



Multiple Selection Signatures in Farmed Atlantic Salmon Adapted to Different Environments Across Hemispheres

María Eugenia López^{1,2}, Tyler Linderoth^{3†}, Ashie Norris⁴, Jean Paul Lhorente⁵, Roberto Neira⁶ and José Manuel Yáñez^{1,7*}

¹ Facultad de Ciencias Veterinarias y Pecuarias, Universidad de Chile, Santiago, Chile, ² Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden, ³ Department of Integrative Biology, University of California, Berkeley, CA, United States, ⁴ Marine Harvest, Kindrum, Fanad, C. Donegal, Ireland, ⁵ Benchmark Genetics Chile, Puerto Montt, Chile, ⁶ Facultad de Ciencias Agronómicas, Universidad de Chile, Santiago, Chile, ⁷ Núcleo Milenio INVASAL, Concepción, Chile

OPEN ACCESS

Edited by:

María Saura,
Instituto Nacional de Investigación
y Tecnología Agraria y Alimentaria
(INIA), Spain

Reviewed by:

Roger Vallejo,
Cool and Cold Water Aquaculture
Research (USDA-ARS),
United States
Andrés Pérez-Figueroa,
University of Porto,
Portugal

*Correspondence:

José Manuel Yáñez
jmayanez@uchile.cl

†Present address:

Department of Genetics,
University of Cambridge, Cambridge,
United Kingdom

Specialty section:

This article was submitted to
Livestock Genomics,
a section of the journal
Frontiers in Genetics

Received: 19 January 2019

Accepted: 26 August 2019

Published: 01 October 2019

Citation:

López ME, Linderoth T,
Norris A, Lhorente JP, Neira R
and Yáñez JM (2019) Multiple
Selection Signatures in Farmed
Atlantic Salmon Adapted to Different
Environments Across Hemispheres.
Front. Genet. 10:901.
doi: 10.3389/fgene.2019.00901

Domestication of Atlantic salmon started approximately 40 years ago, using artificial selection through genetic improvement programs. Selection is likely to have imposed distinctive signatures on the salmon genome, which are often characterized by high genetic differentiation across population and/or reduction in genetic diversity in regions associated to traits under selection. The identification of such selection signatures may give insights into the candidate genomic regions of biological and commercial interest. Here, we used three complementary statistics to detect selection signatures, two haplotype-based (iHS and XP-EHH), and one F_{ST} -based method (BayeScan) among four populations of Atlantic salmon with a common genetic origin. Several regions were identified for these techniques that harbored genes, such as *kind1* and *chp2*, which have been associated with growth-related traits or the *kcnb2* gene related to immune system in Atlantic salmon, making them particularly relevant in the context of aquaculture. Our results provide candidate genes to inform the evolutionary and biological mechanisms controlling complex selected traits in Atlantic salmon.

Keywords: selection signatures, *Salmo salar*, Domestication, SNP data, artificial selection

BACKGROUND

Domestication is a complex evolutionary process whereby wild animals or plant populations adapt to environmental conditions created by humans and so involves genetic and developmental changes over multiple generations (Price, 1984; Liu et al., 2017). Since the beginning of domestication, humans have exploited the genetic diversity of various species to model them according to their needs (Driscoll et al., 2009). This has been amplified since the establishment of explicit genetic improvement objectives. As a result of intense selection pressure, dramatic phenotypic changes (Rubin et al., 2012) and substantial and continued genetic improvement have been made in domestic populations over the past decades (Hill and Bunger, 2004).

Domestication in most fish is relatively recent compared with terrestrial animals (Teletchea and Fontaine, 2014; López et al., 2015), but has expanded rapidly over the last decades (Lorenzen et al., 2012), and several breeding programs have been implemented in different aquatic species, such as

tilapia (*Oreochromis niloticus* L), rainbow trout (*Oncorhynchus mykiss* W), coho salmon (*Oncorhynchus kisutch* W), and Atlantic salmon (*Salmo salar* L) among others (Gjedrem, 2010; Gjedrem, 2012; Yáñez et al., 2014). The latter has become one of the most important aquaculture species (FAO, 2016), since it was first farmed in Norway during the 1960s. Despite a generation interval of 3 to 4 years, breeding programs have achieved rapid improvement of economically important traits, such as growth, sexual maturation, and disease resistance (Gjedrem et al., 2012). Domestication and subsequent artificial selection have produced stark phenotypic changes in farmed Atlantic salmon populations (Glover et al., 2017), as evidenced by differences in traits, such as growth and predator awareness, between wild and farmed populations (Thodesen et al., 1999; Glover et al., 2009; Solberg et al., 2012) (Einum and Fleming, 1997).

Positive selection pressures (natural and artificial) experienced by population undergoing selection will cause the frequency of alleles underlying favorable traits to increase rapidly. Linkage disequilibrium (LD) between favorable mutations and neighboring loci will increase and spread, given that there is little opportunity for recombination over the brief time since the onset of intense selection (Sabeti et al., 2002). Analyses of these selection signatures in domestic animals can provide further insights into the genetic basis of adaptation to diverse environments and genotype/phenotype relationships (Oleksyk et al., 2010; Andersson, 2012). Access to genomic data through next-generation sequencing and high-throughput genotyping technologies have made the comparison of genomic patterns of single nucleotide polymorphism (SNP) variation between different livestock breeds possible, allowing for the identification of putative genomic regions and genes under selection in several terrestrial domestic species, including cattle (e.g., Taye et al., 2017), horses (e.g., Avila et al., 2018), sheep (e.g., Ruiz-Larrazaga et al., 2018), and pigs (e.g., Gurgul et al., 2018).

There are several approaches for detecting genomic selection signatures, one of which relies on the length or variability of haplotypes. Directional selection acting on a new, beneficial mutation causes the haplotype harboring the mutation to increase in frequency and to be longer than average. To exploit this pattern for detecting positive selection, Sabeti et al. (2002) proposed the extended haplotype homozygosity (EHH) statistic, which is specifically the probability that two randomly selected haplotypes are identical-by-descent over their entire length around a core SNP (Sabeti et al., 2002). This concept forms the basis for other haplotype homozygosity-based metrics, such as the relative EHH (REHH) (Sabeti et al., 2002) and the widely used integrated haplotype score (iHS) (Voight et al., 2006). iHS compares EHH between derived and ancestral alleles within a population and has the most power to detect selection when the selected allele is at intermediate frequencies in the population (Sabeti et al., 2006; Voight et al., 2006). To detect selection signatures between populations, the cross-population extended haplotype homozygosity test (XP-EHH) compares the integrated EHH profiles between the two populations in the same SNP. This test was designed to detect ongoing or nearly complete selective sweeps in one population (Sabeti et al., 2007). An alternative approach for identifying

selection signatures when there are multiple populations for comparison is divergence-based methods, which focus on identifying outlier loci with either higher or lower allele frequency differences among populations than expected without selection (Beaumont and Balding, 2004; Foll and Gaggiotti, 2008; Excoffier et al., 2009). One common approach for quantifying the degree of genetic differentiation between populations is through the fixation index, F_{ST} (Wright, 1951). An unusually high F_{ST} value at a given locus can be indicative of directional selection. Divergence approaches to identify signals of selection have been successful in several domestic species including swine (Cesconeto et al., 2017), sheep (Manunza et al., 2016), and cattle (Maiorano et al., 2018) among others.

Although previous studies have already been carried out to detect selection signatures in Atlantic salmon (Mäkinen et al., 2014; Gutierrez et al., 2016; Liu et al., 2017; López et al., 2018), using multiple different strains adapted to different culture conditions across hemispheres, to explore how genetic variation among them differs, has not been done yet. Herein, we used an Affymetrix 200K SNP array data set to investigate selection signatures in farmed Atlantic salmon populations from the same origin, and subsequently cultivated in Ireland and Chile. We found evidence of selection using two haplotype-based approaches iHS and XP-EHH and one F_{ST} -based method, BayeScan, in the genomes of four Atlantic salmon populations. These findings are important because they highlight regions of the genome that might benefit economically relevant attributes, such as growth, resistance to local diseases, and adaptation to specific environmental conditions.

MATERIALS AND METHODS

Samples, Genotyping, and Quality Control

This study was performed using a total of 270 individuals from four populations (Pop-A, $n = 40$; Pop-B, $n = 71$; Pop-C, $n = 85$; Pop-D, $n = 74$) derived from the Mowi strain. This strain comes from one of the first farmed Atlantic salmon populations, which was established with fish from west coast rivers in Norway, with major contributions from River Bolstad in the Vosso watercourse, River Årøy, and possibly from the Maurangerfjord area (Verspoor et al., 2007). Salmon from the Vosso and Årøy rivers are characterized by large size and late maturity (Verspoor et al., 2007). Phenotypic selection for growth, late maturation and fillet quality was the focus in this population until 1999 (Glover et al., 2009). Ova from this population were imported into the Fanad Peninsula, Ireland, between 1982 and 1986 to establish an Irish-farmed population (Norris et al., 1999). Individuals from this population comprise Pop-A, which we estimate had been under artificial selection for growth for at least 10 generations prior to sampling. Similarly, ova from this farmed, Irish population were introduced into Chile in the early 1990s to establish separate farmed populations in the Los Lagos Region (42°S 72°O) and the Magallanes Region (53°S 70°O). Pop-B and Pop-C correspond to samples from two different populations in the Los Lagos Region that were initially founded with fish from different year-classes. Samples from Pop-D represent one population founded

in the Magallanes Region. The three Chilean populations were subsequently adapted to the biotic and abiotic conditions present in southern hemisphere. These populations experienced four generations of selective breeding for growth in Chilean farming conditions prior to sampling, which occurred at the same time that Pop-A was sampled in 2014.

All populations were genotyped using Affymetrix's Atlantic salmon 200K SNP Chip described in Yáñez et al. (2016). We performed SNP quality control using the Axiom Genotyping Console (GTC, Affymetrix) and SNPfisher (an R package developed by Affymetrix), which i) removed SNPs that did not conform high-quality clustering patterns as outlined by Affymetrix, ii) removed SNPs with genotype call rate lower than 95%, and iii) discarded individuals with genotyping call rate under 90%. As part of the validation of the SNPs chip used in this study, Yáñez et al. (2016) identified loci significantly deviating from Hardy–Weinberg equilibrium in eight populations separately and removed these sites if they were deviating from Hardy–Weinberg equilibrium among all populations. In addition, we limited our analyses to SNPs that mapped to chromosomes in the newest version of the Atlantic salmon reference genome, ICSAG_v2 (GenBank: GCA_000233375.4), which comprised 149,060 SNPs.

