

Article

Detecting Local Adaptation between North and South European Atlantic Salmon Populations

María Gabián ¹, Paloma Morán ¹, María Saura ² and Antonio Carvajal-Rodríguez ^{1,*}

¹ Centro de Investigación Mariña (CIM), Departamento de Bioquímica, Genética e Inmunología, Universidade de Vigo, 36310 Vigo, Spain; gabanmaria@gmail.com (M.G.); paloma@uvigo.es (P.M.)

² Departamento de Mejora Genética Animal, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA-CSIC), 28040 Madrid, Spain; saura.maria@inia.csic.es

* Correspondence: acraaj@uvigo.es

Simple Summary: Atlantic salmon is a diadromous species that has the particularity that it returns to reproduce to its river of birth. This means that, in most cases, each river can be considered a population and that these populations can be adapted to the characteristics of the river. The Atlantic Salmon in the rivers of the Iberian Peninsula has been in decline since the 80's of the last century. At that time, to alleviate the decline in the number of individuals, the rivers were stocked with salmon from Scotland and other northern European rivers, but the result was not positive. In this work, we studied eight salmon populations using a 220,000 high-density SNP array (Affymetrix) specific to Atlantic salmon. Our results revealed potential evidence of local adaptation for North Spanish and Scottish populations, which may be the cause of the failure of restocking programs. In addition, by performing a review of previous studies, we demonstrated that, as suggested, NADP-dependent malic enzyme-2 is a candidate for local adaptation in salmon from the Iberian Peninsula.

Abstract: Pollution and other anthropogenic effects have driven a decrease in Atlantic salmon (*Salmo salar*) in the Iberian Peninsula. The restocking effort carried out in the 1980s, with salmon from northern latitudes with the aim of mitigating the decline of native populations, failed, probably due to the deficiency in adaptation of foreign salmon from northern Europe to the warm waters of the Iberian Peninsula. This result would imply that the Iberian populations of Atlantic salmon have experienced local adaptation in their past evolutionary history, as has been described for other populations of this species and other salmonids. Local adaptation can occur by divergent selections between environments, favoring the fixation of alleles that increase the fitness of a population in the environment it inhabits relative to other alleles favored in another population. In this work, we compared the genomes of different populations from the Iberian Peninsula (Atlantic and Cantabric basins) and Scotland in order to provide tentative evidence of candidate SNPs responsible for the adaptive differences between populations, which may explain the failures of restocking carried out during the 1980s. For this purpose, the samples were genotyped with a 220,000 high-density SNP array (Affymetrix) specific to Atlantic salmon. Our results revealed potential evidence of local adaptation for North Spanish and Scottish populations. As expected, most differences concerned the comparison of the Iberian Peninsula with Scotland, although there were also differences between Atlantic and Cantabric populations. A high proportion of the genes identified are related to development and cellular metabolism, DNA transcription and anatomical structure. A particular SNP was identified within the NADP-dependent malic enzyme-2 (*mMEP-2**), previously reported by independent studies as a candidate for local adaptation in salmon from the Iberian Peninsula. Interestingly, the corresponding SNP within the *mMEP-2** region was consistent with a genomic pattern of divergent selection.

Keywords: Atlantic salmon; divergent selection; local adaptation; NADP-dependent malic enzyme-2; SNP array



Citation: Gabián, M.; Morán, P.; Saura, M.; Carvajal-Rodríguez, A. Detecting Local Adaptation between North and South European Atlantic Salmon Populations. *Biology* **2022**, *11*, 933. <https://doi.org/10.3390/biology11060933>

Academic Editor: Brian Dixon

Received: 6 May 2022

Accepted: 16 June 2022

Published: 19 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Wild populations of Atlantic salmon (*Salmo salar*) from the Iberian Peninsula suffered a drastic decrease in the last three decades of the 20th century, mainly due to anthropogenic actions such as pollution [1], the construction of physical barriers that reduced the habitat, and the increase in overfishing [2,3].

To reduce this drastic decline, efforts were focused on restocking from 1981 to 1994, not only with eggs from the wild but also purchased from fish farms, mainly from rivers located in Scotland and Norway. Restocking was done by planting eggs into the river or by releasing fry previously hatched in local fish hatcheries (a practice known as supportive breeding (see [4] and references therein). Nevertheless, there was no increase in the effective population size, suggesting that the introduction of foreign genomes may have contributed even more to the decline of some populations [5,6].

The restocking fail was probably due to a deficiency in the adaptation of salmon from northern Europe to the warm waters of the Iberian Peninsula since this area constitutes the southern limit of distribution of this species in Europe [7]. In fact, local adaptation has been described in previous studies for other populations of this species and salmonids in general [8–11]. Therefore, it seems that the Iberian populations of Atlantic salmon may have also experienced local adaptation at some point in their evolutionary history.

To investigate the causes of the deficient adaptation of the northern European populations to the rivers of the Iberian Peninsula, several translocation experiments were carried out from 1992 to 1993, comparing salmon from River Ulla (northwestern Spain) and River Shin (Scotland), the former being one of the rivers used in the restocking [12]. The results of these experiments, carried out in a hatchery and in an experimental channel located in northwest Spain, indicated that while salmon from River Shin reached a higher condition factor than salmon from River Ulla, mortality was significantly higher in salmon from River Shin during the early stages of development. These findings suggest that salmon from Scotland are not well adapted to the conditions of the northern Iberian Peninsula and may explain the failure of restocking with foreign salmon [13], which contrasts with the success of the supportive breeding program carried out with local populations since 1998 [14].

The fact that salmon populations located in different areas are genetically different is related to an innate preference to return to their river of origin, a behavior known as homing [15]. In line with this, a recent work by Jeffery et al. [16] studied the occurrence of regional differentiation in Atlantic salmon populations from both sides of the Atlantic using a panel of 96 SNPs. The authors were able to assign salmon populations to their areas of origin with high precision. Although the main cause of this strong differentiation is the reproductive behavior of the species, as gene flow is highly restricted by the tendency to return to the river of birth, other differentiating factors, such as local adaptation, may be involved [17,18]. Local adaptation can occur by divergent selection across environments, which favors the fixation of alleles that increase the fitness of a population in the environment it inhabits relative to other alleles favored in another population [19–23]. Consequently, the number of homozygous individuals for favorable alleles in each environment increases because of positive selection [24,25], thus contributing to population differentiation.

For example, in Atlantic salmon, the NADP-dependent malic enzyme malic ($mMEP-2^*$) has been identified as a candidate for local adaptation, particularly for populations in the Iberian Peninsula, as its allelic frequencies vary at macro- and microgeographic scales in relation to freshwater temperatures and the at-sea age of returning adults [13,26,27].

The combined effect of natural selection and interpopulation isolation can produce a fingerprint in genomes. This effect of selection on the genomes of different populations of the same species can be detected by following at least two different methodologies [28,29], namely:

- (1) Detection of atypically high values of differentiation (F_{ST} outliers) associated with particular markers/loci.
- (2) Detection of unusual haplotypic patterns associated with a selective carryover effect, such as increased homozygosity in certain regions.

Regardless of the method, genome-wide fingerprinting simply provides indirect evidence of the possible selective effect, so it is necessary to annotate candidate genomic regions and address the consequent study of candidate genes potentially involved in the selective process. Identification of these selective signatures may reveal genomic regions of biological and commercial interest. In addition, understanding local adaptation could be important for interpreting how salmonid populations respond to habitat alterations, climate change and fisheries.

Therefore, the aim of this work is to compare the genomes of different populations from the Iberian Peninsula and Scotland to identify SNPs and candidate regions associated with high values of interpopulation differentiation and unusual haplotype patterns. The processes in which these genes are involved may explain the adaptive differences between populations and the failures of the restocking carried out during the 1980s.

2. Materials and Methods

A total of 282 returning adult individuals (165 females and 117 males) sampled between 2008 and 2009 were considered. To cover the entire distribution range of salmon in the Iberian Peninsula, 250 individuals from six Spanish rivers (Miño, Ulla, Eo, Sella, Urumea and Bidasoa) were analyzed. This included the Atlantic coast (77 females and 80 males) and the Cantabric coast (70 females and 23 males). The vast majority of the salmon analyzed came from recreational fishing. A few salmon came from sampling conducted by Atlantic salmon program management. For both the salmon caught by the rod (and dead) and those caught at the capture stations, the adipose fin is cut off and stored in ethanol by the management staff responsible for the salmon program. Additionally, 32 individuals from Scotland (18 females and 14 males) coming from two Scottish tributaries (Baddoch and Girnock) of the Dee River (see Figure 1) were available for analysis. To identify the population structure, the samples were grouped using principal component analysis (PCA) with Adegnet [30] based on the Bayesian Information Criterion (BIC).

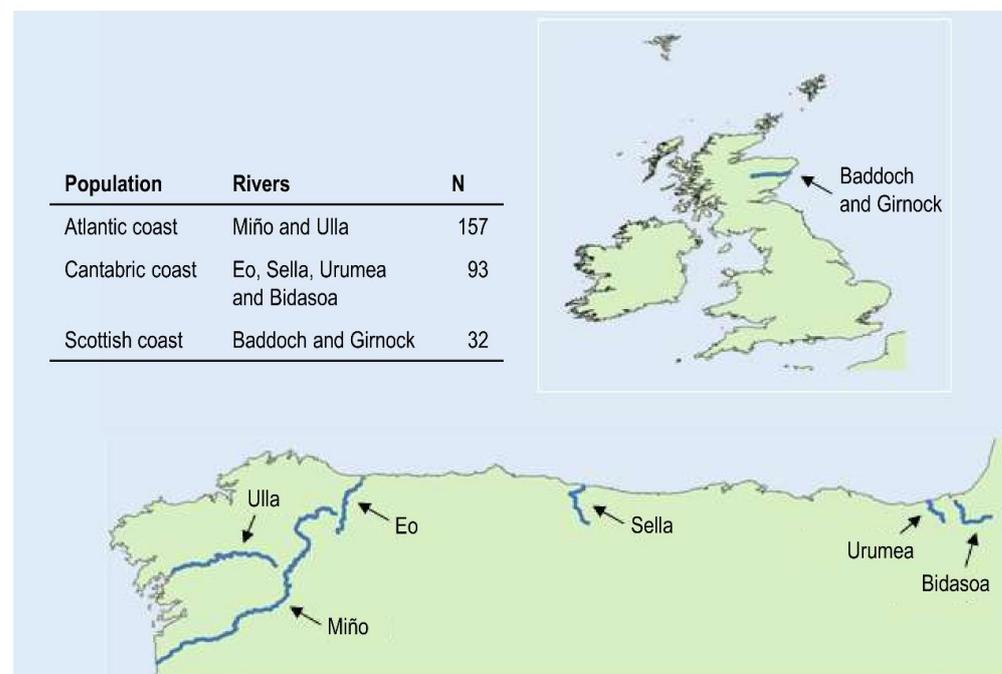


Figure 1. Origin and number of samples analysed: 250 individuals from six rivers in the Iberian Peninsula (Miño, Ulla, Eo, Sella, Urumea and Bidasoa), covering the entire distribution of salmon in its southern boundary, and 32 individuals from two Scottish rivers (Baddoch and Girnock), tributaries of the Dee River.

