Wild Atlantic Salmon (*Salmo Salar*)

Landmarks	Order	1	2	3	4	5	6
Tip of snout	1X	-.787	-.494	.616	2.499	.383	.796
Tip of snout	1Y	-.601	.473	1.697	.513	-.637	.210
Anterior beginning of lateral line	2X	.807	-1.526	-.233	-3.912	5.368	-5.625
Anterior beginning of lateral line	2Y	-.226	2.049	-.849	.876	-.747	.175
Anterior insertion of dorsal fin	3X	1.460	1.791	-1.519	-.107	-2.674	.214
Anterior insertion of dorsal fin	3Y	-.943	-.483	.170	.662	.377	-.499
Anterior insertion of adipose fin	4X	-.428	-.564	-.077	-.422	2.448	.009
Anterior insertion of adipose fin	4Y	.206	-4.035	-3.671	2.482	8.000	-.422
Dorsal insertion of caudal fin	5X	-.431	1.704	1.019	-1.721	-1.031	-.828
Posterior mid-point of caudal peduncle	6X	.938	-1.024	.369	1.873	-1.447	.456
Posterior mid-point of caudal peduncle	6Y	-4.492	4.247	9.932	-1.609	.786	2.489
Fork point in caudal peduncle	7Y	.524	.231	-.622	.151	-.214	-.398
Ventral insertion of caudal fin	8X	2.110	-2.706	-3.223	-.946	-6.717	1.260
Ventral insertion of caudal fin	8Y	-.143	-1.143	-.428	.458	.130	.311
Anterior insertion of anal fin	9X	-.120	2.770	.490	-.714	-2.320	2.929
Anterior insertion of anal fin	9Y	1.081	-1.143	-3.660	-1.528	-1.606	1.065
Anterior insertion of pelvic fin	10X	.890	.903	-1.637	2.213	-.880	-.717
Ventral insertion of pectoral fin	12X	.845	-.940	1.236	-1.585	.741	-.444
Ventral insertion of pectoral fin	12Y	.004	-.103	.464	1.000	.329	-.764
Eigen Value		2.503	1.821	0.896	0.489	0.265	0.056
Percent of variation explained		41.5	30.2	14.9	8.1	4.4	0.9

Note: Eigenvalues from each discriminant function are listed below the column of coefficients. Variables 5Y, 7X, 10Y, 11X, 11Y, 13X, 13Y, 14X, 14Y, 15X, 15Y, UniX and UniY not used in the analyses because they failed the tolerance test.

Table 8S. Four Discriminant Functions based on Either the Five and Seven Population Set Using the Partial Warp Scores from the Geometric Morphometrics. Percentage in Variability of the Scores that was Accounted for by Differences Among the Populations. The CHA and STC Populations were Initially Removed from the Five Population Dataset because the Former is so Differentiated it Masks Subtle Differences between the Other Populations and the Latter Because the Crosses Measured here were Largely Hybrid in Origin

	Partial Warp Coordinates	
Function	**5 Populations**	**7 Populations**
1	66.60%	66.91%
2	48.40%	65.61%
3	26.50%	46.10%
4	13.98%	24.70%
5		12.82%
6		10.89%

Table 9S. Classification of Each Population based on Discriminant Function Analyses of Five and Seven Population Sets from the Partial Warps. The Percentage of Correct Classification for Each Population is Represented by Each Cell. The CHA and STC Populations were Initially Removed from the Initial Analysis Because the Former is so Differentiated it Masks Subtle Differences Between the Other Populations and the Latter Because the Crosses Measured here were Largely Hybrid in Origin

	Partial Warp Coordinates	
Population	**5 Populations**	**7 Populations**
HAM	57.58%	69.70%
SER	64.10%	66.70%
CHA		91.70%
STC		78.10%
USR	80.65%	77.40%
BSR	45.16%	61.30%
TOB	85.00%	82.50%

