Molecular ecology

Genome-wide detection of positive and balancing selection signatures shared by four domesticated rainbow trout populations (*Oncorhynchus mykiss)*

Paul, K.[1], Restoux, G. [1], Phocas, F. [1]

1. Université Paris-Saclay, INRAE, AgroParisTech, GABI, 78350 Jouy-en-Josas, France

Corresponding author: Florence Phocas - florence.phocas@inrae.fr

Abstract

Evolutionary processes leave footprints across the genome over time. Highly homozygous regions may correspond to positive selection of favourable alleles, while maintenance of heterozygous regions may be due to balancing selection phenomena. We analyzed 176 genomes coming from 20 sequenced US fish and 156 fish from three different French lines that were genotyped using a HD Axiom Trout Genotyping 665K SNP Array. Using methods based on either Run of Homozygosity or Extended Haplotype Homozygosity, we detected selection signals in four domesticated rainbow trout populations. Nine genomic regions composed of 253 genes, mainly located on chromosome 2 but also on chromosomes 12, 15, 16, and 20, were identified under positive selection in all four populations. In addition, four heterozygous regions containing 29 genes putatively under balancing selection were also shared by the four populations and located on chromosomes 10, 13, and 19. Whatever the homozygous or heterozygous nature of the region, we always found some genes highly conserved among vertebrates due to their critical roles in cellular and nuclear organisation, embryonic development or immunity. We identify new promising candidate genes involved in rainbow

trout fitness, as well as genes already detected under positive selection in other fishes (*auts2, atp1b3, zp4, znf135, igf-1α, brd2, col9a2, mrap2, pbx1, emilin-3*). These findings represent a genome-wide map of signatures of selection common over rainbow trout populations, which is the foundation to understand the processes in action and to identify what kind of diversity should be preserved, or conversely avoided in breeding programs, in order to maintain or improve essential biological functions in domesticated rainbow trout populations.

Keywords: Runs of Homozygosity, Extended Haplotype Homozygosity, domestication, fitness, selection, fish.

1| Introduction

Any population, whether animal or plant, wild or domesticated, evolved through continuous and cumulative changes over time (Wright, 1931). It relies on various evolutionary forces, mutation, migration, selection, and genetic drift, whose relative effects may vary depending on population history and structure. For example, genetic drift is more substantial when the effective population size is small and randomly induces fixation of alleles, which may lead to degeneration and extinction due to the fixation of deleterious alleles in small populations (Smith & Haigh, 1974). When modifications of environmental conditions occur, allele frequencies will change to a new relevant equilibrium, as a result of natural selection. Indeed, favorable alleles in a particular environment due to either new mutations or standing variation, will be positively selected. In wild populations, favourable alleles are generally affecting fitness through individual survival, mating, or fertility (East, 1918; Fisher, 1958). Natural selection can also act by negative (or purifying) selection that hinders the spread of deleterious alleles (Charlesworth et al., 1995). These two processes tend to reduce the genetic diversity at the target genes but had different effect on the genome, positive selection leading to stronger selection signatures (selective sweep) than negative one. Conversely, the population's polymorphism can be actively maintained in some rare genomic regions through balancing selection that keeps an equilibrium in the frequencies of alleles. The two main biological causes of balancing selection are heterozygote advantage at a single locus, known as overdominance effect, and frequency-dependent selection with a rare-allele advantage that tends to restore a frequency equilibrium between alleles at the population level (Charlesworth, 2006, Fijarczyk & Babik, 2015).

Domestication is the evolutionary process of genetic adaptation over generations of a wild population to handling by humans and breeding in captive environments (Darwin, 1859, 1868; Price, 1984). During domestication, humans exert artificial selection pressure on the initial

population by choosing and organizing the reproduction of the most adapted individuals to cohabitation or more globally to those whose aptitudes correspond the best to their expectations (Price, 1999; Russell, 2002), such as a less fearfulness of humans (Price, 2002; Harri et al., 2007). Domestication induces severe genetic bottlenecks due to the selection and reproduction of only a few adapted animals from the wild population. Thus, many genetic evolutionary processes, such as selection, genetic drift, and inbreeding, have a significant role in the evolution of farmed animal populations (Helmer, 1992; Mignon-Grasteau et al., 2005). The domestication process affects life history traits due to changes in morphological, physiological, reproductive, behavioural, and immune functions (Mignon-Grasteau et al., 2005; Pulcini et al., 2013; Milla et al., 2021 for review in fishes) compared to their wild relatives (Darwin, 1859, 1868). Wilkins et al. (2014) suggest that these specific modifications, called domestication syndrome, may be due to mild deficit of neural-crest cells during embryonic development in domesticated animals. In addition, both natural and artificial selection in domesticated species leaves footprints across the genome, known as selection signatures, which can point to regions harboring essential genes for domestication or natural fitness (Dobney & Larson, 2006; Qanbari & Simianer, 2014; Wright, 2015).

Compared to domestication in terrestrial animals (Mignon-Grasteau et al., 2005), fish domestication is recent and was first documented with carp about 2,000 years ago. The precise date and location (Neolithic China or at the Roman period in Central and East Europe) of the carp domestication are still debated (Balon, 1995; Balon, 2004). However, most farmed fish species have only been domesticated since the last century. The rainbow trout is native to the Pacific drainages of North America and to Kamchatka in Russia and its domestication started in the 1870s in California (Hershberger, 1992; Gall & Crandel, 1992). It was then introduced in Western Europe at the beginning of the 20th century (Fabrice, 2018).

Numerous studies have been carried out over the last ten years to detect signatures of selection in farmed fish species (Channel Catfish: Sun et al., 2014; Atlantic salmon: Mäkinen et al., 2015; Gutierrez et al., 2016; Liu et al., 2016; Pritchard et al., 2018; López et al., 2019; Carp: Su et al., 2018; Nile Tilapia: Hong et al., 2015; Cádiz et al. 2020; Yu et al., 2022; Rainbow trout: Cádiz et al., 2021; Coho salmon : López et al., 2021; Australasian snapper: Baesjou & Wellenreuther, 2021; Brown trout: Magris et al., 2022) in order to identify genomic regions involved in recent adaptation or domestication processes (Smith & Haigh, 1974; Pennings & Hermisson, 2006). In this study, we were interested in farmed rainbow trout populations as it is one of the oldest farmed fish and the analysis of genes under either positive or balancing subsequent selection in. Indeed, only a few studies on selection signatures were performed in rainbow trout. Three of them only focused on wild populations and showed signatures of selection linked to life-history variation, egg development, spawning time (Martínez et al., 2011), immune response (Limborg et al., 2012), and smoltification (Weinstein et al., 2019). The first study in domesticated rainbow trout was performed on a single Chilean population (Cádiz et al., 2021) genotyped with a 57K SNP array; identified signatures of selection were associated with early development, growth, reproduction and immune system. Recently, a high-density array (665K SNPs) was developed for rainbow trout (Bernard et al., 2022), allowing us to potentially more accurately detect signatures of selection and to compare them across various domesticated rainbow trout populations. The existence of signatures of selection shared by farmed populations from different geographical areas is essential to understand the importance of genetic diversity in several genomic regions in rainbow trout and then to identify genes having key roles in either the domestication process or fitness because conserved by all populations (Bruford et al., 2003; Yáñez et al., 2022).

Various approaches have been developed to reveal selection signatures within population based on site frequency spectrum, linkage disequilibrium (LD), or reduction of local variability (Vitti et al., 2013; Saravanan et al., 2020). Among these approaches, we will use two strategies, one based on the reduction of local variability using Run of Homozygosity (ROH) metrics and the second one relying on allele frequencies and the extent of LD based on Extended Haplotype Homozygosity (EHH). ROH is a large homozygous stretch in the genome of an individual inherited from a common ancestor to his parents (McQuillan et al., 2008; Purfield et al., 2012), while EHH measures the extent of shared haplotypes through the association between a single core haplotype and multiple loci at various distances from the core region (Sabeti et al., 2002). In our study, we considered four populations: one INRAE experimental line (with no intentional selection), two French selected lines from two different breeding companies, and a pooled American population gathering samples from one wild river and four hatchery populations, all coming from the North-West of the USA and closely genetically linked (Gao et al., 2018). Our work aimed to discover the main genomic regions sharing strong homozygosity (positive selection) or heterozygosity (balancing selection) across the four rainbow trout populations and to get further insights into the nature of genes spanning these regions.

2 | Material and methods

2.1 | Populations

Three French populations were considered: 14 breeding females from the INRAE synthetic line SY and, 90 and 72 females from two selected lines LB and LC from the breeding companies "Bretagne Truite" (Plouigneau, France) and "Viviers de Sarrance" (Sarrance, France)

138 respectively. The SY was developed by intercrossing several domesticated lines of rainbow
139 trout, in order to create a population with a large diversity (D'Ambrosio et al., 2019).
140 In addition, we considered an American pooled population, hereafter named HA, using the
141 whole genome sequence data of 20 fishes obtained by Gao et al. (2018). The sampling strategy
142 consisted in collecting DNA from 4 individuals in each of five locations from the North-West
143 of the USA: wild fish from Elwha River, and farmed fish from Dworshak, L. Quinault,
144 Quinault, and Shamania hatcheries. We pooled together the 20 individuals, as these five
145 populations were genetically close to each other (Supplementary Figure 1; Gao et al., 2018) and
146 greatly distant from the three French populations (Figure 1).
147
148 2.2 | Genotyping and quality control
149 High-density genotypes were obtained at the INRAE genotyping Platform Gentyane
150 (Clermont-Ferrand, France) for all the 176 French samples using the Affymetrix 665K SNP
151 array recently developed for rainbow trout (Bernard et al., 2022). We only considered the
152 genotypes for the 576,118 SNPs of the Rainbow Trout Axiom® 665K SNP array that were
153 positioned on the Arlee genome (GCA_013265735.3, Gao et al., 2021; Bernard et al., 2022).
154 From the whole-genome sequence information of the 20 American samples (Gao et al., 2018),
155 we extracted the genotypes for the same 576,118 SNPs of the HD chip.
156 Among the 177 French genotyped fish, 19 individuals with more than 30% identity-by-state
157 (IBS) with other individuals were removed from the dataset. We thus kept for the analysis 76,
158 67, 20, and 14 fish sampled from LB, LC, HA, and SY populations, respectively.
159 Then, SNP quality control was performed using PLINK v1.9 software (Chang et al., 2015).
160 Note that, to avoid limitations due to the low number of individuals in SY, quality filters were
161 made considering LC and SY together, as both populations were genotyped on the same SNP

plate and are close genetically (D'Ambrosio et al., 2019). About 4,000 SNPs randomly distributed over all the genome were removed for all populations due to extreme deviation from Hardy-Weinberg equilibrium (p-value < 10-7). It allowed us to discared SNPs with high risk of wrong genotyping, in addition to the edit for SNP call rate lower than 97%. We retained 571,319 SNPs, 569,030 SNPs, and 573,793 SNPs on LB, 'LC- SY', and HA populations, respectively. Finally, crossing the three SNP lists, we kept the 546,903 common SNPs for the analysis.

2.3 | Genetic structure of the populations

Genetic differentiation between populations was measured with a pairwise Fst estimate using the VCFtools v0.1.13 software (Danecek et al., 2011). In addition, a principal component analysis (PCA) was performed with the R package *Adegenet* (function *glPca*) (Jombart & Ahmed, 2011) to visualize the genetic structure of the populations.

2.4 | Runs of homozygosity

Runs of homozygosity (ROH) were identified for each fish using the PLINK v1.9 *homozyg* function (Chang et al., 2015) with the following options *'--homozyg-kb 500 –homozyg-window-snp 40 –homozyg-snp 40 –homozyg-gap 500 –homozyg-density 40 –homozyg-het 1*'. ROH was defined by a sliding window with a minimum length of 500 kb containing at least 40 homozygous SNPs. This minimum number of homozygous SNPs was chosen using the formula described by Purfield et al. (2012) in order to limit the number of ROHs that might only occur by chance.

2.4.1 | Estimation of inbreeding coefficients

The individual inbreeding coefficients (F_{ROH}) were calculated according McQuillian et al (2008) as $F_{i,ROH} = \frac{\sum length(ROH_i)}{LGenome}$

With $\sum length(ROH_i)$ the sum of ROH length in an individual i and $LGenome$ the total length of the autosomal genome covered by SNPs.

2.4.2 | Identification of ROH islands

For each SNP, the number of individuals with this SNP included in a ROH was calculated in order to identify the regions of the genome that were frequently homozygous in each population, i.e. constituting ROH islands (Nothnagel et al., 2010). These ROH hotspots may then be considered as signatures of positive selection (Saravanan et al., 2021).

To allow the comparison of ROH islands across populations, we implemented population-specific thresholds based on the ROH occurrence to define ROH islands, as proposed in many studies (Purfield et al., 2017; Mastrangelo et al., 2017; Zhang et al., 2018; Peripolli et al., 2018; Grilz-Seger et al., 2018; Gorssen et al., 2021; Illa et al., 2022). The number of individuals corresponding to the top 5% of SNPs most often found in a ROH within each population was chosen as a threshold to define a ROH island.

These top 5% values were equivalent to 35, 27, 5, and 10 individuals for LB, LC, SY, and HA, respectively. Values chosen within each population corresponded to 48.6%, 40.3%, 35.7%, and 50% of individuals with a ROH in LB, LC, SY, and HA, respectively. In addition, two close SNPs in the top 5% were considered in the same ROH island if there were separated by a distance lower than 500 kb with less than 40 SNPs in the gap stretch. The ROH island was delimited by a number of individuals, with ROH falling below the top 10% of the SNPs, which correspond to 30, 22, 3, and 7 individuals for LB, LC, SY, and HA populations, respectively.

208 2.4.3 | Detection of balancing selection signals based on regions without ROH

209 We used the ROH occurrence information per SNP to detect extreme heterozygous regions, i.e.
210 without ROH. In these regions, we have an enrichment of heterozygous SNP relative to the
211 genome-wide prevalence that may be due to balancing selection phenomena (Szpiech et al.,
212 2013).

213 Applying the same criteria as for defining ROH, the minimal size and number of SNPs to define
214 a heterozygous region were fixed to 500 kb and 40 SNPs, respectively. Moreover, two
215 successive SNPs were considered in the same region if they were separated by a distance lower
216 than 50 kb. A region was detected in extreme heterozygosity if less than 5% of individuals (per
217 population) have SNPs in ROH in the region, corresponding to a maximum of respectively 4
218 and 3 individuals with a ROH in LB and LC populations, and to 0 individual with a ROH in
219 SY and HA.

220

221 2.5 | Detection of signatures of selection based on Extended Haplotype Homozygosity (EHH)

222 For a given core allele, the EHH is defined as 'the probability that two randomly chosen
223 chromosomes carrying the core haplotype of interest are identical by descent for the entire
224 interval from the core region to the point x' (Sabeti et al., 2002). EHH measures the association
225 between a single allele from the study locus (the core region) with multiple loci at various
226 distances x (Sabeti et al., 2002). The iHS (Integrated Haplotype Homozygosity Score) proposed
227 by Voight et al. (2006) aims to compare the integrated EHH profiles obtained for a SNP in the
228 ancestral versus derived states. An extreme value of iHS corresponds to a positive selection
229 because a core haplotype with unusually high EHH and high frequency in the population

indicates the presence of a mutation that has spread through the population faster than the haplotype broke down.

EHH methodology requires haplotype information. Thus, genotype data must be phased before their calculation. We used FImpute3 (Sargolzaei et al., 2014) to phase the genotypes of the study females, considering all parents (including our study females) and offspring genotyped in LB, LC, and SY populations for different purposes (see respectively Prchal et al., 2022, Lagarde et al., 2022 and Paul et al., 2022). All parents (except 8 SY sires) were genotyped with the HD chip (Bernard et al., 2022), while offsprings (and 8 SY sires) were genotyped with a 57K chip only (Palti et al., 2015). Information used for phasing is given in Table 1. Due to the lack of genotyped offsprings, only the HD genotypes information was used to phase the genotypes of the HA population.

Once phasing was performed, the *rehh* R package (Gautier & Vitalis, 2012; Gautier et al., 2017) was used to conduct EHH-based analyses. EHH detection was stopped when the EHH value declined under 0.1 or when the gap between two consecutive SNPs was higher than 20 kb (*scan_hh* function with the following options: limehh = 0.1; maxgap=20 kb).

2.5.1 | Cross population Extended Haplotype Homozygosity (XP-EHH)

From EHH information, we used the XP-EHH statistic (*ies2xpehh* function) to compared the integrated EHH profiles (iES), two by two, between a French (popA) and the HA (popB) populations at the same focal SNP (Sabeti et al., 2007) as:

$$XP\_EHH = \frac{\ln\left(\frac{iES_{popA}}{iES_{popB}}\right) - Med\left[\ln\left(\frac{iES_{popA}}{iES_{popB}}\right)\right]}{SD\left[\ln\left(\frac{iES_{popA}}{iES_{popB}}\right)\right]}$$

The median (*Med*) and standard deviation (*SD*) of $\ln(iES_A/iES_B)$ were computed over all the analysed SNPs.