Genetic Diversity, LD, and Population Structure

We evaluated genetic diversity in terms of the observed heterozygosity (H_o) and expected heterozygosity (H_e) calculated with PLINK v1.09 (Purcell et al., 2007). We calculated the pairwise LD as the Pearson's squared correlation coefficient (r^2) for each population and within chromosomes using PLINK v1.09 (Purcell et al., 2007). For each SNP pair, bins of 100 kb were created based pairwise distance. To investigate population structure, we performed a principal component analysis (PCA) based on genotypes as implemented in PLINK v1.09 and inferred individual ancestry proportions with ADMIXTURE 1.2.2 (Alexander et al., 2009). For the admixture analysis, we performed 200 bootstraps with a number of ancestral lineages (K) ranging from 1 to 20. Ten-fold cross validation (CV = 10) was specified, and we retained results from the K having the lowest cross-validation error. The aforementioned analyses were conducted using a total of 21,950 SNPs, which had a minor allele frequency (MAF) larger than 0.05, were in Hardy–Weinberg equilibrium, and which had LD values of at most 0.4 (to minimize possible confounding effects of LD on the patterns of genetic structure).

Selection Signatures, Gene Annotation, and Functional Analyses

To identify genomic regions harboring selection signatures, we used one within population iHS and two between-population methods (XP-EHH and BayeScan) over a subset of 120,316 SNPs that had MAF > 0.05 among all populations.

(1) iHS. The iHS score for detecting selection is based on the ratio of EHH for haplotypes anchored with the ancestral versus derived allele. The ancestral allele state for our Atlantic salmon

populations is unknown and so to avoid losing SNPs by trying to polarize them from publicly available outgroup references, we assumed that the major allele represented the ancestral state as in Bahbahani et al. (2015). We phased the haplotypes using Beagle v.5.0 (Browning and Browning, 2009). Single-site iHS values across the genome were calculated for each populations using the REHH package (Gautier and Vitalis, 2012). These per site iHS values were standardized so that they were approximately distributed according to a standard normal distribution. We required candidate-selected regions to have at least two SNPs \leq 500 kb apart, each with iHS scores with $-\log_{10}(p \text{ value})$ of at least three ($p \text{ value} \leq 0.001$) based on a one-tailed test assuming that the standardized iHS $\sim N(0,1)$.

(2) XP-EHH. The XP-EHH statistic compares the integrated EHH between two populations at the same SNP, to identify selection based on overrepresented haplotypes in one of the populations (Sabeti et al., 2007). We evaluated three different pairs of populations with this method Pop-B/Pop-A, Pop-C/Pop-A, and Pop-D/Pop-A. This design was used because of the main objective of this study was to assess how selective pressures have affected populations cultivated in Chile, relative to their founding population, Pop-A, which was used as the reference population. Therefore, we excluded the comparisons between Chilean populations. The XP-EHH statistics were calculated as $\ln(I_{\text{PopO}}/I_{\text{PopR}})$, where I_{PopO} is the integrated EHH for the observed populations and I_{PopR} is the integrated EHH value of the reference population. Negative XP-EHH scores suggest selection in the “reference” population, whereas positive scores suggest selection acting in the “observed” population. A $-\log_{10}(p \text{ value})$ of three ($p \text{ value} \leq 0.001$) was used as the lower threshold for considering XP-EHH score as significant evidence of selection and at least two SNPs \leq 500 kb apart.

(3) BayeScan. We used the Bayesian likelihood method implemented in BayeSCAN v.2.1 to estimate the posterior probability that loci are experiencing selection (Foll and Gaggiotti 2008). This method models allele frequencies in subpopulations derived from a single ancestral population using Dirichlet distributions, which allows for estimating the degree of coancestry within each of these subpopulations through the sum of population-specific, β , and locus-specific, α , effects, making outlier detection robust to confounding complex demographic histories. By estimating the posterior probabilities for both the model including both effects and the model omitting the locus-specific effect, the posterior probability (and posterior odds) for selection at a specific locus can be obtained. When $\alpha > 0$ for a specific locus, it is evidence of directional selection acting on that locus, whereas $\alpha < 0$ suggests balancing or purifying selection. This method was run with 5,000 burn-in iterations, followed by 10,000 iterations with a thinning interval of 10. We evaluated the same three pairs of populations of XP-EHH method: Pop-B/Pop-A, Pop-C/Pop-A, and Pop-D/Pop-A. We considered candidate loci under selection as those having a Bayes factor of at least 32 ($-\log_{10} = 1.5$) and a positive value of α (directional selection), corresponding to a posterior probability of 0.97 and considered as being “very strong” evidence of selection and as in iHS and XP-EHH, we required the candidate selected regions to have at least two SNPs \leq 500 kb apart.

Gene Functional Annotation

Genomic regions harboring SNPs showing evidence of selection were annotated based on the ICSAG_v2 reference genome (Lien et al., 2016). We defined the position of the first and last SNP as boundaries of regions putatively under selection using BedTools (Quinlan and Hall, 2010). Gene transcripts from these candidate regions were aligned (using blastx) (Altschul et al., 1990) to the zebra fish (*Danio rerio*) peptide reference database (downloaded from <http://www.ensembl.org/>) to determine gene identifiers. As evidence of homology, we used an e-value ≈ 0 and then retrieved the zebra fish gene identifiers information from the ensemble biomaRt database (<http://www.ensembl.org/index.html>). Functional annotation of detected genes was performed using DAVID (Huang et al., 2009) with gene list of zebra fish (*Danio rerio*) as reference in Gene Ontology (GO) analysis.

RESULTS

Genetic Diversity and Structure

We performed PCA based on genotypes to look at the genetic relationship among individuals in our sample. The first and second components accounted for 14.2% and 10.3% of the genetic variation, respectively (Figure 1). Pop-A and Pop-C showed close genetic relationship to each other and were most distant to Pop-D from the Magallanes Region along PC1. Pop-B lies between the Pop-A/Pop-C cluster and Pop-D along PC1, with some overlap with Pop-C, which was introduced into the same Los Lagos Region as Pop-B. Overall, principal components showed low genetic variation between populations, but higher within populations, especially in Pop-D that exhibits the most difference among individuals along PC1. Also noteworthy is

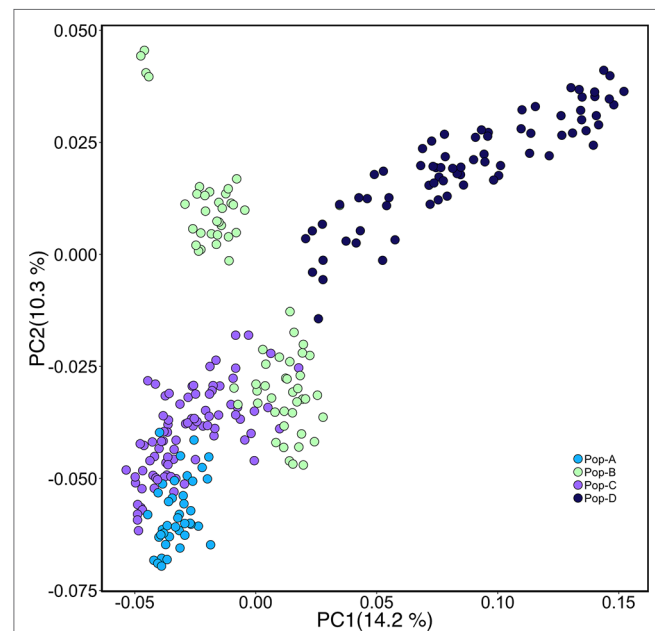


FIGURE 1 | Principal components analysis (PCA) of genetic differentiation among individuals. Each point represents one individual, and different colors represent populations.

that Pop-D, with the highest observed heterozygosity (Table 1), is uniformly farther to the other farmed populations, except for some individuals from Pop-B. We also performed an Admixture analysis to determine the composition of ancestral lineages among individuals. We found that 11 ancestral lineages were optimal for describing the ancestry of the individuals across the four populations (Figure 2). Consistent with the PCA and having the lowest heterozygosity, Pop-A individuals are all relatively the most similar among the populations in terms of their ancestral proportions, being dominated by one ancestral lineage. In contrast, Pop-D individuals tend to be dominated by a single ancestral lineage, but among individuals, the represented lineages are quite different, which is consistent with Pop-D individuals being quite different from each other in the PCA. Pop-B and Pop-C show similar degrees of mixed ancestry, though the dominant lineage is different between the two.

Observed heterozygosity levels were similar across the four domestic populations and were slightly higher than expected for populations A, B, and C, and even more so for population D. All these genetic diversity measures were statistically significant ($p < 0.05$, Kruskal–Wallis test) (see Table 1). Overall LD results revealed similar patterns for Pop-A and Pop-D, which presented longer range of LD and slower decay in comparison with Pop-B and Pop-C, that also presented similarity between them and a substantial faster LD decay (Figure 3). LD measures (r^2) of each chromosome and population are shown in Table S1 and Figure S1. Similar patterns were observed when the chromosomes were analyzed separately. Nevertheless, LD decay in Pop-A was noticeably stronger in chromosomes 2, 9, 19, and 29, whereas LD decay in Pop-D was stronger in chromosomes 13, 17, and 26 (Figure S1).

