Single-step genome-wide association study for resistance to *Piscirickettsia salmonis* in rainbow trout (*Oncorhynchus mykiss*)

Rodrigo Marín-Nahuelpi[*,¶], Agustín Barría[*,¶], Pablo Cáceres[*], María E. López[*,1], Liane N. Bassini[*,†], Jean P. Lhorente[‡], José M. Yáñez[*,§]

[*] Facultad de Ciencias Veterinarias y Pecuarias, Universidad de Chile, La Pintana, Santiago, Chile

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

[‡] Benchmark Genetics Chile S.A., Puerto Montt, Chile

[§] Núcleo Milenio INVASAL Concepción, Chile

ORCID IDs: 0000-0002-6890-9144 (R.M.); 0000-0002-2813-4559 (A.B.); 0000-0003-1955-214X (P.C.); 0000-0001-9272-0694 (ME.L); 0000-0003-2487-5058 (L.B); 0000-0002-9157-4231 (JP.L.); 0000-0002-6612-4087 (JM.Y.)

[¶] Both authors contributed equally to this work

[1] Present address: Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden.

24 **GWAS for *Piscirickettsia salmonis* resistance in rainbow trout**

25

26 *Keywords:* GWAS; *Oncorhynchus mykiss*; disease resistance; heritability, SRS

27 Corresponding author: Facultad de Ciencias Veterinarias y Pecuarias, Universidad de Chile,

28 Avenida Santa Rosa 11735, 8820808, La Pintana, Santiago, Chile

29 E-mail: jmayanez@uchile.cl

30

31

32

33

34

35

36

37

38

39

40

41

42

43

ABSTRACT

One of the main pathogens affecting rainbow trout (*Oncorhynchus mykiss*) farming is the facultative intracellular bacteria *Piscirickettsia salmonis*. Current treatments, such as antibiotics and vaccines, have not had the expected effectiveness in field conditions. Genetic improvement by means of selection for resistance is proposed as a viable alternative for control. Genomic information can be used to identify the genomic regions associated with resistance and enhance the genetic evaluation methods to speed up the genetic improvement for the trait. The objectives of this study were to i) identify the genomic regions associated with resistance to *P. salmonis*; and ii) identify candidate genes associated with the trait. We experimentally challenged 2,130 rainbow trout with *P. salmonis* and genotyped them with a 57 K SNP array. Resistance to *P. salmonis* was defined as time to death (TD) and as binary survival (BS). Significant heritabilities were estimated for TD and BS (0.48 $\pm$ 0.04 and 0.34 $\pm$ 0.04, respectively). A total of 2,047 fish and 26,068 SNPs passed quality control for samples and genotypes. Using a single-step genome wide association analysis (ssGWAS) we identified four genomic regions explaining over 1% of the genetic variance for TD and three for BS. Interestingly, the same genomic region located on *Omy27* was found to explain the highest proportion of genetic variance for both traits (2.4 and 1.5% for TD and BS, respectively). The identified SNP in this region is located within an exon of a gene related with actin cytoskeletal organization, a protein exploited by *P. salmonis* during infection. Other important candidate genes identified are related with innate immune response and oxidative stress. The moderate heritability values estimated in the present study show it is possible to improve resistance to *P. salmonis* through artificial selection in the current rainbow trout population. Furthermore, our results suggest a polygenic genetic architecture

and provide novel insights into the candidate genes underpinning resistance to *P. salmonis* in *O. mykiss*.

INTRODUCTION

As in any intensive animal production system, infectious diseases are one of the main threats affecting the success and sustainability of aquaculture (Yáñez *et al.* 2014a). In the case of salmonid production, one of the major pathogens affecting productivity is the facultative intracellular bacteria *Piscicricketssia salmonis,* etiological agent of salmonid rickettsial syndrome (SRS). This bacterium was first identified in 1989 in Chile, in a farmed coho salmon (*Oncorhynchus kisutch*) population (Cvitanich *et al.* 1991). Since then, mortalities resulting from SRS have been also identified in Atlantic salmon (*Salmo salar)* and rainbow trout (*Oncorhynchus mykiss*) in several countries, such as Scotland, Ireland, Norway and Chile (Fryer and Hedrick 2003). In Chile, SRS was responsible for 20.7, 67.9 and 92.6% of the mortalities associated with infectious diseases in *S. salar*, *O. kisutch* and *O. mykiss*, species respectively (Sernapesca 2018). To date, strategies for *P. salmonis* control and treatment are mainly based on vaccines and antibiotics. The effectiveness of both approaches has not been adequate (Rozas and Enríquez 2014). Therefore, it has been estimated that economic losses due SRS mortalities, reached up to US$450 million in Chile in 2012 (Camussetti *et al.* 2015). However, variables such as laboratory diagnosis screening expenses or loss of quality of the harvested fish and products were not considered, implying that the economic impact could be even higher.

