Genome-scale comparative analysis for host resistance against sea lice between Atlantic salmon and rainbow trout

Pablo Cáceres[1], Agustín Barría[1], Kris A. Christensen[2], Liane N. Bassini[3], Katharina Correa[4], Jean P. Lhorente[4], José M. Yáñez[1,5]

[1] Facultad de Ciencias Veterinarias y Pecuarias, Universidad de Chile, Av. Santa Rosa 11735, La Pintana, Santiago 8820808, Chile

[2] Fisheries and Oceans Canada 4160 Marine Drive, West Vancouver, BC, Canada

[3] Escuela de Medicina Veterinaria, Facultad de Ciencias de la Vida, Universidad Andres Bello, Santiago, Chile.

[4] Benchmark Genetics Chile, Ruta 7 Carretera Austral Km 35, Puerto Montt, Chile.

[5] Núcleo Milenio INVASAL, Concepción Chile

jmayanez@uchile.cl

Abstract

Sea lice (*Caligus rogercresseyi*) are ectoparasites that cause major production losses in the salmon aquaculture industry worldwide. Atlantic salmon (*Salmo salar*) and rainbow trout *(Oncorhynchus mykiss*) are two of the most susceptible salmonid species to sea lice infestation. The goal of this study was to identify common candidate genes involved in resistance against sea lice. For this, 2,626 Atlantic salmon and 2,643 rainbow trout from breeding populations were challenged with sea lice and genotyped with a 50k and 57k SNP panel. We ran two independent genome-wide association studies for sea lice resistance on each species and identified 7 and 13 windows explaining 3% and 2.7% respectively the genetic variance. Heritabilities were observed with values of 0.19 for salmon and 0.08 for trout. We identified genes associated with immune responses, cytoskeletal factors and cell migration. We found 15 orthogroups which allowed us to identify *dust8* and *dust10* as candidate genes in orthogroup 13. This suggests that similar mechanisms can regulate resistance in different species; however, they most likely do not share the same standing variation within the genomic regions and genes that regulate resistance. Our results provide further knowledge and may help establish better control for sea lice in fish populations.

Keywords: *Caligus rogercresseyi*, *Salmo salar*, *Oncorhynchus mykiss,* GWAS, Parasite Resistance, Comparative Genomics.

Introduction

Sea lice, *Caligus rogercresseyi*, first described in 1997 by Boxshall & Bravo[1] are currently the most harmful parasite in salmon farming worldwide[2], and are also the main parasite affecting Chilean salmon production. To date, the economic losses due to this parasite are mainly associated with the reduction of feed conversion and fish growth, indirect mortality, loss of product value and treatment costs. It has been estimated that the global costs for the control of this parasite has reached $436 million (USD) annually[3]. The parasite life cycle is comprised of eight stages of development[4]: two states of nauplii, one copepod state, four chalimus states and the adult state. The stages of nauplii (n1-2) and the stage of copepods (infectious stage) are planktonic stages. The four stages of chalimus (1-4) are sessile stages and the adult is a mobile stage[5].

It has been observed that *Caligus* primarily affects Atlantic salmon (*Salmo salar*) and rainbow trout *(Oncorhynchus mykiss*), while coho salmon (*Oncorhynchus kisutch*) has an innate lower susceptibility to the parasite[6]. The clinical signs attributed to infection by sea lice include skin lesions, osmotic imbalance and greater susceptibility to bacterial and viral infections through the suppression of immune responses by damage to the skin of the host[7].

Recent studies have estimated significant low to moderate genetic variation for resistance to *C. rogercresseyi*, with heritability values ranging between 0.12 and 0.32 in Atlantic salmon when resistance is defined as the number of parasites fixed to all the fins[5,7,8]. Similarly, when sea lice resistance is defined as the logarithm of the parasite load density, heritability values range between 0.13 and 0.33 in Atlantic salmon[9,10]. Although significant, a heritability value of 0.09 for sea lice resistance have been recently reported in rainbow trout (Bassini et al., submitted)[11]. These results support the feasibility of including

resistance to sea lice into breeding programs for both Atlantic salmon and rainbow trout[5,8,12].

Due to increased research efforts focusing on salmonid genomics, there has been a rapid increase in the discovery of Atlantic salmon and rainbow trout SNP markers and high-density SNP genotyping panels have been developed for both species[13–15]. The availability of a reference genome for rainbow trout[16] and Atlantic salmon[17] have allowed genomic regions associated with resistance to diseases to be identified and annotated. For example, it has been possible to identify regions associated with resistance to bacterial diseases such as *Piscirickettsia salmonis*[18], *Flavobacterium psychrophilum*[19,20] and also parasitic diseases[7,10] in both species.

Comparative genomic approaches[21] allow the identification of genomic similarities between evolutionarily distant species, such as: conserved genes, traces of genome duplication and the comparison of gene function in different species[22]. Traditionally, comparative genomics analyses have focused on orthologous genes (genes related to each other resulting from direct transmission from a common ancestor)[22]. Comparative genomic studies between salmonids have mainly focused on finding evolutionary similarities, such as quantitative trait loci (QTL) related to growth or sexual differentiation[23–25]. To date, there have been no studies comparing genomic regions associated with resistance to *Caligus* in salmonid species.

The objective of the present study was to: i) identify genomic regions associated with resistance to *Caligus rogercresseyi* in Atlantic salmon and rainbow trout through GWAS, and ii) identify functional candidate genes potentially related to trait variation through a comparative genomics approach based on exploring orthologous genes within the associated regions across species.

Results and Discussion

The comparative genomics analysis used in this study, allowed us to identify groups of ortholog genes and several candidate genes among adjacent SNP that explained more than 1% of the significant genetic variance for resistance to *Caligus rogercresseyi*. This first study which compared resistance in both salmonid species, focused on candidate genes which function either to directly defend against the parasite or that participate in defense mechanisms against this parasite.

There was no difference between the average number of sea lice found on Atlantic salmon or rainbow trout in the experimental challenge (Table 1). An average of 5.9 ± 6.6 and 6.1 ± 4.2 was estimated for Atlantic salmon and rainbow trout, respectively. In terms of the maximum number of parasites, this value varied from 106 parasites in Atlantic salmon to 28 in rainbow trout. For both species, there were animals that did not present with any parasites. The average weight at the end of the experimental challenge was 278.1 ± 90.3 g (ranging from 104 to 569 g) and 173.1 ± 31.4 g (ranging from 86 to 265 g), for Atlantic salmon and rainbow trout, respectively. Although the average number of parasites was not significantly different between the two species, the difference in the average final weight at the end of each challenge could explain the difference in the maximum number of parasites found (~4 times more parasites in the larger Atlantic salmon). The number of parasites counted in each species after the challenge is below the range determined in previous studies. For instance, Ødegård *et al*. (2014)[9] obtained an average of 20.96 ± 19.68, while Robledo *et al*. (2018)[26] reported an average of 38 ± 16.

To measure resistance to *C. royercresseyi* we used the logarithm of lice density (LogLD), which allows for correction to the number of parasites based on the body weight of each

fish[9]. The empirical LogLD distribution for both species is shown in Figure 1. The range of LogLD for Atlantic salmon was greater than for rainbow trout, varying from -4.18 to 1.023 and from -3.69 to 0.03, respectively. The average LogLD distribution in the present study was 2.12 ± 0.89 and 1.64 ± 0.65 for Atlantic salmon and rainbow trout, respectively, which are similar values to those reported in a previous study in a different Atlantic salmon population (between -1.66 ± 0.73 and -2.55 ± 0.58)[9].