Candidate Regions Under Selection—iHS

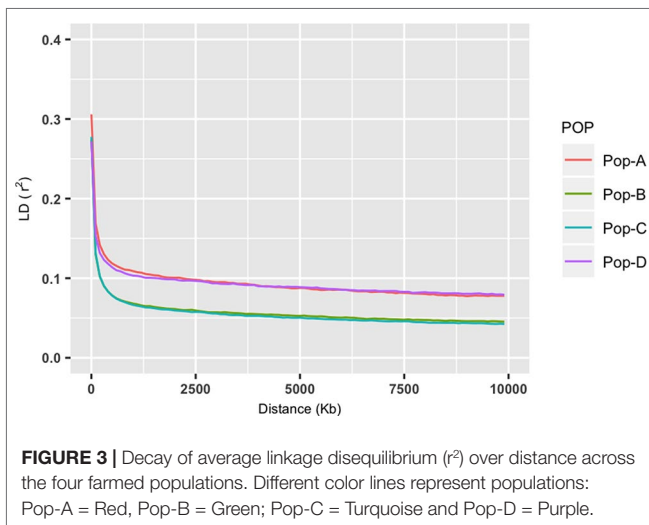
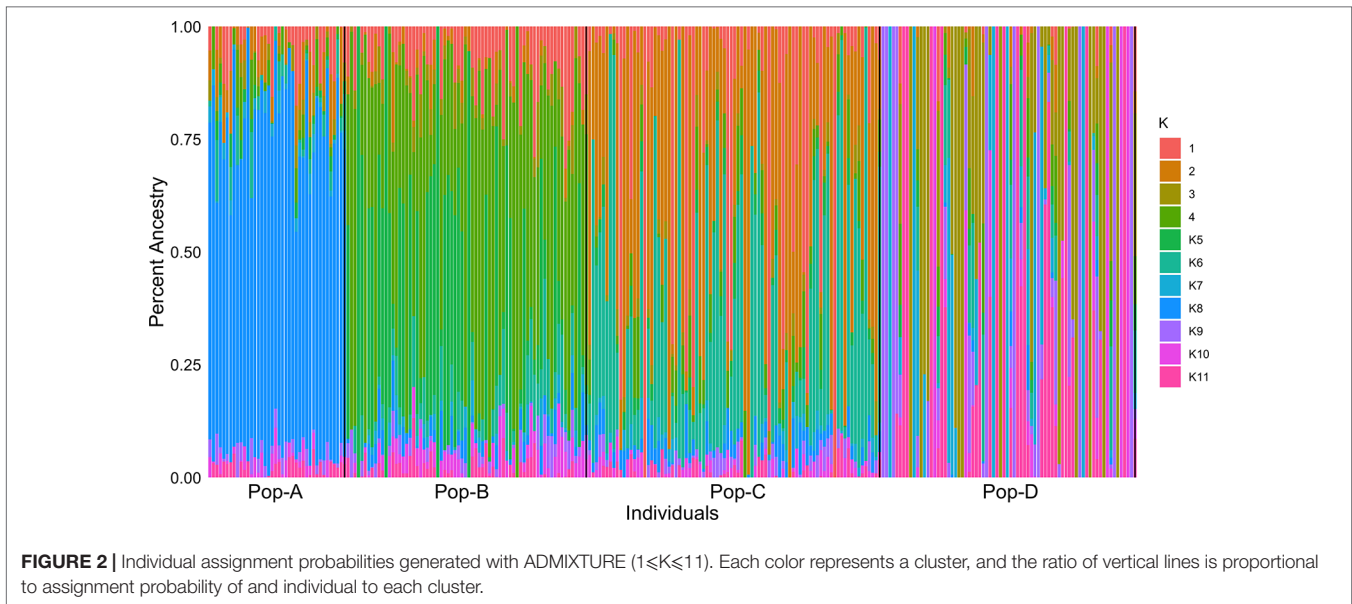
We looked for evidence of selection by comparing the decay of association between alleles from the major versus minor allele at core SNPs using iHS. We found 115, 63, 142, and 467 core SNPs with significant iHS statistics ($p \leq 0.001$) for Pop-A, -B, -C, and -D respectively (Figure 4, Table 2). We find 27, 12, 23, and 83 regions in these respective populations with at least two significant SNPs that are ≤ 500 kb apart, which we classify as putatively, selected regions.

Candidate regions for Pop-A were on Ssa01, Ssa05, and Ssa22. The candidate regions having SNPs with the most significant

TABLE 1 | Genetic diversity values in terms of Observed heterozygosity (H_o) and Expected heterozygosity (H_e) across four Atlantic salmon populations used in this study.

Population	H_o	H_e
Pop-A	0.4 ± 0.13	0.39 ± 0.11
Pop-B	0.41 ± 0.11	0.41 ± 0.1
Pop-C	0.41 ± 0.11	0.41 ± 0.1
Pop-D	0.47 ± 0.17	0.39 ± 0.11

All these genetic diversity measures were statistically significant ($p < 0.05$, Kruskal–Wallis test).

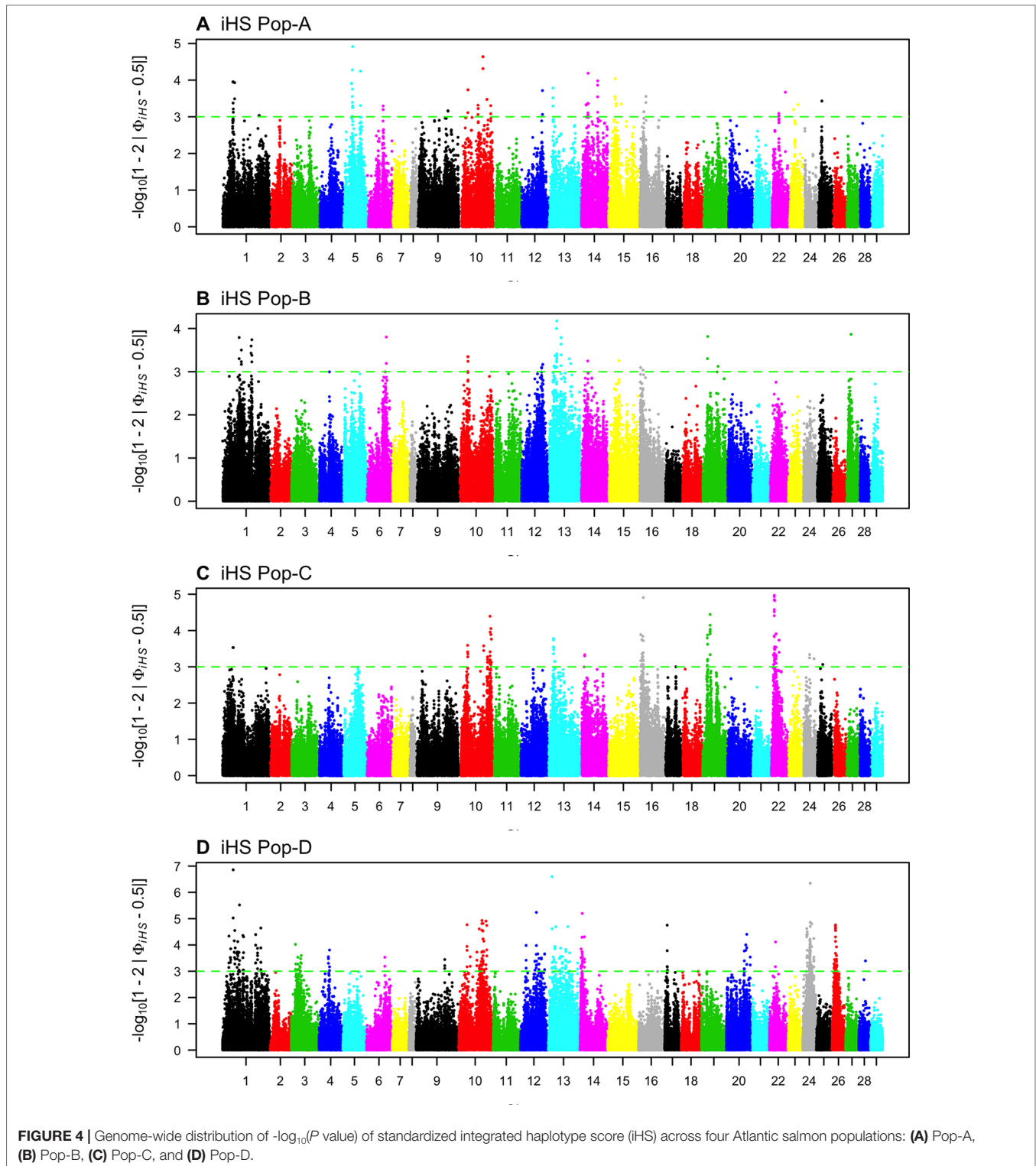


iHS scores were on Ssa05, Ssa10, and Ssa14, which contained the genes *igfbpl1* and *mipol1*.

Pop-B had 12 regions with an average length of ~ 250 kb putatively under selection distributed among five chromosomes. The highest iHS score was for a region found on Ssa13 [$-\log(p \text{ value}) = 4.17$] containing 26 genes including the *soga1* gene. Pop-C had 23 candidate regions that were on average ~ 370 kb long, and which spanned a total of 165 genes. The 1,570-kb-long region with one of the most significant iHS score was on Ssa22, and spanned the genes *kcnkf*, *sc61a*, and *mstn1*. Pop-D had the most significant number of SNPs (467) and had 83 putatively selected genomic regions under our criteria. Most of these regions were located on Ssa01, Ssa10, Ssa13, and Ssa26 and spanned genes, such as *haus2*, *itfg1*, and *phkb*. Details of the total regions and genes can be found in **Supplementary Tables S2** and **S5**, respectively.

Candidate Regions Under Selection—XP-EHH

We compared the decay of LD from a core SNP as measured by EHH between the Norwegian source population and the three derived Chilean populations (Pop-B/Pop-A, Pop-C/Pop-A, Pop-D/Pop-A) to detect regions having unusually high EHH and overrepresented haplotypes consistent with selection. In total, we detected 482 (Pop-B/Pop-A), 800 (Pop-C/Pop-A), and 207 (Pop-D/Pop-A) XP-EHH outlier SNPs indicative of selection (**Figure 5**, **Table 3**). The sign of the XP-EHH score indicates which population selection is acting on. Here, negative scores suggest selection in Pop-A. Most significant SNPs, which we considered as those with XP-EHH score $p \leq 0.001$, had negative scores, suggestive of selection in the Irish source population. The Pop-C/Pop-A and Pop-D/Pop-A comparisons yielded 38 and 3 significant SNPs with positive scores respectively, suggesting that the C and D populations underwent selection after their introduction into Chile. The significant, positive scores suggesting selection in Pop-C were found on Ssa16 within two regions spanning a total of 664.2 kb and which harbored 17 genes. The significant SNPs pointing to selection in Pop-D were located on Ssa14 in an 18.4-kb region, which contained the gene *agla*. XP-EHH did not detect selection signatures in Pop-B, as all significant scores for the Pop-B/Pop-A pair were negative. We classified potential genomic regions under selection as those containing two or more significant, adjacent SNPs less than 500 kb apart. After merging overlapping regions, we identified 34, 28, and 23 candidate regions from the Pop-B/Pop-A, Pop-C/Pop-A, and Pop-D/Pop-A comparisons respectively, which were all suggestive of selection in Pop-A. The average lengths of the candidate regions are approximately 338 kb for Pop-B/Pop-A, 546.5 kb for Pop-C/Pop-A, and 139 kb for Pop-D/Pop-A. Together, these regions span a total of 667 genes. Details of the total regions and genes detected by XP-EHH can be found in **Supplementary Table S3** and **S6**, respectively.



Candidate Regions Under Selection—*BayeScan*

We used the Bayesian approach for estimating the posterior odds of selection acting at particular loci based on pairwise divergence between ancestral and derived populations implemented in

BayeScan. By applying the BayeScan method to Pop-B/Pop-A, Pop-C/Pop-A, and Pop-D/Pop-A population pairs we, respectively, found 167, 155, and 193 SNPs with posterior odds ratios above 32, which was our threshold for showing significant evidence of selection (**Figure 6, Table 4**). F_{ST} -based methods do not directly

TABLE 2 | Ten genome regions spanning the strongest detected selection signatures by iHS in each population.