Therefore, selective breeding could be a feasible alternative to enhance disease resistance; reducing mortality rates from *P. salmonis*, as well as improving animal health and productivity (Bishop and Woolliams 2014: Yáñez and Martínez 2010). However, the main requisite to include a trait into a genetic program is the presence of significant additive genetic variance within the population (Falconer and Mackay 1996). Previous studies estimated heritability values ranging from 0.11 to 0.41 for *P. salmonis* resistance in Atlantic salmon and coho salmon (Yáñez *et al.* 2013; Yáñez *et al.* 2016a; Barría *et al.* 2018). In the case of rainbow trout, Yoshida *et al.* (2018a) estimated heritabilities ranging from 0.39 to 0.57 for resistance to *P. salmonis* using day of death and 0.54 to 0.62 for binary survival as trait definitions. Altogether, these results demonstrate the possibility of improving this trait by means of artificial selection in different salmonid species.

The development of next generation sequencing technologies has facilitated the identification of thousands of single nucleotide polymorphisms (SNPs) segregating along the genome of several animals, including aquaculture species (Yáñez *et al.* 2015). Thus, using a genotyping by sequencing (GBS) approach in conjunction with genome-wide association studies, some authors evaluated genomic regions associated with resistance to bacterial infections in aquaculture species (Liu *et al.* 2015; Palti e*t al.* 2015a; Palaiokostas *et al.* 2016; Barría *et al.* 2018). However, in salmonid species, the use of SNP panels has been the most used alternative for genotyping a high number of individuals with thousands of genetic variants simultaneously. This has been made simpler by the development of high density SNP arrays for Atlantic salmon (Houston *et al.* 2014; Yáñez *et al.* 2016b) and rainbow trout (Palti *et al.* 2015b). The use of these SNP panels have also allowed the comparison of the accuracy of estimated breeding values (EBV) using genomic selection to pedigree-based genetic evaluations for resistance to infectious diseases in Atlantic salmon (Ødegård *et al.*

2014; Tsai *et al.* 2016; Bangera *et al.* 2017; Correa *et al.* 2017), coho salmon (Barría *et al.* 2018) and rainbow trout (Vallejo *et al.* 2016; Vallejo *et al.* 2017a; Yoshida *et al.* 2018a; 2018b). SNP arrays have also enabled the dissection of the genetic architecture of resistance to bacterial diseases in salmonids. For instance, genomic regions and candidate genes associated with resistance to *P. salmonis* in Atlantic and coho salmon (Correa *et al.* 2015; Barría *et al.* 2018), and bacterial cold water disease (BCWD) in rainbow trout (Vallejo *et al.* 2017b) have been identified.

To date there are no studies aimed at identifying genomic regions or candidate genes associated with resistance to *P. salmonis* in rainbow trout populations. Therefore, the main objectives of the current study were to i) identify genomic regions associated with resistance to *P. salmonis* in a farmed rainbow trout population, and ii) identify candidate genes associated with the trait.

MATERIALS AND METHODS

Population and experimental challenge

The population used in this study was rainbow trout (*Oncorhynchus mykiss*) year-class 2011 bloodstock, owned by Aguas Claras (Puerto Montt, Chile) and belonging to a genetic improvement program run by Aquainnovo (Puerto Montt, Chile). This population was artificially selected for growth, appearance-related traits and carcass quality for three generations. For a detailed description about rearing conditions and population management please see Flores-Mara *et al.* (2017), Rodríguez *et al.* (2018) and Neto *et al.* (2019)