A total of 2,040 (77.6%) Atlantic salmon, and 45,117 (96.7%) SNPs passed genotyping quality control, while in rainbow trout, 2,466 (93.3%) individuals and 27,146 (67.4%) SNP remained. In both species, significant genetic variation for resistance to *C. rogercresseyi* was estimated by using genomic information, with heritability values of 0.19 ± 0.03 and 0.08 ± 0.01 for Atlantic salmon and rainbow trout, respectively (Table 2). Tsai *et al*. (2016)[10], estimated genomic heritability values for resistance against *L. salmonis* of 0.22 ± 0.08 and 0.33 ± 0.08, while Ødegård *et al*. (2014)[9], observed heritability values of 0.14 ± 0.03 and 0.13 ± 0.03 in Atlantic salmon. Similarly, Yañez *et al* (2014) and Correa *et al* (2017)[8,12] estimated values ranging from 0.10 to 0.12 when defining resistance as the total number of parasites found on all fins using genomic and pedigree information, and Lhorente et al (2012)[5] estimated heritability values of 0.22 ± 0.06 in Atlantic salmon with traits of total count of sessile sea lice per fish and body weight.

We found 5 chromosomes harboring 7 loci explaining more than 1% of the genetic variance for sea lice resistance in Atlantic salmon (Figure 2). In general, these regions explained a low percentage of the total genetic variation with a maximum of 3% explained by a single locus. Thus, two SNP windows (a window was defined as 20 contiguous SNPs) in *Ssa3* explained 1% and 1.4% of the genetic variance. In *Ssa6* there was a window that explained up to 1.9% of the genetic variance, while two windows that explained 1.7% and 3% were found in *Ssa9*. In addition, in *Ssa20* and *Ssa25* we found windows that explained

1.05% and 1.33%, respectively. Table 3 shows the variance explained by each window of SNP in both species.

For rainbow trout the GWAS for LogLD identified 13 regions located in 13 different chromosomes that exceeded 1% of the total genetic variance (Figure 2). Similar to Atlantic salmon, these windows explained a low percentage of the total variance with a maximum of 2.77%; nevertheless, the number of regions surpassing 1% of the genetic variance explained was almost double in rainbow trout compared to Atlantic salmon. The important genomic regions in rainbow trout were located on chromosomes *Omy2, Omy3, Omy4, Omy7, Omy9, Omy10, Omy14, Omy15, Omy16, Omy21, Omy26, Omy28 and Omy29,* and explained 1.6%, 1.7%, 1.03%, 1.17%, 1.6%, 1.2%, 1.4%, 2.77%, 1.07%, 1.3%, 2.3%, 1.7% and 1.2% of the genetic variance for LogLD, respectively (Table 3).

As has been previously reported for other disease resistance traits in aquaculture species[18,20,27–29], our results suggest that sea lice resistance is mainly of polygenic nature (i.e. many genes with small effect are involved in the trait). These results agree with previous studies on sea lice resistance, where a similar genetic architecture was suggested by Tsai *et al*. (2016), Rochus *et al* (2018) and Correa *et al*. (2017)[7,10,30] for *Lepeophtheirus salmonis* and *C. royercress*eyi resistance. Recently Robledo et al (2018)[31] described three QTL in Atlantic salmon related to sea lice resistance, using RNA-seq and WGS data. Since sea lice resistance is polygenic, the genetic improvement of sea lice resistance would most likely be best accomplished by means of genomic selection instead of marker assisted selection. For instance, Correa *et al* (2017) and Tsai *et al* (2016)[10,12] have shown an increase in the accuracy of estimated breeding values (EBVs) using genomic selection, over the use of pedigree-based models[9].

The exploration of the genes within the windows that explained over 1% of the genetic variance for LogLD showed a series of possible candidate genes that were classified into

three groups: related to the immune response, cytoskeleton or metalloproteases. The genes are listed in Table 3 and Table 4 for Atlantic salmon and rainbow trout, respectively.

In salmonids, the main response of the immune system to parasites is mediated by T-Helper 1 and T-Helper 2[32] cells. Thus, genes related with immune response, either by promoting leukocyte growth or favoring migration or activation are strong candidate genes. For instance, in Atlantic salmon we found *T-cell activation Rho GTPase-activating protein* (TAGAP) which participates in the activation and recruitment of T cells by cytokines[33], and *tenascin R* (TNR) which is an extracellular matrix protein, present in bone marrow, thymus, spleen and lymph nodes[34]. The latter has been described as having an adhesin function favoring the mobility of lymphocytes and lymphoblasts[34,35]. In rainbow trout, we found candidate genes with similar functions, such as, T-box 21 (*tbx21*), also known as T-bet (*T-box expressed in T cells*). This gene belongs to the sub Tbr1 family[36], and generates type 1 immunity and participates in the maturation and migration of T-helper 1 (Th1) cells, which in turn produce interferon-gamma (IFN-γ). Studies have described T-bet expression in NK cells (natural killer), dendritic cells and T CD8+ cells[37,38]. A recent study[26] on gene expression with *C. rogercresseyi* infestation in susceptible and resistant Atlantic salmon indicated that several components of the immune system (inflammatory response, cytokine production, TNF and NF-kappa B signaling and complement activation) and tissue repair are upregulated during infection.

Forkhead box protein N1-like (FOXN1) present on *Ssa9* of Atlantic salmon is part of a family of genes widely studied in humans, which are related to various functions including cell growth, lymph node development and T cell differentiation[39]. In addition, it has been proposed that FOXN1 has a role in the activation of fibroblast growth factor receptors[39].

Meanwhile in trout on *Omy21*, *serine/threonine-protein phosphatase 2A 56 kDa* was identified, which is described as having participated in cell growth and signaling[40]. Robledo

199 *et al*. (2018)[26] recently found that in Atlantic salmon, this protein showed the most
200 significant change in the expression ratio between healthy skin and skin where sea lice
201 were found[26]. In Atlantic salmon, we identified *Tripartite motif-containing protein 45* (TRIM)
202 on *Ssa25* which belongs to a large family of proteins present in diverse organisms that can
203 function as a ligase, and can modify ubiquitins and proteins stimulated by interferon of 15
204 kDa (ISG15)[41].

205 Warm-water fish, such as Zebrafish (*Danio rerio*) and rita catfish (*Rita rita*), lower
206 infiltration of neutrophils, favoring wound closing by means of accelerated growth of the
207 epidermis which can take place in a few hours[42]. In coho salmon (*Oncorhynchus kisutch*) it
208 has been observed that a neutrophil infiltration occurs until the second day after sea lice
209 infestation, together with an inflammatory reaction and hyperplasia in the zone[43], with
210 posterior leukocyte recruitment and migration.

211 Several metalloproteases were found in both species, but for the interest of this study, we
212 focused on GEM-interacting protein which interacts with Rab27a or its effector in
213 leucocytes. Rab is a large family of small GTPasas responsible for vesicle cellular
214 transport[44]. Deficiencies of this molecule or the related human protein (?), is correlated
215 with immune deficiencies due to the malfunction of cytotoxic activity of T-lymphocytes,
216 natural killer cells and neutrophils[45].

217 Considering the importance of cell growth and movement in response to sea lice
218 infestation, the cytoskeleton may play a considerable role in this response as well. Genes
219 related to the cytoskeleton were identified, such as epidermal growth factor (EGF), found
220 on *Ssa9.* This gene is part of a superfamily of receptors with tyrosine kinase activity that
221 have been described in a variety of organs with growth promoter functions, cellular
222 differentiation[46] and could participate in tissue repair by promoting cell growth[36]. In
223 rainbow trout, fibroblast growth factors (*fgf11 - Omy10, fgf13 - Omy29*) have similar

functions (angiogenesis and pro-inflammatory response), and were identified as important genes involved in sea lice resistance by Skugor *et al*. (2009) and Robledo *et al*. (2018) in Atlantic salmon[26,43].