2.5.2 | Integrated Haplotype Homozygosity Score (iHS)

In the same way, we used the iHS test (Voight et al., 2006) to evaluate the evidence of positive selection based on haplotype frequencies in a single population, using the *ihh2ihs* function of the package *rehh*. This statistic was based on the log-ratio of the integrated EHH (iHH) for haplotypes with the ancestral (A) *versus* the derived (D) alleles and was computed for each autosomal SNP as $iHS = \frac{\ln\left(\frac{iHH_A}{iHH_D}\right) - Mean_p\left[\ln\left(\frac{iHH_A}{iHH_D}\right)\right]}{SD_p\left[\ln\left(\frac{iHH_A}{iHH_D}\right)\right]}$

The average ($Mean_p$) and standard deviation (SD_p) of $\ln(iHH_A/iHH_D)$ were computed over all the SNPs with a derived allele frequency *p* similar to that of the core SNP. In our study, the ancestral allele state is unknown. Therefore, we assumed that the most frequent allele represents the ancestral state as proposed by Bahbahani et al. (2015).

2.5.3 | Detection of candidate regions

To detect candidate regions for signatures of selection based on the iHS test, we used *the calc_candidate_region* function of the R package *rehh* (Gautier & Vitalis, 2012). We considered windows of 500 kb across the genome containing at least 30 SNPs, with 10 kb of overlapping. A region was considered as under positive selection if at least one SNP had a log(p-value) > 4 and extreme iHS value *i.e.* $|iHS| \geq 2.5$.

2.6 | Identification of common regions under positive selection

ROH islands and regions identified by iHS were pooled within each population. Then, the intersection set of the regions identified by one or another method across the four studied populations was established. We eliminated an intersection from the study if one population does not have at least one SNP with an |iHS|≥ 2.5 or enough individuals with an ROH in the intersected region. So, only intersections containing either ROH island or extreme iHS (|iHS| ≥ 2.5) for the four populations were thus further analyzed.

2.7 | Gene analysis

The genes annotated in the regions under positive or balacing selection were identified from the NCBI *Oncorhynchus mykiss* genome assembly (GCA_013265735.3). Gene symbols were checked, and, if necessary, familiar names were added using the information available from GeneCards (https://www.genecards.org/).

Gene ontology (GO) terms study was performed with 'g:profiler' (Raudvere et al., 2019; https://biit.cs.ut.ee/gprofiler/gost) for the list of genes identified in the regions of interest.

Percent identity of rainbow trout proteins with nine other vertebrate species (human, mouse, cow, goat, pig, chicken, zebrafish, medaka, and Atlantic salmon) was established using the blastp tool (Basic Local Alignment Search Tool on proteins).

3| Results

3.1 | Genetic diversity within and across populations

The ROH statistics and inbreeding coefficients are presented in **Table 2** for all the populations. The average number of ROH per individual varied between 141 (SY) and 168 (LB). French selected lines had larger average sizes of ROH than populations SY and HA. The average inbreeding coefficients of HA individuals were between three (compared to SY) and five (compared to LB) times lower than those of the French lines.

Based on genome-wide Fst values, large differentiation of around 0.28 was observed between HA and any of the French populations (Table 3). In the PCA figure (Figure 1), the three French lines differed strongly from the American pooled populations, and the first two PCA axes explained 29% of the total genetic variation. In addition, Fst values indicated that all the French lines were moderately differentiated (0.104 – 0.122).

Using the XP-EHH statistic, we identified 93, 105 and, 135 regions that strongly discriminated HA from LB, LC and SY, respectively. Among these regions, 34 regions were shared, spanning about 32 Mb in total over 21 chromosomes, and differentiated any of the French lines from the American HA pooled population (Supplementary information S1).

The distribution of the proportion of individuals having a ROH at each SNP position is presented in Figure 2. In average, ROH were more shared between individuals for selected lines (LB and LC, on average, 23.39% and 19.82% of individuals respectively) than for other populations (SY and HA, on average, 13.67% and 8.91% of individuals respectively). Probably linked to the composite nature of the HA population (5 sub-groups of 4 individuals), HA contained the lowest number of shared ROH among the individuals but also showed the most shared ROH among individuals.

3.2 | Signatures of positive selection

3.2.1 | ROH islands

The sharing of ROH among individuals, regardless of the population considered, was presented Figure 3. Eight ROH islands were shared by at least 2 populations, and a minimum of 50% of individuals concerned. However, only three of these regions were defined as ROH island in each of the 4 populations.

We listed all ROH islands within each population which resulted in the identification of 270 ROH islands distributed among the four populations (Supplementary Informations S2 to S5, for LB, LC, SY, and HA, respectively). The ROH islands were not evenly distributed across populations and chromosomes. The average ROH island size was 2,737 kb, varying from 1,593 kb to 4,465 kb, depending on the population. The longest ROH island was observed in SY (21.4 Mb), while the shortest one was observed in LC (16.1 kb).

3.2.2 | Identifying selection signatures using iHS

The log(p-values) of iHS calculated along the genome are presented in Figure 4 for each population (all regions identified with *calc-candiate_region* are described in Supplementary Informations S6-S9). While numerous regions have been detected as under positive selection overall, fewer candidate regions were detected for the French lines (LB, LC and, SY) than for the Amercian pooled population (HA). The genome-wide highest estimated values of |iHS| were 8.97, 7.24, 5.67, and 9.09 for LB, LC, SY, and HA, respectively (with log(p-values) > 7.8).

In total, 72, 68, 76, and 54 ROH islands were identified in LB, LC, SY, and HA populations respectively (Figure 5). Using iHS statistics, 55, 69, 73, and 362 signatures of selection were detected for LB, LC, SY, and HA populations, respectively. Only 10.4%, 8.7%, 8.0%, and 5.6%

of the regions were detected by both methods (ROH + iHS) for LB, LC, SY, and HA populations, respectively.

3.2.3 | Regions under positive selection shared by the four studied populations

Among the numerous regions identified for each population by either ROH or iHS methods, only nine regions were shared by the four studied populations (Table 4). The average size of these shared regions was 1135 kb. Five of them were located on chromosome 2, and the four other regions were on chromosomes 12, 15, 16, and 20, respectively.

Depending on the population, six regions were identified by either ROH or iHS metrics. Two regions, chr2_c and chr15_a, were only detected by ROH in all four populations, while a single region, chr16_a, was only identified through significant iHS statistics in all the four populations (Supplementary Information S10). The list of genes annotated in the nine shared genomic regions is given in Supplementary Informations S11.

3.3 | Signatures of balancing selection

3.3.1. Regions under balancing selection detected within population

In total, 14, 24, 158, and 265 hot spots of polymorphism (i.e. without ROH) were identified in LB, LC, SY, and HA populations, respectively. The numbers of heterozygous regions detected for SY and HA populations were drastically larger than those observed for the LB and LC selected lines. The average size of the detected heterozygous regions was 1,400 kb, varying from 1,086 kb to 1,828 kb, depending on the population.

Tables listing all heterozygous regions within each population are presented in Supplementary Informations S12 to S15, for LB, LC, SY, and HA, respectively.

364 3.3.2. Regions under balancing selection shared by the four studied populations

365 A substantial lack of ROH was observed in four regions of all studied rainbow trout populations
366 (Table 5). Two of them, chr10_a and chr19_a, were particularly small (53 kb and 70 kb,
367 respectively), but still contained at least 20 SNPs. The region chr10_a only encoded one of the
368 introns of the *ctnna2* (= catenin alpha 2*)* gene while chr19_a was composed of two genes,
369 *smarca5* (=SWI/SNF-related matrix-associated actin-dependent regulator of chromatin
370 subfamily A member 5) and *frem2* (= FRAS1-related extracellular matrix protein 2). A second
371 heterozygous region on chromosome 19 was larger (163 kb) but contained a single annotated
372 gene, *pou4f2* (= POU domain, class 4, transcription factor 2-like). A last region chr13_a
373 spanned over 1,100 kb) on chromosome 13 and was composed of 25 genes. The list of genes
374 annotated in the four shared genomic regions is given in Supplementary Informations S16.
375

376 3.4 | Identification and roles of genes underlying the regions under selection across all
377 populations

378 3.4.1. Common homozygous regions under positive selection

379 The nine common homozygous regions contained a total of 253 genes (listed in Supplementary
380 Information S11). A gene ontology (GO) study was performed and showed a significant over-
381 representation (p-value < 0.01) among the 253 genes of functions related to the following GO
382 terms: membrane (GO:0016020, CC: Cellular Component, *p-value = 1.3e10*$^{-5}$), intrinsic and
383 integral component of membrane (GO:0031224; GO:0016021, CC, *p-value = 0.001/0.005*), ion
384 binding (GO:0043167, MF: Molecular Function, *p-value = 0.002*), and nuclear speck
385 (GO:0016607, CC, *p-value = 0.008*).

Among the nine studied regions, the three regions chr2_a, chr2_c, and chr15_a, that contain less than ten genes annotated in each, were analyzed in further detail to accurately define the roles of underlying genes. The 17 genes located in these three regions are listed in Table 6 with their associated biological functions. These genes play key roles in protein transduction/maturation, genome stability, embryonic development, growth, energetic function, reproduction, or immune function. In addition to this list of genes, a subset of 15 genes in the six other homozygous regions already identified as signatures of selection in the literature were further studied in terms of their biological functions. Detailed information for these genes is also given in Table 6.

We studied the degree of protein identity among 10 vertebrate species for all the 17 genes of regions chr2_a, chr2_c, and chr15_a (Table 7), considering a protein as highly conserved if its identity between rainbow trout and other species was higher than 85%. Except for the proteins linked to *cep162* (centrosomal protein of 162 kDa) and *zp4* (zona pellucida sperm-binding protein 4-like) genes, all other proteins were highly conserved at least between the two studied salmonids. In each of the three regions, one or two genes were highly conserved across the all ten study species: in chr2_a, rainbow trout *cdk14* (cyclin-dependent kinase 14) protein had a percent identity between 86 and 99.6% with the other species; in chr2_c, rainbow trout *brsk2a (*serine/threonine-protein kinase *brsk2*) protein had between 92 and 96.3 % of percent identity with the other species; in chr15_a, two genes, *chn1*(n-chimaerin) and *atp5mc1* (ATP synthase lipid-binding protein, mitochondrial), also had protein percent identity ranging from 85% to 98% depending of the species.

Some other rainbow trout proteins (*tsnare1*, t-SNARE domain-containing protein 1*; pttg1IP*, pituitary tumor-transforming gene 1 protein-interacting protein) were conserved to a lesser extent (minimum 65% of percent identity) with the three other fish species, some being also

conserved at least with chicken (*adgrb1*, adhesion G protein-coupled receptor B1*;b4galnt4a,* N-acetyl-beta-glucosaminyl-glycoprotein 4-beta-N-acetylgalactosaminyltransferase 1) or even with all the nine study species (*zc3h15*, zinc finger CCCH domain-containing protein 15).

3.4.2. Common heterozygous regions under balancing selection

The four common heterozygous regions (Table 5) contained 29 genes (listed in Supplementary Information S16). A gene ontology (GO) terms study showed no significant over-representation of specific GO terms.

The degree of protein percent identity among 10 various vertebrate species for these 29 genes are presented in supplementary information S17.

Regions chr10_a, chr19_a, and chr19_b contained only a few genes and were then analyzed in further detail to accurately determine the role of underlying genes (Table 8). These genes play key roles in cellular and nuclear organisation and in embryonic development.

4| Discussion

The objective of our study was to detect signatures of selection into domestic rainbow trout. To reach that goal we studied four genetically distinct populations coming from different locations either in France or in the North-West of the USA. We used two different approaches, ROH and EHH, to detect the genomic regions shared by all populations using a HD SNP. We were able to detect 9 very conserved regions and 4 hotspots of polymorphism, corresponding to 253 and 29 annotated genes, respectively.

433 ## 4.1 | Genetic structure

434 First, we described the genetic structure of the populations under scrutiny. The three French
435 lines were moderately differentiated with Fst ranging from 0.10 to 0.12. These estimations were
436 congruent with those computed by D'Ambrosio et al.'study (2019) with the same populations
437 that ranged between 0.09 and 0.14, but were estimated using a 38K SNPs array. These moderate
438 differences between the 3 French populations were consistent with the PCA we performed and
439 the history of these populations with a partly common INRAE origin (D'Ambrosio et al.,
440 2019). This trend is shared between European populations with for instance an average Fst of
441 0.13 between 12 European rainbow trout strains (Gross et al., 2007) or 0.12 among 9 Norwegian
442 populations (Glover, 2009). Similarly, US farmed populations are also weakly to moderately
443 differenciated with an average Fst of about 0.09 (Silverstein et al., 2004) or 0.13 (Liu et al.,
444 2017) and pairwise values ranging from 0.06 and 0.16. We observed a similar pattern in the
445 present study with the HA population that consisted in samples from 5 locations, which all
446 clustered together in the PCA. Reversely, we observed a large differentiation between our
447 French and US populations revealed by large Fst values (0.27-0.29). This is likely the result of
448 numerous factors, including selection, genetic drift and absence of gene flow between these
449 very geographically distant populations. In addition, the European farmed populations
450 originated from Californian domesticated strains, that have been shown to differ from strains
451 of North-Western USA (Stanković et al., 2015). We found 34 haplotypes distributed over 21
452 chromosomes that differed between the American pooled population (HA) and all French
453 populations (Supplementary information S1).
454 Due to the moderate to large differentiation between the 4 populations, the conserved regions
455 across all populations are likely to be the result of ancient natural selection traces.
456

4.2 | Comparison of methods to detect common signatures of positive selection

We used a double check of positive selection traces in the genome by using both ROH and EHH approaches. However, for each population, only a few regions were identified by both methods. These regions detected by more than one method represent stronger evidence of selection signatures since outlier markers detected by various genome scan methods help to uncover true selection signatures by reducing the number of false positives.

Even if both methods evaluate the homozygous large stretches in the genome, iHS also considers information based on haplotypic version and linkage disequilibrium from a core SNP. ROH approach detects homozygous regions regardless of their haplotypic versions, contrary to iHS. Thus, it may detect a signature of positive selection even if various haplotypes were present at the homozygous state in the population. In addition, while the ROH approach only detects the homozygous large stretches (at least 500 kb in the present study), iHS can detect small regions under positive selection as the only limitation in EHH region size is based on a threshold value for a minimum LD (0.10). Consequently, the sizes of the detected homozygous region varied between 1,065 kb and 2,857 kb based on ROH metrics and between 1,000 kb and 1,600 kb with iHS statistics.

The high number of regions (55, 69, 73, and 362) detected by iHS in our study was consistent with numbers detected in either Atlantic salmon (López et al., 2019) or cattle (Saravanan et al., 2021). However, these two previous studies used a lower threshold than ours (log(p-value) = 3 and 2, respectively vs 4 in the present study). Lower numbers of regions were previously detected by iHS in rainbow trout by Cádiz et al. (2021) and in Coho salmon by López et al. (2021). We speculate that these differences in the numbers of detected signals may be linked to the lower density of SNPs they used in both studies (57K or 200K chip versus 665K chip for

our study) and the subsequent lower ability to detect LD and haplotypes at fine scale. Indeed, in the Chilean rainbow trout study (Cádiz et al., 2021**)**, only one signal of positive selection was detected by iHS located at 6.398-14.936 Mb on chromosome 20 of the Swanson reference genome, which corresponds to the region 7.488-16.111 Mb on chromosome 20 of the Arlee reference genome. Nevertheless, we also detected by iHS signals of selection in each of our four studied populations, located at 10.5-16.5 Mb for LB, 11.2 – 13.3 Mb for LC, 13.0 – 14.2 Mb for SY and 12.3-13.2 Mb for HA (Supplementary Informations S6 to S9 for LB, LC, SY, and HA, respectively). Thus, all these signals were consistent with the larger region identified by Cádiz et al. (2021).

A common putative selection signature located at 13.0-13.2Mb could also be shared by all studied populations. In this 200kb-region, we observed at least one iHS value over |2.5| for LB, LC and SY lines, but not for HA population. In this region of 200 kb on chromosome 20, six genes were identified (*lgi1, noc3l, plce1, slc35g1, fra10ac1, tbc1d12*). Among these genes, Cádiz et al. (2021) identified two candidates genes associated with domestication, *noc3l* (nucleolar complex protein 3 homolog**)** and *plce1* (1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase epsilon-1). Both are related to early development traits in zebrafish (*noc3l*: Walter et al., 2009; *plce1*: Zhou et al., 2009).

4.3 | Biological functions of genes under positive or balancing selection

Among the 282 genes in the 13 regions detected under either positive or balancing selection, most genes seem to play essential roles in fitness as expected with such a dataset comprising both European and US populations. They are related to all main biological functions (genome stability, cell organization, neuronal and embryonic development, energy metabolism, growth, reproduction, and immunity). All identified biological functions were already described in other

studies of signatures of selection in farmed rainbow trout (Cádiz et al., 2021) and other domesticated species (López et al., 2018, 2018; Naval-Sanchez et al., 2020; Baesjou & Wellenreuther, 2021; Signer-Hasler et al., 2022).

4.3.1 Hotspots of heterozygosity and balancing selection for fitness traits

In livestock species, many variants under balancing selection are known to improve performance in heterozygote state but cause defect in homozygous state (Hedrick, 2015; Georges et al., 2019). However, in such cases of balancing selection, there is generally only one homozygous state, which is deleterious at a locus level, while the alternative homozygous state is observed in the population. In our study, we highlight four regions potentially involved in balancing selection for which we observed a lack of any kind of long streches of homozygosity. Even if these regions are extremely heterozygous, the proteins associated with these genes are highly conserved among vertebrates (Supplementary information S17). Many processes may explain these surprising observations at first glance. First of all, these regions may concentrate polymorphism in non-coding parts of the genome. Polymorphism in intronic regions may promote various proteins by allowing alternative splicing. We may also observe an excess of synonymous polymorphism in exons without effects on proteins. Further analyses must be conducted to better understand the mechanisms underlying the maintenance of extreme polymorphism, whether to validate the hypothesis of balancing selection or the existence of high mutation and recombination rates in these specific regions.