2.1. Genotyping Quality Analyses

To purify the genomic DNA from ethanol-preserved adipose fins, an NZY Tissue gDNA Isolation Kit (NZYtech) was used. Quantification and purity were assessed using a Nanodrop-1000 spectrophotometer. DNA samples were adjusted to a final concentration of 100 ng/ μ L and frozen until use. Morphological sex was confirmed by the presence/absence of the SDY intron gene (~200 bp) successfully amplified in males and absent in females, by using the primers SDY E1S1 and SDY E2AS4 [31,32]. Samples were genotyped using a 220 K Affymetrix genotyping array (Thermo Scientific) according to the manufacturer's recommendations [33]. Genotypes from samples showing a dish quality control (DQC) < 0.82 or call rate < 0.97 were discarded. Only those SNPs classified as Poly High Resolution, with a call rate > 0.97, were used in our analysis. Unmapped SNPs and those with a minor allele frequency (MAF) < 0.01 were also removed (Table 1).

Table 1. Quality control was performed on a 220,000 customized array.

| Quality Filter | N° Removed SNP |
|----------------------------|----------------|
| Not mapped | 1112 |
| Low quality | 47,432 |
| MAF < 0.01 | 5679 |
| Erratic genotypes | 3 |
| Total analysed SNPs | 165,774 |

After applying these filters, a total of 165,774 SNPs remained allocated across the genome and were proportionally distributed according to the size of each chromosome (Figure 2).

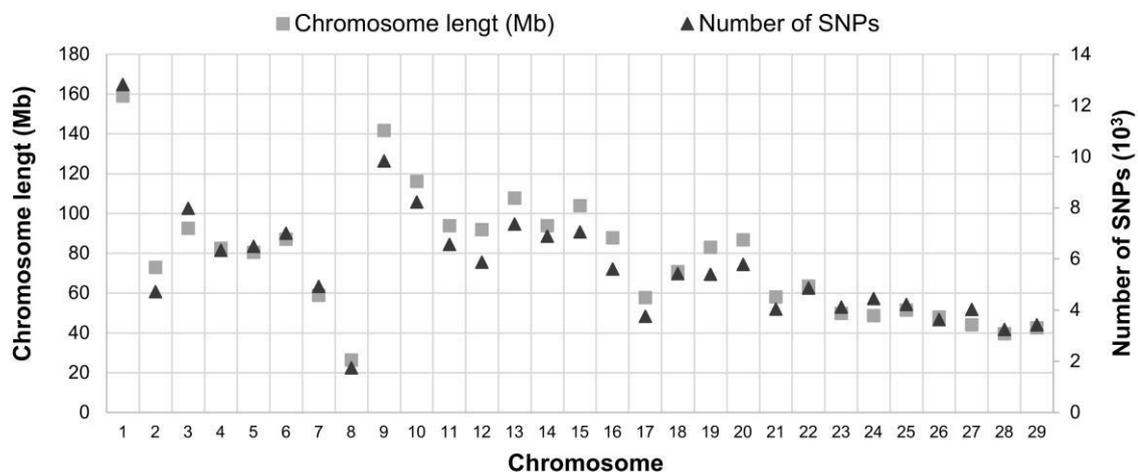


Figure 2. SNP distribution after quality control. A total of 165,774 SNPs remained allocated across the genome and were proportionally distributed according to the size of each chromosome.

2.2. Selection Signatures and Gene Functional Annotation

To evaluate the occurrence of selection, we applied two methods based on the identification of F_{ST} outliers and three haplotype-based methods.

The first family of methods is one of the most common ways to detect local adaptation in non-model organisms. These methods rely on allelic frequencies, assuming that loci under selection will provide unusually high F_{ST} values [34–36]. The outlier-based methods were assessed using BayeScan v.2.1 [37] and the EOS test implemented in HacDivSel v.1.4 software [21]. BayeScan uses a Bayesian likelihood method to estimate the posterior probability of loci experiencing selection. We used default settings, but to protect from false positives, we increased prior odds for the neutral model to 5000 and considered as

significant only SNPs detected with Bayes factor between 32 and 100 ($\log_{10} = 1.5\text{--}2.0$), as these indicate a very strong or decisive signal of divergent selection [36,38]. On the other hand, EOS implements a two-step conservative heuristic strategy for detecting extreme F_{ST} outliers [21]. We performed the EOS analysis using the default parameters. Input files were converted to the correct format using PGDSpider v.2.1.1.2 [39].

Nevertheless, outliers' strategy has some drawbacks and challenges that require the methods to be used with caution; for example, because the markers are considered independent, there is a loss of power when their number is large. In addition, the assumed null model can be wrong, leading to false positives if the population suffered demographic or historical events [38,40–43].

To overcome these drawbacks, we also applied haplotype-based methods for detecting diversifying selection and performed three different methods: (1) a cross-population extended haplotype homozygosity model (XP-EHH) [44], as implemented in selscan v. 1.1.0 [45], (2) the $nvdF_{ST}$ test [21], and (3) a test for multiloci F_{ST} variance as implemented in SmileFinder [46]. Haplotypes were inferred from genotypes with fastPHASE software [47].

Although all three methods require haplotypes, the methodology for searching for selective patterns is different.

- (1) XP-EHH. The cross-population extended haplotype homozygosity model is based on the inspection of patterns of linkage disequilibrium decay around selected loci and detects selection based on an excess of specific haplotypes in one of the populations. This method requires a linkage map. Since no map was available for our 220 K array, we used the physical distance.
- (2) $nvdF_{ST}$. The $nvdF_{ST}$ statistic combines two measures: a normalized variance difference (nvd) and an F_{ST} index. The nvd measure divides the haplotypes into two sets for each candidate SNP: one set carrying the major allele for the SNP and the other set carrying the minor allele. Only SNPs shared by both populations were considered. A variance of mutational distances is computed within each set and a normalized difference between variances defines the statistic that will increase under selection. The F_{ST} measure takes advantage of the fact that if selection acts on a SNP pointed by a high nvd value, then the F_{ST} at that site will be higher when compared to the overall F_{ST} assuming equilibrium in the presence of migration. A resampling method is used to compute the p -value under the hypothesis of panmixia and the final candidate SNPs are those with highest nvd values that additionally reject panmixia [21]. We used different windows sizes (1000, 500, 250, 125 and 62) for computing nvd and considered as potential candidates the 1% of the SNPs with highest nvd , which were also significant for the F_{ST} test under any window size.
- (3) SmileFinder. This method uses a resampling-based strategy to infer the significance of multiloci F_{ST} variance using sliding windows of haplotypes of increasing size. In this case, the highest values of variance indicate the presence of selection.

The gene content within genomic regions identified by candidate SNPs was explored using SalmoBase [48] and gene function annotated with the DAVID tool [49], using as reference and prioritizing by phylogenetic proximity, other salmonids, *Danio rerio*, *Xenopus laevis*, *Mus musculus* and *Homo sapiens*.

3. Results

3.1. Population Structure

Genotyped samples were classified into three groups after clustering with PCA according to samples from Scotland (Baddoch and Girnock) and both slopes of the Iberian Peninsula: rivers from the Atlantic coast (Miño and Ulla) and from the Cantabric coast (Eo, Sella, Urumea and Bidasoa) (Figure 3). From now on, these three pools of data will be listed as Atlantic, Cantabric and Scotland, with 157, 93 and 32 individuals, respectively. Although 9% of individuals from Atlantic rivers were bound with Cantabric rivers, they were considered Atlantic data. This finding is probably related to the fact that during the

80s and until 1992, supportive breeding was performed in Galicia with an initial stock originating from different Galician rivers, including River Eo. However, since 1992, supportive breeding has been exclusively performed with individuals from the same river.

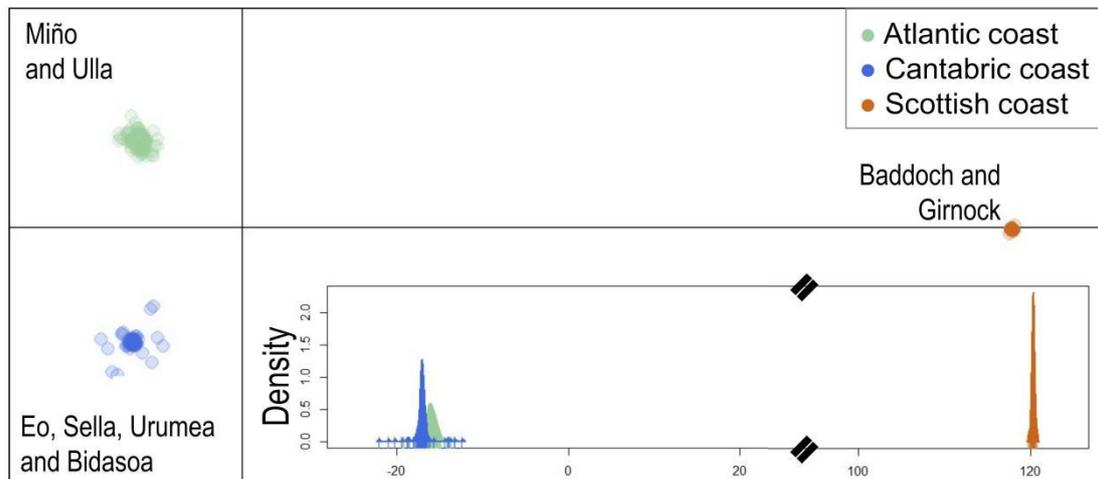


Figure 3. Samples were clustered through principal component analysis with Adegenet (Jombart 2008) based on the Bayesian Information Criterion (BIC). Three pools of data were clustered, listed as Atlantic (Miño and Ulla), Cantabric (Eo, Sella, Urumea and Bidasoa) and Scotland (Baddoch and Girnock), with 157, 93 and 32 individuals, respectively.

3.2. Selection Signatures: Outlier Methods

Bayescan and EOS methods were applied for the comparison of population pairs. For Bayescan, we required a posterior probability of at least 0.9 ($\log BF \geq 1.5$) as evidence for local adaptation, but only the comparison between Atlantic and Scotland showed evidence of positive selection and only for two significant unannotated SNPs (ctg718000187612_7079_SGT and ctg7180001695888_3098_SCT) in chromosome *Ssa22* in a region of around 20 Mb. These same two SNPs were also detected by EOS, which found 142 significant SNPs in the comparison Cantabric–Scotland and 412 in Atlantic–Scotland. There were no significant SNPs in the comparison Atlantic versus Cantabric (Table 2).