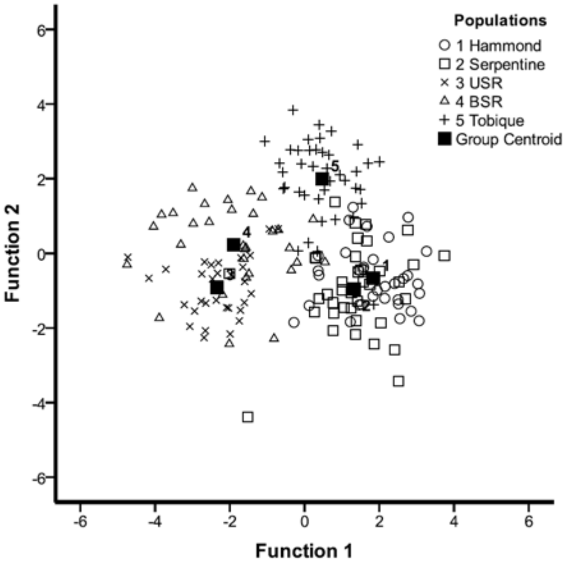


Fig. (3S). Canonical variates ordination of five Atlantic salmon populations from discriminant analysis of partial warp scores. The two canonical variates axes represent 66.6% and 48.4% of the shape variation for the five population analyses.

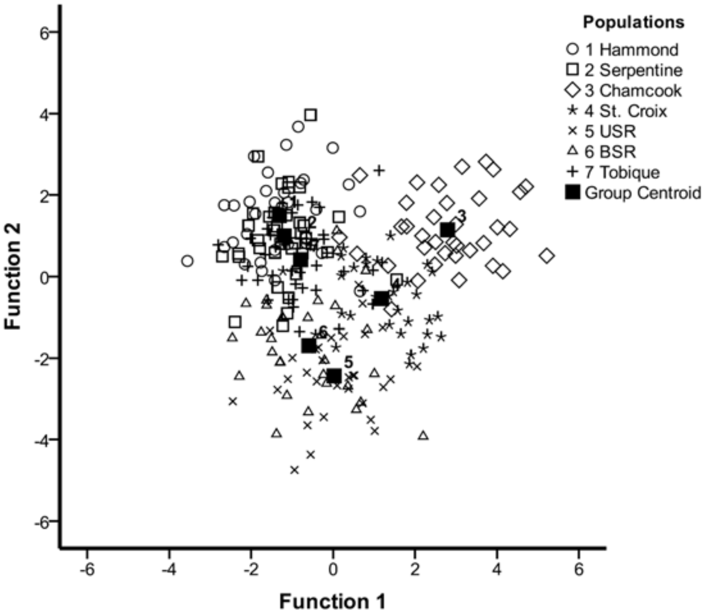
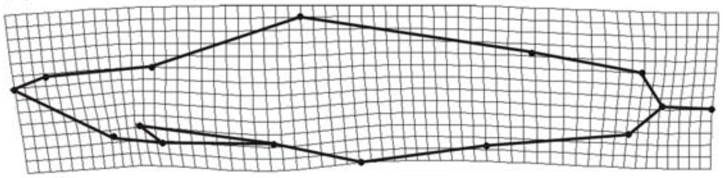


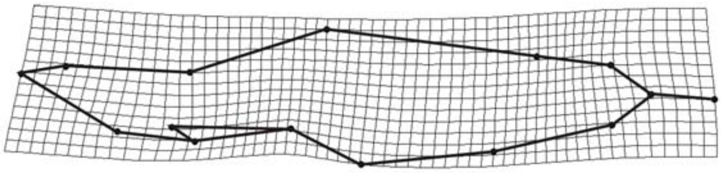
Fig. (4S). Canonical variates ordination of seven Atlantic salmon populations from discriminant analysis of partial warp scores. The two canonical variates axes represent 66.9%, and 65.6% of the shape variation respectively for the seven population analyses.

(a)



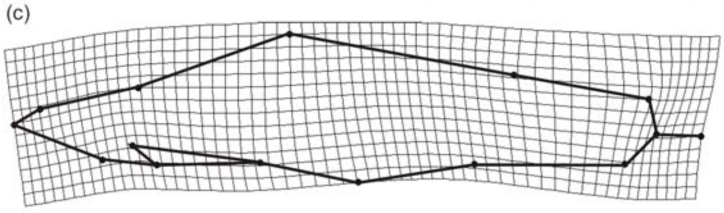
Canonical 1, Negative extreme (X3)

(b)



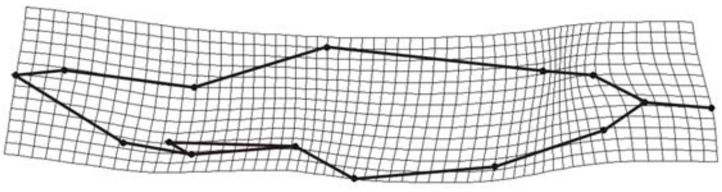
Canonical 1, Positive extreme (X3)

(c)



Canonical 2, Negative extreme (X3)

(d)



Canonical 2, Negative extreme (X3)

Fig. (7S). Geometric morphometric deformation grids depicting shape differences from the extremes of the first two functions of the canonical variates ordination in Fig. (**4S**) of seven wild Atlantic salmon populations (at 3x magnification to improve visualization). Deformations of the grid indicate differentiation from the overall consensus configuration. Refer to Figure (**2**) (main paper) for the landmark locations. The deformation grids were estimated by regressing the partial warps against the scores on the first root of the DFA using TpsRegr [37].