In the heterozygous region chr10_a, the gene *ctnna2 (*Table 8) enables actin filament binding activity and is involved in the regulation of neuron migration and neuron projection development. Thus, *ctnna2* plays an essential role in brain development among vertebrates

(Uvarov et al., 2014). In yonlong grouper, *ctnna2* seems implicated in vertebral development, because significantly differentially expressed between normal and fish with lordosis (Li et al., 2022). In mice, a homozygous for a mutation of *ctnna2* reduced body weight, male fertility, and induced brain abnormalities (hypoplastic cerebellum, abnormal foliation pattern, ectopic Purkinje cells, and abnormal pyramidal cells in the hippocampus). While the protein associated to this gene is highly conserved among vertebrates (Uvarov et al., 2014; Supplementary information S17), the gene exhibits a strong polymorphism in all the four studied rainbow trout populations. However, a large part of its polymorphism is located in one intronic region (intron 6-7) of *ctnna2*. In the zfin database, five transcripts of this gene were identified (three mRNA and two non-coding RNA). We hypothezise that the polymorphism in the intronic region of *ctnna2* is essential for alternative splicing.

In the heterozygous region chr13_a (Supplementary information S16), *mmd* and *map2k4* are identified as highly conserved across vertebrates (Supplementary information S17). The gene *mmd* plays an important role in maturing macrophages, which is essential for immune response as observed in mice (Lin et al., 2021). The gene *map2k4* is implicated in a variety of cellular processes (proliferation, differentiation, transcription regulation, development), seems to play a role in liver organogenesis and embryonic development during gastrulation, as demonstrated by morpholino-mediated knockdown in zebrafish (Seo et al., 2010), and implicated in immune response in yellow catfish **(**Zheng et al., 2022). The inflammatory process in immune response seems linked to the polymorphism of the *map2k4* gene, which is consistent with our hypothesis of balancing selection, and more precisely potential ancestral trans-species polymorphism in this genomic region (Gu et al., 2016; Fijarczyk & Babik, 2015). Trans-species polymorphism is a crucial evolutionary mechanism for sharing adaptative genetic variation across taxa (Klein et al., 1998). The study of this mechanism has primarly

552 concentrated on major histocompatibility complex genes, but a few studies described this
553 process for other immune genes (Ferrer-Admetlla., et al., 2008; Leffler et al., 2013; Těšický &
554 Vinkler, 2015). Maintaining genetic diversity in regions related to the immune system may be
555 essential to resilience against various pathogens. In addition, this region of chromosome 13 has
556 been recently detected as a significant QTL playing a role on resistance to temperature (Lagarde
557 et al., 2022).
558 In the heterozygous region chr19_a (Table 8), the protein encoded by *smarca5* is a component
559 of chromatin remodeling and spacing factor RSF, a facilitator of the transcription of class II
560 genes by RNA polymerase II (zebrafish: Armas et al., 2013; Ding et al., 2021; mice: Limi et
561 al., 2018). The protein is highly conserved among vertebrates (Supplementary information
562 S17), which is consistent with its essential role thought to regulate the transcription of many
563 genes by altering the chromatin structure around those genes. In the same region
564 chr19_a, *frem2* codes for an extracellular matrix protein required for maintenance of the
565 integrity of skin and renal epithelia in zebrafish (Gautier et al., 2008). This protein is moderately
566 conserved across vertebrates (Supplementary information S17). In a study searching for
567 genomic regions with ancestral trans-species polymorphism shared between humans and
568 chimpanzees (Leffler et al., 2013), *frem*3, an important paralog of *frem2*, was identified under
569 balancing selection. However, further studies should test the hypothesis of trans-species
570 conservation of *map2k4* and *frem2* genes that may help to understand the various cellular
571 processes in which the gene is implicated.
572 In the heterozygous region chr19_b (Table 8), *pou4f2* protein is highly conserved among
573 vertebrates (Supplementary information S17) and is a tissue-specific DNA-binding
574 transcription factor involved in the development and differentiation of specific cells. It

maintains the visual system neurons in the retina and the lateral line (DeCarvalho et al., 2004) and seems also related to cardiac development in zebrafish (Maskell et al., 2017).

4.3.2 Hotspots of homozygosity and positive selection for essential biological functions

4.3.2.1 | Regions and genes involved in cellular and nuclear organization, and embryonic development

In homozygous region chr2_a, three genes plays important roles in embryonic development and then fitness (*cep162*, *tsnare1*, *mrap2*, Table 6). In the homozygous region chr2_b (Table 6), the gene *pbx1* (pre-B-cell leukemia transcription factor 1) is related to early development in zebrafish (Teoh et al., 2010). Mutations in this gene generally cause major malformations, which seem to play an essential role in survival in various species (zebrafish: Teoh et al., 2010; mouse: Selleri et al., 2001; human: Le Tanno et al., 2017). It was detected under positive selection in a Chilean farmed rainbow trout population (Cádiz et al., 2021). However only moderate percent identity (> 65%) is observed between *pbx1* proteins across vertebrates.

In the homozygous region chr15_a, many genes *(chn1, atp5mc1, zc3h15, nid2 and brca2)* were playing essential roles in cell functioning and early development (Table 6). However only two of them were highly conserved among vertebrates (> 85%; *chn1* and *atp5mc1*). The gene *atp5mc1* is a crucial gene for mitochondrial cristae morphology, and plays important roles in metabolic processes associated to growth (Table 6; Palmer et al., 2011; Miller et al., 2019; Wang et al., 2020). In zebrafish, a morpholino knockdown of *chn1* reveals its crucial role in early development, revealing severe abnormalities (development of round somites, lack of yolf extension, and kinkled posterior notochord) (Leskow et al., 2006).

Three genes located in close vicinity in region chr16_a (between 46.42 and 46.53 Mb; Supplementary information S11), *samd10* (sterile alpha motif domain-containing protein 10-

like*)*, d*najc5* (dnaJ homolog subfamily C member 5-like*)*, and *tpd54* (tumor protein D54) were also detected in close chromosomal vicinity and under positive selection in ten modern goat breeds and one wild Bezoar goat (Signer-Hasler et al., 2022). This cluster of genes has a significant role in survival and cellular processes (Table 6). In addition, in this region chr16_a, the protein of the gene *magi2* (membrane-associated guanylate kinase, WW and PDZ domain-containing protein 2, Table 6) plays a vital role in embryogenesis in zebrafish (Borah et al., 2016). The gene *magi2* was also identified under positive selection in a domesticated sheep breed compared to the wild Asiatic mouflon (Cumer et al., 2021).

4.3.2.2 | Regions and genes involved in neural and brain development, and behaviour

In total, we identified 7 genes as primarly associated to neural and brain development in both regions detected under positive selection (*tsnare1, cdk14, brsk2a, auts2, brd2, znf135,* and *grxcr1*). Some genes (*brsk2a, znf135, grxcr1, auts2; Table 6*), related to brain development may induce behavior modifications in farmed animals, that may be related to domestication processes (Pasquet, 2018; Milla et al., 2021; Deng et al., 2022; Liu et al., 2022). This is in line with Żarski et al. (2020) study showing that domestication modulates gene expression involved in neurogenesis.

In particular, the gene *auts2* gene was previously identified under positive selection both in cattle (Consortium, bovine Hapmap, 2009) and in domesticated Atlantic salmon populations from Canada and Scotland compared to their wild Atlantic salmon counterpart (López et al., 2018). The gene *znf135* was also detected under positive selection in a farmed population of Atlantic salmon compared to a wild-type population (Gutierrez et al., 2016). The gene *grxcr1* was detected under positive selection in the Tharparkar cattle (Saravanan et al., 2021). It

strongly suggests that all these genes play a key role in domestication processes and may act on essential behaviours in both terrestrial and aquatic farmed animals.

4.3.2.3 | Regions and genes involved in growth metabolism

Genes related to growth metabolism were only identified in four regions under positive selection, none of them were detected in high heterozygosity regions.

In the homozygous region chr2_a (Table 6), the protein *mrap2* (melanocortin-2 receptor accessory protein 2A) is associated to growth. A lack of this gene exhibit severe obesity in many species (human, zebrafish, rodent: Liu et al., 2013; sea lamprey: Zhu et al., 2019; snakehead: Wang et al., 2021). Yoshida et al. (2017) detected a growth-QTL in Atlantic salmon and considered *mrap2* as a candidate gene for growth up to 25 months. In addition, *mrap2* was identified in the Chilean farmed rainbow trout population as under positive selection (Cádiz et al., 2021). A QTL related to sea lice resistance in rainbow trout (Cáceres et al., 2021) was also detected in the region chr2_a (located from 10.43 Mb to 11.81 Mb of the Swanson reference genome, which corresponds to 26.69 Mb – 28.07 Mb of the Arlee reference genome**).** Cáceres et al. (2021) explained that having a high potential for growth seem essential for sea lice resistance. Indeed, proteomic investigations allow to establish a link between growth and immune function in salmonids (Causey, 2018).

In homozygous region chr2_b (Table 6), the *col9a2* (collagen alpha-2(IX) chain) gene is a component of cartilage and seems also related to growth (Xu et al., 2022).-This gene is detected under positive selection in a Scottish farmed population of Atlantic salmon (López et al., 2018). In addition, the gene *scap* (sterol regulatory element-binding protein cleavage-activating

644 protein) was already identified under positive selection in six farmed Pacific white Shrimp
645 populations (Wang et al., 2022).
646 In the homozygous region chr2_d (Table 6), the gene *igf-1α* (insuline like growth factor receptor
647 1a) plays an important role in growth and transformation events. In salmonids, expressions of
648 *igf-1α* and growth hormone were demonstrated to be modified between domesticated and wild
649 populations of rainbow trout and coho salmon (Tymchuk et al., 2009). The same observation
650 was made with a higher expression of *igf-1α* between larvae from domesticated spawners than
651 larvae from wild spawners of the Eurasian perch (Palińska-Żarska et al., 2021). In addition, *igf-*
652 *1* was also observed as a marker of domestication in dogs (Wayne & vonHoldt, 2012).
653 In the homozygous region chr16_a, the *emilin-3a* gene (elastin microfibril interfacer 3a, Table
654 6) plays a role in extracellular matrix organization and elastic fiber formation. Its gene
655 expression is related to embryonic development and involved in muscle fiber development in
656 zebrafish (Milanetto et al., 2008). *Emilin-3a* had already been identified as under positive
657 selection in one population of F2 Australian snapper farmed population compared to the first
658 generation (F1) of domestication of a wild population (Baesjou & Wellenreuther, 2021). Thus,
659 this signature of selection can be considered as a result of the domestication process.
660 All identified growth-related genes seem associated with domestication. This assertion is
661 confirmed for five genes (*mrap2*, *col9a2*, *scap*, *igf-1α*, *emilin-3*) that were also identified under
662 positive selection in various farmed populations with favorable alleles linked to better growth
663 phenotypes (Table 6).

664

665 4.3.2.4 | Regions and genes involved in reproduction

Very few genes directly related to reproduction traits were only identified in highly homozygous regions.

In the homozygous region chr2_b, the *brd2* (bromodomain-containing protein 2, Table 6) gene is implicated in several biological process (see section 4.4.1.). It seems related to oogenesis and egg-to-embryo transition in zebrafish (DiBenedetto et al., 2008), which is consistent with a QTL detected for egg size in this region in rainbow trout (D'Ambrosio et al., 2020). Moreover, it seems that *brd2* is involved in spermatogenesis or folliculogenesis, as demonstrated in situ on mice cells (Rhee et al., 1998). Khendek et al. (2017) compared the reproductive performances (egg size, gonadal histology, hormonal levels) between domesticated and F1 with wild broodstock of Eurasian perch populations. They showed that domestication may have increased the oocyte diameter and the level of 17β-Estradiol, but decreased the embryo survival of domesticated fish. This gene was also identified under positive selection in a selected Canadian population of Atlantic salmon (López et al., 2018).

In the homozygous region chr15_a (Table 6), the gene *zp4* has already been identified under positive selection in a farmed Scottish population of Atlantic salmon compared to a wild population (López et al., 2018), and may be related to domestication process.

4.3.2.5 | Regions and genes involved in immunity

Magris et al. (2022) observed that regions under positive selection revealed an enrichment of KEGG terms related to viral infection in farmed brown trout. However, it should be noticed that in our study, few genes related to immune function were identified and no enrichment in immune terms was observed in GO analysis.

688 Genes related to immune function were mainly identified in three different regions detected as
689 putatively under positive selection for two of them and under balancing selection for the last
690 one.
691 In the homozygous region chr2_b (Table 6), genes *tnfaip8l2b* (tumor necrosis factor, alpha-
692 induced protein 8-like protein 2 B) and *atg5* (autophagy protein 5) are related to immune
693 functions. Note that atg5 is well conserved across vertebrates (> 80%).
694 In the region chr15_a, the gene *zc3h15* (Table 6) seems to have an inhibitory effect on HIV-1
695 replication and then on HIV infection in vitro (mice cells) (Capalbo et al., 2010).

696 A last gene in the region chr16_a, the gene *atp1b3* (sodium/potassium-transporting ATPase
697 subunit beta-1-interacting protein 3) was also identified under positive selection in farmed
698 Atlantic Salmon (Naval-Sanchez et al., 2020). In the Senegalese sole, atp1b3a and atp1b3b
699 paralogs have been hypothesized to be involved in response to low salinity (Armesto et al.,
700 2015). In addition, this gene is involved in some immune responses. It was shown in cell culture
701 study, that *atp1b3* inhibits hepatitis B virus replication via inducing NF-kappa B activation
702 (human: Zhang et al., 2021) and is involved in numerous viral propagation such as HIV and
703 EV71 (Zheng et al., 2020**)** in cell culture experiments.
704

705 4.4. Conclusion

706 To sum up, we identified 13 regions under selection with numerous genes strongly involved in
707 essential biological functions. By identifying signatures of selection shared by our four studied
708 populations, we have focused our detection on regions related to ancient evolutionary processes
709 that are essentially important for species survival. We only identified nine homozygous regions
710 presumably under positive selection and four heterozygous regions putatively under balancing

selection in four different rainbow populations. While common homozygous regions may be associated with important biological functions underlying both fish fitness and domestication, the heterozygous regions seem mainly linked to fitness functions (cell organization, embryonic development, and immunity) which are involved at different developmental stages or to cope with various pathogens or abiotic stressors. Maintaining genetic diversity in these regions could be essential for the species survival.

This study allows us to confirm the importance of a large set of 17 genes already detected as under positive selection in previous studies, among which 10 genes were identified in fishes (*auts2, atp1b3, zp4, znf135, igf-1α brd2, col9a2, mrap2, pbx1* and *emilin-3*). We also identify new promising candidate genes as important for rainbow trout fitness. In our opinion, this study substantially increases knowledge of evolutionary processes and helps to determine the genomic location and the nature of the genetic variation that must be maintained in rainbow trout populations for domestication and selection purposes**.**

Acknowledgements

We thank the two breeding compagnies “Viviers de Sarrance” and "Milin Nevez” that allow us to use their HD genotypes to perform the study. This study was partly supported by the European Maritime and Fisheries Fund and FranceAgrimer (Hypotemp project, n◦ P FEA470019FA1000016).

References

1) Armas, P., Margarit, E., Mouguelar, V. S., Allende, M. L., & Calcaterra, N. B. (2013). Beyond the binding site: in vivo identification of tbx2, smarca5 and wnt5b as molecular targets of CNBP during embryonic development. *PloS one*, *8*(5), e63234. https://doi.org/10.1371/journal.pone.0063234

2) Armesto, P., Infante, C., Cousin, X., Ponce, M., & Manchado, M. (2015). Molecular and functional characterization of seven Na+/K+-ATPase β subunit paralogs in Senegalese sole (Solea senegalensis Kaup, 1858). *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, *182*, 14-26. https://doi.org/10.1016/j.cbpb.2014.11.011

3) Baesjou, J. P., & Wellenreuther, M. (2021). Genomic signatures of domestication selection in the Australasian snapper (Chrysophrys auratus). *Genes*, *12*(11), 1737. https://doi.org/10.3390/genes12111737

4) Bahbahani, H., Clifford, H., Wragg, D., Mbole-Kariuki, M. N., Van Tassell, C., Sonstegard, T., ... & Hanotte, O. (2015). Signatures of positive selection in East African Shorthorn Zebu: A genome-wide single nucleotide polymorphism analysis. *Scientific reports*, *5*(1), 1-13. https://doi.org/10.1038/srep11729

5) Baker, S., Booth, C., Fillman, C., Shapiro, M., Blair, M. P., Hyland, J. C., & Ala-Kokko, L. (2011). A loss of function mutation in the COL9A2 gene causes autosomal recessive Stickler syndrome. *American Journal of Medical Genetics Part A*, *155*(7), 1668-1672. https://doi.org/10.1002/ajmg.a.34071

6) Balon, E. K. (1995). Origin and domestication of the wild carp, Cyprinus carpio: from Roman gourmets to the swimming flowers. *Aquaculture*, *129*(1-4), 3-48. https://doi.org/10.1016/0044-8486(94)00227-F

7) Balon, E. K. (2004). About the oldest domesticates among fishes. *Journal of fish Biology*, *65*, 1-27. https://doi.org/10.1111/j.0022-1112.2004.00563.x

8) Bernard, M., Dehaullon, A., Gao, G., Paul, K., Lagarde, H., Charles, M., ... & Phocas, F. (2022). Development of a high-density 665K SNP array for rainbow trout genome-wide genotyping. Frontiers in Genetics, 13, 1664-8021 https://doi.org/10.3389/fgene.2022.941340

9) Borah, S., Barrodia, P., & Swain, R. K. (2016). Nucleolar protein 4-like has a complex expression pattern in zebrafish embryos. *International Journal of Developmental Biology*, *60*(1-2-3), 53-56. https://doi.org/10.1387/ijdb.150307rs

10) Bovine HapMap Consortium, Gibbs, R. A., Taylor, J. F., Van Tassell, C. P., Barendse, W., Eversole, K. A., ... & Dodds, K. G. (2009). Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science*, *324*(5926), 528-532.