POP	CHR	START	END	$-\log(p)$	iHS	N SNPs	SIZE (kb)
Pop-A	1	35662318	35684677	3.9515	3.8634	4	22.4
	1	40728165	40728699	3.9306	3.8516	2	0.5
	5	25918328	25932901	3.9169	3.8439	2	14.6
	5	28372137	29065939	4.2779	4.0432	5	693.8
	5	29574408	29842752	4.9158	4.3751	4	268.3
	5	55278111	55732536	4.2454	4.0256	2	454.4
	10	79382450	79401333	4.6374	4.2331	3	18.9
	14	24674586	25715785	4.188	3.9944	6	1041.2
	14	56736246	57120611	3.98	3.8794	4	384.4
	15	22773166	23073140	4.0388	3.9122	6	300
Pop-B	1	55995699	56003301	3.7914	3.7725	2	7.6
	1	63381907	63519535	3.5006	3.602	3	137.6
	1	95895287	95964374	3.5929	3.6568	2	69.1
	1	98490168	98568189	3.7394	3.7425	3	78
	6	65448844	65496448	3.8025	3.7788	2	47.6
	10	29948442	30759289	3.347	3.509	4	810.8
	12	71810541	71833088	3.1679	3.3978	2	22.5
	13	22110373	22139660	3.371	3.5237	3	29.3
	13	27127510	28267153	4.1737	3.9866	10	1139.6
	13	41965178	42139618	3.8649	3.8144	15	174.4
Pop-C	10	104686655	105233083	4.3935	4.1051	10	546.4
	10	107544485	107633657	4.0535	3.9204	4	89.2
	16	6030690	6249119	3.8867	3.8268	3	218.4
	16	12985501	13367987	3.8395	3.7999	4	382.5
	16	14071951	14921512	4.9062	4.3703	7	849.6
	19	16762570	17099010	3.7932	3.7735	3	336.4
	19	17814180	18048311	3.877	3.8213	4	234.1
	19	26510295	26774029	4.4423	4.131	7	263.7
	22	16108465	17678398	4.9676	4.401	23	1569.9
	22	21553029	22158359	3.9096	3.8398	6	605.3
Pop-D	1	36254384	36351335	6.8581	5.267	2	97
	1	57451430	57811304	5.5208	4.6699	2	359.9
	10	80280055	81084212	4.9294	4.3819	6	804.2
	10	83621134	84353833	4.8176	4.3255	9	732.7
	10	93572235	93962188	4.9072	4.3708	3	390
	12	53402445	54250457	5.2386	4.5345	6	848
	13	14998846	15196522	6.6012	5.1573	5	197.7
	14	8208052	9149527	5.1987	4.5151	10	941.5
	24	24577538	26845034	6.3462	5.0462	54	2267.5

indicate in which population selection is acting; therefore, we describe our findings in terms of the population pairs. Since we expect regions that are truly under selection to have clusters of highly diverged SNPs in LD, we considered only regions containing at least two significant SNPs that were less than 500 kb adjacent to each other as being strong selection candidates. Under this criterion 104, 98, and 121 SNPs with posterior odds ratios of selection above 32 remain of interest for the Pop-B/Pop-A, Pop-C/Pop-A, and Pop-D/Pop-A comparisons, respectively. Clusters of SNPs identified as being in or adjacent to putatively selected regions from the Pop-B/Pop-A comparison represent 31 regions that are, on average, ~96.8 kb long and which harbored 58 genes. The Pop-C/Pop-A comparison showed 98 highly diverged regions among 29 regions that were, on average, ~220.7 kb long and which spanned 200 genes. Finally, the Pop-D/Pop-A comparison revealed 28 candidate regions that were, on average, ~153.6 kb long and contained 130 genes. Only two SNPs among these candidate regions showed evidence of selection among the three population pairs, which were located on Ssa29 in association with the *kmt2ca*

gene. Twenty SNPs suggestive of selection were shared between Pop-B/Pop-A and Pop-C/Pop-A and were associated to regions that harbor 12 genes *rabac1*, *znf1030*, *tpi1b*, *si:ch211-206a7.2*, *znf1041*, *lpcat3*, *atp1a3b*, *zgc:158654*, and *myh10* on Ssa02, *znf385d* on Ssa05, *agbl4* on Ssa10, and CR388166.1, and *kmt2ca* on Ssa29. Four candidate SNPs were common to Pop-C/Pop-A and Pop-D/Pop-A and two between Pop-B/Pop-A and Pop-D/Pop-A, which correspond to the *kmt2ca* gene shared among three population pairs. Details of the total regions and genes detected by BayeScan can be found in **Supplementary Tables S4** and **S7**, respectively.

Gene Ontology for Candidate Genes Under Selection

To further explore the functions of the candidate genes spanned by regions showing evidence of selection from the iHS, XP-EHH, and BayeScan analyses, we annotated the candidate genes using the DAVID browser (<https://david-d.ncicrf.gov>). The candidate genes were enriched in 37 gene ontology (GO) terms overall, most of them

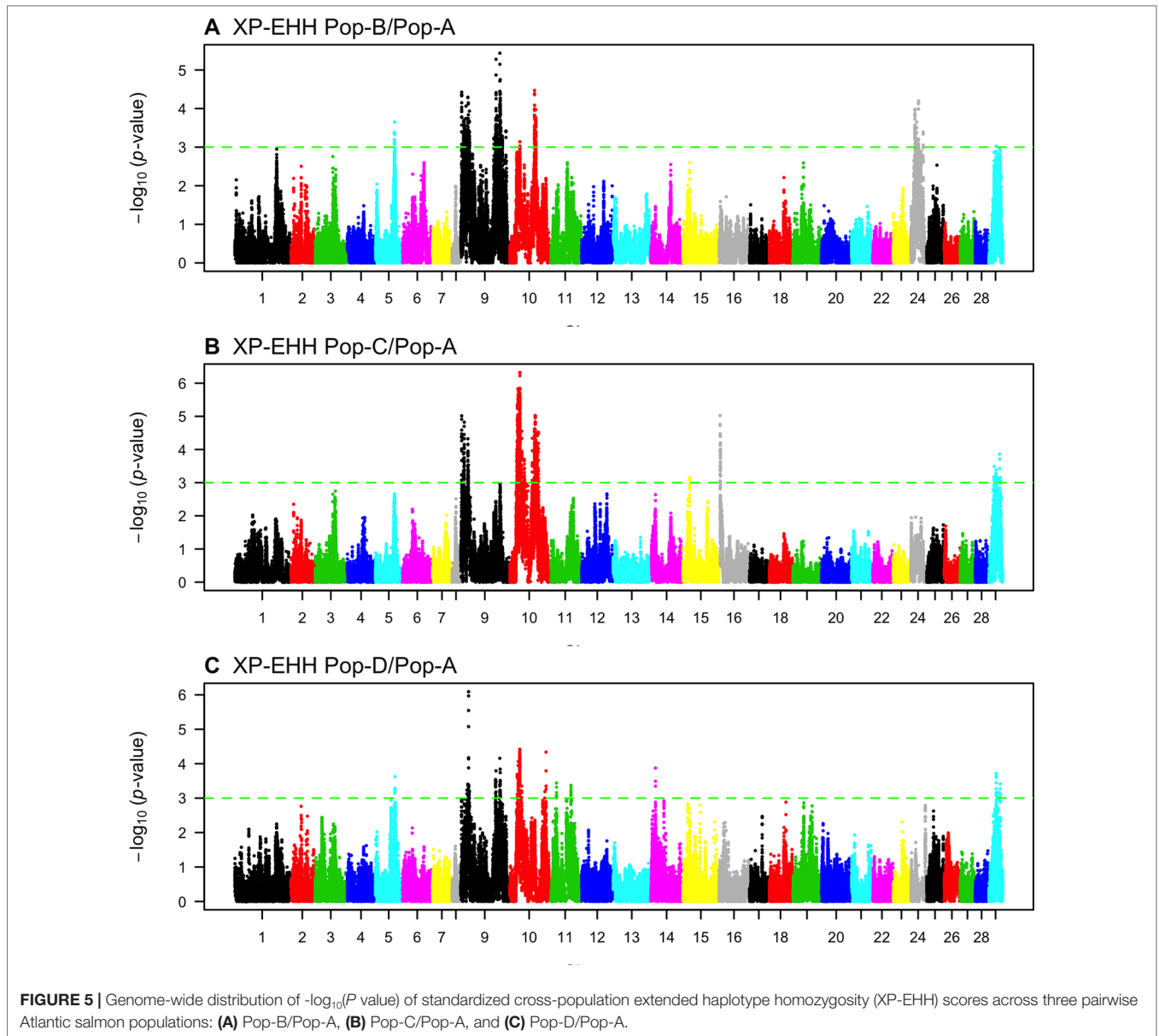


FIGURE 5 | Genome-wide distribution of $-\log_{10}(P \text{ value})$ of standardized cross-population extended haplotype homozygosity (XP-EHH) scores across three pairwise Atlantic salmon populations: **(A)** Pop-B/Pop-A, **(B)** Pop-C/Pop-A, and **(C)** Pop-D/Pop-A.

TABLE 3 | Genome regions spanning the strongest detected selection signatures by XP-EHH in populations A, C and D.

POP	CHR	START	END	$-\log(P)$	XP-EHH	N SNPs	Size (kB)
Pop-A	10	28741972	31140475	6.324	-5.0365	153	2398.5
	9	23728364	24144245	6.0912	-4.9328	15	415.9
	10	24160722	26099914	5.8281	-4.8132	127	1939.2
	9	113910288	114187655	5.4365	-4.6298	30	277.4
	10	21739331	23180890	5.3423	-4.5847	54	1441.6
	9	101786257	103293781	5.2766	-4.553	56	1507.5
	10	73472292	74738689	5.021	-4.4276	73	1266.4
	9	3674860	4026195	5.0148	-4.4245	14	351.3
	9	11161334	11559014	4.823	-4.3282	19	397.7
	9	114997862	115904242	4.7511	-4.2916	44	906.4
Pop-C	16	3564058	3808523	5.0253	3.2964	13	244.5
	16	4345514	4765204	4.4036	3.2915	25	419.7
Pop-D	14	14636389	14654809	3.872	3.5093	3	18.4