Appendices C-F: Supplementary Appendices C-F for the Paper: Signatures of Selection on Growth, Shape, Parr marks, and SNPs among seven Canadian Atlantic salmon (*Salmo salar*) Populations

M. Culling, H. Freamo, K. Patterson, P. Berg, S. Lien and E.G. Boulding

TABLE OF CONTENTS: APPENDICES C-F

Appendix C: Methods and Results for Estimation of Family Sizes with Colony 2.0

Multilocus SNP genotypes used to estimate the number of full-sib families and mean full-sib family size for all seven populations using COLONY 2.0 with the full likelihood, high precision option [86, 87] We assumed the monogamous mating option for all populations except for STC because the Live Gene Banking program uses single pair matings whenever possible [3], (T. Goff, pers. comm.).

The COLONY 2.0 analysis of the two new populations, CHA and STC suggested that there was generally minimal family structuring in most of the sample collections analyzed (Table 1) as previously found by Freamo *et al.* [29]. The population with the largest estimated full-sib family size (9.6 individuals) was CHA which is consistent with it being sampled from a spawning that created only four full-sib families (Table **1**).

Appendix D: Population Genetic and Phenotypic Distances for seven populations

We found statistically significant pairwise differentiation for fifteen out of twenty one (71.4%) comparisons between the seven Atlantic salmon populations using the 195 non-outlier SNP markers (Table **10S**; mean $F_{ST} = 0.12$). The CHA population had the highest six values ($F_{ST} > 0.22$) for all the pairwise comparisons that were made between all populations. The CHA and both Inner Bay of Fundy populations (USR and BSR), were significantly differentiated from the three Outer Bay of Fundy populations.

Table 10S. Pairwise F_{ST} Values (Below Diagonal) and P values (Above Diagonal) for all 195 Non-outlier SNPs in Seven Atlantic Salmon Populations (Table 1)

Populations	BSR	CHA	HAM	SER	STC	TOB	USR
BSR		*	*	*	*	*	*
CHA	0.2200		*	*	*	*	*
HAM	0.0683	0.2574		0.02933	0.01367	0.01529	*
SER	0.0520	0.2316	0.0677		0.02710	0.08857	*
STC	0.0942	0.3231	0.0862	0.0585		0.01624	*
TOB	0.0504	0.2344	0.0361	0.0304	0.0489		
USR	0.0590	0.2729	0.0849	0.0988	0.1210	0.0922	

Note: Exact P values obtained after 21,000 permutations. * indicates a significant pairwise difference after corrected for multiple comparisons $P < 0.002381$.

Table 11S. Pairwise Procrustes Distances in Common Geometric Shape Space for Seven Atlantic Salmon Populations (Table 1). Mean Distances after Removal of Centroid Size Using the Residuals from a LN Regression with Standard7 module from the CoordGen7a Module of IMP 7.0 [67]

Populations	BSR	CHA	HAM	SER	STC	TOB	USR
BSR							
CHA	0.0226						
HAM	0.0192	0.0215					
SER	0.0196	0.0228	0.0083				
STC	0.0201	0.0190	0.0152	0.0161			
TOB	0.0327	0.0439	0.0378	0.0341	0.0413		
USR	0.0152	0.0226	0.0147	0.0156	0.0187	0.0367	

Table 12S. Mean ±SE Parr Mark Values Along Diagonal and Absolute Value of Difference in Parr Mark Number below Diagonal for Seven Atlantic Salmon Populations (Table 1)

Populations	BSR	CHA	HAM	SER	STC	TOB	USR
BSR	7.68±0.18						
CHA	1.705	5.97±0.20					
HAM	1.314	0.391	6.36±0.20				
SER	1.883	0.177	0.569	5.79±0.27			
STC	0.323	2.028	1.636	2.205	8.00±0.18		
TOB	0.173	1.878	1.486	2.055	0.150	7.85±0.15	
USR	0.097	1.802	1.411	1.979	0.226	0.076	7.77±0.16