11) Browning, S. R., & Browning, B. L. (2011). Haplotype phasing: existing methods and new developments. *Nature Reviews Genetics*, *12*(10), 703-714. https://doi.org/10.1038/nrg3054

12) Bruford, M. W., Bradley, D. G., & Luikart, G. (2003). DNA markers reveal the complexity of livestock domestication. *Nature Reviews Genetics*, *4*(11), 900-910. https://doi.org/10.1038/nrg1203

13) Cáceres, P., Barría, A., Christensen, K. A., Bassini, L. N., Correa, K., Garcia, B., ... & Yáñez, J. M. (2021). Genome-scale comparative analysis for host resistance against sea lice between Atlantic salmon and rainbow trout. *Scientific reports*, *11*(1), 1-11. https://doi.org/10.1038/s41598-021-92425-3

14) Cadieux-Dion, M., Andermann, E., Lachance-Touchette, P., Ansorge, O., Meloche, C., Barnabe, A., ... & Cossette, P. (2013). Recurrent mutations in DNAJC5 cause autosomal dominant Kufs disease. *Clinical genetics*, *83*(6), 571-575.

15) Cádiz, M. I., López, M. E., Díaz-Domínguez, D., Cáceres, G., Yoshida, G. M., Gomez-Uchida, D., & Yáñez, J. M. (2020). Whole genome re-sequencing reveals recent signatures of selection in three strains of farmed Nile tilapia (Oreochromis niloticus). *Scientific reports*, *10*(1), 1-14. https://doi.org/10.1038/s41598-020-68064-5

16) Cádiz, M. I., López, M. E., Díaz-Domínguez, D., Cáceres, G., Marin-Nahuelpi, R., Gomez-Uchida, D., ... & Yáñez, J. M. (2021). Detection of selection signatures in the genome of a farmed population of anadromous rainbow trout (Oncorhynchus mykiss). *Genomics*, *113*(5), 3395-3404. https://doi.org/10.1016/j.ygeno.2021.07.027

17) Capalbo, G., Müller-Kuller, T., Dietrich, U., Hoelzer, D., Ottmann, O. G., & Scheuring, U. J. (2010). Inhibition of X4-tropic HIV type 1 replication by knockdown of the cellular protein LEREPO4. *AIDS research and human retroviruses*, *26*(10), 1155-1161. https://doi.org/10.1089/aid.2010.0041

18) Causey, D. R. (2018). Proteomic and molecular investigations of links between growth and immune function in salmonids (Doctoral dissertation, University of Aberdeen).

19) Cazorla-Vázquez, S., & Engel, F. B. (2018). Adhesion GPCRs in kidney development and disease. *Frontiers in Cell and Developmental Biology*, *6*, 9. https://doi.org/10.3389/fcell.2018.00009

20) Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*, *4*(1), s13742-015. https://doi.org/10.1186/s13742-015-0047-8

21) Charlesworth, D., Charlesworth, B., & Morgan, M. T. (1995). The pattern of neutral molecular variation under the background selection model. *Genetics*, *141*(4), 1619-1632. https://doi.org/10.1093/genetics/141.4.1619

22) Charlesworth, D. (2006). Balancing selection and its effects on sequences in nearby genome regions. *PLoS genetics*, *2*(4), e64. https://doi.org/10.1371/journal.pgen.0020064

23) Chen, C. C., Feng, W., Lim, P. X., Kass, E. M., & Jasin, M. (2018). Homology-directed repair and the role of BRCA1, BRCA2, and related proteins in genome integrity and cancer. *Annual review of cancer biology*, *2*, 313-336. https://doi.org/10.1146/annurev-cancerbio-030617-050502

24) Christie, M. R., Marine, M. L., Fox, S. E., French, R. A., & Blouin, M. S. (2016). A single generation of domestication heritably alters the expression of hundreds of genes. *Nature communications*, *7*(1), 1-6. https://doi.org/10.1038/ncomms10676

25) Cumer, T., Boyer, F., & Pompanon, F. (2021). Genome-wide detection of structural variations reveals new regions associated with domestication in small ruminants. *Genome biology and evolution*, *13*(8), evab165. https://doi.org/10.1093/gbe/evab165

26) Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156-2158. https://doi.org/10.1093/bioinformatics/btr330

27) Darwin, C. (1859). *The Origin of Species by Means of Natural Selection*, Murray, London

28) Darwin, C. (1868). The variation of animals and plants under domestication. Volume 2 (London: J. Murray).

29) D'Ambrosio, J., Phocas, F., Haffray, P., Bestin, A., Brard-Fudulea, S., Poncet, C., ... & Dupont-Nivet, M. (2019). Genome-wide estimates of genetic diversity, inbreeding and effective size of experimental and commercial rainbow trout lines undergoing selective breeding. *Genetics Selection Evolution*, *51*(1), 1-15. https://doi.org/10.1186/s12711-019-0468-4

30) D'Ambrosio, J., Morvezen, R., Brard-Fudulea, S., Bestin, A., Acin Perez, A., Guéméné, D., ... & Phocas, F. (2020). Genetic architecture and genomic selection of female reproduction traits in rainbow trout. *BMC genomics*, *21*(1), 1-14. https://doi.org/10.1186/s12864-020-06955-7

31) DeCarvalho, A. C., Cappendijk, S. L., & Fadool, J. M. (2004). Developmental expression of the POU domain transcription factor Brn-3b (Pou4f2) in the lateral line and visual system of zebrafish. *Developmental dynamics: an official publication of the American Association of Anatomists*, *229*(4), 869-876. https://doi.org/10.1002/dvdy.10475

32) Deng, J., Wang, Y., Hu, M., Lin, J., Li, Q., Liu, C., & Xu, X. (2022). Deleterious Variation in BR Serine/Threonine Kinase 2 Classified a Subtype of Autism. *Frontiers in Molecular Neuroscience*, *15*. https://doi.org/1 0.3389/fnmol.2022.904935

33) DiBenedetto, A. J., Guinto, J. B., Ebert, T. D., Bee, K. J., Schmidt, M. M., & Jackman, T. R. (2008). Zebrafish brd2a and brd2bare paralogous members of the bromodomain-ET (BET) family of transcriptional coregulators that show structural and expression divergence. *BMC developmental biology*, *8*(1), 1-19. https://doi.org/10.1186/1471-213X-8-39

34) Ding, Y., Wang, W., Ma, D., Liang, G., Kang, Z., Xue, Y., ... & Liu, F. (2021). Smarca5-mediated epigenetic programming facilitates fetal HSPC development in vertebrates. *Blood*, *137*(2), 190-202. https://doi.org/10.1182/blood.2020005219

35) Dobney, K., & Larson, G. (2006). Genetics and animal domestication: new windows on an elusive process. *Journal of zoology*, *269*(2), 261-271. https://doi.org/10.1111/j.1469-7998.2006.00042.x

36) Durland, E., De Wit, P., & Langdon, C. (2021). Temporally balanced selection during development of larval Pacific oysters (Crassostrea gigas) inherently preserves genetic diversity within offspring. *Proceedings of the Royal Society B*, *288*(1958), 20203223. https://doi.org/10.1098/rspb.2020.3223

37) East, E. M. (1918). The role of reproduction in evolution. *The American Naturalist*, *52*(618/619), 273-289. https://www.journals.uchicago.edu/doi/pdf/10.1086/279670

38) Fabrice, T. (2018). Fish domestication: an overview. *Animal domestication*. https://doi.org/10.5772/intechopen.79628

39) Ferrer-Admetlla, A., Bosch, E., Sikora, M., Marquès-Bonet, T., Ramírez-Soriano, A., Muntasell, A., ... & Casals, F. (2008). Balancing selection is the main force shaping the evolution of innate immunity genes. *The Journal of Immunology*, *181*(2), 1315-1322. https://doi.org/10.4049/jimmunol.181.2.1315

40) Fijarczyk, A., & Babik, W. (2015). Detecting balancing selection in genomes: limits and prospects. *Molecular ecology*, *24*(14), 3529-3545. https://doi.org/10.1111/mec.13226

41) Fisher, RA. (1958). The genetical theory of natual selection.

42) Fromer, M., Roussos, P., Sieberts, S. K., Johnson, J. S., Kavanagh, D. H., Perumal, T. M., ... & Sklar, P. (2016). Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nature neuroscience*, *19*(11), 1442-1453. https://doi.org/10.1038/nn.4399

43) Gall, G.A.E., Crandell, P.A. (1992). The rainbow trout. *Aquaculture*, *100*(1-3), 1-10. https://doi.org/10.1016/0044-8486(92)90333-G

44) Gao, G., Nome, T., Pearse, D. E., Moen, T., Naish, K. A., Thorgaard, G. H., ... & Palti, Y. (2018). A new single nucleotide polymorphism database for rainbow trout generated through whole genome resequencing. *Frontiers in genetics*, *9*, 147. https://doi.org/10.3389/fgene.2018.00147

45) Gao, G., Magadan, S., Waldbieser, G. C., Youngblood, R. C., Wheeler, P. A., Scheffler, B. E., ... & Palti, Y. (2021). A long reads-based de-novo assembly of the genome of the Arlee homozygous line reveals chromosomal rearrangements in rainbow trout. *G3*, *11*(4), jkab052. https://doi.org/10.1093/g3journal/jkab052

46) Gautier, M., & Vitalis, R. (2012). rehh: an R package to detect footprints of selection in genome-wide SNP data from haplotype structure. *Bioinformatics*, *28*(8), 1176-1177. https://doi.org/10.1093/bioinformatics/bts115

47) Gautier, M., Klassmann, A., & Vitalis, R. (2017). rehh 2.0: a reimplementation of the R package rehh to detect positive selection from haplotype structure. *Molecular ecology resources*, *17*(1), 78-90. https://doi.org/10.1111/1755-0998.12634

48) Gautier, P., Naranjo-Golborne, C., Taylor, M. S., Jackson, I. J., & Smyth, I. (2008). Expression of the fras1/frem gene family during zebrafish development and fin morphogenesis. *Developmental dynamics: an official publication of the American Association of Anatomists*, *237*(11), 3295-3304. https://doi.org/10.1002/dvdy.21729

924 49) Georges, M., Charlier, C., & Hayes, B. (2019). Harnessing genomic information for livestock
925 improvement. *Nature Reviews Genetics*, *20*(3), 135-156.
926
927 50) Glover, K. A., Otterå, H., Olsen, R. E., Slinde, E., Taranger, G. L., & Skaala, Ø. (2009). A
928 comparison of farmed, wild and hybrid Atlantic salmon (Salmo salar L.) reared under farming
929 conditions. *Aquaculture*, *286*(3-4), 203-210. https://doi.org/10.1016/j.aquaculture.2008.09.023
930
931 51) Gorssen, W., Meyermans, R., Janssens, S., & Buys, N. (2021). A publicly available repository of
932 ROH islands reveals signatures of selection in different livestock and pet species. *Genetics Selection*
933 *Evolution*, *53*(1), 1-10. https://doi.org/10.1186/s12711-020-00599-7
934
935 52) Grilz-Seger, G., Mesarič, M., Cotman, M., Neuditschko, M., Druml, T., & Brem, G. (2018). Runs
936 of homozygosity and population history of three horse breeds with small population size. *Journal*
937 *of Equine Veterinary Science*, *71*, 27-34. https://doi.org/10.1016/j.jevs.2018.09.004
938
939 53) Gross, M. R. (1998). One species with two biologies: Atlantic salmon (Salmo salar) in the wild and
940 in aquaculture. *Canadian Journal of Fisheries and Aquatic Sciences*, *55*(S1), 131-144.
941
942 54) Gross, R., Lulla, P., & Paaver, T. (2007). Genetic variability and differentiation of rainbow trout
943 (Oncorhynchus mykiss) strains in northern and Eastern Europe. *Aquaculture*, *272*, S139-S146.
944 https://doi.org/10.1016/j.aquaculture.2007.08.004
945
946 55) Gu, L., Wu, Y., Hu, S., Chen, Q., Tan, J., Yan, Y., ... & Tang, N. (2016). Analysis of association
947 between MAP2K4 gene polymorphism rs3826392 and IL-1b serum level in southern Chinese Han
948 Ischemic Stroke Patients. *Journal of Stroke and Cerebrovascular Diseases*, *25*(5), 1096-1101.
949 https://doi.org/10.1016/j.jstrokecerebrovasdis.2015.12.035
950 56) Gutierrez, A. P., Yáñez, J. M., & Davidson, W. S. (2016). Evidence of recent signatures of selection
951 during domestication in an Atlantic salmon population. *Marine genomics*, *26*, 41-50.
952 https://doi.org/10.1016/j.margen.2015.12.007
953
954 57) Hansen, M. M., & Mensberg, K. L. D. (1998). Genetic differentiation and relationship between
955 genetic and geographical distance in Danish sea trout (Salmo trutta L.) populations. *Heredity*, *81*(5),
956 493-504. https://doi.org/10.1046/j.1365-2540.1998.00408.x
957
958 58) Harris, D. R. (2007). Agriculture, cultivation and domestication: exploring the conceptual
959 framework of early food production. *Rethinking agriculture: archaeological and*
960 *ethnoarchaeological perspectives*, *16*, 35.
961
962 59) Hedrick, P. W. (2015). Heterozygote advantage: the effect of artificial selection in livestock and
963 pets. *Journal of Heredity*, *106*(2), 141-154.
964
965 60) Helmer, D. (1992). *La domestication des animaux par les hommes préhistoriques*. FeniXX.
966 61) Hershberger, W. K. (1992). Genetic variability in rainbow trout populations. *Aquaculture*, *100*(1-
967 3), 51-71. https://doi.org/10.1016/0044-8486(92)90339-M
968
969 62) Hiatt, S. M., Thompson, M. L., Prokop, J. W., Lawlor, J. M., Gray, D. E., Bebin, E. M., ... & Cooper,
970 G. M. (2019). Deleterious variation in brsk2 associates with a neurodevelopmental disorder. *The*
971 *American Journal of Human Genetics*, *104*(4), 701-708. https://doi.org/10.1016/j.ajhg.2019.02.002

63) Hong Xia, J., Bai, Z., Meng, Z., Zhang, Y., Wang, L., Liu, F., ... & Hua Yue, G. (2015). Signatures of selection in tilapia revealed by whole genome resequencing. *Scientific reports*, *5*(1), 1-10. https://doi.org/10.1038/srep14168

64) Hori, K., Shimaoka, K., & Hoshino, M. (2021). AUTS2 gene: keys to understanding the pathogenesis of neurodevelopmental disorders. *Cells*, *11*(1), 11. https://doi.org/10.3390/cells11010011

65) Howarth, D. L., Yin, C., Yeh, K., & Sadler, K. C. (2013). Defining hepatic dysfunction parameters in two models of fatty liver disease in zebrafish larvae. *Zebrafish*, *10*(2), 199-210. https://doi.org/10.1089/zeb.2012.0821

66) Illa, S. K., Mumtaz, S., Mukherjee, S., Nath, S., & Mukherjee, A. (2022). Genome-wide evaluation of runs of homozygosity in Indian Sahiwal cattle. *The 12th World Congress on Genetic Applied to Livestock*.

67) Ioerger, T. R., Clark, A. G., & Kao, T. H. (1990). Polymorphism at the self-incompatibility locus in Solanaceae predates speciation. *Proceedings of the National Academy of Sciences*, *87*(24), 9732-9735. https://doi.org/10.1073/pnas.87.24.9732

68) Ip, J. P., Shi, L., Chen, Y., Itoh, Y., Fu, W. Y., Betz, A., ... & Ip, N. Y. (2012). α2-chimaerin controls neuronal migration and functioning of the cerebral cortex through CRMP-2. *Nature neuroscience*, *15*(1), 39-47. https://doi.org/10.1038/nn.2972

69) Ishikawa, K., Azuma, S., Ikawa, S., Semba, K., & Inoue, J. I. (2005). Identification of DRG family regulatory proteins (DFRPs): specific regulation of DRG1 and DRG2. *Genes to Cells*, *10*(2), 139-150.

70) Jombart, T., & Ahmed, I. (2011). Adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*(21), 3070-3071.

71) Jombart, T., & Ahmed, I. (2011). Adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*(21), 3070-3071. https://doi.org/10.1093/bioinformatics/btr521

72) Khendek, A., Alix, M., Viot, S., Ledoré, Y., Rousseau, C., Mandiki, R., ... & Milla, S. (2017). How does a domestication process modulate oogenesis and reproduction performance in Eurasian perch?. *Aquaculture*, *473*, 206-214. https://doi.org/10.1016/j.aquaculture.2017.02.003

73) Kishi, M., Pan, Y. A., Crump, J. G., and Sanes, J. R. (2005). Mammalian SAD kinases are required for neuronal polarization. *Science* 307, 929–932. https://doi.org/10.1126/science.1107403

74) Klein, J., Sato, A., Nagl, S., & O'hUigín, C. (1998). Molecular trans-species polymorphism. *Annual Review of Ecology and Systematics*, 1-C1. https://www.jstor.org/stable/221700

75) Leffler, E. M., Gao, Z., Pfeifer, S., Ségurel, L., Auton, A., Venn, O., ... & Przeworski, M. (2013). Multiple instances of ancient balancing selection shared between humans and chimpanzees. *Science*, *339*(6127), 1578-1582.