*ELMO/CED-12 domain-containing prot 1 w*as identified on *Omy10* of the trout. This protein is characterized mainly from research in the model organism *C. elegans. ELMO/CED-12 domain-containing prot 1* participates in the phagocytosis of apoptotic cells, and in mammals it also has a role in cell migration[47]. Other cytoskeleton related candidate genes include: *Procollagen galactosyltransferase 1* present on *Ssa6*, *collagen alpha-1 (XXVIII) chain-like* on *Ssa25* and *pleckstrin homology domain-containing family H member 1-like*[48].

The top ten SNPs that explained the greatest variance are located on *Ssa9* of the Atlantic salmon, in close proximity to the *breast carcinoma-amplified sequence 3 (bcas3)* gene, which in Atlantic salmon codes for a cell migration factor associated with microtubules that favors cellular mobility[49]. Cell migration is generally induced in response to chemotactic signals, which induces changes in the cytoskeleton and extracellular matrix[50].

We also found in trout, the *tripartite motif-containing protein 16-like* on *Omy15,* which is part of the TRIM superfamily and has functions related to cell differentiation, apoptosis, regulation of transcription and signaling pathways[41] This gene is similar to *Tripartite motif-containing protein 45* present on *Ssa25*. In this region, we also found a locus that codes for interferon-γ 2 (*ifng2*), which is a cytokine that participates in type 1 immune responses and that favors the presentation of antigens and activation of macrophages[51]. On this same chromosome (*Omy15*), we also identified *putative ferric-chelate reductase 1 (frrs1),* which has been described as having functions in the fixation of iron in teleosts[52]. Robledo *et al*. (2018)[26] identified *heme-binding protein* 2 (HEBP2) as a gene involved in Atlantic salmon sea lice resistance, which has an iron-binding function. Different authors[53,54] have

stated that decreasing the availability of iron can be part of a nutritional defense mechanism against sea lice infestation.

The analyses performed show regions of synteny between both species (Figure 3): there are homologous regions that share similarity between the chromosomes across species. However, there was no obvious shared sea lice resistance associations between Atlantic salmon and rainbow trout (Figure 3). It is possible that similar mechanisms regulate resistance between the two species, but the examined populations did not share the same standing variation of the genes regulating resistance. For example, *Ssa03* (Atlantic salmon) shares homology with *Omy28* (rainbow trout) and *Ssa25* with *Omy03*. When performing the search for genes by chromosome (see Table 4), it was not possible to identify genes that were shared in the indicated regions that were related to the trait studied.

We determined 15 orthogroups were shared between both species (Table S1), which we classified according to gene ontology annotations[55]. One the most interesting groups is orthogroup 12 which contained lysophosphatidic acid receptor 2-like (LPA_2) of Atlantic salmon and rainbow trout G-protein coupled receptor 12-like and an uncharacterized protein. This orthogroup shares the same GO categories (GO: 0004930, GO: 0007186, GO: 0016021, GO: 0070915, GO: 0007165, GO: 0016020) related to the receptor signaling pathway associated with protein G. The activation of LPA_2 participates in multiple biological processes, such as cytoskeleton modification via actin fiber formation[56] and have a role in the activation of related adhesion focal tyrosine kinase (RAFTK)[57], which in turn participates like a stimulating factor for monocytes and macrophages[58]. In orthogroup 13, we identified *dual specificity protein phosphatase 10-like* (*dust10*) in Atlantic salmon and *dual specificity protein phosphatase 8* (*dust8*) in rainbow trout. These genes have

similar annotations in both species, which have the function of inactivating p38[59] within the MAPK cascade[60].

Conclusion

The GWAS performed here for Atlantic salmon and rainbow trout made it possible to compare the genetic basis of sea lice resistance in both species. We found novel information about the resistance of Atlantic salmon and rainbow trout to sea lice, which suggests that there might be a response mediated by leukocytes, and at the same time, the cytoskeleton to promote cell mobility and repair of the wound. The analysis of orthologous proteins provided few characterized proteins, therefore, further investigations of these species are needed to better annotate genes and generate advances in the elucidation of genetics behind resistance to *Caligus rogercresseyi* and other important biologically and economic important traits. Although we did not find common genes explaining resistance between species, we found potential functional genes that can be classified under similar mechanisms. These results suggest that it is possible that similar mechanisms regulate resistance between Atlantic salmon and rainbow trout. Our results provide further knowledge to help establish better control and treatment measures for one of the most important parasitic diseases affecting salmon and trout aquaculture.

Material and methods

Rainbow trout

A total of 2,643 rainbow trout were sampled for this study. The fish originated from 105 maternal, 2012 year-class, full-sib families, and belonged to a broodstock population of Aguas Claras S.A company. The fish were separated into three different ponds so that each family was equally represented in each pond. The sea lice infestation was initiated with 105,600 copepodites, on average, an infestation pressure of 40 copepods/fish (produced *in vitro* from ovigerous females). The infestation consisted of depositing the copepodites in each test pond, stopping the flow of water and keeping the pond in darkness for a period of 6 hours. On the sixth day after infestation, parasite counting was performed and caudal fins were sampled for genetic analysis. All the salmon are euthanized and fins were examined for parasite count using a stereoscopic magnifying glass. Wet body weight was recorded for each animal at the end of the challenge.

Atlantic salmon

A total of 2,628 Atlantic salmon smolts belonging to 118 maternal full-sib families from a 2010 year-class of Salmones Chaicas, X Región, Chile, were challenged with *C. rogercresseyi*. The fish were PIT-tagged (Passive Integrated Transponder), acclimated and distributed into three ponds as described in previous studies[5,8]. Infestation with the parasite was carried out using 13 to 24 copepods per fish, stopping the flow of water for 6 hours after the infestation. The challenge lasted 6 days, then the fish were euthanized and the sea lice were counted on all of the fins. A sample of tail fin was taken for genetic analysis and the wet body weight of each fish was measured.

Genotyping

Genomic DNA was extracted from the caudal fin of each challenged fish using the DNeasy Blood & Kit tissue kit (Qiagen), following the manufacturer's instructions. The Atlantic salmon samples were genotyped using an Affymetrix® 50K Axiom® myDesign™ Genotyping Array designed by AquaInnovo and the University of Chile[61], and the rainbow trout samples were genotyped with a 57K SNP array developed by the USDA[13].

Quality control of the genotypes was carried out in PLINK[62]. SNPs with a call rate <= 0.95, a major allele frequency (MAF) < 0.05 and those that were not in Hardy-Weinberg equilibrium ($p < 1x10^{-6}$) were discarded. Individuals were filtered if they had a call rate <= 0.95. All the SNPs and fish that passed quality control, were used for downstream analysis.

Genomic Association Analysis

Resistance to *C. rogercresseyi* was defined as follows, according to Ødegård *et al*. (2014)[9]:

$$LogLD = log_e\left(\frac{LC+1}{\sqrt[3]{BW^2}}\right)$$

Where LD is the Caligus density defined as the Caligus count (LC) on each fish at the end of the experimental challenge divided by the cube root of the body weight of the fish on the same day (BW) squared, is an approximation of the surface of the skin of each fish. The logarithm of LD was used as it has an approximately normal distribution.

Single step genomic BLUP (ssGBLUP) and wide single step genomic BLUP with two iterations (wssGBLUP)[63] was used to identify associations between SNPs and resistance to *Caligus rogercresseyi*, using the BLUPF90 family of programs[64]. Both approaches use

the combination of a genomic and pedigree matrix. Genotype and pedigree information was used to generate the kinship matrix H[65] with the following equation:

$$H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & G^{-1} - A_{22}^{-1} \end{bmatrix}$$

Where A^{-1} is the inverse relationship matrix, for all the animals, constructed from the pedigree, A_{22}^{-1} is the inverse of the pedigree matrix produced from genotyped animals, and G^{-1} is the reverse matrix of genomic relationship. The SNPs were weighted with equal value and assigned the constant 1 to perform the ssGBLUP method. For the wssGBLUP method, the markers used weights estimated by the previous method. The association analysis for both traits were performed using the following mixed linear model y = **X**b + **Z**a + e, where **y** is the vector of phenotypic values (LogLD); **b** is the fixed effects vector (tank); **a** is the vector for random effects considering the structure of covariance between individuals established by matrix **H**; and **e** is the vector for the random residuals; **X** and **Z** are the incidence matrices for fixed and individual effects respectively.