Fish from 105 full-sib families (48 half-sib families) with an average weight of 7.0 ± 1.5 g, were PIT-tagged for individual traceability of families. After tagging, fish were

maintained in a single tank until they were transferred to Aquainnovo's Aquaculture Technology Center Patagonia in August 2012. Fish were acclimated for 20 days in a 15m^3 tank, prior to experimental challenge. A random sample of fish were selected to evaluate the sanitary status of the population, i.e. qRT-PCR for Infectious Salmon Anemia virus (ISAV), Infectious Pancreatic Necrosis virus (IPNV), and *Renibacterium salmoninarum*, and culture for *Flavobacterium spp*. Later, a total of 2,130 juveniles (with an average of 23 individuals per family and ranging from 17 to 27 fish per family), were intraperitoneally (IP) injected with 0.2ml of a lethal dose (LD_{50}) of the LF-89 strain of *P. salmonis* inoculum. Post injection, fish were equally distributed into three different tanks, considering similar family distribution into each replicate (with 5 to 9 fish per family in each tank). Environmental parameters were measured throughout the challenge and the experimental challenge continued until the mortality curve showed a plateau. Daily mortality was recorded, and body weight was measured for each fish at time of death or at the end of the experiment (FW). Surviving fish were euthanized and body weight was also recorded. Fin clips from all fish were sampled and stored in 95% ethanol at -80°C until they were genotyped.

Genotyping

The genomic DNA from the sampled fin clips was extracted using a commercial kit (DNeasy Blood & Tissue Kit, Qiagen), following the manufacturer's instructions. Genotyping was performed using a commercial 57K SNP array (Affymetrix® Axiom® myDesignTM SNP) developed by the National Center for Cool and Cold water Aquaculture at USDA (Palti *et al.* 2015b).

Quality control (QC) was assessed through Affymetrix's Axiom Analysis Software, using default settings. Then, a second QC using Plink software (Purcell *et al.* 2007), was

applied to remove SNPs with a genotype call rate lower than 0.90, minor allele frequency (MAF) < 0.01 and deviated from Hardy-Weinberg Equilibrium ($p < 1x10^{-6}$). Individuals with a call rate lower than 0.90 were also removed from further analyses.

Trait definition

Resistance to *P. salmonis* was defined as time to death (TD), measured in days, with values ranging from 1 until the end of challenge test. Additionally, resistance to *P. salmonis* was also defined as binary survival (BS), with a value of 1 or 0 based on if the fish died or survived until the end of the challenge.

Genomic-Wide association study

A single-step GWAS (ssGWAS) analysis was performed to identify genomic regions associated with resistance to *P. salmonis*. This approach considered fish with both phenotypes and genotypes and also individuals with phenotypes but no genotypes in the analysis (Wang *et al.* 2012). The pedigree and genotypic data in ssGWAS are connected through the H matrix. Thus, the H matrix combines both the pedigree and the genomic relationship matrices (Aguilar *et al.* 2010). Thus, the inverse of the H matrix is:

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

Where A^{-1} is the inverse of the numerator relationship matrix, considering all the phenotyped animals, A_{22}^{-1} is the inverse of the pedigree-based relationship matrix considering only the genotyped animals, and G^{-1} is the inverse genomic relationship matrix. The following model was used for GWAS analysis:

183 $$y = \boldsymbol{X}\beta + \boldsymbol{Z}a + e$$

184 Where *y* is the vector of phenotypes (for TS and BS), β is the vector of fixed effects (tank as
185 factor and final body weight as a covariate), *a* is the vector of random effects, *e* is the vector
186 of residuals, and $\boldsymbol{X}$ and $\boldsymbol{Z}$ are the incidence matrices for fixed and random effects,
187 respectively. A linear model and a threshold model were used for TD and BS, respectively.
188 Both trait definitions were fitted using BLUPF90 statistical software (Misztal *et al.* 2016).
189 Thus, AIREML and THRGIBBS1F90 were used for TD and BS, respectively. For the latter,
190 a total of 200,000 Markov Chain Monte Carlo (MCMC) iterations were used, the first 20,000
191 were discarded as burn-in iterations and from the remaining 180,000 samples, we saved one
192 from every 50. Therefore, the analyses included 3,600 independent samples.