Table 2. Summary of significant SNPs in each comparison between studied populations after the application of different selection detection methods (Atl-Can: Atlantic–Cantabric; Can-Scot: Cantabric–Scotland; Atl-Scot: Atlantic–Scotland).

| Statistic | Software | Number of Significant SNPs | | |
|------------------------------------|------------------------------|----------------------------|----------|----------|
| | | Atl-Can | Can-Scot | Atl-Scot |
| F _{ST} outliers detection | HacDivSel (EOS test) | 0 | 142 | 412 |
| | BayeScan ($\log BF = 1.5$) | 0 | 0 | 2 |
| Haplotype based methods | HacDivSel (nvdFST) | 748 | 1504 | 2607 |
| | SmileFinder | 631 | 1346 | 2786 |
| | selscan (XP-EHH) | 201 | 1863 | 2880 |

The largest number of significant SNPs was found in chromosome *Ssa04*, for the comparison of Atlantic versus Scotland (Figure 4). In addition, additional regions were detected by EOS in both comparisons with Scotland. Notably, the 24.8–24.9 Mb region in *Ssa09* has previously been identified by several studies as carrying a strong diversifying selection signal among European but also North American Atlantic salmon populations [10,33,50,51].

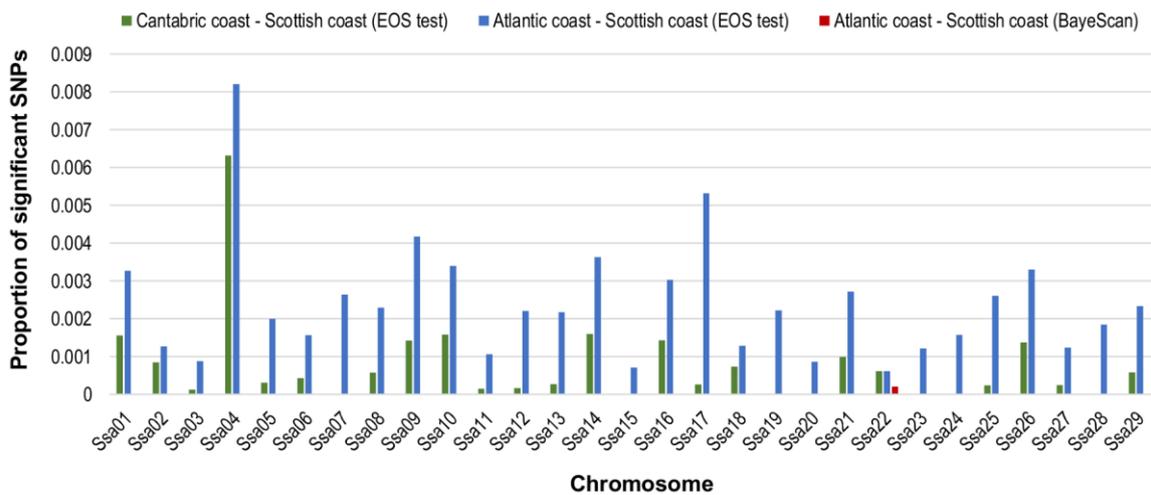


Figure 4. Proportion of significant SNPs of outliers, methods: BayeScan (log BF = 1.5) and EOS-HacDivSel (Atl-Can: Atlantic–Cantabric; Can-Scot: Cantabric–Scotland; Atl-Scot: Atlantic–Scotland).

3.3. Selection Signatures: Haplotype Methods

The haplotype-based methods identified many more candidate SNPs, even for the Atlantic–Cantabric comparison (Table 2). Since the methods assayed imply different strategies to infer selection and it is a recommended practice to combine various approaches [52], we focused on those SNPs detected for at least two of the haplotype-based methods (Figure 5 and Table 3), whether they were also detected by the outlier-based methods. The complete list of these SNPs and the methods for which they were significant are given in Supplementary Tables S1 and S2.

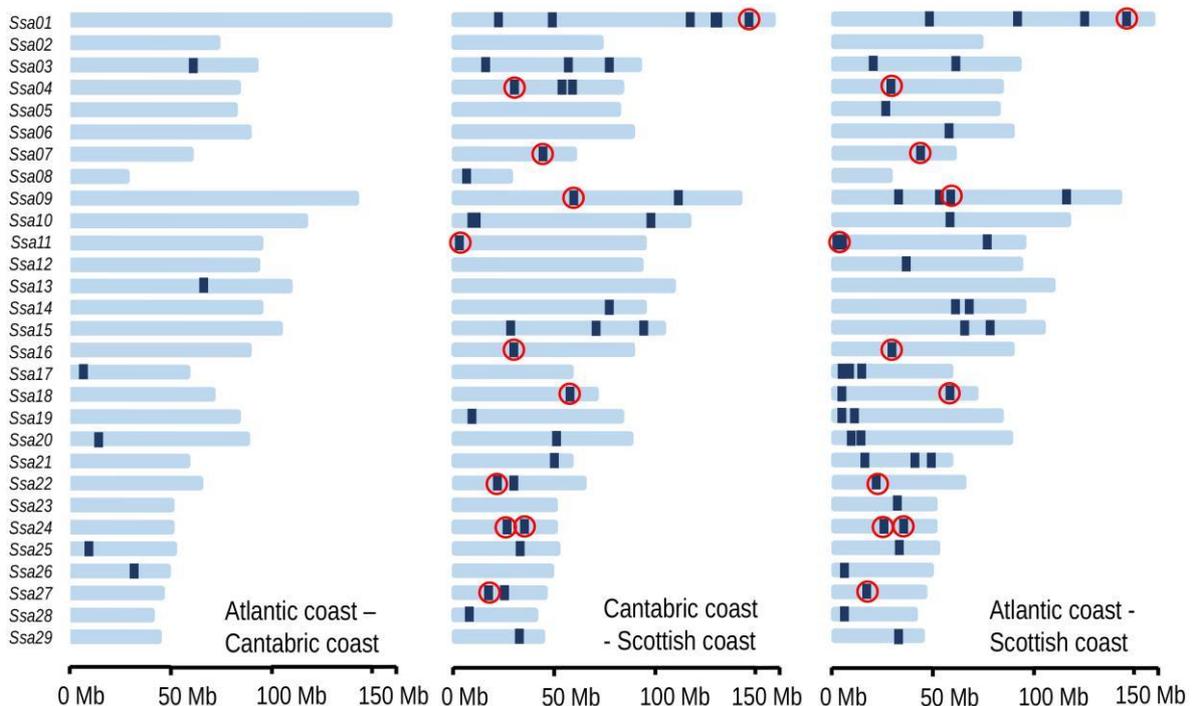


Figure 5. Per chromosome regions with significant SNPs obtained from at least two haplotype-based methods. Regions surrounded by a red circle indicate SNPs detected in both comparisons with Scotland.

Table 3. Number of significant SNPs in at least two haplotype-based methods for each comparison. Comparisons: Atl-Can: Atlantic–Cantabric; Atl-Scot: Atlantic–Scotland; Can-Scot: Cantabric–Scotland. Methods: X: XP-EHH; N: nvdFST; S: SmileFinder. Total: total number of different SNPs detected.

| Comparison | Methods | | | | Total |
|------------|---------|-----|-----|-------|-------|
| | X-N | X-S | N-S | X-N-S | |
| Atl-Can | 0 | 4 | 14 | 0 | 16 |
| Atl-Scot | 210 | 275 | 147 | 64 | 506 |
| Can-Scot | 59 | 31 | 82 | 19 | 170 |

Chromosomes *Ssa09*, *Ssa11* and *Ssa24* had the highest abundance of selection signatures detected by at least two haplotype-based methods, with 90 (Table S1 identifiers H136–H225), 95 (H241–H335) and 119 (H474–H592) SNPs, respectively (Table S1). There were 16 candidate SNPs for local adaptation between the Atlantic and Cantabric rivers, 506 for Atlantic–Scotland and 170 for Cantabric–Scotland (Table 3). Regarding the comparison with Scotland, 62 out of the 170 SNPs detected in Cantabric–Scotland were also detected in the Atlantic–Scotland pair. These common SNPs were located in different chromosomes: *Ssa01*, *04*, *07*, *09*, *11*, *16*, *18*, *22*, *24* and *27* (red circles in Figure 5).

Regarding the aforementioned region, 24.8–24.9 Mb of divergent selection in *Ssa09* was also identified in both comparisons with Scotland by the nvdFST method, although not by selscan or SmileFinder. On the other hand, chromosome *Ssa22* had 13 significant SNPs detected for the Atlantic–Scotland comparison by both XP-EHH and nvdFST methods in a close region (1–2 Mb downstream), where BayeScan and EOS found two significant (non-annotated) SNPs (Figure 4).

Notably, 64 SNPs were significantly detected under all three methods for the Atlantic–Scotland comparison and 19 of these were for the Cantabric–Scotland one.

3.4. Gene Functional Annotation

The complete list of significant SNPs in one or two population comparisons and for two or three of the haplotype-based methods can be found in the supplementary Tables S1 and S2 and the genes containing these SNPs in Table S3.

3.5. Atlantic–Cantabric Comparison

As expected, there were fewer significant SNPs detected by the two haplotype methods in the Atlantic–Cantabric comparison than in either of the two pairwise comparisons with Scotland. These SNPs were located in chromosomes *Ssa03* (identifiers H040–H043 in Table S1), *Ssa13* (identifiers H339–H340 in Table S1), *Ssa17* (identifiers H408–H412 in Table S1), *Ssa20* (identifier H439 in Table S1), *Ssa25* (identifiers H593–H594 in Table S1) and *Ssa26* (identifiers H600–H601 in Table S1).

Notably, SNP H601 within gene *Cyld*, which was detected by nvd and SmileFinder (Table S2), has recently been associated with body weight in rainbow trout [53].

3.6. Comparison with Scotland: SNPs Significant for All Three Haplotype-Based Methods

There were 64 significant SNPs for all three haplotype-based methods. The GO (Gene Ontology) categories of the genes containing those SNPs correspond mainly to developmental and cellular processes. Among the annotated genes (Table 4), we found the *Protocadherin fat4* that pertains to the cadherin gene family and has been associated with amoebic gill disease resistance [54]; the glyoxalase 1 gene, *glo1*, which is important in many physiological processes and diseases and seems to increase its expression in response to stress [55]; MAM domain containing glycosylphosphatidylinositol anchor 1, *mdga1*, which is related to brain development [56], and has been previously identified as selective in the divergence between Norway populations of Atlantic salmon [57]; *WD repeat domain 43*, *wdr43* is related with lipid metabolism [58] and transcriptional response to

contaminant exposure [59] and Zinc finger AN1-type domain 3, *zfang3* which is involved in sex determination and male germ cell maturation in the teleost Nile tilapia [60].