Table 13S. **Mean (±SE) Parr Mark Contrast Along Diagonal and Absolute Value of Difference in Parr Mark Contrast Below Diagonal for Seven Atlantic Salmon Populations (Table 1)**

Populations	BSR	CHA	HAM	SER	STC	TOB	USR
BSR	1.79±0.05						
CHA	0.096	1.69±0.05					
HAM	0.274	0.178	1.51±0.03				
SER	0.469	0.373	0.195	1.32±0.04			
STC	0.191	0.287	0.465	0.660	1.98±0.03		
TOB	0.090	0.186	0.365	0.560	0.101	1.88±0.03	
USR	0.289	0.385	0.563	0.758	0.098	0.199	2.08±0.06

Appendix E: Methods for Detecting Metapopulations with STRUCTURE

We used the program STRUCTURE, which uses a Markov Chain Monte Carlo (MCMC) algorithm to cluster individuals into metapopulations on the basis of multilocus genotype data [88, 89]. STRUCTURE has been successfully applied to problems such as identifying cryptic population structure, detecting migrants or admixed individuals, and inferring historical population admixture [89]. We used STRUCTURE version 2.3.3 [88] with a burn-in of 5 X 10^4 and with 10^6 iterations using prior information (LOCISPOP = 1) about sampling sites, no admixture and the LOCPRIOR model to infer population clusters. True population number (*K*) were made from one to a maximum of 10 values and replicated 20 times for each *K*-value. Populations were ascribed to a particular STRUCTURE group when the highest proportion of sampled individuals was in that group. We used the statistic ΔK based on the rate of change in the log probability of data between successive *K* values to provide the correct estimation of the number of clusters [90].

Results for Detecting Metapopulations with STRUCTURE

In the hierarchical analyses, using prior information (LOCISPOP = 1) about sampling sites, no admixture, and the LOCPRIOR model to infer population clusters, the STRUCTURE program reached a plateau at an apparent *K* estimate of two clusters or more. The estimate of *K* was also two, which extrapolated to iBoF and oBoF clusters.

Appendix F: Outlier Analysis Results for Five and Seven Populations

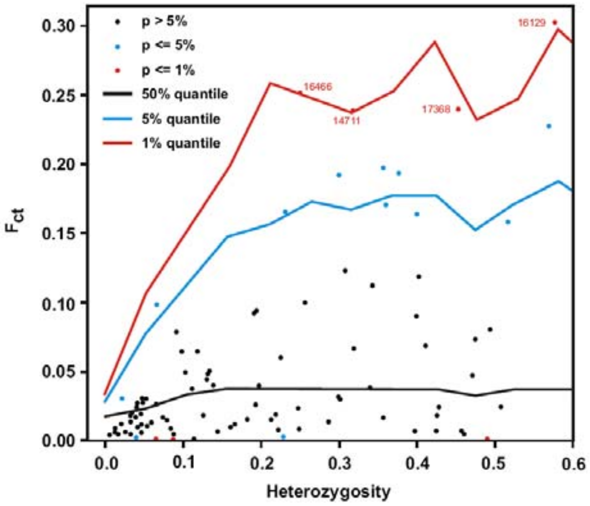
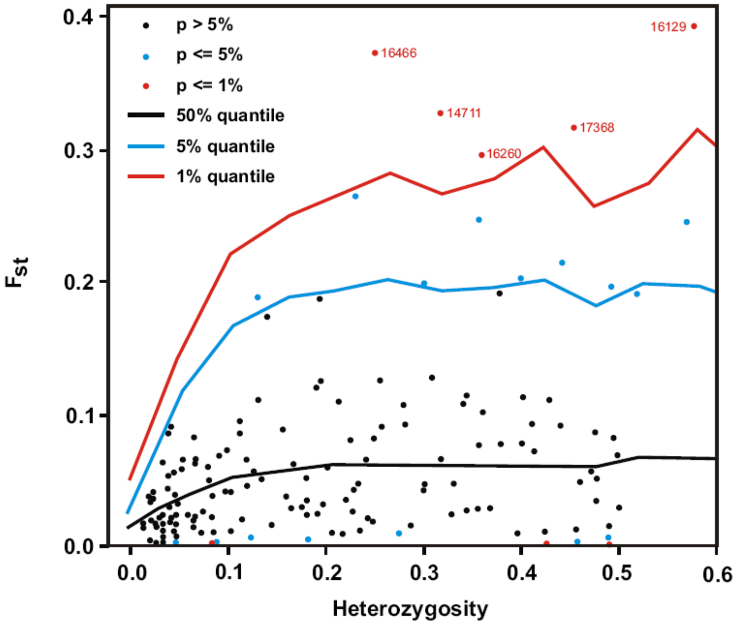