76) Lagarde, H., Lallias, D., Patrice, P., Prchal, M., Francois, Y., D'Ambrosio, J., ... & Phocas, F. (2022). Genetic architecture of acute hyperthermia resistance in juvenile rainbow trout (Oncorhynchus mykiss) and genetic correlations with production traits. *bioRxiv*.

77) Le Tanno, P., Breton, J., Bidart, M., Satre, V., Harbuz, R., Ray, P. F., ... & Coutton, C. (2017). PBX1 haploinsufficiency leads to syndromic congenital anomalies of the kidney and urinary tract (CAKUT) in humans. *Journal of Medical Genetics*, *54*(7), 502-510. http://dx.doi.org/10.1136/jmedgenet-2016-104435

78) Leskow, F. C., Holloway, B. A., Wang, H., Mullins, M. C., & Kazanietz, M. G. (2006). The zebrafish homologue of mammalian chimerin Rac-GAPs is implicated in epiboly progression during development. *Proceedings of the National Academy of Sciences*, *103*(14), 5373-5378. https://doi.org/10.1073/pnas.0508585103

79) Li, Q., Hou, J., Guo, C., Zhang, Y., Liu, L., Hu, H., ... & Cui, H. (2021a). ZC3H15 Correlates with a Poor Prognosis and Tumor Progression in Melanoma. *BioMed Research International*, *2021*. https://doi.org/10.1155/2021/8305299

80) Li, H., Zhu, Q., Chen, R., Liu, M., & Xu, D. (2021b). Identification and Characterization of Dimorphic Expression of Sex-Related Genes in Rock Bream, a Fish With Multiple Sex Chromosomes. *Frontiers in genetics*, *12*. https://doi.org/10.3389/fgene.2021.791179

81) Li, Z., Tian, Y., Wang, L., Li, Z., Chen, S., Li, L., ... & Zhai, J. (2022). Comparative transcriptomics analyses and revealing candidate networks and genes involved in lordosis of the Yunlong grouper (Epinephelus moara♀× Epinephelus lanceolatus♂). *Aquaculture*, *550*, 737866. https://doi.org/10.1016/j.aquaculture.2021.737866

82) Limi, S., Senecal, A., Coleman, R., Lopez-Jones, M., Guo, P., Polumbo, C., ... & Cvekl, A. (2018). Transcriptional burst fraction and size dynamics during lens fiber cell differentiation and detailed insights into the denucleation process. *Journal of Biological Chemistry*, *293*(34), 13176-13190. https://doi.org/10.1074/jbc.RA118.001927

83) Lin, Y., Huang, M., Wang, S., You, X., Zhang, L., & Chen, Y. (2021). PAQR11 modulates monocyte-to-macrophage differentiation and pathogenesis of rheumatoid arthritis. *Immunology*, *163*(1), 60-73. https://doi.org/10.1111/imm.13303

84) Liu T., Elmquist J.K. & Williams K.W. (2013) Mrap2: an accessory protein linked to obesity. Cell Metabolism 18, 309–11. https://doi.org/10.1016/j.cmet.2013.08.016

85) Liu, S., Palti, Y., Martin, K. E., Parsons, J. E., & Rexroad III, C. E. (2017). Assessment of genetic differentiation and genetic assignment of commercial rainbow trout strains using a SNP panel. *Aquaculture*, *468*, 120-125. https://doi.org/10.1016/j.aquaculture.2016.10.004

86) Liu, S., Huang, J., Zhang, Y., Liu, Y., Zuo, S., & Li, R. (2019). MAP2K4 interacts with Vimentin to activate the PI3K/AKT pathway and promotes breast cancer pathogenesis. *Aging (Albany NY)*, *11*(22), 10697. https://doi.org/10.18632/aging.102485

87) Liu, C., & Zhao, B. (2021). Murine GRXCR1 has a different function than GRXCR2 in the morphogenesis of stereocilia. *Frontiers in Cellular Neuroscience*, 276. https://doi.org/10.3389/fncel.2021.714070

88) Liu, Y., Wang, Y., Li, N., and Jiang, S. (2022). Avobenzone and nanoplastics affect the development of zebrafish nervous system and retinal system and inhibit their locomotor behavior. *Sci. Total Environ.* 806:150681. https://doi.org/10.1016/j.scitotenv.2021.150681

89) López, M. E., Neira, R., & Yáñez, J. M. (2015). Applications in the search for genomic selection signatures in fish. *Frontiers in genetics*, *5*, 458. https://doi.org/10.3389/fgene.2014.00458

90) López, M. E., Benestan, L., Moore, J. S., Perrier, C., Gilbey, J., Di Genova, A., ... & Yáñez, J. M. (2018). Comparing genomic signatures of domestication in two Atlantic salmon (Salmo salar L.) populations with different geographical origins. *Evolutionary Applications*, *12*(1), 137-156. https://doi.org/10.1111/eva.12689

91) López, M. E., Linderoth, T., Norris, A., Lhorente, J. P., Neira, R., & Yáñez, J. M. (2019). Multiple selection signatures in farmed atlantic salmon adapted to different environments across hemispheres. *Frontiers in Genetics*, *10*, 901. https://doi.org/10.3389/fgene.2019.00901

92) López, M. E., Cádiz, M. I., Rondeau, E. B., Koop, B. F., & Yáñez, J. M. (2021). Detection of selection signatures in farmed coho salmon (Oncorhynchus kisutch) using dense genome-wide information. *Scientific reports*, *11*(1), 1-13. https://doi.org/10.1038/s41598-021-86154-w

93) Limborg, M. T., Blankenship, S. M., Young, S. F., Utter, F. M., Seeb, L. W., Hansen, M. H., & Seeb, J. E. (2012). Signatures of natural selection among lineages and habitats in Oncorhynchus mykiss. *Ecology and Evolution*, *2*(1), 1-18. https://doi.org/10.1002/ece3.59

94) Lorente-Cánovas, B., Eckrich, S., Lewis, M. A., Johnson, S. L., Marcotti, W., & Steel, K. P. (2022). Grxcr1 regulates hair bundle morphogenesis and is required for normal mechanoelectrical transduction in mouse cochlear hair cells. *Plos one*, *17*(3), e0261530. https://doi.org/10.1371/journal.pone.0261530

95) Magris, G., Marroni, F., D'Agaro, E., Vischi, M., Chiabà, C., Scaglione, D., ... & Morgante, M. (2022). ddRAD-seq reveals the genetic structure and detects signals of selection in Italian brown trout. *Genetics Selection Evolution*, *54*(1), 1-14. https://doi.org/10.1186/s12711-022-00698-7

96) Mäkinen, H., Vasemägi, A., McGinnity, P., Cross, T. F., & Primmer, C. R. (2015). Population genomic analyses of early-phase A tlantic S almon (S almo salar) domestication/captive breeding. *Evolutionary applications*, *8*(1), 93-107. https://doi.org/10.1111/eva.12230

97) Margarit, E., Armas, P., Siburu, N. G., & Calcaterra, N. B. (2014). CNBP modulates the transcription of Wnt signaling pathway components. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, *1839*(11), 1151-1160. https://doi.org/10.1016/j.bbagrm.2014.08.009

98) Martínez, A., Garza, J. C., & Pearse, D. E. (2011). A microsatellite genome screen identifies chromosomal regions under differential selection in steelhead and rainbow trout. *Transactions of the American Fisheries Society*, *140*(3), 829-842. https://doi.org/10.1080/00028487.2011.588094

99) Maskell, L. J., Qamar, K., Babakr, A. A., Hawkins, T. A., Heads, R. J., & Budhram-Mahadeo, V. S. (2017). Essential but partially redundant roles for POU4F1/Brn-3a and POU4F2/Brn-3b transcription factors in the developing heart. *Cell Death & Disease*, *8*(6), e2861-e2861. https://doi.org/10.1038/cddis.2017.185

100) Mastrangelo, S., Tolone, M., Sardina, M. T., Sottile, G., Sutera, A. M., Di Gerlando, R., & Portolano, B. (2017). Genome-wide scan for runs of homozygosity identifies potential candidate genes associated with local adaptation in Valle del Belice sheep. *Genetics Selection Evolution*, *49*(1), 1-10. https://doi.org/10.1186/s12711-017-0360-z

101) McQuillan, R., Leutenegger, A. L., Abdel-Rahman, R., Franklin, C. S., Pericic, M., Barac-Lauc, L., ... & Wilson, J. F. (2008). Runs of homozygosity in European populations. *The American Journal of Human Genetics*, *83*(3), 359-372. https://doi.org/10.1016/j.ajhg.2008.08.007

102) Mignon-Grasteau, S., Boissy, A., Bouix, J., Faure, J. M., Fisher, A. D., Hinch, G. N., ... & Beaumont, C. (2005). Genetics of adaptation and domestication in livestock. *Livestock Production Science*, *93*(1), 3-14. https://doi.org/10.1016/j.livprodsci.2004.11.001

103) Milla, S., Pasquet, A., El Mohajer, L., & Fontaine, P. (2021). How domestication alters fish phenotypes. *Reviews in Aquaculture*, *13*(1), 388-405. https://doi.org/10.1111/raq.12480

104) Milanetto, M., Tiso, N., Braghetta, P., Volpin, D., Argenton, F., & Bonaldo, P. (2008). Emilin genes are duplicated and dynamically expressed during zebrafish embryonic development. *Developmental dynamics: an official publication of the American Association of Anatomists*, *237*(1), 222-232. https://doi.org/10.1002/dvdy.21402

105) Miller, B. C., Zhao, Z., Stephenson, L. M., Cadwell, K., Pua, H. H., Lee, H. K., ... & Virgin, IV, H. W. (2008). The autophagy gene ATG5 plays an essential role in B lymphocyte development. *Autophagy*, *4*(3), 309-314. https://doi.org/10.4161/auto.5474

106) Miller, T. E., Henkels, K. M., Huddleston, M., Salisbury, R., Hussain, S. M., Sasaki, A. T., & Cho, K. J. (2019). Depletion of phosphatidylinositol 4-phosphate at the Golgi translocates K-Ras to mitochondria. *Journal of cell science*, *132*(16), jcs231886. https://doi.org/10.1242/jcs.231886

107) Miyake, N., Chilton, J., Psatha, M., Cheng, L., Andrews, C., Chan, W. M., ... & Engle, E. C. (2008). Human CHN1 mutations hyperactivate α2-chimaerin and cause Duane's retraction syndrome. *Science*, *321*(5890), 839-843.

108) Moynahan, M. E., Pierce, A. J., & Jasin, M. (2001). BRCA2 is required for homology-directed repair of chromosomal breaks. *Molecular cell*, *7*(2), 263-272. https://doi.org/10.1016/S1097-2765(01)00174-5

109) Mukudai, Y., Kondo, S., Fujita, A., Yoshihama, Y., Shirota, T., & Shintani, S. (2013). Tumor protein D54 is a negative regulator of extracellular matrix-dependent migration and attachment in oral squamous cell carcinoma-derived cell lines. *Cellular Oncology*, *36*(3), 233-245. https://doi.org/10.1007/s13402-013-0131-y

110) Muragaki, Y., Mariman, E., van Beersum, S. E., Perälä, M., van Mourik, J., Warman, M. L., ... & Hamel, B. C. (1996). A mutation in the gene encoding the α2 chain of the fibril-associated collagen IX, COL9A2, causes multiple epiphyseal dysplasia (EDM2). *Nature genetics*, *12*(1), 103-105. https://doi.org/10.1038/ng0196-103

111) Nakanishi, K., Niida, H., Tabata, H., Ito, T., Hori, Y., Hattori, M., et al. (2019). Isozyme-Specific Role of SAD-A in Neuronal Migration During Development of Cerebral Cortex. *Cereb. Cortex* 29, 3738–3751. https://doi.org/10.1093/cercor/bhy253

112) Naval-Sanchez, M., McWilliam, S., Evans, B., Yáñez, J. M., Houston, R. D., & Kijas, J. W. (2020). Changed patterns of genomic variation following recent domestication: selection sweeps in farmed Atlantic salmon. *Frontiers in genetics*, *11*, 264. https://doi.org/10.3389/fgene.2020.00264

113) Nosková, L., Stránecký, V., Hartmannová, H., Přistoupilová, A., Barešová, V., Ivánek, R., ... & Kmoch, S. (2011). Mutations in DNAJC5, encoding cysteine-string protein alpha, cause autosomal-dominant adult-onset neuronal ceroid lipofuscinosis. *The American Journal of Human Genetics*, *89*(2), 241-252. https://doi.org/10.1016/j.ajhg.2011.07.003

114) Nothnagel, M., Lu, T. T., Kayser, M., & Krawczak, M. (2010). Genomic and geographic distribution of SNP-defined runs of homozygosity in Europeans. *Human molecular genetics*, *19*(15), 2927-2935. https://doi.org/10.1093/hmg/ddq198

115) Oksenberg, N., Stevison, L., Wall, J. D., & Ahituv, N. (2013). Function and regulation of AUTS2, a gene implicated in autism and human evolution. *PLoS genetics*, *9*(1), e1003221. https://doi.org/10.1371/journal.pgen.1003221

116) Palińska-Żarska, K., Król, J., Woźny, M., Kamaszewski, M., Szudrowicz, H., Wiechetek, W., ... & Żarski, D. (2021). Domestication affected stress and immune response markers in Perca fluviatilis in the early larval stage. *Fish & Shellfish Immunology*, *114*, 184-198. https://doi.org/10.1016/j.fsi.2021.04.028

117) Palmer, C. S., Osellame, L. D., Stojanovski, D., & Ryan, M. T. (2011). The regulation of mitochondrial morphology: intricate mechanisms and dynamic machinery. *Cellular signalling*, *23*(10), 1534-1545. https://doi.org/10.1016/j.cellsig.2011.05.021

118) Palti, Y., Gao, G., Miller, M. R., Vallejo, R. L., Wheeler, P. A., Quillet, E., ... & Rexroad III, C. E. (2014). A resource of single-nucleotide polymorphisms for rainbow trout generated by restriction-site associated DNA sequencing of doubled haploids. *Molecular ecology resources*, *14*(3), 588-596. https://doi.org/10.1111/1755-0998.12204

119) Palti, Y., Gao, G., Liu, S., Kent, M. P., Lien, S., Miller, M. R., ... & Moen, T. (2015). The development and characterization of a 57 K single nucleotide polymorphism array for rainbow trout. *Molecular ecology resources*, *15*(3), 662-672. https://doi.org/10.1111/1755-0998.12337

120) Pasquet, A. (2018). Effects of domestication on fish behaviour. In *Animal domestication*. IntechOpen. https://doi.org/ 10.5772/intechopen.78752

121) Paul, K., Pelissier P., Goardon L., Dechamp N., Danon J., Jaffrelo L., Poncet C., Dupont-Nivet M., Phocas, F. (*In review*) Maternal and genetic effects on embryonic survival from fertilization to swim up stage and reproductive success in a farmed rainbow trout line. *Aquaculture Reports*.

122) Pennings, P. S., & Hermisson, J. (2006). Soft sweeps II—molecular population genetics of adaptation from recurrent mutation or migration. *Molecular biology and evolution*, *23*(5), 1076-1084. https://doi.org/10.1093/molbev/msj117

123) Peripolli, E., Stafuzza, N. B., Munari, D. P., Lima, A. L. F., Irgang, R., Machado, M. A., ... & da Silva, M. V. G. B. (2018). Assessment of runs of homozygosity islands and estimates of genomic inbreeding in Gyr (Bos indicus) dairy cattle. *BMC genomics*, *19*(1), 1-13. https://doi.org/10.1186/s12864-017-4365-3

124) Pulcini, D., Wheeler, P. A., Cataudella, S., Russo, T., & Thorgaard, G. H. (2013). Domestication shapes morphology in rainbow trout Oncorhynchus mykiss. *Journal of Fish Biology*, *82*(2), 390-407. https://doi.org/10.1111/jfb.12002

125) Prchal, M., D'Ambrosio, J., Lagarde, H., Lallias, D., Patrice, P., François, Y., … & Phocas, F. (*In review*). Genome-wide association study and genomic prediction of tolerance to acute hypoxia in rainbow trout. Aquaculture

126) Price, E. O. (1984). Behavioral aspects of animal domestication. *The quarterly review of biology*, *59*(1), 1-32.