To identify the regions of the genome associated with the traits analyzed, we identified windows of 20 adjacent SNPs where 1% or more of the phenotypic variance was explained, similar to Neto *et al* (2019)[66]. The cumulative percentage of variance explained for each trait was visualized using a Manhattan plot in R [67].

Genome comparison

The rainbow trout (GCF_002163495.1)[68] and Atlantic salmon (GCF_000233375.1)[17] genomes were downloaded from the NCBI and subset for chromosomes associated with sea lice resistance was downloaded using samtools[69]. Synteny between the chromosomes was identified by aligning the sequences using the program Symap v3.4[70]. Circos[71] was

363 used to plot the relationships between rainbow trout and Atlantic salmon chromosomes
364 and to plot sea lice resistance associations to their respective locations.

365

366 **Candidate Genes**

367 The flanking sequences surrounding SNPs associated with sea lice resistance were
368 aligned to the most recent reference genomes of rainbow trout and Atlantic salmon using
369 BLASTn[72]. The sequence was saved in FASTA format. BLASTx was then used to identify
370 coding sequences for proteins in these associated windows. Blast2Go[73] was used in
371 parallel with the FASTA file to identify proteins and select them by function.

372 For both species, the reference genome of *Danio rerio* (GenBank Assembly Accession:
373 GCA_000002035.4) was used to annotate proteins that were not characterized in the
374 rainbow trout or Atlantic salmon reference genomes. To identify orthologous
375 proteins/genes between species, the OrthoFinder[74] program was used with the FASTA
376 sequences obtained with BLASTx.

377

378 **Ethics approval and consent to participate**

379 All the experimental challenges were approved by the Comité Institucional de Cuidado y
380 Uso de Animales of the Universidad de Chile (Certificate N 17,041-VET-UCH).

381

382 **Consent for publication**

383 Not applicable

384

Availability of data and material

Atlantic salmon phenotype and genotype data are available at Figshare (10.6084/m9.figshare.7676147). Rainbow trout phenotypic and genotype data are available in the same repository (https://figshare.com/s/ 5219597a19f23873fda3).

Conflict of interest statement

The Authors declare no conflict of interest.

Authors' contributions

PC assessed the analyses and wrote the initial version of the manuscript. AB contributed with results interpretation discussion and writing. KC contributed with genome comparison analysis, writing and discussion. LNB and KC managed samples, performed DNA extraction and performed the quality control of genotypes. JPL contributed with the study design. JMY conceived the study, contributed to results interpretation with a discussion. All authors approved the manuscript.

Acknowledgments

The current work was funded by Period: 2016-2019, Funding Agency: Research Council United Kingdom and CONICYT, Chile. Project Number: MR/N026144/1 PI: UK: Ross Houston; Chile: José M. Yáñez Co-researcher: John Hickey Institutions: The Roslin Institute - The University of Edinburgh; University of Chile; Aquainnovo