Table 4. Genes that include (or are located at less than 1000 kb) significant SNPs for the Atlantic–Scotland comparison detected by all three haplotype-based methods. SNPs were annotated using the online tool SalmoBase (<https://salmobase.org/>, accessed on 20 May 2021). The #SNPs refer to the number of all SNPs detected by all three methods in that chromosome, not only to those included in genes. The latter are indicated by SNP IDs. In parentheses, the Mb position in the corresponding chromosome.

| Chromosome | # SNPs | SNP IDs | Gene (Mb) |
|--------------|--------|---|--|
| <i>Ssa09</i> | 23 | H181–186, H188 | fat4 (59) |
| <i>Ssa11</i> | 37 | H256 *, H259, H261–H264, H269, H275 *, H281 *, (H297–300) *, H304–310 *, H313, H324–325 | wdr43 (1), trmt61b (1), atts-glupro (1), glo1 (1), zfang3 (1), mdga1 (1) |
| <i>Ssa24</i> | 2 | H479 *, H488 * | ppp6c (25), golga1 (25), rpl35 (25), ofml2a (25) |
| <i>Ssa27</i> | 2 | H615–616 | - |

*: was significant in both comparisons Atlantic–Scotland and Cantabric–Scotland.

When the SNPs detected only by two haplotype-based methods were also considered, most of the genes identified were related to metabolic or cellular developmental processes, DNA transcription and anatomical structure. In *Ssa03*, the gene *tpr* is related to response to temperature changes in mice [61] affected by two SNPs in the comparison Cantabric–Scotland. Another polymorphism is close to *pgk*, a locus related to growth in the turbot in *Ssa04* [62]. Other genes involved with growth in Atlantic salmon are close to significant SNPs: *agrn* and *pomt1* in *Ssa15* (1,164,514 bp) and *Ssa20* (31,286 bp) [63]. In addition, other SNPs are near to *e2f4* in *Ssa10* and *fra10ac1* in *Ssa28* and are related to growth and late maturation in salmon [64].

3.7. Malic Enzyme

As the NADP-dependent malic enzyme-2 ($mMEP-2^*$) has been reported to be a candidate for local adaptation in salmon from the Iberian Peninsula [27], there is a priori independent information about local adaptation effects related to that locus (LOC106586750), which is located at 39,878,908–39,909,791 bp in *Ss25* [65]. In this region, there is just one SNP (ctg7180001794010_7928_SCT) at 39,897,822 bp in chromosome *Ssa25* (id 3579 in Table S4). This malic SNP is surrounded by two other SNPs located at –11 kb and +15.6 kb (ids 3578 and 3580, Table S4). However, in chromosome *Ssa25*, there were only 6 SNPs detected as significant by at least two haplotype-based methods and these SNPs are located at 7,175,102 bp and 32,983,986 bp (Table S1), thus leaving out the region coding for $mMEP-2^*$.

In a previous study using the same microarray and populations (from Spain), genomic linkage disequilibrium highly decreased for SNPs separated 1 Mb apart (reaching half of the maximum value between SNPs separated at 0.5 Mb [66]). In our case, if we consider a window size of 51 SNPs centered in the $mMEP-2^*$ gene, SNPs located upstream and downstream of the region correspond to distances of –0.23 Mb (SNP id 3554) at the left and +0.64 Mb (SNP id 3604) at the right. These distances are minimal because, contrary to the SNPs presented in Table S4—HacDivSel does not consider non-shared SNPs, which could increase the distance between the adjacent SNPs in a window of the given size.

Similarly, if we use a window of size 125 SNPs, we will have 62 SNPs on each side, implying a distance of at least 0.61 Mb at the left and 1.5 Mb at the right side from the malic SNP (ctg7180001794010_7928_SCT).

Therefore, we expect that if the malic SNP reflects a selective pattern affecting the variance of haplotype allelic classes, it could be detected under a window size of less than 50 SNPs or even 125 SNPs, since linkage disequilibrium is still appreciable up to 1 Mb away.

Moreover, considering a candidate SNP that has been previously identified, we can perform the test only for this SNP so that multiple testing corrections are not required.

Unfortunately, because the distribution of the *nvd* statistic is not known, we still face the problem that the statistical significance of a SNP that we assigned to the 1% of candidates with the highest *nvd* is meaningless if we have only one candidate. Fortunately, we can still do a F_{ST} test and look for a more direct way of estimating statistical significance by comparing haplotype variances.

Therefore, we modified our program HacDivSel to introduce user-defined candidate SNPs and performed the analysis in a window of SNPs centered at the gene coding *mMEP-2**. For each population, we compared the homogeneity of variances between the partitions having one or other malic allele, i.e., C or T at ctg7180001794010_7928_SCT SNP. These are the same variances compared by *nvd*, and we can use a robust test of variances (e.g., the composite test as defined in Ramsey and Ramsey 2007) to assess whether the variance for the major allele partition is significantly lower, which would indicate the presence of a positive selection effect in the population. We considered divergent selection when the comparison of variances was significant in at least one population and when both populations had a significantly high F_{ST} value (see Appendix A for details of the variance comparison). With this criterion, we analysed the SNP within the *mMEP-2** region using windows of 25, 51 and 125 SNPs to study the possible pattern of divergent selection. The results are presented in Table 5.

Table 5. Homogeneity variance test for divergent selection from the new version of the HacDivSel program. Var test *p* value: The *p*-value of the within-population variance homogeneity *F* test. The within-population variance homogeneity test implies two comparisons, one for each population, and two *p* values are obtained. The minimum *p*-value obtained is given in the table.

| Comparison | WINDOW SIZE | Var Test <i>p</i> -Value | F_{ST} <i>p</i> -Value | Divergence Significance Test |
|------------|-------------|--------------------------|--------------------------|------------------------------|
| Atl-Can | 25 | 4×10^{-8} | 0.014 | * |
| | 51 | 0.003 | 0.004 | * |
| | 125 | 0.164 | 0.002 | ns |
| Can-Scot | 25 | 0.011 | 0.038 | * |
| | 51 | 0.005 | 0.030 | * |
| | 125 | 0.653 | 0.024 | ns |
| Atl-Scot | 25 | 8×10^{-10} | 0 | * |
| | 51 | 0.021 | 0 | * |
| | 125 | 1 | 0 | ns |

*: Significant variance and F_{ST} *p*-values. ns: non-significant.

The malic *mMEP-2** SNP studied in this new method with the new version of HacDivSel software was clearly detected as a candidate for local adaptation in the three comparisons under the windows sizes assayed within 1 Mb. The window size of 125 SNPs was non-significant in the three comparisons, which probably means that the linkage disequilibrium decreased enough to make it difficult to detect the selection pattern under the given sample sizes.

Furthermore, the neighbors of *mMEP-2** are only a few kb apart, which suggests that SNPs in the vicinity should still present a selective pattern, even in the Atlantic–Cantabric comparison (Table 6). In contrast, the SNPs in the extremes of the windows of size 50 and 125 SNPs should not have a selective pattern, at least not one caused by the *mMEP-2** effect, with the exception of SNP ID 3704 (Table 6).

Table 6. Homogeneity variance test for divergent selection for the Atlantic–Cantabric comparison. ID: SNP ID in Table S4 (e.g., ID 3579 corresponds to SNP ctg7180001794010_7928_SCT in the region with the gene symbol LOC106586750 corresponding to *mMEP-2**). Var test *p* value: minimum *p*-value of the two variance homogeneity *F* tests (one for each population).

| ID | Window Size | Var Test <i>p</i> -Value | F_{ST} <i>p</i> -Value | Divergence Significance Test |
|------|-------------|--------------------------|--------------------------|------------------------------|
| 3454 | 25 | 1 | 1 | ns |
| 3529 | 25 | 1 | 1 | ns |
| 3578 | 25 | 0.026 | 3×10^{-9} | * |
| 3579 | 25 | 4×10^{-8} | 0.002 | * |
| 3580 | 25 | 0.002 | 0.013 | * |
| 3629 | 25 | 1 | 1 | ns |
| 3704 | 25 | 2×10^{-8} | 0.008 | * |

*: Significant variance and F_{ST} *p*-values. ns: non-significant.

In addition, we estimated the false-positive rate (FPR) of the new method using simulated data. We note that, at least with the tested window sizes (from 25 to 400 SNPs), the method appears conservative, with an FPR below 5% (see Table A1 in Appendix A), supporting the reliability of our results.

4. Discussion

In this work, we performed a genome-wide scan using a high-density SNP array with the aim of investigating different patterns of selective variation in the Atlantic salmon genome to identify candidate loci that may be involved in the local adaptation of Atlantic salmon in the northern Iberian Peninsula, which is the southern limit of the distribution of this species.

We used different frequency-based (F_{ST} outliers) and haplotype-based selection detection methods and observed little overlap in the regions detected by the two types of strategies. The lack of overlap between frequency-based and haplotype-based methods has already been reported in other studies [67,68]. Regarding the use of outliers, most salmonid phenotypic traits are polygenic, such as growth, body size and fat content, among others [9], for which the effect of selection may involve a subtle change in allele frequency at several loci. This context makes it difficult to detect the selection effect using outlier-based methods that rely on a strong change in allele frequency at a few independent loci [21]. Therefore, as expected, haplotype-based methods have more power to detect recent signatures of selection [69]. We combined the results of three haplotype-based methods to identify 630 SNPs, 55% of which matched 116 genes in 25 out of 29 chromosomes (Tables S1–S3). The chromosomes containing more candidate genes were *Ssa24* (29 genes), *Ssa09* (13 genes), *Ssa11* (9 genes), *Ssa01* and *Ssa17* (8 genes) and *Ssa10* and *Ssa03* (7 and 6 genes, respectively). Genes potentially more affected by divergent selection were those resulting from the comparison between the peninsular slopes and Scotland, as expected, since the intensity of local adaptation appears to be correlated with the geographical distance between salmonid populations [9,57]. Our results support the idea that Atlantic salmon populations show significant genetic differences that may be the result of a combination of diversifying natural selection and the effect of gene flow and drift.

Even with the conservative threshold we used for outlier methods, coupled with the requirement of a positive signal in at least two haplotype-based methods, we detected quite a few signals and genes that could be involved in local adaptation processes. Similar to other genome-wide studies looking for local adaptation, several of the candidate genes we have found are distributed along different chromosomal regions and are related to cellular growth and lipid metabolic processes, among others.

In addition to the coincident SNPs for at least two haplotype-based methods, we found a well-known *Ssa09* haploblock associated with a strong signature of divergent selection among Atlantic salmon populations [10]. This region, located at 24.8–24.9 Mb,

presented a strong signal in the comparisons between Scotland and the Atlantic and Cantabric populations for both the EOS and the $nvdF_{ST}$ methods. The same region has been previously reported as a candidate for selection by several authors and contains the protein phosphatase 1a (*ppm1a*) gene and the SIX homeobox (*six6*), the latter related to circadian timing processes [10]. For the comparison between Atlantic and Scotland, $nvdF_{ST}$ also detected SNPs within the region located at 77–78 Mb on *Ssa13*, corresponding to the SLC25A14 gene, which encodes a brain mitochondrial transporter protein and has also been found to diverge between European and North American populations [70]. Similarly, we found divergence in regions containing immune-related genes already identified in the literature as divergent in comparisons between northern and southern Norwegian populations [57].