127) Price, E. O. (1999). Behavioral development in animals undergoing domestication. *Applied Animal Behaviour Science*, *65*(3), 245-271. https://doi.org/10.1016/S0168-1591(99)00087-8

128) Price, E. O. (2002). *Animal domestication and behavior*. Cabi. http://sherekashmir.informaticspublishing.com/479/1/9780851995977.pdf

129) Pritchard, V. L., Mäkinen, H., Vähä, J. P., Erkinaro, J., Orell, P., & Primmer, C. R. (2018). Genomic signatures of fine-scale local selection in Atlantic salmon suggest involvement of sexual maturation, energy homeostasis and immune defence-related genes. *Molecular Ecology*, *27*(11), 2560-2575. https://doi.org/10.1111/mec.14705

130) Purcell, R. (2017). Activation and regulation of the brain-expressed adhesion G protein-coupled receptors *ADGRB1/BAI1* and *ADGRB2/BAI2*: Implications for human disease (Doctoral dissertation, Emory University). https://etd.library.emory.edu/concern/etds/mc87pq78v?locale=fr

131) Purfield, D. C., Berry, D. P., McParland, S., & Bradley, D. G. (2012). Runs of homozygosity and population history in cattle. *BMC genetics*, *13*(1), 1-11. https://doi.org/10.1186/1471-2156-13-70

132) Purfield, D. C., McParland, S., Wall, E., & Berry, D. P. (2017). The distribution of runs of homozygosity and selection signatures in six commercial meat sheep breeds. *PLoS One*, *12*(5), e0176780. https://doi.org/10.1371/journal.pone.0176780

133) Qanbari, S. & Simianer, H. (2014). Mapping signatures of positive selection in the genome of livestock. *Livest. Sci.* **166**, 133–143. https://doi.org/10.1016/j.livsci.2014.05.003

134) Raghuram, V., Weber, S., Raber, J., Chen, D. H., Bird, T. D., Maylie, J., & Adelman, J. P. (2017). Assessment of mutations in KCNN2 and ZNF135 to patient neurological symptoms. *Neuroreport*, *28*(7), 375. https://doi.org/10.1097/WNR.0000000000000754

135) Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., & Vilo, J. (2019). g: Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic acids research*, *47*(W1), W191-W198. https://doi.org/10.1093/nar/gkz369

136) Repo, H., Gurvits, N., Löyttyniemi, E., Nykänen, M., Lintunen, M., Karra, H., ... & Kronqvist, P. (2017). PTTG1-interacting protein (PTTG1IP/PBF) predicts breast cancer survival. *BMC cancer*, *17*(1), 1-8. https://doi.org/10.1186/s12885-017-3694-6

137) Rhee, K., Brunori, M., Besset, V., Trousdale, R., & Wolgemuth, D. J. (1998). Expression and potential role of Fsrg1, a murine bromodomain-containing homologue of the Drosophila gene female sterile homeotic. *Journal of Cell Science*, *111*(23), 3541-3550. https://doi.org/10.1242/jcs.111.23.3541

138) Rodríguez-Marí, A., Wilson, C., Titus, T. A., Canestro, C., BreMiller, R. A., Yan, Y. L., ... & Postlethwait, J. H. (2011). Roles of brca2 (fancd1) in oocyte nuclear architecture, gametogenesis, gonad tumors, and genome stability in zebrafish. *PLoS genetics*, *7*(3), e1001357. https://doi.org/10.1371/journal.pgen.1001357

139) Russell, N. (2002). The wild side of animal domestication. *Society & Animals*, *10*(3), 285-302. https://doi.org/10.1163/156853002320770083

140) Sabeti, P. C., Reich, D. E., Higgins, J. M., Levine, H. Z., Richter, D. J., Schaffner, S. F., ... & Lander, E. S. (2002). Detecting recent positive selection in the human genome from haplotype structure. *Nature*, *419*(6909), 832-837. https://doi.org/10.1038/nature01140

141) Sabeti, P. C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., ... & Lander, E. S. (2007). Genome-wide detection and characterization of positive selection in human populations. *Nature*, *449*(7164), 913-918. https://doi.org/10.1038/nature06250

142) Saravanan, K. A., Panigrahi, M., Kumar, H., Bhushan, B., Dutt, T., & Mishra, B. P. (2020). Selection signatures in livestock genome: A review of concepts, approaches and applications. *Livestock Science*, *241*, 104257. https://doi.org/10.1016/j.livsci.2020.104257

143) Saravanan, K. A., Panigrahi, M., Kumar, H., Parida, S., Bhushan, B., Gaur, G. K., ... & Singh, R. K. (2021). Genomic scans for selection signatures revealed candidate genes for adaptation and production traits in a variety of cattle breeds. *Genomics*, *113*(3), 955-963. https://doi.org/10.1016/j.ygeno.2021.02.009

144) Sargolzaei, M., Chesnais, J. P., & Schenkel, F. S. (2014). A new approach for efficient genotype imputation using information from relatives. *BMC genomics*, *15*(1), 1-12. https://doi.org/10.1186/1471-2164-15-478

145) Selleri L, Depew MJ, Jacobs Y, Chanda SK, Tsang KY, Cheah KSE, Rubenstein JLR, O'Gorman S, Cleary ML. (2001). Requirement for Pbx1 in skeletal patterning and programming chondrocyte proliferation and differentiation. Development 128: 3543– 3557. https://doi.org/10.1242/dev.128.18.3543

146) Seo, J., Asaoka, Y., Nagai, Y., Hirayama, J., Yamasaki, T., Namae, M., ... & Nishina, H. (2010). Negative regulation of wnt11 expression by Jnk signaling during zebrafish gastrulation. *Journal of cellular biochemistry*, *110*(4), 1022-1037. https://doi.org/10.1002/jcb.22616

147) Shiu, F. H., Wong, J. C., Yamamoto, T., Lala, T., Purcell, R. H., Owino, S., ... & Escayg, A. (2022). Mice lacking full length Adgrb1 (Bai1) exhibit social deficits, increased seizure susceptibility, and altered brain development. *Experimental Neurology*, *351*, 113994. https://doi.org/10.1016/j.expneurol.2022.113994

148) Shive, H. R., West, R. R., Embree, L. J., Azuma, M., Sood, R., Liu, P., & Hickstein, D. D. (2010). brca2 in zebrafish ovarian development, spermatogenesis, and tumorigenesis. *Proceedings of the National Academy of Sciences*, *107*(45), 19350-19355. https://doi.org/10.1073/pnas.1011630107

149) Signer-Hasler, H., Henkel, J., Bangerter, E., Bulut, Z., Drögemüller, C., Leeb, T., & Flury, C. (2022). Runs of homozygosity in Swiss goats reveal genetic changes associated with domestication and modern selection. *Genetics Selection Evolution*, *54*(1), 1-11. https://doi.org/10.1186/s12711-022-00695-w

150) Silverstein, J. T., Rexroad III, C. E., & King, T. L. (2004). Genetic variation measured by microsatellites among three strains of domesticated rainbow trout (Oncorhynchus mykiss, Walbaum). *Aquaculture Research*, *35*(1), 40-48. https://doi.org/10.1111/j.1365-2109.2004.00979.x

151) Smith, J. M., & Haigh, J. (1974). The hitch-hiking effect of a favourable gene. *Genetics Research*, *23*(1), 23-35. https://doi.org/10.1017/S0016672300014634

152) Stanković, D., Crivelli, A. J., & Snoj, A. (2015). Rainbow trout in Europe: introduction, naturalization, and impacts. *Reviews in Fisheries Science & Aquaculture*, *23*(1), 39-71. https://doi.org/10.1080/23308249.2015.1024825

153) Su, S., Li, H., Du, F., Zhang, C., Li, X., Jing, X., ... & Bouzoualegh, R. (2018). Combined QTL and genome scan analyses with the help of 2b-RAD identify growth-associated genetic markers in a new fast-growing carp strain. *Frontiers in genetics*, *9*, 592. https://doi.org/10.3389/fgene.2018.00592

154) Sullivan, C., Lage, C. R., Yoder, J. A., Postlethwait, J. H., & Kim, C. H. (2017). Evolutionary divergence of the vertebrate TNFAIP8 gene family: Applying the spotted gar orthology bridge to understand ohnolog loss in teleosts. *Plos one*, *12*(6), e0179517. https://doi.org/10.1371/journal.pone.0179517

155) Sun, L., Liu, S., Wang, R., Jiang, Y., Zhang, Y., Zhang, J., ... & Liu, Z. (2014). Identification and analysis of genome-wide SNPs provide insight into signatures of selection and domestication in channel catfish (Ictalurus punctatus). *PloS one*, *9*(10), e109666. https://doi.org/10.1371/journal.pone.0109666

156) Teoh, P. H., Shu-Chien, A. C., & Chan, W. K. (2010). Pbx1 is essential for growth of zebrafish swim bladder. *Developmental dynamics: an official publication of the American Association of Anatomists*, *239*(3), 865-874. https://doi.org/10.1002/dvdy.22221

157) Těšický, M., & Vinkler, M. (2015). Trans-species polymorphism in immune genes: general pattern or MHC-restricted phenomenon? *Journal of immunology research*, *2015*. https://doi.org/10.1155/2015/838035

158) Torky, H. A., Sherif, A., Abo-Louz, A., Ali, M., Ahmed, A., & Ali, A. (2018). Evaluation of serum nidogen-2 as a screening and diagnostic tool for ovarian cancer. *Gynecologic and Obstetric Investigation*, *83*(5), 461-465. https://doi.org/10.1159/000481798

159) Tymchuk, W. E., Beckman, B., & Devlin, R. H. (2009). Altered expression of growth hormone/insulin-like growth factor I axis hormones in domesticated fish. *Endocrinology*, *150*(4), 1809-1816. https://doi.org/10.1210/en.2008-0797

160) Umasuthan, N., Revathy, K. S., Whang, I., Kim, E., Oh, M. J., Jung, S. J., ... & Lee, J. (2014). Genomic identification and molecular characterization of a non-mammalian TNFAIP8L2 gene from Oplegnathus fasciatus. *Gene*, *542*(1), 52-63. https://doi.org/10.1016/j.gene.2014.02.047

161) Uvarov, P., Kajander, T., & Airaksinen, M. S. (2014). Origin and loss of nested LRRTM/α-catenin genes during vertebrate evolution. *PloS one*, *9*(2), e89910. https://doi.org/10.1371/journal.pone.0089910

162) Vitti, J. J., Grossman, S. R., & Sabeti, P. C. (2013). Detecting natural selection in genomic data. *Annual review of genetics*. https://doi.org/10.1146/annurev-genet-111212-133526

163) Voight, B. F., Kudaravalli, S., Wen, X., & Pritchard, J. K. (2006). A map of recent positive selection in the human genome. *PLoS biology*, *4*(3), e72. https://doi.org/10.1371/journal.pbio.0040072

164) Wang, W. J., Tay, H. G., Soni, R., Perumal, G. S., Goll, M. G., Macaluso, F. P., ... & Bryan Tsou, M. F. (2013). CEP162 is an axoneme-recognition protein promoting ciliary transition zone assembly at the cilia base. *Nature cell biology*, *15*(6), 591-601. https://doi.org/10.1038/ncb2739

165) Wang, G. D., Xie, H. B., Peng, M. S., Irwin, D., & Zhang, Y. P. (2014). Domestication genomics: evidence from animals. *Annu. Rev. Anim. Biosci.*, *2*(1), 65-84. https://doi.org/10.1146/annurev-animal-022513-114129

166) Wang, Y., Miao, X., Li, H., Su, P., Lin, L., Liu, L., & Li, X. (2020). The correlated expression of immune and energy metabolism related genes in the response to Salmonella enterica serovar Enteritidis inoculation in chicken. *BMC veterinary research*, *16*(1), 1-9. https://doi.org/10.1186/s12917-020-02474-5

167) Wang, H., Teng, M., Liu, P., Zhao, M., Wang, S., Hu, J., ... & Zeng, Q. (2022). Selection Signatures of Pacific White Shrimp Litopenaeus vannamei Revealed by Whole-Genome Resequencing Analysis. *Advances in the Biology, Aquaculture, and Conservation of Threatened Marine Species and their Application in Human Health and Nutrition*. https://doi.org/10.3389/fmars.2022.844597

168) Wassarman, P. M., & Litscher, E. S. (2018). The mouse egg's zona pellucida. *Current topics in developmental biology*, *130*, 331-356. https://doi.org/10.1016/bs.ctdb.2018.01.003

169) Wassarman, P. M., & Litscher, E. S. (2021). Zona pellucida genes and proteins: Essential players in mammalian oogenesis and fertility. *Genes*, *12*(8), 1266. https://doi.org/10.3390/genes12081266

170) Wayne, R. K., & Vonholdt, B. M. (2012). Evolutionary genomics of dog domestication. *Mammalian Genome*, *23*(1), 3-18. https://doi.org/10.1007/s00335-011-9386-7

171) Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 1358-1370.

172) Weinstein, S. Y., Thrower, F. P., Nichols, K. M., & Hale, M. C. (2019). A large-scale chromosomal inversion is not associated with life history development in rainbow trout from Southeast Alaska. *PloS one*, *14*(9), e0223018. https://doi.org/10.1371/journal.pone.0223018

173) Wellenreuther, M., Le Luyer, J., Cook, D., Ritchie, P. A., & Bernatchez, L. (2019). Domestication and temperature modulate gene expression signatures and growth in the Australasian snapper Chrysophrys auratus. *G3: Genes, Genomes, Genetics*, *9*(1), 105-116. https://doi.org/10.1534/g3.118.200647

174) Wen, Z. Y., Liu, T., Qin, C. J., Zou, Y. C., Wang, J., Li, R., & Tao, Y. X. (2021). MRAP2 interaction with melanocortin-4 receptor in snakehead (Channa argus). *Biomolecules*, *11*(3), 481. https://doi.org/10.3390/biom11030481

175) Williams, J. L., Hall, S. J., Del Corvo, M., Ballingall, K. T., Colli, L. I. C. I. A., Ajmone Marsan, P. A. O. L. O., & Biscarini, F. (2016). Inbreeding and purging at the genomic Level: the Chillingham cattle reveal extensive, non-random SNP heterozygosity. *Animal genetics*, *47*(1), 19-27. https://doi.org/10.1111/age.12376

176) Wright, D., Nakamichi, R., Krause, J., & Butlin, R. K. (2006). QTL analysis of behavioral and morphological differentiation between wild and laboratory zebrafish (Danio rerio). *Behavior genetics*, *36*(2), 271-284.

177) Wright, D. (2015). Article commentary: the genetic architecture of domestication in animals. *Bioinformatics and biology insights*, *9*, BBI-S28902. https://doi.org/10.4137/BBI.S28902

178) Wright, S. 1931. Evolution in Mendelian populations. *Genetics*, *16*(2), 97.

179) Xu, H., Dong, R., Zeng, Q., Fang, L., Ge, Q., Xia, C., ... & Tong, P. (2022). Col9a2 gene deletion accelerates the degeneration of intervertebral discs. *Experimental and Therapeutic Medicine*, *23*(3), 1-13. https://doi.org/10.3892/etm.2022.11130

180) Yan, Q., Huang, C., Jiang, Y., Shan, H., Jiang, R., Wang, J., ... & Sun, H. (2018). Calpain7 impairs embryo implantation by downregulating β3-integrin expression via degradation of HOXA10. *Cell death & disease*, *9*(3), 1-12.

181) Yáñez, J. M., Xu, P., Carvalheiro, R., & Hayes, B. (2022). Genomics applied to livestock and aquaculture breeding. *Evolutionary Applications*, *15*(4), 517. https://doi.org/10.1111/eva.13378

182) Yao, X., Liu, X., Zhang, Y., Li, Y., Zhao, C., Yao, S., & Wei, Y. (2017). Gene therapy of adult neuronal ceroid lipofuscinoses with CRISPR/Cas9 in zebrafish. *Human gene therapy*, *28*(7), 588-597.