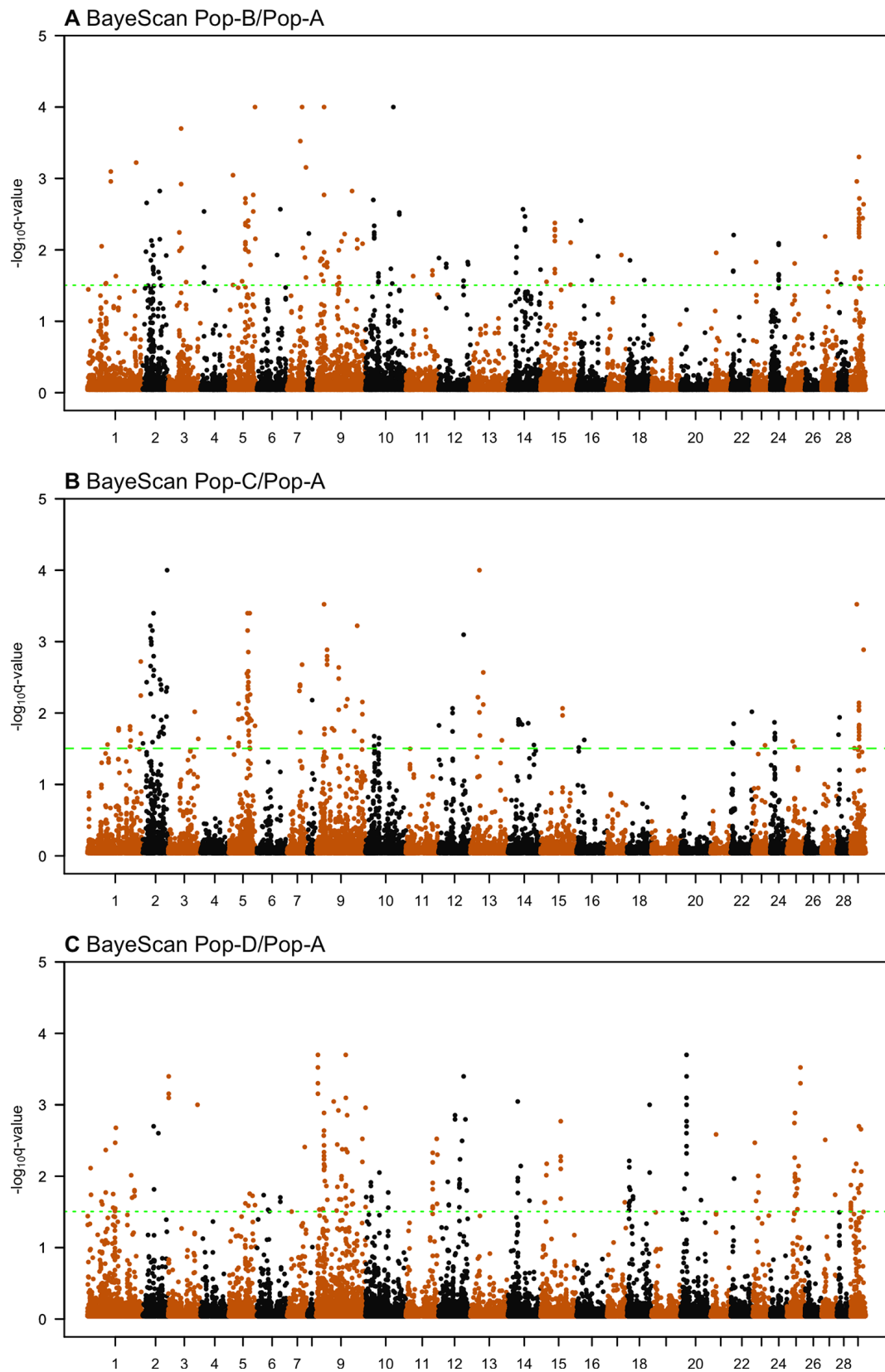


FIGURE 6 | Genome-wide distribution of $-\log_{10}(q \text{ value})$ in BayeScan analysis across three pairwise Atlantic salmon populations: **(A)** Pop-B/Pop-A, **(B)** Pop-C/Pop-A, and **(C)** Pop-D/Pop-A.

TABLE 4 | Ten genome regions spanning the strongest detected selection signatures by BayeScan method in each population pair.

POP	CHR	START	END	$-\log(q \text{ value})$	N SNPs	SIZE (kb)
Pop-B/Pop-A	1	66642439	66648870	3.097	2	6.4
	2	48416445	48567681	2.824	3	151.2
	3	37023703	37052183	3.699	2	28.5
	5	11553116	11556394	3.046	2	3.3
	5	47250892	47494177	2.721	8	243.3
	5	69864532	69865664	2.770	2	1.1
	7	53824902	53839042	3.155	2	14.1
	9	22192894	22527645	4.000	4	334.8
	29	23820153	24379023	3.301	10	558.9
	29	25107616	25137079	2.721	6	29.5
Pop-C/Pop-A	2	21743330	22285719	3.222	4	542.4
	2	24203593	24203644	3.000	2	0.1
	2	27316859	27731651	3.155	2	414.8
	2	30394158	31352454	3.398	5	958.3
	2	69206072	69622942	4.000	2	416.9
	5	52915411	53613743	3.398	5	698.3
	5	59616884	59678559	3.398	4	61.7
	9	30961027	30994613	2.886	4	33.6
	13	25691366	25715347	4.000	2	24.0
	29	23852604	24289616	2.886	12	437.0
Pop-D/Pop-A	3	1316421	1317893	3.398	3	1.5
	9	4536926	4590569	3.699	4	53.6
	9	22080850	22146356	2.886	9	65.5
	9	84229138	85051608	3.699	6	822.5
	9	141700047	141700106	2.959	2	0.1
	14	28094905	28343120	3.046	4	248.2
	18	64503159	64648197	3.000	2	145.0
	20	17158525	17477234	3.699	10	318.7
	25	22716335	22760611	2.886	6	44.3
	25	38351034	38355710	3.523	2	4.7

population specific (Table 5). Four GO categories were common between Pop-A and Pop-B (single-multicellular organism process, single-organism developmental process, regulation of metabolic process, and anatomical structure development) and one between Pop-C and Pop-D (animal organ development). The remaining GO categories were unique to each population.

DISCUSSION

In this study, we used three complementary tests to detect selection signatures within and between four Atlantic salmon populations with Norwegian origin. We used the iHS test to scan for selection signatures within populations and XP-EHH and BayeScan to find evidence of selection in terms of divergence of the Chilean populations to their ancestral Irish population. We detected several genomic regions under putative selection across all of the populations evaluated, which provides insight into the genes contributing to traits of importance to Atlantic salmon farming. It is important to mention that these findings should be interpreted with caution since other evolutionary and demographic process, such as bottlenecks and differences in the amount of genetic drift resulting from different effective populations sizes, can produce patterns of genetic diversity that mimic selection leading to the finding of possible false positives as well. However, the selection detection methods we used have all been shown to be robust to these confounding effects.

Structure and Diversity

To examine genetic population structure and relationships among the major groups of salmon, we conducted an ADMIXTURE analyses based on high-quality SNP data. This analysis revealed that 12 ancestral lineages contribute to the modern gene pool represented by the four farmed populations, which was expected considering the admixed origin of these populations (Verspoor et al., 2007). The four populations used in this study are derived from the Mowi strain, which was created using samples from several rivers along the west coast of Norway (Norris et al., 1999). The population with the lowest level of admixture was Pop-A, which was also the population with the lowest genetic diversity, a condition that could reflect a better culture management, as well as intense artificial selection that erodes genetic variation through mating related individuals (Gjedrem, 2005). Pop-B and Pop-C which were introduced into the same region in Chile have very similar amounts of heterozygosity and similar degrees of admixture though the dominant lineages are different, which was expected due to the similar breeding practices and environmental conditions to which they have been subjected. Pop-D, however, showed the highest level of heterozygosity and a more complex pattern of admixture, whereby a single ancestral lineage is highly represented within individuals but with many ancestral lineages present among individuals. This pattern may, in part, reflect lower artificial selection pressure. Recent genetic introgression cannot be ruled out for Pop-D given the potential for crossing with different strains for management reasons. LD analysis revealed that overall

TABLE 5 | Biological processes enriched in genes detected by iHS and XP-EHH in each Atlantic salmon population.

Population	Biological Process	GO Term	%	p	Benjamini
Pop-A	Cellular metabolic process	GO:0044237	36.8	3.0E-4	3.7E-2
	Organic substance metabolic process	GO:0071704	38.7	9.4E-4	5.6E-2
	Primary metabolic process	GO:0044238	37.1	1.2E-3	4.8E-2
	Catabolic process	GO:0009056	5.7	2.3E-2	5.1E-1
	Single-multicellular organism process	GO:0044707	19.1	4.7E-2	7.0E-1
	Developmental induction	GO:0031128	0.4	5.9E-2	7.1E-1
	Single-organism developmental process	GO:0044767	19.2	6.7E-2	7.1E-1
	Regulation of metabolic process	GO:0019222	14.0	7.6E-2	7.0E-1
	Anatomical structure development	GO:0048856	19.1	9.7E-2	7.5E-1
Pop-B	Regulation of signaling	GO:0023051	14.5	8.8E-3	4.7E-1
	Regulation of cellular process	GO:0050794	45.2	1.4E-2	4.0E-1
	Regulation of metabolic process	GO:0019222	22.6	3.8E-2	6.0E-1
	Anatomical structure morphogenesis	GO:0009653	17.7	4.8E-2	5.9E-1
	Regulation of response to stimulus	GO:0048583	12.9	5.0E-2	5.2E-1
	Cellular component organization	GO:0016043	24.2	5.1E-2	4.7E-1
	Single-organism developmental process	GO:0044767	27.4	6.2E-2	4.8E-1
	Anatomical structure development	GO:0048856	27.4	6.6E-2	4.6E-1
	Single-multicellular organism process	GO:0044707	25.8	9.8E-2	5.6E-1
	Methylation	GO:0032259	4.8	9.9E-2	5.3E-1
Pop-C	Heart development	GO:0007507	5.3	2.4E-2	1.0E0
	Regulation of cell communication	GO:0010646	8.8	2.7E-2	9.7E-1
	Regulation of signal transduction	GO:0009966	8.2	2.8E-2	9.2E-1
	Animal organ development	GO:0048513	14.7	3.0E-2	8.8E-1
	Organ morphogenesis	GO:0009887	6.5	3.3E-2	8.4E-1
	Digestive tract development	GO:0048565	2.4	3.7E-2	8.2E-1
	Muscle system process	GO:0003012	2.4	4.9E-2	8.6E-1
	Tissue development	GO:0009888	9.4	6.3E-2	8.9E-1
	Cellular developmental process	GO:0048869	13.5	6.5E-2	8.7E-1
	Phosphorus metabolic process	GO:0006793	12.4	8.1E-2	9.0E-1
Pop-D	System development	GO:0048731	17.6	9.7E-2	9.2E-1
	Pancreas development	GO:0031016	1.7	5.1E-3	9.0E-1
	Cellular lipid metabolic process	GO:0044255	4.0	6.3E-3	7.6E-1
	Regulation of blood pressure	GO:0008217	1.0	1.1E-2	8.1E-1
	Lipid metabolic process	GO:0006629	4.7	1.6E-2	8.3E-1
	Gland development	GO:0048732	2.2	1.6E-2	7.7E-1
	Forebrain development	GO:0030900	1.5	2.6E-2	8.6E-1
	Small molecule metabolic process	GO:0044281	6.2	4.4E-2	9.5E-1
	Atrioventricular canal development	GO:0036302	0.5	5.0E-2	9.5E-1
	Organic acid metabolic process	GO:0006082	3.5	5.1E-2	9.3E-1
	Embryonic organ development	GO:0048568	3.5	7.0E-2	9.6E-1
	Animal organ development	GO:0048513	11.4	8.6E-2	9.7E-1
	Single-organism biosynthetic process	GO:0044711	4.2	9.2E-2	9.7E-1

LD decays more rapidly in Pop-B and Pop-C over short physical distances and is lower than Pop-A and Pop-D. The pattern of LD in Pop-A is consistent with its lower heterozygosity level. However, similar pattern was observed in Pop-D, likely due to higher level of admixture in this population, where several ancestral lineages can be observed. Chromosomal LD decay followed similar patterns, but in Pop-A, LD decay was noticeably higher in chromosomes 2, 9, 11, 19, and 29, which is agreed with a greater number of regions detected under selection in those chromosomes. Conversely, in chromosome 26, Pop-D showed the highest value of LD ($r^2 = 0.12$), probably related to a larger region under selection detected in this population. The results presented here also reinforce the notion that exposure to different management and environmental conditions over just a few generations (at least four in this particular case) is sufficient to generate large changes in the genetic structure of farmed Atlantic salmon populations with the same genetic origin.