Malic Enzyme

The gene encoding the malic enzyme deserves special mention. As previously explained, this gene located at *Ss25* was specifically explored for candidate SNPs, since early electrophoretic studies on genetic variation in Atlantic salmon referred specifically to this gene [71]. Variation at $mMEP-2^*$ alleles (isozymes 100 and 125) presents strong correlations with environmental temperature, both within and among rivers, and associations with phenotypic performance [71]. Indeed, there is a latitudinal cline in which the frequency of the *125 allele of $mMEP-2^*$ is higher in northern European Atlantic salmon populations than in southern European populations. Moreover, in the Atlantic populations of the Iberian Peninsula, the frequency of the *125 allele is almost zero [72]. In fact, the failure of the restocking carried out in the Iberian Peninsula with salmon from Northern Europe was proven by the low frequency of the *125 allele in the Iberian populations after restocking [13].

However, with the genomic selection detection methods initially assayed, we could not detect any significant SNP within the region coding for $mMEP-2^*$. However, this result was expected. First, the allele frequencies of the malic SNP were 0.02229 in the Atlantic sample, 0.134 in the Cantabric sample and 0.4219 in the Scotland sample, which implies that the frequencies of the three samples are in the same half of the frequency range. Second, since we are analyzing several thousand SNPs, multitest correction implies a loss of statistical power. Thus, it is expected that other more powerful signals will be more easily detected. Fortunately, after modification of the HacDivSel method to consider some candidates a priori, the malic $mMEP-2^*$ SNP, was clearly detected as a candidate for local adaptation.

Although the maladaptation of northern European salmon demonstrated by the garden experiment described by García de Leániz et al. [12] is not exclusively due to the malic enzyme, the role of this gene could be relevant, as suggested by the results obtained in the present work.

5. Conclusions

In this study, we provided additional evidence of local adaptation for north Spanish and Scottish populations of Atlantic salmon using genome-wide information. Defining the spatiotemporal scales over which adaptation operates is important from the point of view of management and conservation. Ignoring the existence of locally adapted populations may pose a greater risk to the conservation management of threatened populations [8]. We modified the program HacDivSel to introduce user-defined candidate SNPs and performed the analysis using a window of SNPs centered at a SNP within the NADP-dependent malic enzyme-2 ($mMEP-2^*$), previously reported by independent studies as a candidate for local adaptation in salmon from the Iberian Peninsula. Interestingly, the corresponding SNP within the $mMEP-2^*$ region was consistent with a genomic pattern of divergent selection.

In summary, our results suggest that Spanish and Scottish populations may differ at the functional gene level due to local adaptation, which would explain the failure to restock Spanish rivers with Norway and Scottish individuals.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/biology11060933/s1>. Gabianetal_SuppM_SNPs_array.xlsx—Subsequent to the quality control, a total of 165,774 SNPs remained from an 220 K Affymetrix genotyping array (Thermo Scientific) [33], customized from the 930 K XHD Ssal array (dbSNP accession numbers ss1867919552-ss1868858426). Gabianetal_SuppM_significant_SNPs.xlsx—Due to the low statistical power of the outlier tests, gene annotation was made from significant SNPs in at least two haplotype-based methods. Gabianetal_SuppM_gene_annotation.xlsx—Gene annotation was made from significant SNPs in at least two haplotype-based methods using the online tools SalmoBase [48] and [49]. Gabianetal_Appendix_MalicSelectionDetection.pdf. Supplementary Tables S1–S4.pdf. Table S1. Significant SNPs for at least two haplotype-based methods. ID: identifier. Chrom: chromosome. Atlantic and Cantabric—Scotland: indicates that both Atlantic–Scotland and Cantabric–Scotland comparisons were significant for the SNP. Table S2. Methods corresponding to the significant SNPs in Table S1 (significant for at least two haplotype-based methods). An asterisk indicates that the corresponding method detected the SNP as significant (SF: SmileFinder; sel: selscan). Table S3. Candidate genes, corresponding significant SNPs and gene ontology (GO). Table S4. SNPs in chromosome *Ssa25*.

Author Contributions: Conceptualization, P.M. and A.C.-R.; Data curation, M.G. and M.S.; Formal analysis, M.G. and A.C.-R.; Funding acquisition, P.M.; Investigation, M.G.; Methodology, P.M. and A.C.-R.; Software, A.C.-R.; Supervision, P.M. and A.C.-R.; Validation, P.M. and A.C.-R.; Writing—original draft, M.G. and A.C.-R.; Writing—review & editing, M.G., P.M., M.S. and A.C.-R. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Ministerio de Economía y Competitividad, Spain (grant RZ2A12-00011-CO2-01/02) and has received financial support from Xunta de Galicia (Grupo de Referencia Competitiva, ED431C 2020/05, Centro Singular de Investigación de Galicia accreditation 2019–2022) and the European Union (European Regional Development Fund—ERDF).

Institutional Review Board Statement: Ethical review and approval was waived for this study because the samples came from salmon caught in recreational fishing, which is regulated by Spanish laws and regulations; we did not sacrifice any animals. We used samples from salmon that are already legally dead.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article and supplementary material.

Acknowledgments: We want to thank Pablo Caballero (Xunta de Galicia), Gerónimo de la Hoz (Principado de Asturias), Íñigo Mendiola (Diputación Foral de Guipúzcoa), Javier Álvarez, Pedro Luenda (GANASA) and John Taggart for kindly providing salmon samples.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Appendix A.1. Malic Enzyme

As explained in the main text, the NADP-dependent malic enzyme-2 ($mMEP-2^*$) has been reported by independent studies as a candidate for local adaptation in salmon from the Iberian Peninsula [13,26,27]. There is *a priori* independent information about local adaptation effects related to that locus (LOC106586750), which is located at 39,878,908–39,909,791 bp in *Ss25* (Davidson et al., 2010). In our samples and after the quality controls, there is in this region just one SNP (ctg7180001794010_7928_SCT) at 39,897,822 bp in chromosome *Ssa25* (id 3579 in Table S4). This malic SNP is surrounded by two other SNPs that are –11 kb and +15.6 kb (ids 3578 and 3580, Table S4). However, after the analysis of chromosome *Ssa25* with 2 outlier and 3 haplotype-based methods (see main text), there were only 6 SNPs detected as significant by at least two haplotype-based methods. These SNPs were between positions 7,175,102 bp and 32,983,986 bp (Table S1), thus leaving out the region of $mMEP-2^*$.

This is not so surprising; first, the sample allele frequencies of ctg7180001794010_7928_SCT were 0.02229 in the Atlantic sample, 0.134 in the Cantabric and 0.4219 in the Scotland sample, which implies the frequencies in the three samples are in the same half of the frequency range.

In addition, since we are analyzing several thousand SNPs and all methods correct for multiple comparisons in some way, which implies a loss of statistical power, it is expected that other more powerful signals will be more easily detected.

Appendix A.2. HAC Variance Method

The HAC variance method was developed by [73] and adapted in [21] for the two-population divergence scenario (*nvd* statistic). As with many other methods that explore the genome looking for characteristic selection footprints, there is a trade-off between power and false-positive control when applying a sliding window across the genome looking for candidate sites. However, if we had a priori information about a candidate locus, we could simply test that site, which would give us more statistical power since we would not have to correct for excess false positives caused by multiple testing. However, we still face the problem that the distribution of the *nvd* statistic is not known, and we cannot deduce as significant the 1% of candidates with the highest *nvd* if we have only one candidate. Therefore, we must look for a more direct way to compare haplotype variances and the simplest and most intuitive way to do this would be to test the homogeneity of variances between the two partitions.

Therefore, the HacDivSel program was modified to enter user-defined candidate SNPs and perform the analysis in a candidate-centered SNP window. For each population, we compared the homogeneity of variances between the partitions having one or the other candidate allele. These are the same variances compared by *nvd* and we can use a robust test of variances (see below) to assess whether the variance of the major allele partition is significantly lower, which could indicate the presence of a positive selection footprint in the population.

Appendix A.3. Divergent Selection Detection Variance Test

Appendix A.3.1. Major Allele Reference Haplotype and Sample Partitions P_1 and P_2

Let us recall the concept of the major-allele-reference haplotype (MARH henceforth denoted by R) as an haplotype of length $L + 1$ constructed with the alleles at highest frequency (major alleles) for each of the $L + 1$ loci in a sample of n haplotypes. Now consider that we have *a priori* information about the selective effect of locus with SNP at position x ($1 \leq x \leq L + 1$). Let us call this the candidate position. We make a partition of the data in two groups, P_1 including the haplotypes having the reference allele at the candidate position (i.e., having the same allele at position x as R) and P_2 including the remaining haplotypes, i.e., having the non-reference allele at position x . Since the candidate allele is fixed in each partition, the haplotype length can henceforth be considered L .

We want to test the influence of the candidate allele on the surrounding positions. We expect that if the candidate position is selective, the distribution in P_1 would be different from the distribution in P_2 because the candidate major frequency allele (which defines P_1) would produce a sweep effect of the surrounding ones. This sweep effect should translate into clustering of major alleles around the candidate allele. We do not expect a similar effect in P_2 . Thus, we want to test how the alleles are distributed around the candidate in P_1 and compare them with P_2 . We can test this in several populations.

Appendix A.3.2. Haplotype Allelic Class (HAC)

We defined the mutational distance between any haplotype and R as the number of sites (SNPs) carrying an allele different from the one in R. Each group of haplotypes with the same mutational distance will constitute a haplotype allelic class (HAC). Therefore, let us note that as h_i the number of minor alleles each haplotype i carries, then every haplotype having the same h number belongs to the same HAC class. Thus, h is the sum of positions that have different alleles than in the reference haplotype. So, in a given haplotype for each position $k \in \{1, L\}$ we count the outcome $X_k = I(s_k \neq r_k)$ where s_k is the allele in the haplotype,

r_k is the allele in the reference and $I(A)$ is the indicator variable taking value 1 if A is true and 0 otherwise. Therefore, the h value of a haplotype of length L is:

$$h = \sum_{k=1}^L X_k \text{ where } X_k = I(s_k \neq r_k) \tag{A1}$$

Recall we have two partitions and for each partition the candidate position x is not included since we know that in P_1 every haplotype i has by definition $X_{ix} = 0$ and similarly $X_{ix} = 1$ in P_2 . We are interested in studying the distribution of X_{ik} in the remaining L positions. From the original sample of n haplotypes, each partition (P_1 and P_2) has n_1 and $n_2 = n - n_1$ haplotypes of length L , respectively.

Appendix A.3.3. HAC Variance

For comparing the distribution of alleles around a candidate site in both samples P_1 and P_2 of haplotypes, we can perform a variance homogeneity test because if the candidate site is selective, we expect that a sweep effect would be translated into lower HAC variance v_1 in P_1 than variance v_2 in P_2 ($v_1 < v_2$). On the contrary, if the site is neutral, we expect similar variances or even $v_1 > v_2$ if the number of haplotypes in P_1 is higher than in P_2 [73]. Therefore, we want to test the neutral null hypothesis $v_1 \geq v_2$ versus the alternative $v_1 < v_2$.