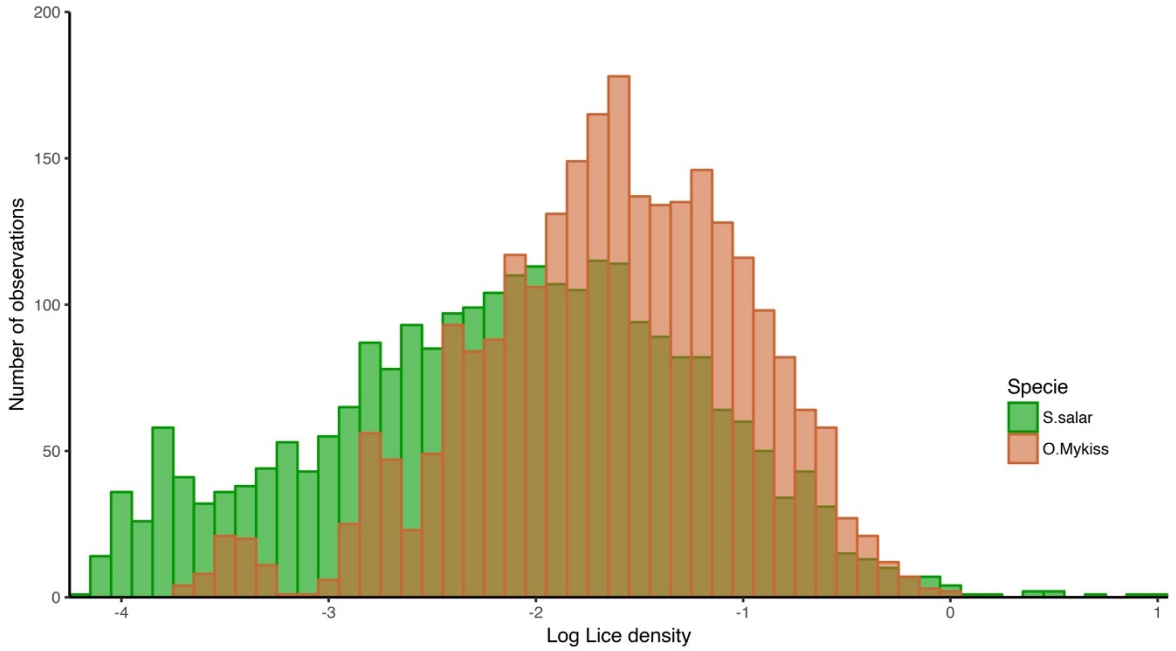
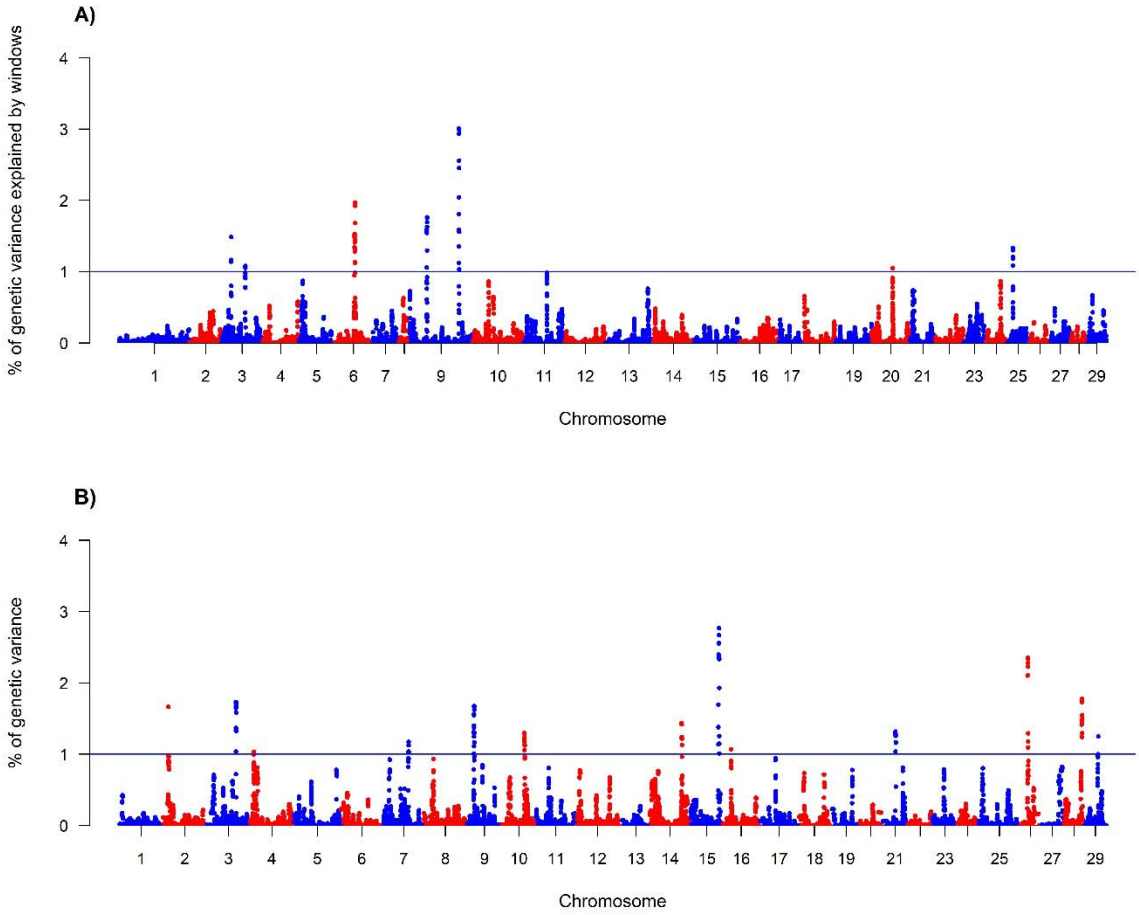
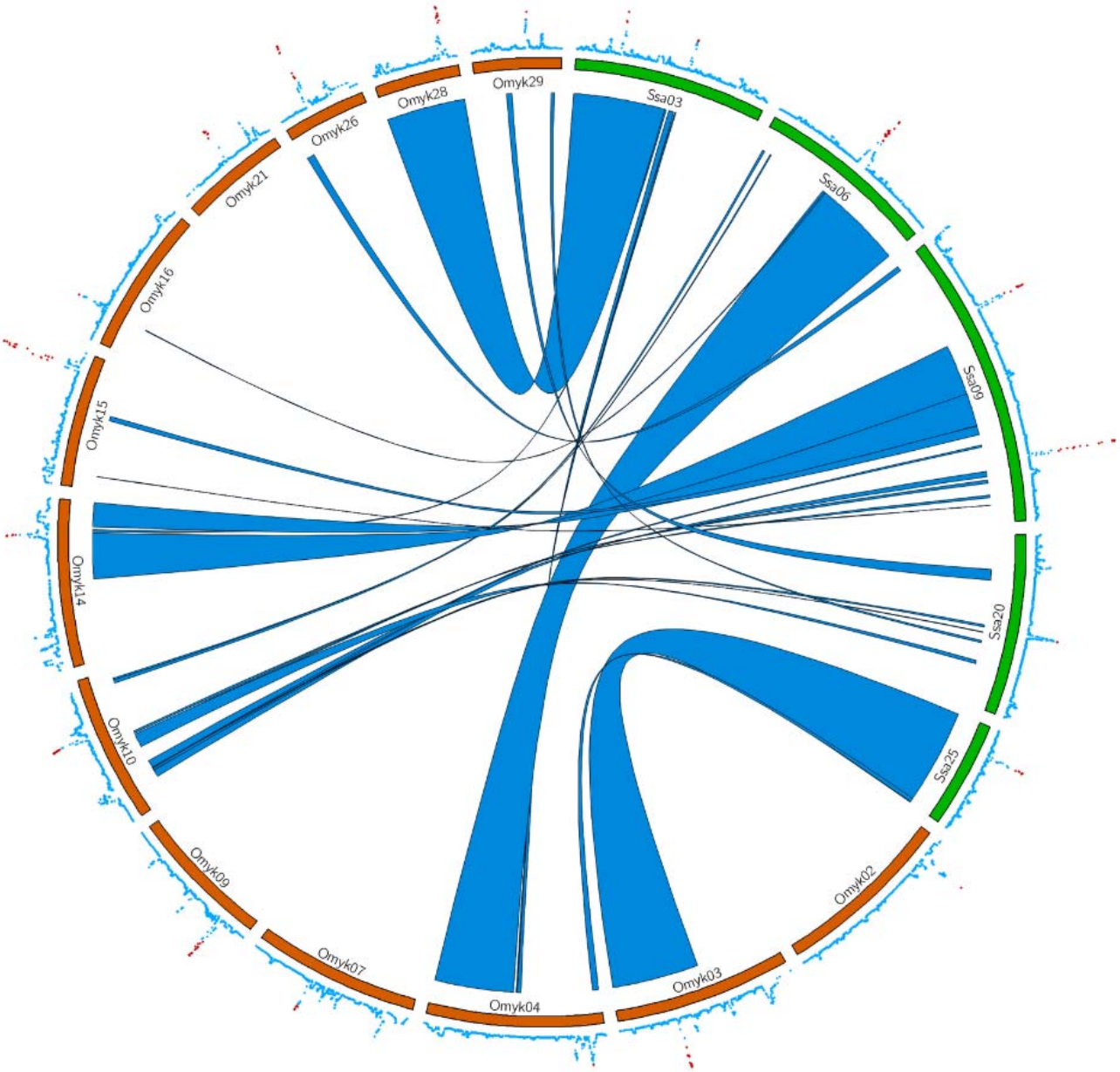


Figure1. Histogram for log lice density (LogLD) for *S. salar* (green) and *O. mykiss* (orange).



411

412 **Figure 2**. Weighted single-step GBLUP (wssGBLUP) for Log sea lice density (LogLD) in
413 Atlantic salmon (A) and rainbow trout (B). Blue line indicates greater than 1% of the
414 variance explained.



415

416 **Figure 3**. A circos plot of sea lice resistance. The inner ribbons mark syntenic regions
417 between Atlantic salmon (green and labeled *Ssa*) and rainbow trout (orange and labeled
418 *Omyk*) chromosomes. Values from the wssGBLUP analysis are plotted on the outer ring,
419 with significant associations plotted in red (values ≥ 1).

420

421 **Table 1**. Summary statistics for body weight (BW) and lice count (LC) in Atlantic salmon
422 and rainbow trout.

Species	Mean BW (g)	SD[1] BW (g)	Min BW(g)	Max BW (g)	LC mean	LC SD	LC min	LC max
S. salar	278.1	90.3	104.0	569.0	5.9	6.61	0	106
O. mykiss	173.1	31.4	86.0	265	6.1	4.22	0	28

423

424 [1] SD (standard deviation)

425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445

446 **Table 2**. Window position and genetic variance of representative SNP.

447

Window Position (bp)[1]	Chr[2]	First and last SNP of window	Var (%)[3]
		Atlantic Salmon	
111398941-112068765	9	Affx-93376378-Affx-93407247	3.00
46812870-48951308	6	Affx-93402585-Affx-93349216	1.97
38862311-39791664	9	Affx-93301094-Affx-93369286	1.76
21387020-22266613	3	Affx-93351937-Affx-93441285	1.48
13302321-13992142	25	Affx-93449253-Affx-93379811	1.33
53395987-54096750	3	Affx-93397375-Affx-93327854	1.08
47471286-48390124	20	Affx-93269128-Affx-93365186	1.05
		Rainbow trout	
55155679-55750071	15	Affx-88957834-Affx-88937607	2.77
14742278- 15862261	26	Affx-88926342-Affx-88913419	2.35
35337013- 36760480	28	Affx-88961757-Affx-88909402	1.77
57059237- 57975684	3	Affx-88911532-Affx-88917594	1.73
14313209- 15091943	9	Affx-88944322-Affx-88927798	1.67
10428810- 11808090	2	Affx-88922063-Affx-88937734	1.66
62615440- 63610148	14	Affx-88942436-Affx-88960365	1.43
27521395-28664751	21	Affx-88932888-Affx-88952962	1.31
43708499-44788897	10	Affx-88912735-Affx-88908041	1.29
26370287-27397531	29	Affx-88925775-Affx-88920967	1.25
50072755-50809742	7	Affx-88956194-Affx-88906442	1.17
15571716-16056405	16	Affx-88940047-Affx-88947010	1.07
7055909-8035171	4	Affx-88935786-Affx-88957720	1.03

448

449 [1]Window position in base pair (BP).