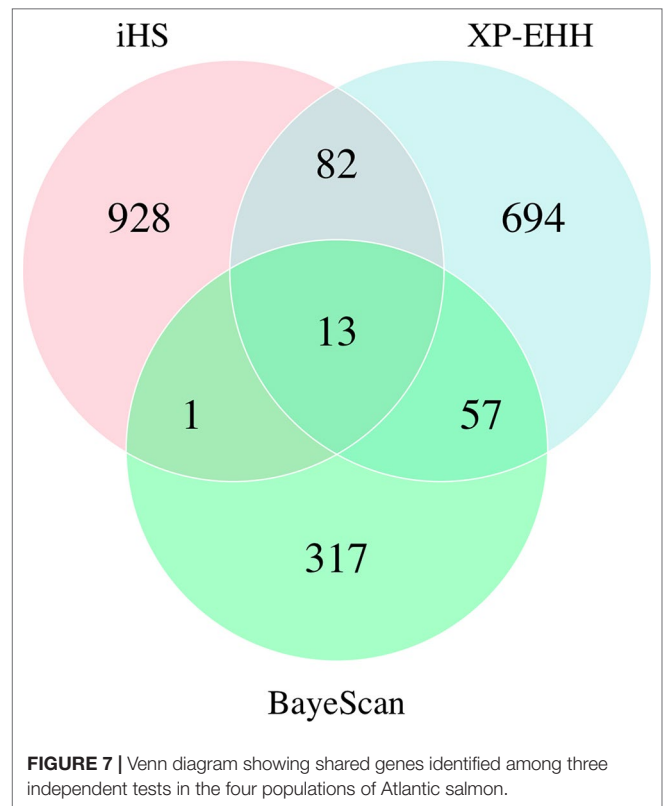
Selection Signatures

Pop-D had regions showing the strongest evidence for selection as well as the most candidate regions according to the iHS test. Although the iHS test has a lower power to detect selection under nearly complete sweeps (Sabeti et al., 2007; Simianer et al., 2010), it has greater power when selected alleles are at intermediate frequencies. Pop-D has experienced weaker artificial selection pressure than the other populations used in this study (Jean Paul Lhorente, personal communication), and so the higher number of putatively selected regions identified in this population by iHS may reflect more sweeps at intermediate frequencies because they are taking relatively longer to complete under weaker selection. In addition, this population is located in the Magallanes Region in Chile, which exposes salmon to more extreme environmental conditions than in the Los Lagos region where Pop-B and Pop-C were introduced. Therefore, the selection imposed by the natural

environmental may also contribute to a relatively high number of selected regions in Pop-D. In contrast to iHS, XP-EHH is powerful at detecting complete or nearly complete selective sweeps (Sabeti et al., 2007). According to the XP-EHH method, Pop-A shows the greatest number of regions under selection across the genome, which is consistent with XP-EHH having greater power to identify selection in regions that experienced older selection events (Sabeti et al., 2007; Klimentidis et al., 2011) than iHS since Pop-A is the oldest population in the present study while also being subjected to more intense artificial selection. We identified several putative directional selection targets using BayeScan, but given the nature of F_{ST} -based methods we are unable to directly identify which population in a pairwise comparison is experiencing selection from the posterior odds alone. Low overlap in selected regions identified with haplotype-based and single-SNP F_{ST} -based approaches have been reported in other studies in Atlantic salmon (Mäkinen et al., 2014; López et al., 2018) and other species (Bahbahani et al., 2015). However, we did find some degree of overlap among genes detected by both haplotype methods and the F_{ST} method as shown in Figure 7 and Table 6.

Biological Function of Candidate Selected Regions

Geographical adaptation and selection in farmed Atlantic salmon has resulted in considerable differences between wild and farmed strains (Glover et al., 2009). Genomic regions detected in this study strongly suggest selection on traits that could be associated with either natural or artificial selection, as they relate to the immune system, growth, and behavior, which are all often altered through domestication. Growth has been the main trait focused on by the breeding programs represented by our focal salmon populations. In agreement with this, we found several genes showing evidence of selection that could be potentially influencing growth such as *chp2* and *ccser1*, which were associated with body weight in a previous genome-wide association study (GWAS) on Atlantic salmon (Yoshida



et al., 2017). We detected the *kind1* gene that is also associated with growth traits in juvenile, farmed Atlantic salmon (Tsai et al., 2015). It has also been shown that insulin growth factors (IGFs), IGF receptors, and IGF binding proteins, play an important role in regulating growth in several teleost fish species (Duan, 1997). We detected the IGF 1-receptor (*igf1r*), IGF binding protein 6 paralog A2 (*igfbp-6a2*), and IGF binding protein-related protein 1 precursor (*igfbprp1*) as being under selection. We hypothesize that these genes are all contributing to weight variation in farmed salmon. The GO analyses for our

TABLE 6 | Genes detected by at least two selection signatures methods. Genes are indicated in the left column and in the right column their corresponding methods.

GENES	METHODS
CRISP3, NOTCHL, GPSM1B, SI : ZFOS-367G9.1, PHF1, FQ976914.1, TAP1, PBX2, DNASE2, RGL2, PLCL2, SYNGAP1B, BRD2A	iHS; XP-EHH; BayeScan
CRISP3, SI : ZFOS-367G9.1, NOTCHL, GPSM1B, PHF1, FQ976914.1, TAP1, PBX2, DNASE2, RGL2, PLCL2, SYNGAP1B, BRD2A, DOCK10, CRK, LRRC75A, SI : CH211-232I5.3, BLOC1S2, SI : DKEYP-51F12.3, CEP120, CABZ01077978.1, SI : CH211-232I5.1, PRKAA1, PLPP3, BX546500.1, DHCR24, USP24, DAB1A, PRDM5, ANAPC4, SLC10A4, FRYL, PALLD, SLAIN2, MOGAT3B, C1QTNF7, FTR14, LRRC66, SGCB, RASL11B, NDNF, ZBTB34, CPEB2, CC2D2A, FBXL5, NEK1, SH3RF1, OCIAD2, DCUN1D4, USP46, OCIAD1, SCFD2, CDKN1BB, YARS2, PPARAB, BX537249.1, JPH3, KLHDC4, SLC7A5, HMCN2, CDH13, RANBP10, NUTF2, EDC4, NRN1LA, MBTPS1, SLC38A8, PNP6, CALB2A, PSKH1, NECAB2, SCAPER, PSTPIP1A, THBS4A, SERINC5, TRAFD1, SMTNB, UBE2G1B, ANAPC7, ADORA2AA, GUCD1, TAS2R200.1, GSTT1A, DERL3, SMARCB1A, ATP2A2A, BCR, SPECC1LA, SI : CH211-191O15.6, SNRPD3, P2RX7, MMP11A, RALGDS, IFT81, MPEG1.1	iHS; XP-EHH
UNC13B, CRISP3, SI : ZFOS-367G9.1, NOTCHL, GPSM1B, PHF1, FQ976914.1, TAP1, PBX2, DNASE2, RGL2, PLCL2, SYNGAP1B, BRD2A	iHS; BayeScan
NXP1, ICA1, MIOS, COL28A1B, TAC1, SEPT7B, NEK10, NR1D2B, PHLPP1, RAB5AB, EFHB, CRISP3, SI : ZFOS-367G9.1, NOTCHL, GPSM1B, PHF1, GLCCI1, COL28A1A, RARGA, UBE2E2, ZNF385D, SATB1, FQ976914.1, TAP1, PBX2, DNASE2, RGL2, PLCL2, SYNGAP1B, BRD2A, CTSS2.1, STARD13A, VASH1, OLFM4, RPS6KL1, AREL1, FCF1, ANGEL1, DLST, ESRRB, GPATCH2, TGFB3, PROX2, TMEM179, ARHGEF18B, CABZ01071407.1, ATXN3, SERPINA10, FOXP1A, SI : DKEY-206P8.1, DDX24, SI : CH1073-416D2.4, PRIMA1, UBR7, ITPK1B, HSPA4L, MRPL35, SI : DKEY-21A6.5, CABZ01052815.1, CABZ01066926.1, CHMP3, REEP1, BTBD7, PLK4, MYO1CB, AGBL4, MYL2B, PPP1CC, MTMR3, CUX2B	XP-EHH; BayeScan

candidate genes also showed enrichment for categories related to metabolic and developmental processes, which could certainly affect growth.

Genes functioning in host–pathogen interactions may be targets of natural selection more often than genes from other functional categories (Schlenke and Begun, 2003). The populations used in this study have not been artificially selected for disease resistance; however, we suspect that the culture environment has imposed natural selection on regions implicated in immune system function. We found evidence of selection in seven genes (*kcnb2*, *rlf*, *synrg*, *snx14*, *fbx15*, *e2f4*, *blm*) that were previously shown to be affected by parasite-driven selection (Zueva et al., 2014). We also identified three genes potentially under selection (*kcnq1*, *lrp5*, and *sh3rf1*) that have been associated with disease resistance in the face of a bacterial disease (*Piscirickettsia salmonis*) in Coho salmon (Barría et al., 2018) and *mettl12* which is associated with immune response to parasites in three-spined stickleback (Huang et al., 2016).