The unbiased variance of HAC values in a sample of n haplotypes is

$$v = \sum_{i=1}^n \frac{(h_i - m)^2}{(n - 1)} = \frac{1}{n - 1} \sum_{i=1}^n (h_i^2 - m^2) \tag{A2}$$

where h_i is the hac value for the haplotype i and $m = \sum_{i=1}^n \frac{h_i}{n}$ is the sample mean.

Appendix A.3.4. Skewness and Kurtosis

Because the distribution of HAC classes is unknown, we should use statistics that are robust to deviations from normality, so we must identify the kind of deviation from normality, as different statistics perform better under different deviations. Therefore, we will measure both skewness and kurtosis in the haplotype samples.

There are various formalizations of the skewness and kurtosis concepts [74]. All these measures are based on the sample r th moment for a sample of size n

$$m_r = \frac{\sum_{i=1}^n (x_i - \hat{x})^r}{n} \tag{A3}$$

Let

$$g_1 = \frac{m_3}{m_2^{3/2}} \text{ and } g_2 = \frac{m_4}{m_2^2} - 3$$

then the sample skewness measures G_1 and b_1 are

$$G_1 = \frac{\sqrt{n(n - 1)}}{n - 2} g_1 \text{ and } b_1 = \frac{m_3}{s^3} = \left(\frac{n - 1}{n} \right)^{3/2} g_1$$

where $s = v^{1/2}$ and the sample kurtosis G_2 and b_2 are

$$G_2 = \frac{n - 1}{(n - 2)(n - 3)} \{(n + 1)g_2 + 6\} \text{ and } b_2 = \frac{m_4}{s^4} - 3 = \left(\frac{n - 1}{n} \right)^2 \frac{m_4}{m_2^2} - 3$$

G_1 and G_2 formulations for skewness and kurtosis respectively, are recommended for small samples from non-normal distributions as they have smaller mean-squared error than other measures [74]. By other side, b_1 and b_2 measures work well for normally distributed data while g_1 and g_2 measures are the basis of the D'Agostino normality test [75]. The variances of these measures are related, and normally distributed transformations are available to perform skewness and kurtosis tests [75].

Appendix A.3.5. Homogeneity Variance Tests

There are a number of procedures for testing homogeneity of variances in two samples; however, for some of them, Type I error control is extremely sensitive to kurtosis [76]. Two robust tests of variance are BF [77] and OB [78]. Based on both, a composite method has been proposed that applies OB when a sample is detected as platykurtic and BF otherwise [76].

BF and OB tests were based on the one-way ANOVA statistic.

$$W(z) = \frac{\sum_{i=1}^K n_i (\bar{z}_i - \hat{z})^2 / (K - 1)}{\sum_i^K \sum_{j=1}^{n_i} (z_{ij} - \bar{z}_i)^2 / \sum_i (n_i - 1)} \tag{A4}$$

where

$$\bar{z}_i = \sum_{j=1}^{n_i} z_{ij} / n_i \text{ and } \hat{z} = \frac{\sum_{i=1}^K \sum_{j=1}^{n_i} z_{ij}}{\sum_{i=1}^K n_i}$$

Brown-Forsythe (BF) test

Depending on the definition of the z values, we have different statistics, for example if $z_{ij} = |x_{ij} - x_i|$, $W(z)$ is the Levene test with $K - 1$ (K is the number of groups, 2 in our case) and $\sum_i (n_i - 1)$ degrees of freedom. Alternatively, if we define $z_{ij} = |x_{ij} - \tilde{x}_i|$ where \tilde{x}_i is the sample median; then we have the BF test [77].

O'Brien (OB) test

Define $y_{ij} = (x_{ij} - \hat{x}_i)^2$ and

$$r_{ij} = \frac{(n_i - 3/2)n_i y_{ij} - v_i(n_i - 1)/2}{(n_i - 1)(n_i - 2)}$$

where v_i is the sample variance of group i . Given the transformation r_{ij} note that for each group, the mean of the r transformed values is equal to the sample variance of that group [78],

$$\bar{r}_i = \sum_{j=1}^{n_i} r_{ij} / n_i = \sum_{j=1}^{n_i} (x_{ij} - \hat{x}_i)^2 / (n_i - 1) = v_i$$

The OB test takes $z_{ij} = r_{ij}$ and then applies $W(r)$ from Equation (A4)

$$W(r) = \frac{\sum_{i=1}^K n_i (v_i - \hat{r})^2 / (K - 1)}{\sum_i^K \sum_{j=1}^{n_i} (r_{ij} - v_i)^2 / \sum_i (n_i - 1)}$$

Given BF or OB, the corresponding W statistic is compared with the F critical value with $K - 1$ and $\sum_i (n_i - 1)$ degrees of freedom in the numerator and denominator, respectively.

Appendix A.4. Algorithm for Divergent Selection Variance Test

The procedure for testing divergent selections in a set of populations is as follows.

1. Decide a candidate SNP x and the number of markers L (window size).
2. For each population, i define its major-allele-reference haplotype R_i , and perform sample partitions P_{i1} and P_{i2} with the haplotypes carrying the major or minor allele of x , respectively.
3. Compute sample variances v_{i1} for the P_{i1} partition of population i and compute v_2 as the sample variance of the accumulated HAC values of the P_2 partitions for all populations. The reason for accumulating HAC values over the different populations for the second partition is that the sample size is smaller in these partitions and the allele distribution is expected to be random (no sweep) around the candidate site.
4. If $v_{i1} \geq v_2$ we conclude that there is no selection at the candidate site and if there are more populations, return to the second step to analyze the next population. Otherwise, if $v_{i1} < v_2$, we continue with step 5.

5. For each partition, test skewness and kurtosis, i.e., test the null of no skewness ($g_1 = 0$) and if kurtosis is not platykurtic ($g_2 \geq 0$, D'Agostino test).
6. If data for both partitions are normally distributed, an F test is performed. If data is marginally normal (one partition is normal and the other has a p value above $\alpha/2$) then a Levene test is performed. Otherwise, we use the composite test BF-OB that applies OB if any of the partitions is platykurtic and BF otherwise. If homogeneity of variances is rejected, we conclude that candidate site x is selective in population i . For the computation of skewness and kurtosis using the D'Agostino test, a sample size of at least 8 is required. If the sample size is below this value, the homogeneity test will be the OB, which is more robust under a low sample size.
7. Repeat the above steps 1–6 for the next population. Divergent selection between populations i and j at site x can be concluded if x was selective in at least one population and the F_{ST} index that compares the F_{ST} for site x with the overall F_{ST} is significant. The significance of the F_{ST} index was computed by resampling, as in [21].

Appendix A.5. Simulations

To study the control of type I error in the method, we use publicly available data from simulations in [21]. Specifically, the files indicated in Table A1 correspond to two populations of 1000 facultative hermaphrodites simulated during 10,000 generations. The mating was random within each population. The between-population migration was $Nm = 10$. Each individual consisted of a diploid chromosome of length 106 biallelic sites. Two different population mutation rates per genome were assayed ($\theta = 4 N\mu = \{12, 60\}$) and four different population recombination rates ($\rho = 4 Nr = \{0, 4, 12, 60\}$). Natural selection was not simulated, so all scenarios were neutral. At the end of each run, 50 haplotypes were sampled from each population.

Analysis of the Simulated Data

We assumed as a candidate a SNP located in the center of the SNPs available in the genome. Different window sizes were used, depending on the number of segregating sites. In general, the minimum window size was 25, and the maximum was the total number of segregating sites. The results indicated that at least for a significance level of 0.05, the test was conservative, and the false positive rate was below 0.05 (Table A1).

Table A1. False positive rate (FPR) for the divergent selection detection variance test using simulated neutral evolution data available from Carvajal-Rodríguez, 2017 [21]. The genome size was 1 Mb. Population size $N = 1000$. Number of generations $T = 10,000$. Population mutation rate $\theta = 4 N\mu$. Population recombination rate $\rho = 4 Nr$. The candidate position is in the middle of the window. The total number of replicates for each case was 1000. $FPR = 100 \times \text{number of replicates with significant test} / 1000$. $MAF = 1\%$. Significance level $\alpha = 0.05$.

| File Name | θ | ρ | Mean Window Size | FPR |
|-----------|----------|--------|------------------|-----|
| C4 | 12 | 0 | 25 | 1% |
| C4 | 12 | 0 | 87 | 1% |
| C5 | 12 | 4 | 25 | 2% |
| C5 | 12 | 4 | 87 | 1% |
| C6 | 12 | 12 | 25 | 2% |
| C6 | 12 | 12 | 86 | 1% |
| C16 | 60 | 0 | 25 | 1% |

Table A1. Cont.

| File Name | θ | ρ | Mean Window Size | FPR |
|-----------|----------|--------|------------------|-----|
| C16 | 60 | 0 | 51 | 1% |
| C16 | 60 | 0 | 125 | 2% |
| C16 | 60 | 0 | 436 | 1% |
| C17 | 60 | 4 | 25 | 2% |
| C17 | 60 | 4 | 51 | 2% |
| C17 | 60 | 4 | 125 | 2% |
| C17 | 60 | 4 | 429 | 2% |
| C18 | 60 | 60 | 25 | 2% |
| C18 | 60 | 60 | 51 | 2% |
| C18 | 60 | 60 | 125 | 2% |
| C18 | 60 | 60 | 431 | 1% |