450 [2]Chromosome (Chr).

451 [3] Variance (Var)

452 **Table 3.** Genes identified as possible candidates for sea lice resistance in Atlantic salmon.

Protein name	Gene ID[1]	Location (BP)[2]	Chr[3]	Function
Tenascin R	*tnr*	*22.113,558 - 22.312,716*	3	Immune response
T-cell activation Rho GTPase-activating protein	*tagap*	47.344,667 - 47.350,357	6	Immune response
Forkhead box protein N1-like isoform X1	*LOC106612922*	*111.631,268 - 111.668,978*	9	Immune response
Immunoglobulin superfamily member 11-like	*igsf11*	47.453,435 - 47.570,948	20	Immune response
Tripartite motif-containing 45	*trim45*	*31.699,662 - 31.705,320*	25	Immune response
pre-B-cell leukemia transcription factor 1-like	*LOC106607421*	43.977,204 - 44.037,235	6	Immune response
Bromodomain-containing protein 4-like isoform	*LOC106600922*	*53.482,098 - 53.526,053*	3	Immune response
PAPPALYSIN-2	*pappa2*	21.972,287 - 22.050,408	3	Metalloprotease
Carboxypeptidase D	*cdpa*	*111.778,091 - 111.828,613*	9	Metalloprotease
GEM-interacting protein-like isoform X3	*LOC106600913*	53.672,635 - 53.709,985	3	Metalloprotease
Epidermal growth factor	*egf*	55.552,926 - 55.584,698	9	Cytoskeletal
Procollagen galactosyltransferase 1	*LOC106607427*	*43.820,342 - 43.851,337*	6	Cytoskeletal
Heat shock protein HSP 90-beta	*hs90b*	48.362,362 - 48.374,359	6	Cytoskeletal
Collagen alpha-1(XXVIII) chain-like	*col28a1*	*13.872,240 - 13.923,778*	25	Cytoskeletal
gap junction alpha-4 protein-like	*LOC106611294*	39.373,024 - 39.373,931	9	Cytoskeletal
Pleckstrin homology domain-containing family H member 1-like	*LOC106611291*	*39.101,628 - 39.165,084*	9	Cytoskeletal
serine/threonine-protein kinase OSR1-like	*LOC106600551*	44.726,186 - 44.781,350	3	Cytoskeletal
rho-related GTP-binding protein RhoB-like	*LOC106607565*	46.401,032 - 46.403,034	6	Cytoskeletal
pleckstrin homology and RhoGEF domain containing G1	*plekhg1*	64.024,693 - 64.178,709		Cytoskeletal

453

454 [1]Gene identification (Gene ID) from the NCBI database GenBank assembly accession:

455 GCA_000233375.4.

456 [2]Location in base pair (BP).

457 [3]Chromosome (Chr).

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477 **Table 4.** Genes identified as possible candidates for sea lice resistance in Rainbow trout.

Protein name	*Gene ID*[1]	*Location (BP)*[2]	*Chr*[3]	*Function*
Nuclear factor of activated T-cells,cytoplasmic 1-like	*LOC110518416*	35.501,191 - 35.592,262	3	Immune response
Thymocyte selection-associated family member 2	*themist2*	33.704,133 - 33.726,370	4	Immune response
C-C motif chemokine receptor 10	*ccr10*	21.961,454 - 21.965,777	16	Immune response
T-box 21	*tbx21*	15.560,804 - 15.578,932	16	Immune response
Leucine-rich repeat-containing protein 15-like	*LOC110539182*	11.715,000 - 11.717,281	2	Immune response
Toll-like receptor 13	*LOC110490289*	48.117,536 - 48.120,927	15	Immune response
Adhesion G protein-coupled receptor L3	*LOC110531730*	14.279,162 - 14.524,070	9	Immune response
Inhibin Beta A Chain	*inhba*	36.734,061 - 36.745,506	28	Immune response
interferon gamma2	*ifng2*	153,268 – 183,367	Unplaced Scaffold	Immune response
heat shock protein family A (Hsp70) member 8	*hspa8*	17.691,060 - 17.715,579	10	Immune response
Fibroblast growth factor	*fgf13*	26.660,561 - 26.715,195	29	Cytoskeletal
Fibroblast growth factor	*fgf11*	38.417,625 - 38.489,806	10	Cytoskeletal
Alpha-actinin-3	*3a, 3b*	43.817,039 - 43.841,681	10	Cytoskeletal
Dipeptidyl peptidase 3	*ddp3*	33.026,452 – 33.044,218	29	Cytoskeletal
ELMO/CED-12 domain-containing 1	*elmod1*	16.419,827 – 16.441,415	10	Cytoskeletal
cysteine-rich protein 2-like	*LOC106611283*	38.639,636 – 38.675,313	9	Cytoskeletal
coagulation factor IX-like	*LOC110534144*	43.984,179 – 43.991,091	10	Cytoskeletal
lysyl oxidase homolog 1	*LOC110526332*	59.312,382 – 59.352,343	26	Cytoskeletal

isoform X2				
AFG3 Like Matrix AAA Peptidase Subunit 2	*afg3l2*	35.488,672 – 35.515,514	28	Metalloprotease
Afg3-Like Protein 1	*LOC110506600*	19.946,590 – 19.959,716	26	Metalloprotease

[1]Gene identification (Gene ID) from the NCBI database GenBank assembly accession: GCA_002163495.1.

[2]Location in base pair (BP).

[3]Chromosome (Chr).

References

1. Boxshall, A. & Bravo, S. On the identity of the common Caligus (Copepoda: *Siphonostomatoida: Caligidae*) from salmonid netpen systems in southern Chile. *Contrib. to Zool.* **69**, 137–146 (2000).

2. Costello, M. J. How sea lice from salmon farms may cause wild salmonid declines in Europe and North America and be a threat to fishes elsewhere. *Proc. R. Soc. B Biol. Sci.* **276**, 3385–3394 (2009).

3. Abolofia, J., Asche, F. & Wilen, J. E. The Cost of Lice: Quantifying the Impacts of Parasitic Sea Lice on Farmed Salmon. *Mar. Resour. Econ.* **32**, 329–349 (2017).

4. Gonzalez Laura & Juan, C. Life cycle of *Caligus rogercresseyi*, (Copepoda: *Caligidae*) parasite of Chilean reared salmonids. *Aquaculture* **220**, 101–117 (2003).

5. Lhorente, J. P. *et al.* Quantitative genetic basis for resistance to *Caligus rogercresseyi* sea lice in a breeding population of Atlantic salmon (*Salmo salar*). *Aquaculture* **324**–**325**, 55–59 (2012).

6. Hamilton-West, C. *et al.* Epidemiological description of the sea lice (*Caligus rogercresseyi*) situation in southern Chile in August 2007. *Prev. Vet. Med.* **104**, 341–345 (2012).

7. Correa, K. *et al.* Genome wide association study for resistance to *Caligus rogercresseyi* in Atlantic salmon (*Salmo salar*) using a 50K SNP genotyping array. *Aquaculture* **472**, 61–65 (2017).