Behavioral traits are among the first traits affected by animal domestication (Kohane and Parsons, 1988), and it has been suggested that domestication may impact behavior even after only one generation (Huntingford, 2004). Among our candidate genes putatively under selection, we identified the endoplasmic reticulum protein 27 (*erp27*) gene, the differential expression of which has been associated to tameness in the red junglefowl (Bélteky et al., 2016). Also, among our candidates were genes, such as *gabbr1*, *scaper*, *clstn3*, and *pex5*, related to mental disorders in humans such as alcoholism and schizophrenia (Glatt et al., 2005; Enoch, 2008; Pettem et al., 2013). We think that these genes may be influencing behavior in the salmon populations we studied, and that the artificial selection and domestication could be acting inadvertently on the traits affected by these genes like those that occur in other domestic animals (Clutton-Brock, 1999).

In salmon culture, early sexual maturation has undesired consequences, such as decreased growth and feed conversion efficiency (Good and Davidson, 2016). To avoid these negative effects, maturation is commonly delayed by exposing fish to continuous light, which affects the perception of seasonality and circannual rhythms (Taranger et al., 2010). We would expect then to find genes underlying traits related to maturation rate as showing signs of selection, which we apparently do. One putatively selected gene that we found that may affect maturation rate is *akap13*, which has been shown to play a role in ovarian development in human (Wu et al., 2015), as well as a gene in the AKAP (*akap11*) family, which was previously associated with age to maturity in Atlantic salmon (Barson et al., 2015).

Other interesting genes spanned by regions showing evidence for selection in this study are *hao1*, which is associated with chicken sexual ornaments (comb size), *myo3a*, which is involved in allowing dogs to sense local environmental stimuli (Wang et al., 2013), and *pgbd4*, which is considered a candidate gene involved in adaptation at the regional scale in Atlantic salmon

(Bourret et al., 2013) and so could be functioning in adaptation to the aquaculture environment.

CONCLUSIONS

To summarize, in this study we used three different but complementary statistical approaches, iHS, XP-EHH, and BayeScan to detect selection signatures in four farmed Atlantic salmon populations with the same geographical origin, but adapted to different environmental conditions. The methods used in this study were useful for detecting selection signals across populations and allowed us to find genes that could be related to growth, immune system function, and behavior in this species, characters that are commonly influenced by domestication. This study provides potential candidate genes for traits with both biological and economic importance for Atlantic salmon and establishes a strong platform for further studies seeking to better understand how particular genomic variants influence the evolution and cultivation of this species.

ETHICS STATEMENT

The sampling protocol was previously approved by The Comité de Bioética Animal, Facultad de Ciencias Veterinarias y Pecuarias, Universidad de Chile (certificate 29-2014).

AUTHOR CONTRIBUTIONS

ML and JY conceived the research idea. ML drafted the manuscript and carried out the analyses. TL supervised the data analyses and contributed to discussion and writing. TL, AN, JL, RN, and JY reviewed the manuscript. All authors read and approved the final manuscript.

FUNDING

This work has been conceived on the frame of the grant CORFO (11IEI-12843 and 12PIE17669), Government of Chile.

ACKNOWLEDGMENTS

ML acknowledges the National Commission of Scientific and Technologic Research (CONICYT) for the funding through the National PhD funding program. JY is supported by Núcleo Milenio INVASAL funded by Chile's government program, Iniciativa Científica Milenio from Ministerio de Economía, Fomento y Turismo.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2019.00901/full#supplementary-material>

REFERENCES

- Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19(9), 1655–1664. doi: 10.1101/gr.094052.109
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2
- Andersson, L. (2012). How selective sweeps in domestic animals provide new insight into biological mechanisms. *J. Intern. Med.* 271, 1–14. doi: 10.1111/j.1365-2796.2011.02450.x
- Avila, F., Mickelson, J. R., Schaefer, R. J., and McCue, M. E. (2018). Genome-wide signatures of selection reveal genes associated with performance in American quarter horse subpopulations. *Front. Genet.* 9, 249. doi: 10.3389/fgene.2018.00249
- Bahbahani, H., H. Clifford, D. Wragg, M.N. Mbole-Kariuki, C. Van Tassell, T. Sonstegard et al. (2015). Signatures of positive selection in East African Shorthorn Zebu: a genome-wide single nucleotide polymorphism analysis. *Sci. Rep.* 5, 11729. doi: 10.1038/srep11729
- Barria, A., Christensen, K. A., Yoshida, G. M., Correa, K., Jedlicki, A., Lhorente, J. P. et al. (2018). Genomic Predictions and Genome-Wide Association Study of Resistance Against *Piscirickettsia salmonis* in Coho Salmon (*Oncorhynchus kisutch*); Using ddRAD Sequencing. *G3: Genes, Genomes, Genet.* 8, 1183. doi: 10.1534/g3.118.200053
- Barson, N. J., Aykanat, T., Hindar, K., Baranski, M., Bolstad, G. H., Fiske, P. et al. (2015). Sex-dependent dominance at a single locus maintains variation in age at maturity in salmon. *Nature* 528, 405–408. doi: 10.1038/nature16062
- Beaumont, M. A., and Balding, D. J. (2004). Identifying adaptive genetic divergence among populations from genome scans. *Mol. Ecol.* 13, 969–980. doi: 10.1111/j.1365-294X.2004.02125.x
- Bélteky, J., Agnvall, B., Johnsson, M., Wright, D., and Jensen, P. (2016). Domestication and tameness: brain gene expression in red junglefowl selected for less fear of humans suggests effects on reproduction and immunology. *R. Soc. Open Sci.* 3, 160033–160033. doi: 10.1098/rsos.160033
- Bourret, V., Dionne, M., Kent, M. P., Lien, S., and Bernatchez, L. (2013). Landscape genomics in Atlantic salmon (*Salmo salar*): searching for gene–environment interactions driving local adaptation. *Evolution* 67, 3469–3487. doi: 10.1111/evo.12139
- Browning, B., and Browning, S. (2009). A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am. J. Hum. Genet.* 84, 210–223. doi: 10.1016/j.ajhg.2009.01.005
- Cesconeto, R. J., Joost, S., McManus, M. C., Paiva, S. R., Cobuci, J. A., and Braccini, J. (2017). Landscape genomic approach to detect selection signatures in locally adapted Brazilian swine genetic groups. *Ecol. Evol.* 7, 9544–9556. doi: 10.1002/ece3.3323
- Clutton-Brock, J. (1999). *A natural history of domesticated mammals*. Cambridge, UK: Cambridge University Press.
- Driscoll, C. A., Macdonald, D. W., and O'Brien, S. J. (2009). From wild animals to domestic pets, an evolutionary view of domestication. *Proc. Natl. Acad. Sci.* 106, 9971–9978. doi: 10.1073/pnas.0901586106
- Duan, C. (1997). The insulin-like growth factor system and its biological actions in fish1. *Integr. Comp. Biol.* 37, 491–503. doi: 10.1093/icb/37.6.491
- Einum, S., and Fleming, I. (1997). Genetic divergence and interactions in the wild among native, farmed and hybrid Atlantic salmon. *J. Fish Biol.* 50 (3), 634–651. doi: 10.1111/j.1095-8649.1997.tb01955.x
- Enoch, M.-A. (2008). The role of GABA(A) receptors in the development of alcoholism. *Pharmacol. Biochem. Behav.* 90, 95–104. doi: 10.1016/j.pbb.2008.03.007
- Excoffier, L., Hofer, T., and Foll, M. (2009). Detecting loci under selection in a hierarchically structured population. *Heredity* 103, 285–298. doi: 10.1038/hdy.2009.74
- FAO (2016). *The State of World Fisheries and Aquaculture 2016*. Contributing to food security and nutrition for all 200 pp.
- Foll, M., and Gaggiotti, O. (2008). A Genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* 180, 977–993. doi: 10.1534/genetics.108.092221
- Gautier, M., and Vitalis, R. (2012). rehh: an R package to detect footprints of selection in genome-wide SNP data from haplotype structure. *Bioinformatics* 28, 1176–1177. doi: 10.1093/bioinformatics/bts115
- Gjedrem, T. (2005). *Selection and Breeding Programs in Aquaculture*. Dordrecht, The Netherlands: Springer. doi: 10.1007/1-4020-3342-7
- Gjedrem, T. (2010). The first family-based breeding program in aquaculture. *Rev. Aquacult.* 2, 2–15. doi: 10.1111/j.1753-5131.2010.01011.x
- Gjedrem, T. (2012). Genetic improvement for the development of efficient global aquaculture: a personal opinion review. *Aquaculture* 344–349, 12–22. doi: 10.1016/j.aquaculture.2012.03.003
- Gjedrem, T., Robinson, N., and Rye, M. (2012). The importance of selective breeding in aquaculture to meet future demands for animal protein: a review. *Aquaculture* 350–353, 117–129. doi: 10.1016/j.aquaculture.2012.04.008
- Glatt, S. J., Everall, I. P., Kremen, W. S., Corbeil, J., Šašik, R., Khanlou, N., et al. (2005). Comparative gene expression analysis of blood and brain provides concurrent validation of SELENBP1 up-regulation in schizophrenia. *Proc. Natl. Acad. Sci. U.S.A.* 102, 15533–15538. doi: 10.1073/pnas.0507666102
- Glover, K., H. Ottera, R. Olsen, E. Slinde, G. Taranger and Ø. Skaala (2009). A comparison of farmed, wild and hybrid Atlantic salmon (*Salmo salar* L.) reared under farming conditions. *Aquaculture* 286 (3–4), 203–210. doi: 10.1016/j.aquaculture.2008.09.023
- Glover, K. A., Solberg, M. F., McGinnity, P., Hindar, K., Verspoor, E., Coulson, M. W., et al. (2017). Half a century of genetic interaction between farmed and wild Atlantic salmon: status of knowledge and unanswered questions. *Fish Fish. (oxf)* 18, 890–927. doi: 10.1111/faf.12214
- Good, C., and Davidson, J. (2016). A review of factors influencing maturation of Atlantic salmon, *Salmo salar*, with focus on water recirculation aquaculture system environments. *J. World Aquacult. Soc.* 47, 605–632. doi: 10.1111/jwas.12342
- Gurgul, A., Jasielczuk, I., Ropka-Molik, K., Semik-Gurgul, E., Pawlina-Tyszko, K., Szmatoła, T., et al. (2018). A genome-wide detection of selection signatures in conserved and commercial pig breeds maintained in Poland. *BMC Genet.* 19, 95–95. doi: 10.1186/s12863-018-0681-0
- Gutierrez, A. P., Yáñez, J. M., and Davidson, W. S. (2016). Evidence of recent signatures of selection during domestication in an Atlantic salmon population. *Mar. Genomics* 26, 41–50. doi: 10.1016/j.margen.2015.12.007
- Hill, W. G., and Bunger, L. (2004). Inferences on the genetics of quantitative traits from long-term selection in laboratory and domestic animals. *Plant Breed. Rev.* 24, 169–210. doi: 10.1002/9780470650288.ch6
- Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 2009;4, 44–57. doi: 10.1038/nprot.2008.211
- Huang, Y., F. J. J. Chain, M. Panchal, C. Eizaguirre, M. Kalbe, T.L. Lenz et al. (2016). Transcriptome profiling of immune tissues reveals habitat-specific gene expression between lake and river sticklebacks. *Mol. Ecol.* 25, 943–958. doi: 10.1111/mec.13520
- Huntingford, F. A. (2004). Implications of domestication and rearing conditions for the behaviour of cultivated fishes. *J. Fish Biol.* 65, 122–142. doi: 10.1111/j.0022-1112.2004.00562.x
- Klimentidis, Y. C., Abrams, M., Wang, J., Fernandez, J. R., and Allison, D. B. (2011). Natural selection at genomic regions associated with obesity and type-2 diabetes: East Asians and sub-Saharan Africans exhibit high levels of differentiation at type-2 diabetes regions. *Hum. Genet.* 129, 407–418. doi: 10.1007/s00439-010-0935-z
- Kohane, M., and Parsons, P. (1988). “Domestication,” in *Evolutionary biology* Springer, Boston, MA. (Springer), 31–48. doi: 10.1007/978-1-4613-1043-3_2
- Lien, S., Koop, B. F., Sandve, S. R., Miller, J. R., Kent, M. P., Nome, T., et al. (2016). The Atlantic salmon genome provides insights into rediploidization. *Nature* 533, 200–205. doi: 10.1038/nature17164
- Liu, L., Ang, K. P., Elliott, J. A., Kent, M. P., Lien, S., MacDonald, D., et al. (2017). A genome scan for selection signatures comparing farmed Atlantic salmon with two wild populations: testing colocalization among outlier markers, candidate genes, and quantitative trait loci for production traits. *Evol. Appl.* 10, 276–296. doi: 10.1111/eva.12450
- López, M. E., Benestan, L., Moore, J. S., Perrier, C., Gilbey, J., Di Genova, A., et al. (2018). Comparing genomic signatures of domestication in two Atlantic salmon (*Salmo salar* L.) populations with different geographical origins. *Evol. Appl.* 12(1): 137–156. doi: 10.1111/eva.12689
- López, M. E., Neira, R., and Yáñez, J. M. (2015). Applications in the search for genomic selection signatures in fish. *Front. Genet.* 5, 458. doi: 10.3389/fgene.2014.00458