193 For TS and BS, the heritability was estimated as follows:

194 $$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_e^2}$$

195 Where σ_a^2 is the additive genetic variance estimated using the H matrix, and σ_e^2 is the residual
196 variance.

197 To identify genomic regions associated with each trait, we estimated the percentage
198 of the genetic variance (PGV) explained by windows of 20 adjacent SNPs. Then, if a 20 SNP
199 window explained more than 1% of the PGV, we considered that region as associated with
200 resistance to *P. salmonis*.

201

202 **Candidate genes**

The candidate genes were identified by searching 500kb up and downstream from the SNP explaining the highest proportion of PGV within each associated window. For this purpose, we used the last version of the *Oncorhynchus mykiss* reference genome (GCA_002163495.1). The criteria for selecting candidate genes lies in the function of the protein that encodes each gene found, mainly related to immune response, DNA repair, stress response and similar pathways.

Data availability

The raw genotypes and phenotype data are available from the online repository figshare (https://figshare.com/s/221a39319b236d46f9fc). Table S1 contains all genes located within 1Mb window surrounding the SNPs explaining the highest proportion of genetic variance and is available at 10.6084/m9.figshare.7883342.

RESULTS

Descriptive statistics and heritabilities

Summary statistics for resistance to *P. salmonis* measured as TD and as BS and for FW are shown in Table 1. The first death was recorded on day 10 post intraperitoneal injection; the last on day 32. Average TD was 23.26 $\pm$ 7.86 days. At the end of the experimental challenge the proportion of non-survivor fish was 0.59 $\pm$ 0.49. Cumulative

mortality among all 105 families ranged from 7.7 to 100%, indicating considerable phenotypic variation for resistance to *P. salmonis* in the current rainbow trout population. Cumulative mortality within each replicate tank was 59.4, 65.1 and 64.7%. Mortality peaked on days 12, 15 and 19 post injection. Average final body weight was 173.80 ± 52.27 g. This trait ranged considerably among challenged fish, with a minimum of 46.10g and maximum 448g.

Variance components for TD and BS are shown in Table 2. Significant heritability values were estimated for both trait definitions. Thus, 0.48 ± 0.04 and 0.34 ± 0.04 were estimated for TD and BS, respectively. Furthermore, a high genetic correlation was found between both traits (-0.96 ± 0.01).

Genome-wide association study

From all genotyped animals, 2,047 passed quality control (representing 97.10% of the total). A total of 26,068 SNPs remained in the set for further analyses (~ 64.68%). The Figure 1 shows the Manhattan plot for resistance to *P. salmonis* measured as TD and BS. We identified four genomic regions associated with resistance as TD. These regions were located on *Omy3*, *Omy14*, *Omy24* and *Omy27*. For BS, we identified three genomic regions associated with the trait. These were found on *Omy5*, *Omy27* and *Omy30*. Interestingly, the genomic region located on *Omy27* was found to be associated with resistance to *P. salmonis* for both TD and BS. In both cases, this common genomic region explains the highest proportion of genetic variance for each trait, with 2.4 and 1.5% for TD and BS, respectively. The SNP explaining the highest proportion of the genetic variance (Affx-88923370) is the same for both TD and BS (Table 3).

Using the *O. mykiss* reference genome (GCA_002163495) we identified candidate genes associated with resistance to *P. salmonis*. Table 3 shows a summary of the genes located proximate to the SNPs explaining the highest proportion of the genetic variance within each genomic region.

Among the candidate genes flanking the most important SNP on *Omy3* for TD, we found *Gluthatione S-transferease kappa 1* (*gstk1*) and Interleukin-11 (*il11*). These genes are involved in the response to oxidative stress and the immune response to bacterial infections, respectively (Oruc *et al*. 2004; Wang *et al*. 2005). On *Omy14,* we found the *Toll-like receptor 4* (*tlr4*) gene, which has been suggested to act as a bacteria sensor (Palti 2011). On *Omy24,* we found *alpha-2-macroglobulin-like* (*a2m*), which is part of a broad-spectrum protease inhibitor, and it has been suggested that plays a role in the defense against *Cryptobia salmositica* on rainbow trout (Zuo and Woo 1997).