References

- Morán, P.; Cal, L.; Cobelo-García, A.; Almécija, C.; Caballero, P.; García de Leániz, C. Historical legacies of river pollution reconstructed from fish scales. *Environ. Pollut.* **2018**, *234*, 253–259. [[CrossRef](#)] [[PubMed](#)]
- Saura, M.; Caballero, A.; Caballero, P.; Morán, P. Impact of precocious male parr on the effective size of a wild population of Atlantic salmon. *Freshw. Biol.* **2008**, *53*, 2375–2384. [[CrossRef](#)]
- Álvarez, J.J.; Antón, A.; Azpiroz, I.; Caballero, P.; Hervella, F.; De la Hoz, J.; Urrizalki, I. *Atlas de los Ríos Salmoneros de la Península Ibérica*; Ekolur SLL: España, Spain, 2010.
- Campos, J.L.; Posada, D.; Morán, P. Introgression and genetic structure in northern Spanish Atlantic salmon (*Salmo salar* L.) populations according to mtDNA data. *Conserv. Genet.* **2008**, *9*, 157–169. [[CrossRef](#)]
- Morán, P.; Pendás, A.M.; García-Vázquez, E.; Izquierdo, J.T.; Rutherford, D.T. Electrophoretic Assessment of the Contribution of Transplanted Scottish Atlantic Salmon (*Salmo salar*) to the Esva River (Northern Spain). *Can. J. Fish. Aquat. Sci.* **1994**, *51*, 248–252. [[CrossRef](#)]
- Morán, P.; Perez, J.; Dumas, J.; Beall, E.; García-Vázquez, E. Stocking-related patterns of genetic variation at enzymatic loci in south European Atlantic salmon populations. *J. Fish Biol.* **2005**, *67*, 185–199. [[CrossRef](#)]
- Horreo, J.L.; Machado-Schiaffino, G.; Ayllon, F.; Griffiths, A.M.; Bright, D.; Stevens, J.R.; Garcia-Vazquez, E. Impact of climate change and human-mediated introgression on southern European Atlantic salmon populations. *Glob. Change Biol.* **2011**, *17*, 1778–1787. [[CrossRef](#)]
- García de Leániz, C.; Fleming, I.A.; Einum, S.; Verspoor, E.; Jordan, W.C.; Consuegra, S.; Aubin-Horth, N.; Lajus, D.; Letcher, B.H.; Youngson, A.F.; et al. A critical review of adaptive genetic variation in Atlantic salmon: Implications for conservation. *Biol. Rev. Camb. Philos. Soc.* **2007**, *82*, 173–211. [[CrossRef](#)]
- Fraser, D.J.; Weir, L.K.; Bernatchez, L.; Hansen, M.M.; Taylor, E.B. Extent and scale of local adaptation in salmonid fishes: Review and meta-analysis. *Heredity* **2011**, *106*, 404–420. [[CrossRef](#)]
- Pritchard, V.L.; Mäkinen, H.; Vähä, J.-P.; Erkinaro, J.; Orell, P.; Primmer, C.R. Genomic signatures of fine-scale local selection in Atlantic salmon suggest involvement of sexual maturation, energy homeostasis and immune defence-related genes. *Mol. Ecol.* **2018**, *27*, 2560–2575. [[CrossRef](#)]
- Mobley, K.B.; Granroth-Wilding, H.; Ellmen, M.; Vähä, J.-P.; Aykanat, T.; Johnston, S.E.; Orell, P.; Erkinaro, J.; Primmer, C.R. Home ground advantage: Local Atlantic salmon have higher reproductive fitness than dispersers in the wild. *Sci. Adv.* **2019**, *5*, eaav1112. [[CrossRef](#)]
- García de Leániz, C.; Morán, P.; Pendás, A.M.; Caballero, C.; López-Beiras, P.; Valero, E.; García-Vázquez, E. Freshwater performance of native and foreign Atlantic salmon populations stocked in two Galician rivers (Spain). In *Fifth International Symposium on Genetics in Aquaculture, Halifax*; Halifax University: Halifax, NS, Canada, 1994.
- Morán, P.; Perez, J.; Garcia-Vazquez, E. The malic enzyme MEP-2* locus in Spanish populations of Atlantic salmon: Sea age and foreign stocking. *Aquat. Sci.* **1998**, *60*, 359–366. [[CrossRef](#)]
- Saavedra-Nieves, P.; Crujeiras, R.M.; Caballero, P.; Cobo, F. Assessing the effect of recovery programs for salmon (*Salmo salar* Linnaeus, 1758) at its Southern limit in Europe: Application of segmented regression models to long-term data from the Ulla River. *Limnetica* **2021**, *40*, 189–203. [[CrossRef](#)]
- Altukhov, Y.P.; Salmenkova, E.A.; Omelchenko, V.T. *Salmonid Fishes: Population Biology, Genetics and Management*; John Wiley & Sons: Hoboken, NJ, USA, 2000.

16. Jeffery, N.W.; Wringe, B.F.; McBride, M.C.; Hamilton, L.C.; Stanley, R.R.E.; Bernatchez, L.; Kent, M.; Clément, M.; Gilbey, J.; Sheehan, T.F.; et al. Range-wide regional assignment of Atlantic salmon (*Salmo salar*) using genome wide single-nucleotide polymorphisms. *Fish. Res.* **2018**, *206*, 163–175. [[CrossRef](#)]
17. Blanquart, F.; Kaltz, O.; Nuismer, S.L.; Gandon, S. A practical guide to measuring local adaptation. *Ecol. Lett.* **2013**, *16*, 1195–1205. [[CrossRef](#)]
18. Savolainen, O.; Lascoux, M.; Merila, J. Ecological genomics of local adaptation. *Nat. Rev. Genet.* **2013**, *14*, 807. [[CrossRef](#)]
19. Nielsen, R. Molecular signatures of natural selection. *Annu. Rev. Genet.* **2005**, *39*, 197–218. [[CrossRef](#)]
20. Barreiro, L.B.; Laval, G.; Quach, H.; Patin, E.; Quintana-Murci, L. Natural selection has driven population differentiation in modern humans. *Nat. Genet.* **2008**, *40*, 340–345. [[CrossRef](#)]
21. Carvajal-Rodríguez, A. HacDivSel: Two new methods (haplotype-based and outlier-based) for the detection of divergent selection in pairs of populations. *PLoS ONE* **2017**, *12*, e0175944. [[CrossRef](#)]
22. Pavlidis, P.; Alachiotis, N. A survey of methods and tools to detect recent and strong positive selection. *J. Biol. Res.-Thessalon.* **2017**, *24*, 7. [[CrossRef](#)]
23. Weigand, H.; Leese, F. Detecting signatures of positive selection in non-model species using genomic data. *Zool. J. Linn. Soc.* **2018**, *187*, 528–583. [[CrossRef](#)]
24. Granka, J.M.; Henn, B.M.; Gignoux, C.R.; Kidd, J.M.; Bustamante, C.D.; Feldman, M.W. Limited evidence for classic selective sweeps in African populations. *Genetics* **2012**, *192*, 1049–1064. [[CrossRef](#)] [[PubMed](#)]
25. Horscroft, C.; Ennis, S.; Pengelly, R.J.; Sluckin, T.J.; Collins, A. Sequencing era methods for identifying signatures of selection in the genome. *Brief. Bioinform.* **2018**, *20*, 1997–2008. [[CrossRef](#)] [[PubMed](#)]
26. Jordan, W.C.; Youngson, A.F. Genetic protein variation and natural selection in Atlantic salmon (*Salmo salar*, L.) parr. *J. Fish Biol.* **1991**, *39*, 185–192. [[CrossRef](#)]
27. Verspoor, E.; Jordan, W.C. Genetic variation at the Me-2 locus in the Atlantic salmon within and between rivers: Evidence for its selective maintenance. *J. Fish Biol.* **1989**, *35*, 205–213. [[CrossRef](#)]
28. Sabeti, P.C.; Schaffner, S.F.; Fry, B.; Lohmueller, J.; Varilly, P.; Shamovsky, O.; Palma, A.; Mikkelsen, T.S.; Altshuler, D.; Lander, E.S. Positive natural selection in the human lineage. *Science* **2006**, *312*, 1614–1620. [[CrossRef](#)] [[PubMed](#)]
29. Williamson, S.H.; Hubisz, M.J.; Clark, A.G.; Payseur, B.A.; Bustamante, C.D.; Nielsen, R. Localizing Recent Adaptive Evolution in the Human Genome. *PLoS Genet.* **2007**, *3*, e90. [[CrossRef](#)] [[PubMed](#)]
30. Jombart, T. adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* **2008**, *24*, 1403–1405. [[CrossRef](#)]
31. Yano, A.; Nicol, B.; Jouanno, E.; Quillet, E.; Fostier, A.; Guyomard, R.; Guiguen, Y. The sexually dimorphic on the Y-chromosome gene (sdY) is a conserved male-specific Y-chromosome sequence in many salmonids. *Evol. Appl.* **2013**, *6*, 486–496. [[CrossRef](#)]
32. Quéméré, E.; Perrier, C.; Besnard, A.-L.; Evanno, G.; Baglinière, J.-L.; Guiguen, Y.; Launey, S. An improved PCR-based method for faster sex determination in brown trout (*Salmo trutta*) and Atlantic salmon (*Salmo salar*). *Conserv. Genet. Resour.* **2014**, *6*, 825–827. [[CrossRef](#)]
33. Barson, N.J.; Aykanat, T.; Hindar, K.; Baranski, M.; Bolstad, G.H.; Fiske, P.; Jacq, C.; Jensen, A.J.; Johnston, S.E.; Karlsson, S.; et al. Sex-dependent dominance at a single locus maintains variation in age at maturity in salmon. *Nature* **2015**, *528*, 405–408. [[CrossRef](#)]
34. Lewontin, R.C.; Krakauer, J. Testing Heterogeneity of F-Values. *Genetics* **1975**, *80*, 397–398. [[CrossRef](#)] [[PubMed](#)]
35. Beaumont, M.A. Adaptation and speciation: What can F(st) tell us? *Trends Ecol. Evol.* **2005**, *20*, 435–440. [[CrossRef](#)] [[PubMed](#)]
36. Whitlock, M.C.; Lotterhos, K.E. Reliable Detection of Loci Responsible for Local Adaptation: Inference of a Null Model through Trimming the Distribution of FST. *Am. Nat.* **2015**, *186*, S24–S36. [[CrossRef](#)] [[PubMed](#)]
37. Foll, M.; Gaggiotti, O. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics* **2008**, *180*, 977–993. [[CrossRef](#)]
38. Lischer, H.E.L.; Excoffier, L. PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinform. Oxf. Engl.* **2012**, *28*, 298–299. [[CrossRef](#)]
39. Lotterhos, K.E.; Whitlock, M.C. Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Mol. Ecol.* **2014**, *23*, 2178. [[CrossRef](#)]
40. Pérez-Figueroa, A.; Garcia-Pereira, M.J.; Saura, M.; Rolan-Alvarez, E.; Caballero, A. Comparing three different methods to detect selective loci using dominant markers. *J. Evol. Biol.* **2010**, *23*, 2267–2276. [[CrossRef](#)]
41. Bierne, N.; Roze, D.; Welch, J.J. Pervasive selection or is it . . . ? why are FST outliers sometimes so frequent? *Mol. Ecol.* **2013**, *22*, 2061–2064. [[CrossRef](#)]
42. De Mita, S.; Thuillet, A.-C.; Gay, L.; Ahmadi, N.; Manel, S.; Ronfort, J.; Vigouroux, Y. Detecting selection along environmental gradients: Analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Mol. Ecol.* **2013**, *22*, 1383. [[CrossRef](#)]
43. Hoban, S.; Kelley, J.L.; Lotterhos, K.E.; Antolin, M.F.; Bradburd, G.; Lowry, D.B.; Poss, M.L.; Reed, L.K.; Storfer, A.; Whitlock, M.C. Finding the Genomic Basis of Local Adaptation: Pitfalls, Practical Solutions, and Future Directions. *Am. Nat.* **2016**, *188*, 379–397. [[CrossRef](#)]
44. Sabeti, P.C.; Varilly, P.; Fry, B.; Lohmueller, J.; Hostetter, E.; Cotsapas, C.; Xie, X.; Byrne, E.H.; McCarroll, S.A.; Gaudet, R.; et al. Genome-wide detection and characterization of positive selection in human populations. *Nature* **2007**, *449*, 913–918. [[CrossRef](#)] [[PubMed](#)]