8. Yáñez, J. M. *et al.* Genetic co-variation between resistance against both *Caligus rogercresseyi* and *Piscirickettsia salmonis*, and body weight in Atlantic salmon (*Salmo salar*). *Aquaculture* **433**, 295–298 (2014).

9. Ødegård, J. *et al.* Genomic prediction in an admixed population of Atlantic salmon (*Salmo salar*). *Front. Genet.* **5**, 1–8 (2014).

10. Tsai, H. Y. *et al.* Genomic prediction of host resistance to sea lice in farmed Atlantic salmon populations. *Genet. Sel. Evol.* **48**, 1–11 (2016).

11. Bassini, N. *et al.* Genetic parameters for resistance to *Caligus rogercresseyi*, *Piscirickettsia salmonis* and body weight in rainbow trout *(Oncorhynchus mykiss*). *Aquaculture* **472**, 127–127 (2017).

12. Correa, K., Bangera, R., Figueroa, R., Lhorente, J. P. & Yáñez, J. M. The use of genomic information increases the accuracy of breeding value predictions for sea louse (*Caligus rogercresseyi*) resistance in Atlantic salmon (*Salmo salar*). *Genet. Sel. Evol.* **49**, 1–5 (2017).

13. Palti, Y. *et al.* The development and characterization of a 57K single nucleotide polymorphism array for rainbow trout. *Mol. Ecol. Resour.* **15**, 662–672 (2015).

14. Houston, R. D. *et al.* Development and validation of a high density SNP genotyping array for Atlantic salmon (*Salmo salar*). *BMC Genomics* **15**, 1–13 (2014).

15. Yáñez, J. . *et al.* Genomewide single nucleotide polymorphism discovery in Atlantic salmon (*Salmo salar*): validation in wild and farmed American and European populations. **16**, 1002–1011 (2016).

16. Berthelot, C. *et al.* The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. *Nat. Commun.* **5**, (2014).

17. Lien, S. *et al.* The Atlantic salmon genome provides insights into rediploidization. *Nature* **533**, 200–205 (2016).

18. Correa, K. *et al.* Genome-wide association analysis reveals loci associated with

529 resistance against *Piscirickettsia salmonis* in two Atlantic salmon (*Salmo salar*)
530 chromosomes. *BMC Genomics* **16**, 1–9 (2015).

531 19. Rochat, T. *et al.* Genomic characterization of *Flavobacterium psychrophilum*
532 serotypes and development of a multiplex PCR-based serotyping scheme. *Front.*
533 *Microbiol.* **8**, 1–10 (2017).

534 20. Vallejo, R. L. *et al.* Similar genetic architecture with shared and unique quantitative
535 trait loci for bacterial cold water disease resistance in two rainbow trout breeding
536 populations. *Front. Genet.* **8**, 1–15 (2017).

537 21. Hardisson, R. C. Comparative Genomics. *Brenner's Encycl. Genet. Second Ed.* **1**,
538 102–105 (2013).

539 22. Fang, G., Bhardwaj, N., Robilotto, R. & Gerstein, M. B. Getting started in gene
540 orthology and functional analysis. *PLoS Comput. Biol.* **6**, (2010).

541 23. Reid, D. P., Szanto, A., Glebe, B., Danzmann, R. G. & Ferguson, M. M. QTL for
542 body weight and condition factor in Atlantic salmon (*Salmo salar*): Comparative
543 analysis with rainbow trout *(Oncorhynchus mykiss*) and Arctic charr (*Salvelinus*
544 *alpinus*). *Heredity (Edinb).* **94**, 166–172 (2005).

545 24. Moghadam, H. K., Ferguson, M. M. & Danzmann, R. G. Evolution of Hox clusters in
546 salmonidae: A comparative analysis between atlantic salmon (*Salmo salar*) and
547 rainbow trout *(Oncorhynchus mykiss*). *J. Mol. Evol.* **61**, 636–649 (2005).

548 25. Woram, R. A. Comparative Genome Analysis of the Primary Sex-Determining Locus
549 in Salmonid Fishes. *Genome Res.* **13**, 272–280 (2003).

550 26. Robledo, D., Gutiérrez, A. P., Barría, A., Yáñez, J. M. & Houston, R. D. Gene
551 expression response to sea lice in Atlantic salmon skin: RNA sequencing

552 comparison between resistant and susceptible animals. *Front. Genet.* **9**, 1–10
553 (2018).

554 27. Barría, A. *et al.* Genomic Predictions and Genome-Wide Association Study of
555 Resistance Against *Piscirickettsia salmonis* in Coho Salmon (*Oncorhynchus*
556 *kisutch*) Using ddRAD Sequencing. *G3; Genes|Genomes|Genetics* **8**,
557 g3.200053.2018 (2018).

558 28. Palaiokostas, C., Ferraresso, S., Franch, R., Houston, R. D. & Bargelloni, L.
559 Genomic Prediction of Resistance to Pasteurellosis in Gilthead Sea Bream (Sparus
560 aurata) Using 2b-RAD Sequencing. *G3; Genes|Genomes|Genetics* **6**, 3693–3700
561 (2016).

562 29. Palaiokostas, C. *et al.* Genome-wide association and genomic prediction of
563 resistance to viral nervous necrosis in European sea bass (*Dicentrarchus labrax*)
564 using RAD sequencing. *Genet. Sel. Evol.* **50**, 1–11 (2018).

565 30. Rochus, C. M. *et al.* Genome-wide association analysis of salmon lice
566 (*Lepeophtheirus salmonis*) resistance in a North American Atlantic salmon
567 population. *Aquac. Res.* **49**, 1329–1338 (2018).

568 31. Robledo, D., Gutierrez, A. P., Barria, A., Lhorente, J. P. & Houston, R. Discovery
569 and functional annotation of quantitative trait loci affecting resistance to sea lice in
570 Atlantic salmon. (2018). doi:10.1101/455626

571 32. Wiegertjes, G. F., Wentzel, A. S., Spaink, H. P., Elks, P. M. & Fink, I. R. Polarization
572 of immune responses in fish: The 'macrophages first' point of view. *Mol. Immunol.*
573 **69**, 146–156 (2016).

574 33. Tamehiro, N. *et al.* T-cell activation RhoGTPase-activating protein plays an

important role in T H 17-cell differentiation. *Immunol. Cell Biol.* **95**, 729–735 (2017).

34. Ruegg, C. R., Chiquet-Ehrismann, R. & Alkan, S. S. Tenascin, an extracellular matrix protein, exerts immunomodulatory activities. *Proc. Natl. Acad. Sci.* **86**, 7437–7441 (1989).

35. Clark, R. A., Erickson, H. P. & Springer, T. A. Tenascin supports lymphocyte rolling. *J. Cell Biol.* **137**, 755–765 (1997).

36. Fast, M. D. Fish immune responses to parasitic copepod (namely sea lice) infection. *Dev. Comp. Immunol.* **43**, 300–312 (2014).

37. Berry, G. *et al.* Distinct Effects of T-bet in TH1 Lineage Commitment and IFN-gamma Production in CD4 and CD8 T Cells. *Science (80-.).* **295**, 338–342 (2002).

38. Townsend, M. J. *et al.* T-bet regulates the terminal maturation and homeostasis of NK and Vα14i NKT cells. *Immunity* **20**, 477–494 (2004).

39. Coffer, P. J. & Burgering, B. M. T. Forkhead-box transcription factors and their role in the immune system. *Nat. Rev. Immunol.* **4**, 889–899 (2004).

40. Veerle, J. & Jozef, G. Protein phosphatase 2A: a high regulated family of serine/threonine phosphatases implicated in cell growth and signalling. *Biochem. J.* **353**, 417–439 (2001).

41. McNab, F. W., Rajsbaum, R., Stoye, J. P. & O'Garra, A. Tripartite-motif proteins and innate immune regulation. *Curr. Opin. Immunol.* **23**, 46–56 (2011).