183) Ye, X., Zhou, X. J., & Zhang, H. (2018). Exploring the role of autophagy-related gene 5 (ATG5) yields important insights into autophagy in autoimmune/autoinflammatory diseases. *Frontiers in immunology*, *9*, 2334. https://doi.org/10.3389/fimmu.2018.02334

184) Yin, X., Huo, Z., Yan, S., Wang, Z., Yang, T., Wu, H., & Zhang, Z. (2021). MiR-205 inhibits sporadic vestibular schwannoma cell proliferation by targeting cyclin-dependent kinase 14. *World Neurosurgery*, *147*, e25-e31. https://doi.org/10.1016/j.wneu.2020.11.043

185) Yoshida, G. M., Lhorente, J. P., Carvalheiro, R., & Yáñez, J. M. (2017). Bayesian genome-wide association analysis for body weight in farmed Atlantic salmon (Salmo salar L.). *Animal Genetics*, *48*(6), 698-703. https://doi.org/10.1111/age.12621

186) Yu, X., Setyawan, P., Bastiaansen, J. W., Liu, L., Imron, I., Groenen, M. A., ... & Megens, H. J. (2022). Genomic analysis of a Nile tilapia strain selected for salinity tolerance shows signatures of selection and hybridization with blue tilapia (Oreochromis aureus). *Aquaculture*, *560*, 738527. https://doi.org/10.1016/j.aquaculture.2022.738527

187) Żarski, D., Le Cam, A., Nynca, J., Klopp, C., Ciesielski, S., Sarosiek, B., ... & Bobe, J. (2020). Domestication modulates the expression of genes involved in neurogenesis in high-quality eggs of Sander lucioperca. *Molecular Reproduction and Development*, *87*(9), 934-951. https://doi.org/10.1002/mrd.23414

188) Zhang, Z., Zhang, Q., Xiao, Q., Sun, H., Gao, H., Yang, Y., ... & Pan, Y. (2018). Distribution of runs of homozygosity in Chinese and Western pig breeds evaluated by reduced-representation sequencing data. *Animal genetics*, *49*(6), 579-591. https://doi.org/10.1111/age.12730

189) Zhang, S. H., Liu, D. X., Wang, L., Li, Y. H., Wang, Y. H., Zhang, H., ... & Chen, Y. H. (2019). A CASPR1–ATP1B3 protein interaction modulates plasma membrane localization of Na+/K+-ATPase in brain microvascular endothelial cells. *Journal of Biological Chemistry*, *294*(16), 6375-6386. https://doi.org/10.1074/jbc.RA118.006263

190) Zhang, J., Zheng, T., Zhou, X., Wang, H., Li, Z., Huan, C., ... & Zhang, W. (2021). ATP1B3 Restricts Hepatitis B Virus Replication Via Reducing the Expression of the Envelope Proteins. *Virologica Sinica*, *36*(4), 678-691. https://doi.org/10.1007/s12250-021-00346-2

191) Zhang, J. L., Richetti, S., Ramezani, T., Welcker, D., Lütke, S., Pogoda, H. M., ... & Hammerschmidt, M. (2022). Vertebrate extracellular matrix protein hemicentin-1 interacts physically and genetically with basement membrane protein nidogen-2. *Matrix Biology*. https://doi.org/10.1016/j.matbio.2022.08.009

192) Zheng, B., Zhang, J., Zheng, T., Wang, H., Li, Z., Huan, C., ... & Zhang, W. (2020). ATP1B3 cooperates with BST-2 to promote hepatitis B virus restriction. *Journal of medical virology*, *92*(2), 201-209. https://doi.org/10.1002/jmv.25599

193) Zheng, T., Song, Z., Tao, Y., Qiang, J., Ma, J., Lu, S., & Xu, P. (2022). Transport stress induces innate immunity responses through TLR and NLR signaling pathways and increases mucus cell number in gills of hybrid yellow catfish (Tachysurus fulvidraco♀× Pseudobagrus vachellii♂). *Fish & Shellfish Immunology*, *127*, 166-175. https://doi.org/10.1016/j.fsi.2022.06.015

194) Zhou, W., & Hildebrandt, F. (2009). Molecular cloning and expression of phospholipase C epsilon 1 in zebrafish. *Gene Expression Patterns*, *9*(5), 282-288. https://doi.org/10.1016/j.gep.2009.03.003

195) Zhu, M., Xu, B., Wang, M., Liu, S., Zhang, Y., & Zhang, C. (2019). Pharmacological modulation of MRAP2 protein on melanocortin receptors in the sea lamprey. *Endocrine connections*, *8*(4), 378-388. https://doi.org/10.1530/EC-19-0019

196) Zhuang, Y., Ly, R. C., Frazier, C. V., Yu, J., Qin, S., Fan, X. Y., ... & Wang, L. (2019). The novel function of tumor protein D54 in regulating pyruvate dehydrogenase and metformin cytotoxicity in breast cancer. *Cancer & metabolism*, *7*(1), 1-15. https://doi.org/10.1186/s40170-018-0193-4

Data Accessibility and Benefit-Sharing

Restrictions applied to the availability of the data that support the findings of this study, which were used under license and so are not publicly available. The data can be made available for reproduction of the results from Florence Phocas (florence.phocas@inrae.fr) on request via a material transfer agreement and with permission of the two breeding companies “Viviers de Sarrance” (Sarrance, France) and "Milin Nevez” (Plouigneau, France).

Author contributions

Katy Paul: Investigations, Methodology, Formal analysis, Writing - Original Draft;

Gwendal Restoux: Conceptualization, Methodology, Draft Reviewing;

Florence Phocas: Supervision, Conceptualization, Methodology, Investigation, Formal analysis, Resources, Writing - Original Draft.

Tables and Figures (with captions)

TABLE 1. Data information used to phase the HD genotypes of the study females that belong to the parental cohorts. Number of individuals and SNPs available after quality control

Line	Status of individuals	Number of individuals	Number of SNP used
LB	parents	288	571,319
	offsprings	1,297	29,091
LC	parents	173	569,03
	offsprings	1,350	30,379
SY	parents (dams + 1 sire)	16	569,03
	offsprings (+ 8 sires)	866	32,725

TABLE 2. ROH statistics and inbreeding coefficients of the four studied populations (Standard deviations are indicated in brackets).

Population	Average number of ROH	Average size of ROH (in kb)	Average F_{ROH}
LB	168 (14.6)	2,770 (270.8)	0.20 (0.02)
LC	157 (15.9)	2,485 (326.8)	0.17 (0.03)
SY	141 (33.5)	1,860 (291.2)	0.12 (0.05)
HA	167 (65.6)	1,433 (145.6)	0.04 (0.03)

TABLE 3. Genome-wide Fst statistics derived two-by-two between the four populations.

	LC	LB	HA
SY	0.104	0.122	0.275
LC		0.122	0.275
LB			0.289

1580 **TABLE 4.** Homozygous regions under positive selection in the four studied populations.

Region	CHR	Start (Mb)	End (Mb)	Size (kb)
chr2_a	2	25.40	26.30	900
chr2_b	2	31.60	34.20	2600
chr2_c	2	46.00	46.66	664
chr2_d	2	69.70	71.20	1500
chr2_e	2	88.46	89.34	878
chr12_a	12	57.97	59.10	1138
chr15_a	15	38.96	39.57	610
chr16_a	16	45.80	47.00	1200
chr20_a	20	19.10	19.83	726

1581

1582 **TABLE 5.** Highly heterozygous regions shared by the four studied populations.

Region	CHR	Start (Mb)	End (Mb)	Size (kb)	SNP number	SNP density per Mb
chr10_a	10	56.314	56.366	53	20	379
chr13_a	13	46.959	48.071	1,112	446	401
chr19_a	19	10.753	10.823	70	24	342
chr19_b	19	11.354	11.517	163	52	319

1583

1584

1585

1586

1587

1588

1589

1590

1591

1592

1593

1594

1595

1596

1597

TABLE 6. List and functions of the 17 genes annotated in three homozygous regions (chr2_a, chr2_c and chr15_a) shared by the four rainbow trout populations, and the 15 genes in the six other regions already identified as signatures of selection in the literature. **SS : Identify by signature of selection in that study*

Region	Gene name	Protein name		General functions	References
chr2_a	*mrap2*	melanocortin-2 receptor accessory protein 2A	Cellular organization and growth	May regulate both receptor trafficking and activation in response to ligands. Link to energy homeostasis control and body weight regulation. Linked to severe obesity in many species	Liu et al., 2013 (human, zebrafish, rodent); Zhu et al., 2019 (sea lamprey); Wang et al., 2021 (snakehead); **SS** : Cadiz et al., 2021 (rainbow trout), Cumer et al., 2021 (goat)
	cep162	centrosomal protein of 162 kDa	Cellular and nuclear organization	Involved in cilium assembly (promote transition at the cilia base). Acts by specifically recognizing and binding the axonemal microtubule.	Wang et al., 2013
	uncharacterized LOC110539089				
	adgrb1	adhesion G protein-coupled receptor B1	Neuronal and embryonic development	Essential for growth and metastasis of solid tumors (zebrafish). Plays a role during brain/neuron development, associated with autism in mice and human (BAI1 synonymous of adgrb1).	Purcell, 2017 (human); Cazorla-Vázquez & Engel, 2018 (from zebrafish to human); Shiu et al., 2022 (mice)
	tsnare1	t-SNARE domain-containing protein 1	Cellular organization and neuronal developement	Predicted to be involved in intracellular protein transport; vesicle docking; vesicle fusion ; and integral component of membrane. Neurodevelopment function.	Fromer et al., 2016 (zebrafish and human)
	pttg1ip	pituitary tumor-transforming gene 1 protein-interacting protein	Cellular organization and growth	Participates in metaphase-anaphase transition of the cell cycle and facilitates translocation of pttg1 into the nucleus + allow to predict breast cancer survival + induced transcriptional activation of transcriptional basic fibroblast growth factor (when coexpressing with pttg1).	Repo et al., 2017 (human)
	Cdk14	cyclin-dependent kinase 14	Neuronal and embryonic development	Regulator of cell cycle progression and proliferation + role in meiosis, neuron differentiation/craniofacial development (Wnt signaling pathway)	Margarit et al., 2014 (zebrafish); Yin et al., 2021 (human)
chr2_b	*pbx1*	pre-B-cell leukemia transcription factor 1	Neuronal and embryonic development	Related to early development in zebrafish. Mutations in this gene generally cause major malformations, which seem to play an essential role in survival in various species.	Teoh et al., 2010 (zebrafish); Selleri et al., 2001 (mouse); Le Tanno et al., 2017 (human); **SS:** Cadiz et al., 2021 (rainbow trout)
	col9a2	collagen alpha-2(IX) chain	Neuronal and embryonic development	Component of cartilage, implicated in human intervertebral disc degeneration (IVDD) and seems also related to growth. Mutations in this gene may cause diverse syndromes, such as multiple epiphyseal dysplasias and ocular, skeletal, orofacial, and auditory abnormalities in humans.	Muragaki et al., 1996; Baker et al., 2011; Xu et al., 2022 (human); **SS:** Lopez et al., 2018 (atlantic salmon)
	brd2	bromodomain containing 2	Nuclear and cellular organization, neuronal and embryonic development	Associated with transcription complexes and acetylated chromatin during mitosis. Potential role in oogenesis, egg-to-embryo transition, and proper development of the digestive and central nervous systems (zebrafish). And involved in spermatogenesis or folliculogenesis, as demonstrated in situ on mice cells.	DiBenedetto et al., 2008 (zebrafish); Rhee et al., 1998 (mouse); **SS:** Lopez et al., 2018 (atlantic salmon)
	scap	sterol regulatory element-binding protein cleavage-activating protein	Cellular organization	Binds to sterol regulatory element binding proteins (SREBPs) and transports them from the ER to the Golgi.	Howarth et al., 2013 (zebrafish); **SS:** Wang et al., 2022 (Pacific White Shrimp)

Table 6 (continued)

Region	Gene name	Protein name	General functions		References
	tnf- α - ip8l2b	tumor necrosis factor, alpha-induced protein 8-like protein 2 B	Immunity	Predicted to be involved in the negative regulation of T-cell activation, inflammatory response, innate and adaptative immunity by maintaining immune homeostasis.	Umasuthan et al., 2014 (Oplegnathus fasciatus); Sullivan et al., 2017 (vertebrates)
	atg5	autophagy protein 5	Immunity	Involved in several cellular processes linked to the immune response, such as autophagic vesicle formation, innate antiviral immune response, lymphocyte development and prolifération in mice.	Miller et al., 2008; Ye et al., 2018 (mouse)
chr2_c	*brsk2a*	serine/threonine-protein kinase BRSK2	cellular organization, neuronal and embryonic development	Enable in several functions: ATP/ATPase binding activity, proteine kinase activity. Key role in polarization of neurons and axonogenesis, cell cycle progress (apoptotic signaling pathway) and insulin secretion (metabolic process). This gene is related to autism spectrum disorder (social deficit) and locomotor defects (larval phase and adulthood) in zebrafish	Hiatt et al., 2019 (human); Deng et al., 2022 (human); Liu et al., 2022 (zebrafish)
	abtb2b	Ankyrin Repeat And BTB Domain Containing 2b	Cellular organization	Predicted to be involved in SMAD protein signal transduction., heterodimerization activity. Act upstream of or within cellular response to toxic substance.	
	b4galnt4a	N-acetyl-beta-glucosaminyl-glycoprotein 4-beta-N-acetylgalactosaminyltransferase 1	Cellular organization	Enables acetylgalactosaminyltransferase activity. Predicted to be located in Golgi cisterna membrane. Predicted to be integral component of membrane.	
chr2_d	*igf-1α*	insulin-like growth factor 1a receptor	Growth	Plays a critical role in transformation events. Cleavage of the precursor generates alpha and beta subunits. It is highly overexpressed in most malignant tissues where it functions as an anti-apoptotic agent by enhancing cell survival.	**SS :** Wayne & vonHoldt, 2012 (dog); Lopez et al., 2019 (atlantic salmon)
chr2_e	*znf135*	gastrula zinc finger protein XlCGF26.1	Neuronal development and cellular organization	Involved in cytoskeleton organization, regulation of cell morphogenesis, and RNA-binding. A mutation of znf135 is related to neurological symptoms in humans.	Raghuram et al., 2017 (human); **SS** : Gutierrez et al., 2016 (atlantic salmon)
chr12_a	*grxcr1*	glutaredoxin domain-containing cysteine-rich protein 1-like	Neuronal development and cellular organization	Involved in actin organization in hair cells and is associated with a non-syndromic hearing impairment and the regulation of hair bundle morphogenesis in mouse. A mutant for this gene was identified in mice and linked to hyperactivity (modifies behaviour).	Liu et al., 2021, Lorente-Cánovas et al., 2022 (mouse); **SS:** Saravanan et al., 2021 (cattle)
chr15_a	*chn1*	N-chimaerin	Embryonic development	Encodes GTPase-activating protein. Plays an important role in neuronal signal-transduction mechanisms. Implication during embryonic development: cell polarity and lack of yolk extension. In zebrafish, a morpholino knockdown of chn1 reveals its crucial role in early development, revealing severe abnormalities (development of round somites, lack of yolf extension, and kinkled posterior notochord).	Leskow et al., 2006 (zebrafish); Miyake et al., 2008; Ip et al., 2012 (human)
	atp5mc1	ATP synthase lipid-binding protein, mitochondrial	Energetic function	Loss of ATP synthase -> aberrant mitochondria cristae morphology + energy metabolism correlated to immune system	Palmer et al., 2011; Miller et al., 2019 (human); Wang et al., 2020 (chicken)

Table 6 (continued)

Region	Gene name	Protein name	General functions		References
	zc3h15	zinc finger CCCH domain-containing protein 15	Embryonic development and cellular organization	Embryonic development (positive regulation of GTPase activity) / Elongation processivity (high tumor progression in melanoma). In addition, in vitro (mice cells) that zc3h15 knockdown had an inhibitory effect on HIV-1 replication and then on HIV infection.	Capalbo et al., 2010 (mice); Li et al 2021 (human)
	zp4	zona pellucida sperm-binding protein 4-like	Reproduction	Extracellular matrix that surrounds the oocytes and early embryo. Plays vital roles during oogenesis, gamete development, fertilization and preimplantation development. Mutation in this gene induces infertility in both males and females in mammals.	Wassarman & Litscher, 2018 (fish); Li et al., 2021 (zebrafish); SS : Lopez et al., 2019 (atlantic salmon)
	uncharacterized protein LOC110490841				
	nid2	nidogen-2-like	Cellular and nuclear organization	Cell-adhesion protein that binds collagens I and IV and laminin and may be involved in maintaining the structure of the basement membrane. Linked to ovarian cancer	Torky et al., 2018 (human); Zhang et al., 2022 (zebrafish, mouse)
	brca2	breast cancer type 2 susceptibility protein	Genome stability and cellular organization	Essential for efficient homology-directed ADN repair. Impaired homology-directed repair caused by brca2 deficiency leads to chromosomal instability and tumorigenesis through lack of repair or misrepair of DNA damage. plays an essential role in ovarian development and tumorigenesis of reproductive tissues	Shive et al., 2010 (zebrafish); Rodriguez-Mari et al., 2011 (zebrafish); Moynahan et al., 2001; Chen et al., 2018 (human)
chr16_a	*atp1b3*	sodium/potassium-transporting ATPase subunit beta-1-interacting protein 3	Cellular and nuclear organization	ATPase responsible for establishing and maintaining the electrochemical gradient of Na+ and K+ ions across the plasma membrane, essential for osmoregulation.	Zhang et al., 2019 (human); **SS:** Naval-Sanchez et al., 2020 (atlantic salmon)
	Dnajc5	dnaJ homolog subfamily C member 5-like	Cellular and nuclear organization	Regulated the ATPase activity of 70kDa heat shock proteins and plays a role in membrane trafficking and protein folding. This protein has been shown to have also anti-neurodegenerative properties in human with a gene expression study.	Nosková et al., 2011 (human); **SS:** Signer-Hasler et al., 2022 (goat)
	samD10	sterile alpha motif domain-containing protein 10-like	Cellular and nuclear organization	Linked to binding activity and transmembranaire pathway	**SS:** Signer-Hasler et al., 2022 (goat)
	nol4	nucleolar protein 4-like	Cellular and nuclear organization	Predicted to enable RNA binding activity	**SS:** Signer-Hasler et al., 2022 (goat)
	tpd54 (=TPD52L2)	tumor protein D54	Cellular and nuclear organization	Related to cellular organization, are characterized by an N-terminal coiled-coil motif that forms homo and heteromeric complexes and affects cell proliferation, adhesion, and invasion.	Mukudai et al., 2013; Zhuang et al., 2019 (human); **SS**: Signer-Hasler et al., 2022 (goat)
	magi2	membrane-associated guanylate kinase, WW and PDZ domain-containing protein 2	Neuronal and embryonic development	Plays a role in regulating activin-mediated signaling in neuronal cells. In zebrafish, the protein of this gene plays a vital role in embryogenesis.	Borah et al., 2016 (zebrafish); SS: Cumer et al., 2021 (sheep); Hou et al., 2012 (cattle)
	emilin3 (=emilin-2)	EMILIN-3	Embryonic development and growth	Played a role in extracellular matrix organization and elastic fiber formation. Its gene expression was related to embryonic development and involved in muscle fiber development in zebrafish.	Milanetto et al., 2008 (zebrafish); **SS:** Baesjou & Wellenreuther, 2021 (australasian snapper)
chr20_a	*auts2*	autism susceptibility gene 2 protein homolog	Neuronal development	Related to central nervous system development and is associated with autism in humans.	Oksenberg et al., 2013; Hori et al., 2021 (human); SS: Lopez et al., 2018 (atlantic salmon); Consortium, bovine hapmap, 2009 (cattle)

1602 **TABLE 7.** Percentage of protein identity between rainbow trout and nine other vertebrate
1603 species for all genes annotated in homozygous regions chr2_a, chr2_c and chr15_a.