Also, we found *POU class 2 associating factor 1* (*pou2af1*), which has been described as a coactivator of transcription factors that regulate Ig expression of B cells (Teitell 2003), and *NF-kappa-B inhibitor zeta-like (nfkbiz),* a regulator of process like the pathogen recognition, phagocytosis and production of cytokines by dendritic cells (Rozas-Serri *et al.* 2017).

For BS, on *Omy5* we found *fas ligand* (*faslg*), whose protein has been suggested as an important mediator of anti-bacterial innate immune response, by inducing apoptosis of target cells and recruiting phagocytic cells (Kaur *et al.* 2004). On the same chromosome we found *Peroxiredoxin-6-like* (*prdx6*), one of the six different isoforms that conforms the peroxiredoxins group, which are antioxidants proteins that protect cells from oxidative damage and is likely to be involved in protective response against a bacterial infection in *Scophthalmus maximus* (Zheng *et al*. 2010).

On *Omy29*, *MAPK12* was found; previous studies described that MAPK12 is involved on the signaling pathways responsible for TNF-α secretion from rainbow trout macrophages, there for in innate immunity (Roher *et al*. 2011). *Glutaminase kidney isoform, mitochondrial-like* (*gls*) was also found on *Omy29*, which family proteins, generally forms a part of enzymes that plays a role in nucleotide, amino acid and urea biosynthesis (Kumada *et al.* 1993).

On *Omy27* we found genes related with innate immune response regulation, *NF-kB* activation by *TNFα*, and some molecules related with metabolic process and apoptosis. However, the SNP explaining the highest proportion of genetic variance is located within an exon of the gene *Smoothelin protein 2* (*Smtnl2*) which remains poorly characterized both in humans and fishes, but it is believed that participates in actin cytoskeleton organization.

The complete list of genes located within the 1Mb window flanking the SNPs explaining the highest proportion of genetic variance, within each genomic region associated with resistance to *P. salmonis,* is shown in Table S1.

DISCUSSION

In the current study we show significant genetic variation for resistance to *P. salmonis* in a farmed rainbow trout population. A moderate to high heritability was estimated for resistance as TD (0.48) and BS (0.34). These estimates are higher than those reported in previous studies carried out for resistance to other bacterial diseases in aquaculture species, with heritabilities ranging from 0.22 to 0.38 (Ødegård *et al.* 2006; Palaiokostas *et al.* 2016; Vallejo, *et al.* 2017b). In the case of *P. salmonis* resistance, several studies have evaluated the presence of genetic variation in different salmonid species. Thus, similar estimates have

been shown for Atlantic salmon, when using pedigree or genomic data, with values ranging from 0.19 to 0.39 (Yáñez *et al.* 2013; Yáñez, *et al.* 2014b; Correa *et al.* 2015; Bangera *et al.* 2017). In the case of coho salmon, heritability estimates range from 0.16 to 0.27 when resistance is defined as a linear or binary trait (Yáñez, *et al.* 2016a; Barría, *et al.* 2018).

Recent studies in rainbow trout, using different pedigree and genome-based genetic evaluation approaches, estimate heritabilities ranging from 0.39 to 0.57 for TD and from 0.54 to 0.62 for BS (Yoshida, *et al.* 2018a); values which are within the range of our estimations. Moreover, our results suggest a higher effect of the additive genetic component on the phenotypic variance for resistance to *P. salmonis* in rainbow trout when compared to *S. salar* and *O. kisutch*, which would imply potentially faster genetic progress for the improvement of resistance to *P. salmonis* by means of artificial selection in the rainbow trout population used in the present study.

The effect of the genetic architecture of a trait (among other variables) on the accuracy of breeding values obtained through genomic selection (GS) is widely known (Daetwyler *et al.* 2008; Goddard 2009). Previous studies in salmonid species (Atlantic salmon and coho salmon), suggest that resistance to *P. salmonis* is a polygenic trait (Correa *et al.* 2015; Barría, *et al.* 2018). Based on the 26K SNPs which passed QC, our study similarly suggests a polygenic nature for resistance to *P. salmonis* resistance in rainbow trout (*i.e.* no QTL explaining >= 10% of the genetic variance). Thus, it is expected that, when compared with a pedigree-based Best Linear Unbiased Predictor (BLUP) method, a genomic BLUP approach for GS would have an increase in accuracy of breeding values over a Bayesian approach (Habier *et al.* 2007; Hayes *et al.* 2009) for the current rainbow trout population. Nonetheless, as predicted by Yoshida, *et al.* (2018b) this was true only at low SNP densities (*i.e.* 0.5 to 10 K). When 20K and 27K were used, Bayes C outperformed GBLUP accuracies. The authors

319 suggested that this could be due to an oligogenic architecture of the resistance trait, or that
320 Bayes C had higher effectiveness in capturing the linkage disequilibrium between the SNPs
321 and a QTL when more SNPs were used.