Fig. (8S). Outlier loci under diversifying selection detected in a total of 207 polymorphic loci using the Arlequin 3.5 hierarchical island model [55] comparing two inner Bay of Fundy populations (BSR, USR) and three outer Bay of Fundy populations (HAM, SER, TOB). For full names of populations refer to Table **1**. All loci above the upper 1% quantile would be classified as outlier loci under diversifying selection by Beaumont and Nichols [48]. See Table **4** for identity of outlier loci below 1% quantile but above 5% quantile in at least some of the five replicate runs.

a. F_{CT} for five populations: Four outlier Loci (Table **4**): Contig 14711_157 (Ssa0075ECIG, Galactosyltransferase), Contig 16129_0239 (Ssa0156ECIG, trypsin inhibitor), Contig 16466_1044 (Ssa0181ECIG, peptidase D), Contig 17368_0088 (Ssa0252ECIG, triosephosphate isomerase 1b).



b. F_{ST} for five populations: five highly significant ($p < 0.01$) outlier loci: Contig 14711_157 (Ssa0075ECIG, Galactosyltransferase), Contig 16129_0239 (Ssa0156ECIG, trypsin inhibitor), Contig 16260_0757 (Ssa0168ECIG, Beta-crystallin Bp), Contig 16466_1044 (Ssa0181ECIG, peptidase D), Contig 17368_0088 (Ssa0252ECIG, triosephosphate isomerase 1b).

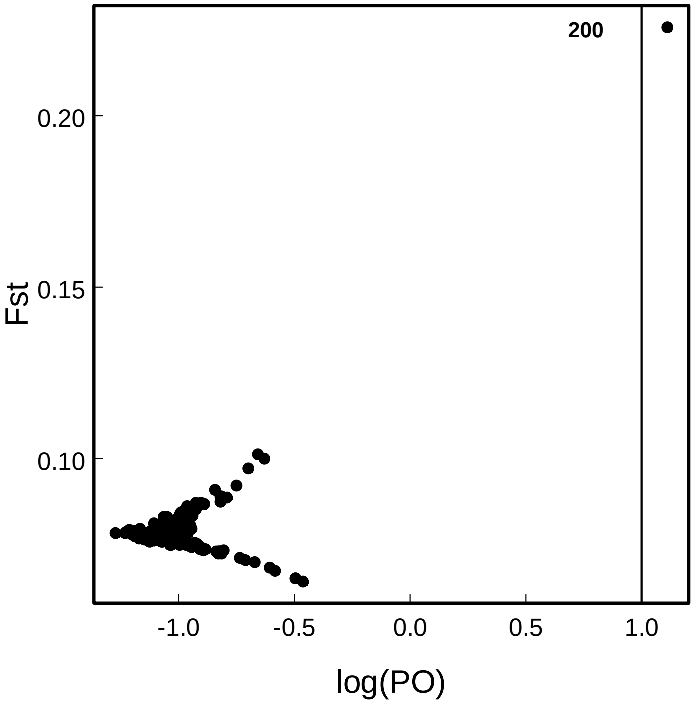


Fig. (9S). Single F_{ST} outlier locus under diversifying selection detected in a total of 207 polymorphic SNP loci in five Bay of Fundy populations using Bayescan version 2 that incorporates False Discovery Rates and Posterior Odds [47, 50, 91]. Outlier locus Contig 17368_0088 (Ssa0252ECIG, triosephosphate isomerase 1b) had "strong" support (N=207, P= 0.928, $\log_{10}(PO)=1.11$, $\alpha= 1.23$, $F_{ST} = 0.226$, FDR=0.07).