45. Szpiech, Z.A.; Hernandez, R.D. selscan: An Efficient Multithreaded Program to Perform EHH-Based Scans for Positive Selection. *Mol. Biol. Evol.* **2014**, *31*, 2824–2827. [[CrossRef](#)] [[PubMed](#)]
46. Guiblet, W.M.; Zhao, K.; O'Brien, S.J.; Massey, S.E.; Roca, A.L.; Oleksyk, T.K. SmileFinder: A resampling-based approach to evaluate signatures of selection from genome-wide sets of matching allele frequency data in two or more diploid populations. *GigaScience* **2015**, *4*, 1. [[CrossRef](#)]
47. Scheet, P.; Stephens, M. A Fast and Flexible Statistical Model for Large-Scale Population Genotype Data: Applications to Inferring Missing Genotypes and Haplotypic Phase. *Am. J. Hum. Genet.* **2006**, *78*, 629–644. [[CrossRef](#)] [[PubMed](#)]
48. Samy, J.K.A.; Mulugeta, T.D.; Nome, T.; Sandve, S.R.; Grammes, F.; Kent, M.P.; Lien, S.; Våge, D.I. SalmoBase: An integrated molecular data resource for Salmonid species. *BMC Genom.* **2017**, *18*, 482.
49. Huang, D.W.; Sherman, B.T.; Tan, Q.; Collins, J.R.; Alvord, W.G.; Roayaei, J.; Stephens, R.; Baseler, M.W.; Lane, H.C.; Lempicki, R.A. The DAVID Gene Functional Classification Tool: A novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biol.* **2007**, *8*, R183. [[CrossRef](#)] [[PubMed](#)]
50. Bourret, V.; Kent, M.P.; Primmer, C.R.; Vasemägi, A.; Karlsson, S.; Hindar, K.; McGinnity, P.; Verspoor, E.; Bernatchez, L.; Lien, S. SNP-array reveals genome-wide patterns of geographical and potential adaptive divergence across the natural range of Atlantic salmon (*Salmo salar*). *Mol. Ecol.* **2013**, *22*, 532–551. [[CrossRef](#)]
51. Moore, J.-S.; Bourret, V.; Dionne, M.; Bradbury, I.; O'Reilly, P.; Kent, M.; Chaput, G.; Bernatchez, L. Conservation genomics of anadromous Atlantic salmon across its North American range: Outlier loci identify the same patterns of population structure as neutral loci. *Mol. Ecol.* **2014**, *23*, 5680–5697. [[CrossRef](#)]
52. De Villemereuil, P.; Frichot, É.; Bazin, É.; François, O.; Gaggiotti, O.E. Genome scan methods against more complex models: When and how much should we trust them? *Mol. Ecol.* **2014**, *23*, 2006–2019. [[CrossRef](#)]
53. Reis Neto, R.V.; Yoshida, G.M.; Lhorente, J.P.; Yáñez, J.M. Genome-wide association analysis for body weight identifies candidate genes related to development and metabolism in rainbow trout (*Oncorhynchus mykiss*). *Mol. Genet. Genom. MGG* **2019**, *294*, 563–571. [[CrossRef](#)]
54. Boison, S.A.; Gjerde, B.; Hillestad, B.; Makvandi-Nejad, S.; Moghadam, H.K. Genomic and Transcriptomic Analysis of Amoebic Gill Disease Resistance in Atlantic Salmon (*Salmo salar* L.). *Front. Genet.* **2019**, *10*, 68. [[CrossRef](#)] [[PubMed](#)]
55. Du, F.; Xu, G.; Li, Y.; Nie, Z.; Xu, P. Glyoxalase 1 gene of *Coilia nasus*: Molecular characterization and differential expression during transport stress. *Fish. Sci.* **2016**, *82*, 719–728. [[CrossRef](#)]
56. Sano, S.; Takashima, S.; Niwa, H.; Yokoi, H.; Shimada, A.; Arenz, A.; Wittbrodt, J.; Takeda, H. Characterization of teleost Mdgal1 using a gene-trap approach in medaka (*Oryzias latipes*). *genesis* **2009**, *47*, 505–513. [[CrossRef](#)] [[PubMed](#)]
57. Kjærner-Semb, E.; Ayllon, F.; Furmanek, T.; Wennevik, V.; Dahle, G.; Niemelä, E.; Ozerov, M.; Vähä, J.-P.; Glover, K.A.; Rubin, C.J.; et al. Atlantic salmon populations reveal adaptive divergence of immune related genes—A duplicated genome under selection. *BMC Genom.* **2016**, *17*, 610. [[CrossRef](#)] [[PubMed](#)]
58. Ning, L.-J.; He, A.-Y.; Li, J.-M.; Lu, D.-L.; Jiao, J.-G.; Li, L.-Y.; Li, D.-L.; Zhang, M.-L.; Chen, L.-Q.; Du, Z.-Y. Mechanisms and metabolic regulation of PPAR α activation in Nile tilapia (*Oreochromis niloticus*). *Biochim. Biophys. Acta BBA—Mol. Cell Biol. Lipids* **2016**, *1861*, 1036–1048. [[CrossRef](#)] [[PubMed](#)]
59. Hasenbein, M.; Werner, I.; Deanovic, L.A.; Geist, J.; Fritsch, E.B.; Javidmehr, A.; Foe, C.; Fangue, N.A.; Connon, R.E. Transcriptomic profiling permits the identification of pollutant sources and effects in ambient water samples. *Sci. Total Environ.* **2014**, *468–469*, 688–698. [[CrossRef](#)]
60. Ma, K.; Liao, M.; Liu, F.; Ye, B.; Sun, F.; Yue, G.H. Characterizing the ZFAND3 gene mapped in the sex-determining locus in hybrid tilapia (*Oreochromis* spp.). *Sci. Rep.* **2016**, *6*, 25471. [[CrossRef](#)]
61. Skaggs, H.S.; Xing, H.; Wilkerson, D.C.; Murphy, L.A.; Hong, Y.; Mayhew, C.N.; Sarge, K.D. HSF1-TPR interaction facilitates export of stress-induced HSP70 mRNA. *J. Biol. Chem.* **2007**, *282*, 33902–33907. [[CrossRef](#)]
62. Robledo, D.; Rubiolo, J.A.; Cabaleiro, S.; Martínez, P.; Bouza, C. Differential gene expression and SNP association between fast- and slow-growing turbot (*Scophthalmus maximus*). *Sci. Rep.* **2017**, *7*, 12105. [[CrossRef](#)]
63. Tsai, H.-Y.; Hamilton, A.; Tinch, A.E.; Guy, D.R.; Gharbi, K.; Stear, M.J.; Matika, O.; Bishop, S.C.; Houston, R.D. Genome wide association and genomic prediction for growth traits in juvenile farmed Atlantic salmon using a high density SNP array. *BMC Genom.* **2015**, *16*, 969.
64. Gutierrez, A.P.; Yáñez, J.M.; Fukui, S.; Swift, B.; Davidson, W.S. Genome-wide association study (GWAS) for growth rate and age at sexual maturation in Atlantic salmon (*Salmo salar*). *PLoS ONE* **2015**, *10*, e0119730. [[CrossRef](#)] [[PubMed](#)]
65. Davidson, W.S.; Koop, B.F.; Jones, S.J.; Iturra, P.; Vidal, R.; Maass, A.; Jonassen, I.; Lien, S.; Omholt, S.W. Sequencing the genome of the Atlantic salmon (*Salmo salar*). *Genome Biol.* **2010**, *11*, 403. [[CrossRef](#)] [[PubMed](#)]
66. Gabián, M.; Morán, P.; Fernández, A.L.; Villanueva, B.; Chtioui, A.; Kent, M.P.; Covelo-Soto, L.; Fernández, A.; Saura, M. Identification of genomic regions regulating sex determination in Atlantic salmon using high density SNP data. *BMC Genom.* **2019**, *20*, 764. [[CrossRef](#)] [[PubMed](#)]
67. López, M.E.; Benestan, L.; Moore, J.-S.; Perrier, C.; Gilbey, J.; Genova, A.D.; Maass, A.; Diaz, D.; Lhorente, J.-P.; Correa, K.; et al. Comparing genomic signatures of domestication in two Atlantic salmon (*Salmo salar* L.) populations with different geographical origins. *Evol. Appl.* **2019**, *12*, 137–156. [[CrossRef](#)]
68. Mäkinen, H.; Vasemägi, A.; McGinnity, P.; Cross, T.F.; Primmer, C.R. Population genomic analyses of early-phase Atlantic Salmon (*Salmo salar*) domestication/captive breeding. *Evol. Appl.* **2015**, *8*, 93–107. [[CrossRef](#)]

69. Oleksyk, T.K.; Smith, M.W.; O'Brien, S.J. Genome-wide scans for footprints of natural selection. *Philos. Trans. R. Soc. B Biol. Sci.* **2010**, *365*, 185–205. [[CrossRef](#)]
70. Lehnert, S.J.; Kess, T.; Bentzen, P.; Clément, M.; Bradbury, I.R. Divergent and linked selection shape patterns of genomic differentiation between European and North American Atlantic salmon (*Salmo salar*). *Mol. Ecol.* **2020**, *29*, 2160–2175. [[CrossRef](#)]
71. Verspoor, E.; Beardmore, J.A.; Consuegra, S.; García de Leániz, C.; Hindar, K.; Jordan, W.C.; Koljonen, M.-L.; Mahkrov, A.A.; Paaver, T.; Sánchez, J.A.; et al. Population structure in the Atlantic salmon: Insights from 40 years of research into genetic protein variation. *J. Fish Biol.* **2005**, *67*, 3–54. [[CrossRef](#)]
72. Morán, P.; Pérez, J.; Garcia-Vazquez, E. Genetic variation at enzyme loci in the southernmost European populations of Atlantic salmon. *J. Fish Biol.* **2005**, *67*, 206–212. [[CrossRef](#)]
73. Hussin, J.; Nadeau, P.; Lefebvre, J.-F.; Labuda, D. Haplotype allelic classes for detecting ongoing positive selection. *BMC Bioinform.* **2010**, *11*, 65. [[CrossRef](#)]
74. Joanes, D.N.; Gill, C.A. Comparing measures of sample skewness and kurtosis. *J. R. Stat. Soc. Ser. Stat.* **1998**, *47*, 183–189. [[CrossRef](#)]
75. D'Agostino, R.B.; Belanger, A. A Suggestion for Using Powerful and Informative Tests of Normality. *Am. Stat.* **1990**, *44*, 316–321.
76. Ramsey, P.H.; Ramsey, P.P. Testing Variability in the Two-Sample Case. *Commun. Stat.—Simul. Comput.* **2007**, *36*, 233–248. [[CrossRef](#)]
77. Brown, M.B.; Forsythe, A.B. Robust Tests for the Equality of Variances. *J. Am. Stat. Assoc.* **1974**, *69*, 364. [[CrossRef](#)]
78. O'Brien, R.G. A simple test for variance effects in experimental designs. *Psychol. Bull.* **1981**, *89*, 570–574. [[CrossRef](#)]