322 Resistance to bacterial infections implies a modulation of the host immune response
323 to inhibit or reduce the replication rate of the pathogen (Doeschl-Wilson and Kyriazakis
324 2012). The infection process carried out by *P. salmonis* uses clathrin for internalization and
325 then the actin cytoskeleton for vacuole generation (Ramírez *et al.* 2015). Similar pathways
326 have been observed in other mammalian intracellular gram-negative bacteria (Manon *et al.*
327 2012; Valencia-gallardo *et al.* 2015). Within the region associated with TD on *Omy3* we
328 identified a gene coding for the receptor DC-SIGN related with the immune response and
329 expressed on macrophage and dendritic-cell surfaces (Ahmed *et al.* 2015). It has been
330 previously described that *Mycobacterium tuberculosis*, interferes with the Toll-like receptor
331 signaling by DC-SIGN, inhibiting interleukin-12 production (Gorvel *et al.* 2014), a
332 proinflammatory cytokine, which plays a key role in the performance of phagocytes in teleost
333 fish (Alvarez *et al.* 2016).

334 As mentioned before, endocytosis mediated by clathrin is the main pathway used by
335 *P. salmonis* for cell invasion. Clathrin recruits, among other cell components, AP-2; which
336 is regulated by NECAP-1 (Ritter *et al.* 2013), a gene flanking the SNP explaining the highest
337 proportion of genetic variance in *Omy3* for resistance measured as TD. Similarly, on this
338 chromosome we also found the gene *glutathione S-transferase kappa 1* (*gstk1*) (GTS), which
339 is member of the glutathione S-transferase family (GST), involved in cellular detoxification,
340 and expressed in cells to reduce oxidative stress-related damage (Morel and Aninat 2011), a
341 consequence of *P. salmonis* infection (Rozas and Enríquez 2014), and differentially
342 expressed in Atlantic salmon after *P. salmonis* exposure (Rise *et al.* 2004). A candidate gene

343 related to resistance as measured by BS, was found on *Omy5*, the *fas ligand* gene (*faslg*) is a
344 member of the TNF superfamily. The Fas/FasL pathway is essential for immune system
345 regulation, including apoptosis induced by T cell activation and by cytotoxic T lymphocytes
346 (Siegel *et al.* 2000).

347 For both resistance trait definitions, the same chromosome and identical SNP was
348 identified as the marker explaining the highest genetic variation for resistance, which makes
349 this QTL as an interesting region in rainbow trout. Within this region we found the gene
350 *phosphatidylinositol transfer protein alpha* (*pitpna*), which belongs to the
351 phosphatidylinositol family (ptdlns) (Piscatelli *et al.* 2016), and is responsible for
352 phospholipid transfer between cellular membranes (Thornbrough *et al.* 2016), which in turn
353 are regulators of cell signal transduction, membrane trafficking and cytoskeleton
354 organization (Hilbi and Haas 2012). The latter process is affected by *P. salmonis* once inside
355 the macrophages (Ramírez *et al.* 2015). Similar to *P. salmonis*, *Legionella pneumophila* also
356 replicates inside macrophages, and manipulates the vesicle generation inside the cell by
357 joining with ptdlns 5 (Hilbi and Haas 2012).

358 Additionally, in this region we found the gene *nlr family card domain containing 3*
359 (*nlrc3*). Previously, Álvarez *et al.* (2017), described a higher differential expression of *nlrc3*
360 in rainbow trout in response to bacterial lipopolysaccharides (lps), specifically in the skin,
361 liver and gills. This pattern has also been observed in Atlantic salmon during an infection
362 with *P. salmonis* (Tacchi *et al.* 2011), and is therefore a likely mechanism used by this
363 bacteria to evade the immune response.