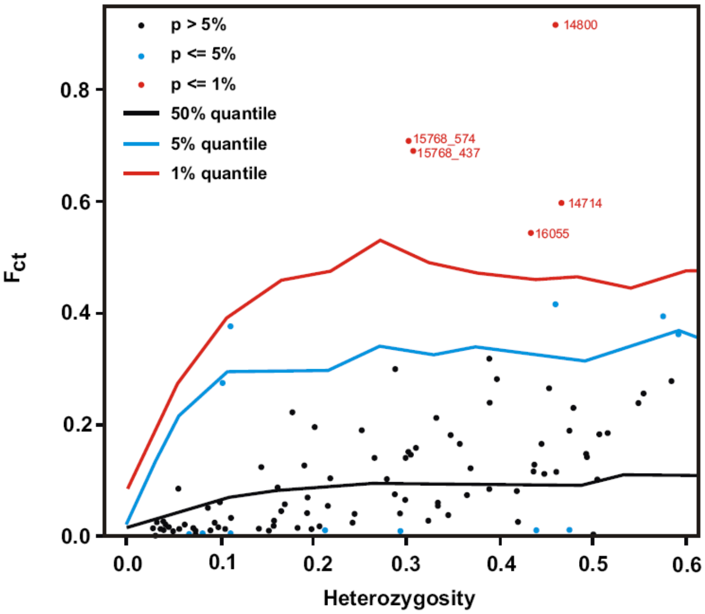


Fig. (10S). Outlier loci under diversifying selection detected in a total of 208 polymorphic loci using the Arlequin 3.5 hierarchical island model [55] comparing six populations: two from the inner Bay of Fundy (BSR, USR), three from the outer Bay of Fundy (HAM, SER, TOB) and a landlocked lake population (CHAM). For full names of populations refer to Table **1**.

a. F_{CT} for six populations: five highly significant loci: Contig 14714_122 (Ssa0076ECIG unknown protein), Contig 14800_360 (Ssa0081ECIG, WD repeat protein 43), Contig 15768_437 (Ssa0134aECIG, NADH dehydrogenase 4), Contig 15768_574 (Ssa0134bECIG, NADH dehydrogenase 4), Contig 16055_561 (Ssa0151ECIG, Hyaluron glucosamide 2).

b. F_{ST} for six populations: Five highly significant outlier loci: Contig 14714_122 (Ssa0076ECIG unknown protein), Contig 14800_360 (Ssa0081ECIG, WD repeat protein 43), Contig 15768_437 (Ssa0134aECIG, NADH dehydrogenase 4), Contig 15768_574 (Ssa0134bECIG, NADH dehydrogenase 4), Contig 16055_561 (Ssa0151ECIG, Hyaluron glucosamide 2).

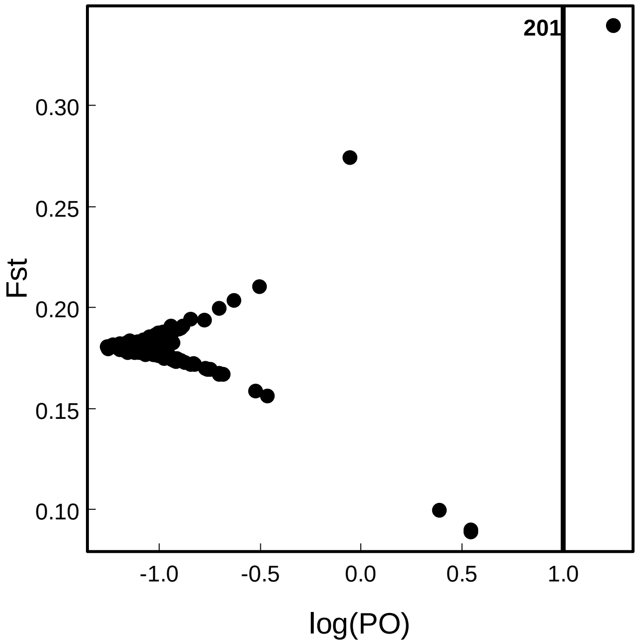


Fig. (11S). Single F_{ST} outlier locus under diversifying selection detected in a total of 208 polymorphic SNP loci in comparing six populations: two from the inner Bay of Fundy (BSR, USR), three from the outer Bay of Fundy (HAM, SER, TOB) and a landlocked lake population (CHAM) using Bayescan version 2 that incorporates False Discovery Rates and Posterior odds [47, 50, 91]. Outlier locus 201= Contig 17368_0088 (Ssa0252ECIG, triosephosphate isomerase 1b) had “strong” support (N=208, P= 0.944, $\log_{10}$(PO)=1.23, α= 1.26, F_{ST} = 0.338, FDR= 0.053).