- Lorenzen, K., Beveridge, M. C. M., and Mangel, M. (2012). Cultured fish: integrative biology and management of domestication and interactions with wild fish. *Biol. Rev.* 87, 639–660. doi: 10.1111/j.1469-185X.2011.00215.x
- Maiorano, A. M., Lourenco, D. L., Tsuruta, S., Ospina, A. M. T., Stafuzza, N. B., Masuda, Y., et al. (2018). Assessing genetic architecture and signatures of selection of dual purpose Gir cattle populations using genomic information. *PLOS One* 13, e0200694. doi: 10.1371/journal.pone.0200694
- Mäkinen, H., Vasemägi, A., McGinnity, P., Cross, T. F., and Primmer, C. R. (2014). Population genomic analyses of early-phase Atlantic Salmon (*Salmo salar*) domestication/captive breeding. *Evol. Appl.* 8, 93–107. doi: 10.1111/eva.12230
- Manunza, A., Cardoso, T. F., Noce, A., Martínez, A., Pons, A., Bermejo, L. A., et al. (2016). Population structure of eleven Spanish ovine breeds and detection of selective sweeps with BayeScan and hapFLK. *Sci. Rep.* 6, 27296. doi: 10.1038/srep27296
- Norris, A. T., Bradley, D. G., and Cunningham, E. P. (1999). Microsatellite genetic variation between and within farmed and wild Atlantic salmon (*Salmo salar*) populations. *Aquaculture* 180 (3–4), 247–264. doi: 10.1016/S0044-8486(99)00212-4
- Oleksyk, T. K., Smith, M. W., and O'Brien, S. J. (2010). Genome-wide scans for footprints of natural selection. *Philos. Trans. R. Soc B: Biol. Sci.* 365, 185–205. doi: 10.1098/rstb.2009.0219
- Pettem, K. L., Yokomaku, D., Luo, L., Linhoff, M. W., Prasad, T., Connor, S. A., et al. (2013). The specific α -neurexin interactor calyntenin-3 promotes excitatory and inhibitory synapse development. *Neuron* 80, 113–128. doi: 10.1016/j.neuron.2013.07.016
- Price, E. O. (1984). Behavioral aspects of animal domestication. *Q. Rev. Biol.* 59, 1–32. doi: 10.1086/413673
- Purcell, S., Neale, K., Todd-Brown, L., Thomas, M. A., Ferreira, D., Bender, et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81 (3), 559–575.
- Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi: 10.1093/bioinformatics/btq033
- Rubin, C.-J., H.-J. Megens, A. M. Barrio, K. Maqbool, S. Sayyab, D. Schwochow et al. (2012). Strong signatures of selection in the domestic pig genome. *Proc. Natl. Acad. Sci.* 109 (48), 19529–19536. doi: 10.1073/pnas.1217149109
- Ruiz-Larrañaga, O., Langa, J., Rendo, F., Manzano, C., Iriondo, M., and Estonba, A. (2018). Genomic selection signatures in sheep from the Western Pyrenees. *Genet. Sel. Evol.* 50, 9. doi: 10.1186/s12711-018-0378-x
- Sabeti, P. C., Reich, D. E., Higgins, J. M., Levine, H. Z., Richter, D. J., Schaffner, S. F., et al. (2002). Detecting recent positive selection in the human genome from haplotype structure. *Nature* 419, 832–837. doi: 10.1038/nature01140
- Sabeti, P. C., Schaffner, S. F., Fry, B., Lohmueller, J., Varilly, P., Shamovsky, O., et al. (2006). Positive natural selection in the human lineage. *Science* 3129, 1614–1620. doi: 10.1126/science.1124309
- Sabeti, P. C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., et al. (2007). Genome-wide detection and characterization of positive selection in human populations. *Nature* 449, 913–918. doi: 10.1038/nature06250
- Schlenke, T. A., and Begun, D. J. (2003). Natural selection drives *Drosophila* immune system evolution. *Genetics* 164, 1471–1480.
- Simianer, H., Qanbari, S., and Gianola, D. (2010). “Detection of selection signatures within and between cattle populations,” in *Proceedings of 9th World Congress on Genetics Applied to Livestock Production*.
- Solberg, M. F., Kvamme, B. O., Nilsen, F., and Glover, K. (2012). Effects of environmental stress on mRNA expression levels of seven genes related to oxidative stress and growth in Atlantic salmon *Salmo salar* L. of farmed, hybrid and wild origin. *BMC Res. Notes* 5 (672), 1–15. doi: 10.1186/1756-0500-5-672
- Taranger, G. L., Carrillo, M., Schulz, R. W., Fontaine, P., Zanuy, S., Felip, A., et al. (2010). Control of puberty in farmed fish. *Gen. Comp. Endocrinol.* 165, 483–515. doi: 10.1016/j.ygcen.2009.05.004
- Taye, M., Lee, W., Jeon, S., Yoon, J., Dessie, T., Hanotte, O., et al. (2017). Exploring evidence of positive selection signatures in cattle breeds selected for different traits. *Mamm. Genome* 28, 528–541. doi: 10.1007/s00335-017-9715-6
- Teletchea, F., and Fontaine, P. (2014). Levels of domestication in fish: implications for the sustainable future of aquaculture. *Fish Fish.* 15, 181–195. doi: 10.1111/faf.12006
- Thodesen, J., Grisdale-Helland, B., Helland, S. J., and Gjerde, B. (1999). Feed intake, growth and feed utilization of offspring from wild and selected Atlantic salmon (*Salmo salar*). *Aquaculture* 180, 237–246. doi: 10.1016/S0044-8486(99)00204-5
- Tsai, H.-Y., Hamilton, A., Tinch, A. E., Guy, D. R., Gharbi, K., Stear, M. J., et al. (2015). Genome wide association and genomic prediction for growth traits in juvenile farmed Atlantic salmon using a high density SNP array. *BMC Genomics* 16, 969–969. doi: 10.1186/s12864-015-2117-9
- Verspoor, E., Stradmeyer, L., and Nielsen, J. L. (2007). *The Atlantic salmon: genetics, conservation and management*. Blackwell Publishing Ltd, Oxford, UK. John Wiley & Sons. doi: 10.1002/9780470995846
- Voight, B., Kudravalli, S., Wen, X., and Pritchard, J. (2006). A map of recent positive selection in the human genome. *PLoS Biol.* 4, e72. doi: 10.1371/journal.pbio.0040072
- Wang, G. D., Zhai, W., Yang, H. C., Fan, R. X., Cao, X., Zhong, L., et al. (2013). The genomics of selection in dogs and the parallel evolution between dogs and humans. *Nat. Commun.* 4, 1860. doi: 10.1038/ncomms2814
- Wright, S. (1951). The genetical structure of populations. *Ann. Eugen.* 15, 323–354. doi: 10.1111/j.1469-1809.1949.tb02451.x
- Wu, X., Devine, K., Quagliari, C., Driggers, P., and Segars, J. (2015). AKAP13 is required for normal murine ovarian development. *Fertil. Steril.* 10, e134. doi: 10.1016/j.fertnstert.2015.07.415
- Yáñez, J. M., Naswa, S., López, M. E., Bassini, L., Correa, K., Gilbey, J., et al. (2014). Inbreeding and effective population size in a coho salmon (*Oncorhynchus kisutch*) breeding nucleus in Chile. *Aquaculture*. 420, S15–S19. doi: 10.1016/j.aquaculture.2013.05.028
- Yáñez, J. M., S. Naswa, M. E. López, L. Bassini, K. Correa, J. Gilbey et al. (2016). Genomewide single nucleotide polymorphism discovery in Atlantic salmon (*Salmo salar*): validation in wild and farmed American and European populations. *Mol. Ecol. Resour.* 2016 (4), 1002–1–11. doi: 10.1111/1755-0998.12503
- Yoshida, G. M., Lhorente, J. P., Carvalheiro, R., and Yáñez, J. M. (2017). Bayesian genome-wide association analysis for body weight in farmed Atlantic salmon (*Salmo salar* L.). *Anim. Genet.* 48, 698–703. doi: 10.1111/age.12621
- Zueva, K. J., Lumme, J., Veselov, A. E., Kent, M. P., Lien, S., and Primmer, C. R. et al. (2014). Footprints of directional selection in wild Atlantic salmon populations: evidence for parasite-driven evolution? *PLoS ONE* 9, e91672. doi: 10.1371/journal.pone.0091672

Conflict of Interest: AN was employed by Marine Harvest, Kindrum, Fanad, C. Donegal, Ireland. JL was employed by company Benchmark Genetics Chile, Puerto Montt, Chile. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 López, Linderoth, Norris, Lhorente, Neira and Yáñez. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.