364 The gene *tapsain* (*tap*) is also involved in the immune response, transporting cytosolic
365 peptides generated by the proteasome to load on MHC class I (Procko *et al.* 2005). On

366 *Omy27*, we found a gene that encodes a protein related to tapsain (TAPBPR), which
367 negatively regulates *tap*; generating a reduction in immune response efficiency (Boyle *et al.*
368 2013).

369 We expect that in the near future, the identification and validation of causative
370 mutations affecting some of the candidate genes presented here, by means of functional
371 studies, will provide a better understanding of resistance against this and other infectious
372 diseases in rainbow trout and other salmonid species. These studies will be facilitated through
373 international collaborative initiatives such as the Functional Annotation of All Salmonid
374 Genomes, FAASG (Macqueen *et al.* 2017).

375

376 ## CONCLUSIONS

377 To the best of our knowledge this is the first report identifying candidate genes related
378 to resistance to *P. salmonis* in a farmed rainbow trout population. Genes likely related with
379 resistance were identified close to SNPs explaining the highest proportion of genetic
380 variance. Furthermore, we identified the same genomic region associated with resistance
381 using both a linear and binary trait. Our results show that this trait is controlled by multiple
382 genes each with a small effect. Therefore, a genomic selection approach is suggested as the
383 best method to improve this trait by means of artificial selection.

384

385

386 ## ACKNOWLEDGMENTS

387 We would like to thank Aguas Claras S.A. for providing funding for the experimental
388 challenge and fish used in this study. This work was also partially funded by the grant

CORFO Innova-Chile (11IEI-12843) and FONDEF NEWTON-PICARTE (IT14I10100), funded by CONICYT (Government of Chile) and the Newton Fund - The British Council (Government of United Kingdom). JMY is supported by Núcleo Milenio INVASAL funded by Chile's government program, Iniciativa Científica Milenio from Ministerio de Economía, Fomento y Turismo.

AUTHORS' CONTRBUTIONS

RM-N assessed the GWAS analyses, genes identification and contributed with discussion. AB wrote the initial version of the manuscript and contributed with discussion. PC contributed with discussion. MEL contributed with initial analysis. LB performed DNA extraction. JPL contributed with study design. JMY conceived and designed the study and supervised the work of RM-N. All authors reviewed and approved the manuscript.

Animal ethics approval

Experimental challenge was approved by the Comité de Bioética Animal from University of Chile (Certificate Number 17041-VET-UCH)

References

Aguilar, I., I. Misztal, D. L. Johnson, A. Legarra, S. Tsuruta *et al.*, 2010 Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. J. Dairy Sci. 93: 743–752.

Ahmed, Z., T. Kawamura, S. Shimada, and V. Piguet, 2015 The Role of Human Dendritic Cells in HIV-1 Infection. J. Invest. Dermatol. 135: 1225–1233.

Alvarez, C., F. A. Gomez, L. Mercado, R. Ramirez, and H. Marshall, 2016 Piscirickettsia

413 salmonis Imbalances the Innate Immune Response to Succeed in a Productive
414 Infection in a Salmonid Cell Line Model. PLoS One 11: 1–14.
415 Álvarez, C. A., F. Ramírez-cepeda, P. Santana, E. Torres, J. Cortés *et al.*, 2017 Insights into
416 the diversity of NOD-like receptors : Identification and expression analysis of NLRC3
417 , NLRC5 and NLRX1 in rainbow trout. Mol. Immunol. 87: 102–113.
418 Bangera, R., K. Correa, J. P. Lhorente, R. Figueroa, and J. M. Yáñez, 2017 Genomic
419 predictions can accelerate selection for resistance against Piscirickettsia salmonis in
420 Atlantic salmon (Salmo salar). BMC Genomics 18: 121.
421 Barría, A., K. A. Christensen, G. M. Yoshida, K. Correa, A. Jedlicki *et al.*, 2018 Genomic
422 predictions and genome-wide association study of resistance against Piscirickettsia
423 salmonis in coho salmon (Oncorhynchus kisutch) using ddRAD sequencing. G3
424 Genes Genomes Genet. 4231: g3.200053.2018.
425 Bishop, S. C., and J. A. Woolliams, 2014 Genomics and disease resistance studies in
426 livestock. Livest. Sci. 166: 190–198.
427 Boyle, L. H., C. Hermann, J. M. Boname, K. M. Porter, P. A. Patel *et al.*, 2013 Tapasin-
428 related protein TAPBPR is an additional component of the MHC class I presentation
429 pathway. Proc. Natl. Acad. Sci. 110: 3465–3470.
430 Camussetti, M., A. Gallardo, D. Aguilar, and J. Larenas, 2015 Análisis de los costos por la
431 utilización de quimioterápicos y vacunas en la salmonicultura. Salmonexpert 4: 46–49.
432 Correa, K., R. Bangera, R. Figueroa, J. P. Lhorente, and J. M. Yáñez, 2017 The use of
433 genomic information increases the accuracy of breeding value predictions for sea
434 louse (Caligus rogercresseyi) resistance in Atlantic salmon (Salmo salar). Genet. Sel.
435 Evol. 49: 15.
436 Correa, K., J. Lhorente, M. Lopez, L. Bassini, S. Naswa *et al.*, 2015 Genome-wide