Region	gene_ID	Human	Mouse	Goat	Cattle	Pig	Chicken	Zebrafish	Medaka	Atlantic salmon
chr2_a	mrap2a	45.71	43.52	43.87	44.98	45.45	42.20	56.22	50.45	87.55
	cep162	36.61	37.50	40.95	40.95	50.00	53.73	38.61	62.50	80.66
	adgrb1	62.37	63.44	62.12	61.95	61.74	67.72	84.02	80.94	98.16
	tsnare1	54.78	31.50	56.99	55.79	56.02	60.74	85.62	78.55	98.29
	pttg1IP	60.00	57.89	57.04	57.04	58.82	59.74	70.92	66.03	93.89
	cdk14	87.05	87.05	86.44	86.02	85.99	87.24	88.96	91.08	99.58
chr2_c	brsk2a	92.12	92.50	92.66	92.19	92.66	93.82	96.14	92.05	96.26
	abtb2b	71.54	70.76	71.93	71.74	71.74	72.46	81.05	61.68	97.35
	b4galnt4a	63.35	65.38	57.14	64.63	64.95	66.37	66.06	79.96	96.55
chr15_a	chn1	88.80	86.59	88.04	87.32	88.04	88.10	85.29	85.01	98.04
	atp5mc1	97.37	87.10	91.76	94.44	93.33	86.17	91.30	97.87	94.12
	zc3h15	69.35	68.57	67.55	67.55	67.55	66.90	74.33	71.57	97.30
	zp4	29.67	29.61	31.87	37.47	30.21	31.46	45.60	49.74	74.74
	nid2	52.00	51.16	51.11	51.11	51.22	55.00	58.14	52.78	97.87
	brca2	46.24	43.72	37.12	32.42	45.61	45.85	38.68	54.08	92.05

1604

1605 **TABLE 8.** List and functions of the 4 genes annotated in three heterozygous regions (chr10_a,
1606 chr19_a, and chr19_b) shared by the four rainbow trout populations.

Region	Gene name	Protein name	General function	References
chr10_a	*ctnna2*	catenin alpha 2	Enables actin filament binding activity, and involved in negative regulation of Arp2/3 complex-mediated actin nucleation. Regulation of neuron migration and of neuron projection development. Implicated in brain malformations. Seems implicated in vertebral development(/deformities) in Yunlong grouper.	Uvarov et al., 2014 (vertebrates) ; Li et al., 2022 (Yunlong grouper)
chr19_a	*smarca5*	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 5	The protein encoded by this gene is a member of the SWI/SNF family of proteins. Members of this family have helicase and ATPase activities and are thought to regulate transcription of certain genes by altering the chromatin structure around those genes. The protein encoded by this gene is a component of the chromatin remodeling and spacing factor RSF, a facilitator of the transcription of class II genes by RNA polymerase II	Ding et al., 2021 (zebrafish); Limi et al., 2018 (mice) : Armas et al., 2013 (zebrafish)
	frem2	FRAS1-related extracellular matrix protein 2	Plays a role in epidermal-dermal interactions -> important for the integrity of skin and renal epithelia.	Gautier et al., 2008 (zebrafish)
chr19_b	*pou4f2*	POU domain, class 4, transcription factor 2-like	May be involved in maintaining visual system neurons in the retina, and in the lateral line. The level of the encoded protein is also elevated in a majority of breast cancers, resulting in accelerated tumor growth. Seems link to cardiac development in zebrafish	DeCarvalho et al., 2004 (zebrafish); Maskell et al. 2017 (zebrafish)

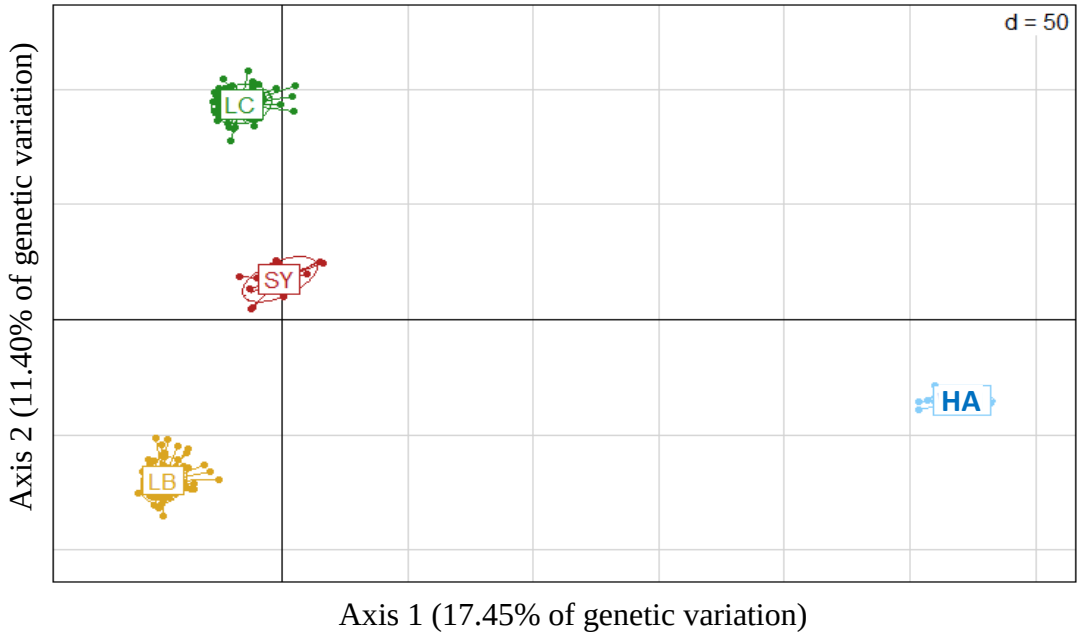


FIGURE 1. Principal component analysis (PCA) of the genetic diversity of the four rainbow trout populations (LB, LC, SY, and HA) based on 546,903 SNPs.

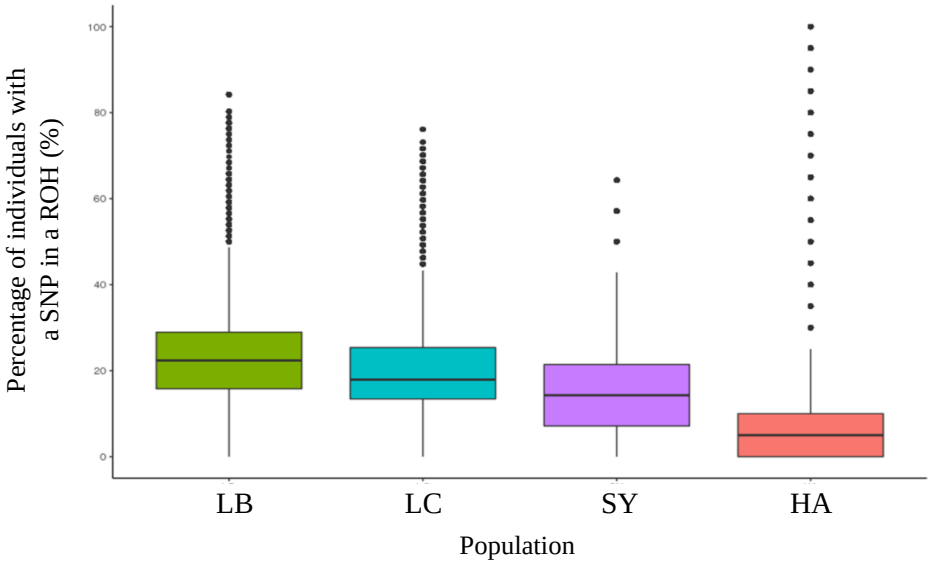


FIGURE 2. Box-plots of the occurrence of ROH (number of individuals having this ROH) per SNP for each rainbow trout population LB, LC, SY, and HA.

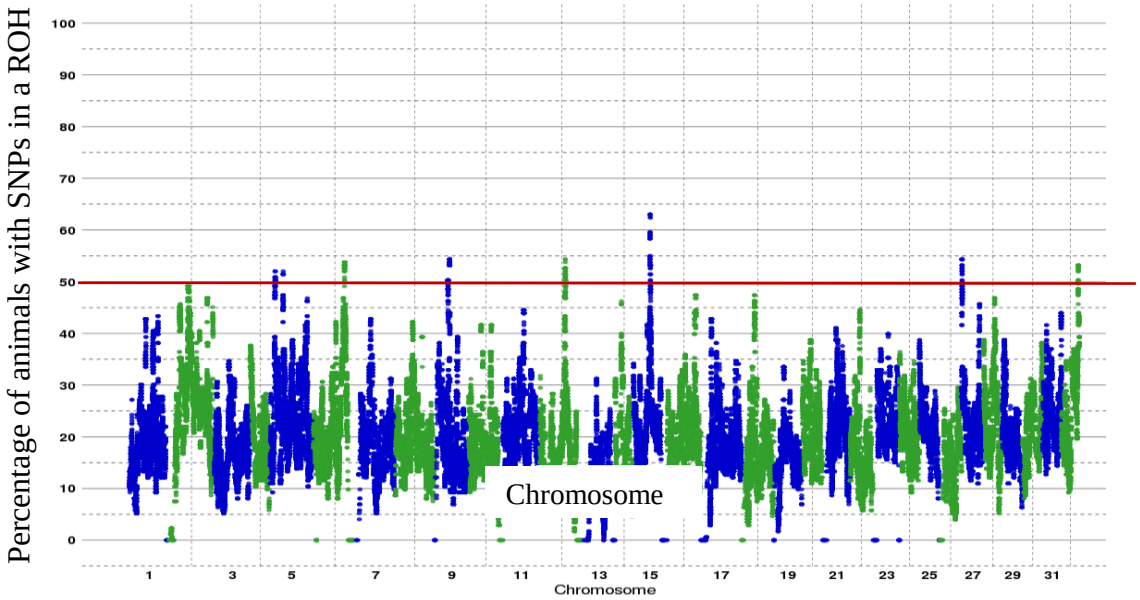


FIGURE 3. Manhattan plot of the occurrence of ROH per SNP across chromosomes (gathering all rainbow trout populations). The red line highlights the ROH islands.

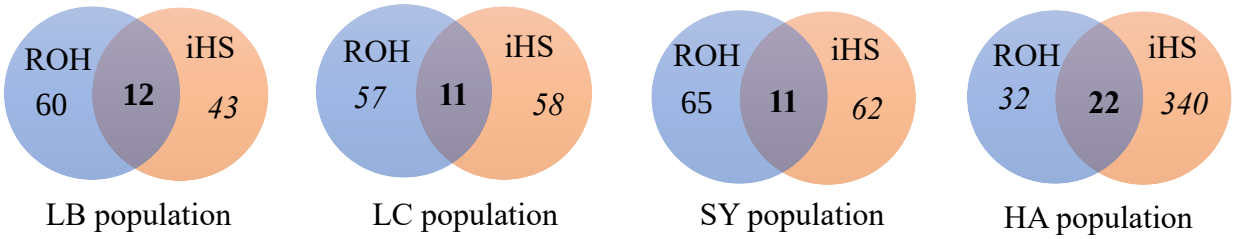


FIGURE 5. Venn Diagram of the number of regions identified as ROH island or iHS signature of selection for each rainbow trout population.

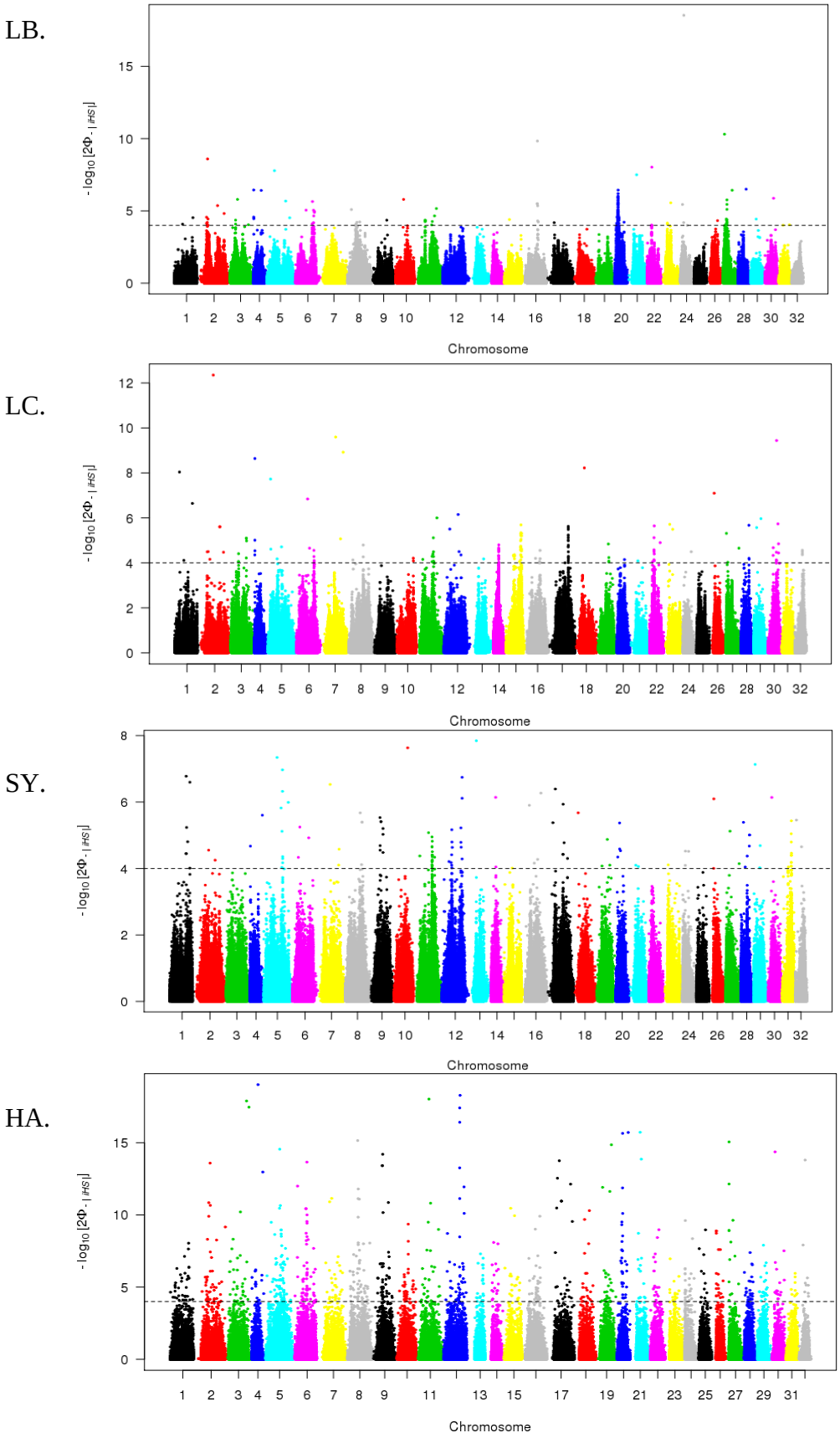
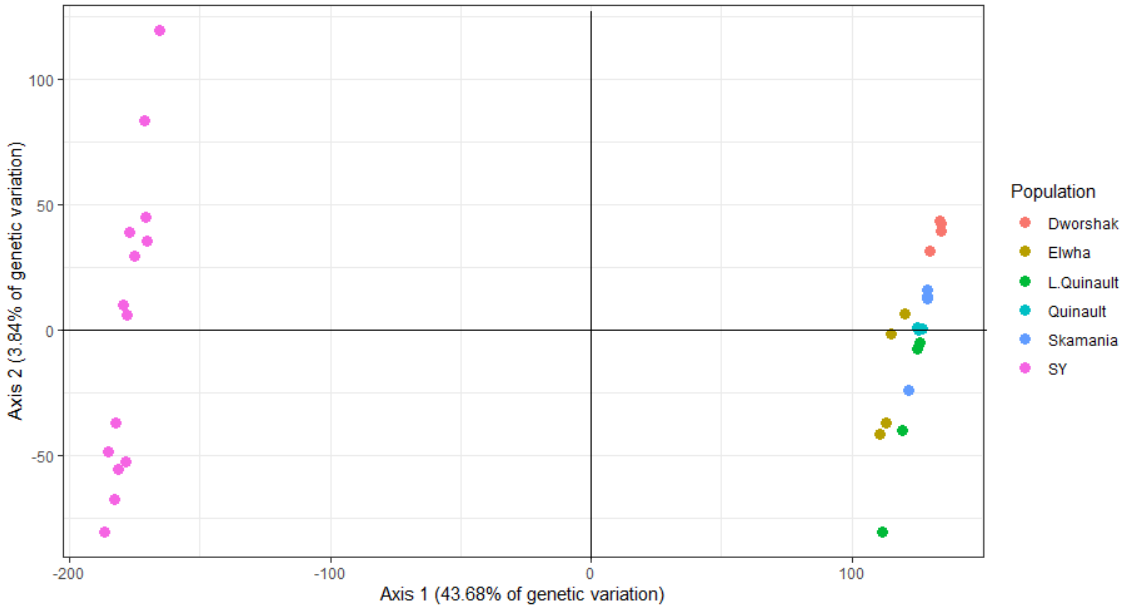


FIGURE 4. Genome-wide distribution of log(p-value) for standardized iHS for each population (LB, LC, SY, HA). The dashed line indicates the log(p-value) significance threshold set to 4 to identify regions under positive selection

Appendices



Supplementary figure 1. Principal component analysis (PCA) of the genetic diversity of SY, and HA sub-populations based on 546,903 SNPs.

File: *Supplementary_Tables.xlsx*