42. Fast, M. D., Muise, D. M., Easy, R. E., Ross, N. W. & Johnson, S. C. The effects of *Lepeophtheirus salmonis* infections on the stress response and immunological status of Atlantic salmon (*Salmo salar*). *Fish Shellfish Immunol.* **21**, 228–241 (2006).

43. Skugor, S., Glover, K. A., Nilsen, F. & Krasnov, A. Local and systemic gene

598 expression responses of Atlantic salmon (*Salmo salar* L.) to infection with the
599 salmon louse (*Lepeophtheirus salmonis*). *BMC Genomics* **9**, 1–18 (2008).

600 44. Stenmark, H. Rab GTPases as coordinators of vesicle traffic. *Nat. Rev. Mol. Cell*
601 *Biol.* **10**, 513–525 (2009).

602 45. Johnson, J. L., Monfregola, J., Napolitano, G., Kiosses, W. B. & Catz, S. D.
603 Vesicular trafficking through cortical actin during exocytosis is regulated by the
604 Rab27a effector JFC1/Slp1 and the RhoA-GTPase-activating protein Gem-
605 interacting protein. *Mol. Biol. Cell* **23**, 1902–1916 (2012).

606 46. Chen, J. *et al.* Expression and Function of the Epidermal Growth Factor Receptor in
607 Physiology and Disease. *Physiol. Rev.* **96**, 1025–1069 (2016).

608 47. Brzostowski, J. A., Fey, P., Yan, J., Isik, N. & Jin, T. The Elmo family forms an
609 ancient group of actin-regulating proteins. *Commun. Integr. Biol.* **2**, 337–340 (2009).

610 48. Lemmon, M. A., Ferguson, K. M. & Abrams, C. S. Pleckstrin homology domains and
611 the cytoskeleton. *FEBS Lett.* **513**, 71–76 (2002).

612 49. Wang, T., Holland, J. W., Martin, S. A. M. & Secombes, C. J. Sequence and
613 expression analysis of two T helper master transcription factors, T-bet and GATA3,
614 in rainbow trout *Oncorhynchus mykiss* and analysis of their expression during
615 bacterial and parasitic infection. *Fish Shellfish Immunol.* **29**, 705–715 (2010).

616 50. Lamalice, L., Le Boeuf, F. & Huot, J. Endothelial cell migration during angiogenesis.
617 *Circ. Res.* **100**, 782–794 (2007).

618 51. Zou, J. *et al.* Identification and Bioactivities of IFN- in Rainbow Trout *Oncorhynchus*
619 *mykiss*: The First Th1-Type Cytokine Characterized Functionally in Fish. *J.*
620 *Immunol.* **175**, 2484–2494 (2005).

621 52. Bury, N. & Grosell, M. Iron acquisition by teleost fish. *Comp. Biochem. Physiol. - C*
622 **135**, 76490 (2003).

623 53. Jones, S. R. M., Fast, M. D., Johnson, S. C. & Groman, D. B. Differential rejection of
624 salmon lice by pink and chum salmon: Disease consequences and expression of
625 proinflammatory genes. *Dis. Aquat. Organ.* **75**, 229–238 (2007).

626 54. Valenzuela-Muñoz, V. & Gallardo-Escárate, C. Iron metabolism modulation in
627 Atlantic salmon infested with the sea lice *Lepeophtheirus salmonis* and *Caligus*
628 *rogercresseyi*: A matter of nutritional immunity? *Fish Shellfish Immunol.* **60**, 97–102
629 (2017).

630 55. Bateman, A. *et al.* UniProt: The universal protein knowledgebase. *Nucleic Acids*
631 *Res.* **45**, D158–D169 (2017).

632 56. Moolenaar, W. H., Kranenburg, O., Postma, F. R. & Zondag, G. C. M.
633 Lysophosphatidic acid: G-protein signalling and cellular responses. *Curr. Opin. Cell*
634 *Biol.* **9**, 168–173 (1997).

635 57. Park, S. Y., Schinkmann, K. A. & Avraham, S. RAFTK/Pyk2 mediates LPA-induced
636 PC12 cell migration. *Cell. Signal.* **18**, 1063–1071 (2006).

637 58. Hatch, W. C. *et al.* The related adhesion focal tyrosine kinase (RAFTK) is tyrosine
638 phosphorylated and participates in colony-stimulating factor-1/macrophage colony-
639 stimulating factor signaling in monocyte-macrophages. *Blood* **91**, 3967–73 (1998).

640 59. Keyse, S. M. Dual-specificity MAP kinase phosphatases (MKPs) and cancer.
641 *Cancer Metastasis Rev.* **27**, 253–261 (2008).

642 60. Owens, D. M. & Keyse, S. M. Differential regulation of MAP kinase signalling by
643 dual-specificity protein phosphatases. *Oncogene* **26**, 3203–3213 (2007).

644 61. J.M. Yáñez,, S. Naswa, M.E. López, L. Bassini, M.E. Cabrejos, J. Gilbey, L.
645 Bernatchez, A. Norris, C. Soto, J. Eisenhart, B. Simpson, R. Neira, J.P. Lhorente, P.
646 Schnable, S. Newman, A. Mileham, N. D. Development of a 200K SNP Array for
647 Atlantic Salmon: Exploiting Across Continents Genetic Variation. *Proc. 10th World*
648 *Congr. Genet. Appl. to Livest. Prod.* 11–13 (2014).

649 62. Purcell, S. *et al.* PLINK: A Tool Set for Whole-Genome Association and Population-
650 Based Linkage Analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).

651 63. Legarra, A., Christensen, O. F. & Aguilar, I. M. Single Step, A General Approach For
652 GenomicSelection. *Livest. Sci.* (2014). doi:10.1016/j.livsci.2014.04.029

653 64. Misztal, I., Aggrey, S. E. & Muir, W. M. Experiences with a single-step genome
654 evaluation. *Poult. Sci.* **92**, 2530–2534 (2013).

655 65. Aguilar, I. *et al.* Hot topic: A unified approach to utilize phenotypic, full pedigree, and
656 genomic information for genetic evaluation of Holstein final score1. *J. Dairy Sci.* **93**,
657 743–752 (2010).

658 66. Reis Neto, R. V., Yoshida, G. M., Lhorente, J. P. & Yáñez, J. M. Genome-wide
659 association analysis for body weight identifies candidate genes related to
660 development and metabolism in rainbow trout *(Oncorhynchus mykiss*). *Mol. Genet.*
661 *Genomics* **0**, 0 (2019).

662 67. Team, R. D. C. R: A language and environment for statistical computing. (2008).

663 68. Pearse, D. E. *et al.* Sex-dependent dominance maintains migration supergene in
664 rainbow trout. *bioRxiv* 504621 (2018). doi:10.1101/504621

665 69. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**,
666 2078–2079 (2009).

667 70. Soderlund, C., Bomhoff, M. & Nelson, W. M. SyMAP v3.4: A turnkey synteny
668 system with application to plant genomes. *Nucleic Acids Res.* **39**, (2011).

669 71. Connors, J. *et al.* Circos□: An information aesthetic for comparative genomics.
670 *Genome Res.* **19**, 1639–1645 (2009).

671 72. Madden, T. The BLAST Sequence Analysis Tool. in *The NCBI Handbook* (eds.
672 McEntyre J & Ostell J.) (Bethesda (MD): National Center for Biotechnology
673 Information (US), 2002).

674 73. Conesa, A. *et al.* Blast2GO□: a universal tool for annotation , visualization and
675 analysis in functional genomics research. **21**, 3674–3676 (2005).

676 74. Emms, D. M. & Kelly, S. OrthoFinder: solving fundamental biases in whole genome
677 comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.*
678 **16**, 1–14 (2015).

679

680