association analysis reveals loci associated with resistance against Piscirickettsia salmonis in two Atlantic salmon (Salmo salar L.) chromosomes. BMC Genomics 16: 854.

Cvitanich, J., O. Garate, and C. E. Smith, 1991 The isolation of a rickettsia-like organism causing disease and mortality in Chilean salmonids and its confirmation by Koch's postulate. J. Fish Dis. 14: 121–146.

Daetwyler, H. D., B. Villanueva, and J. A. Woolliams, 2008 Accuracy of Predicting the Genetic Risk of Disease Using a Genome-Wide Approach. PlosOne 3:.

Doeschl-Wilson, A. B., and I. Kyriazakis, 2012 Should we aim for genetic improvement in host resistance or tolerance to infectious pathogens? Front. Genet. 3: 1–2.

Falconer, D. S., and T. F. C. Mackay, 1996 *Quantitative Genetics*.

Flores-mara, R., F. H. Rodríguez, R. Bangera, J. P. Lhorente, R. Neira *et al.*, 2017 Resistance against infectious pancreatic necrosis exhibits signi fi cant genetic variation and is not genetically correlated with harvest weight in rainbow trout (Oncorhynchus mykiss). Aquaculture 479: 155–160.

Fryer, J. L., and R. P. Hedrick, 2003 Piscirickettsia salmonis : a Gram-negative intracellular bacterial pathogen of fish. J. Fish Dis. 247: 251–262.

Goddard, M., 2009 Genomic selection: Prediction of accuracy and maximisation of long term response. Genetica 136: 245–257.

Gorvel, L., J. Textoris, R. Banchereau, A. Ben Amara, W. Tantibhedhyangkul *et al.*, 2014 Intracellular Bacteria Interfere with Dendritic Cell Functions : Role of the Type I Interferon Pathway. PLoS One 9: 1–11.

Habier, D., R. L. Fernando, and J. C. M. Dekkers, 2007 The impact of genetic relationship information on genome-assisted breeding values. Genetics 177: 2389–2397.

461 Hayes, B. J., P. M. Visscher, and M. E. Goddard, 2009 Increased accuracy of artificial
462 selection by using the realized relationship matrix. Genet. Res. (Camb). 91: 47–60.
463 Hilbi, H., and A. Haas, 2012 Secretive Bacterial Pathogens and the Secretory Pathway.
464 Traffic 13: 1187–1197.
465 Houston, R. D., J. B. Taggart, T. Cézard, M. Bekaert, N. R. Lowe *et al.*, 2014 Development
466 and validation of a high density SNP genotyping array for Atlantic salmon (Salmo
467 salar). BMC Genomics 15: 90.
468 Kaur, H., L. Jaso-Friedmann, and D. L. Evans, 2004 Single base oligodeoxyguanosine
469 upregulates Fas ligand release by nonspecific cytotoxic cells. Develpmental and
470 Comparative Immunology 28: 571-579.
471 Kumada, Y., D. R. Benson, D. Hillemann, T. J. Hosted, D. A. Rochefort and C. J.
472 Thompson, 1993 Evolution of the glutamine synthetase gene, one of the oldest
473 existing and functioning genes. Proc. Natl. Acad. Sci. 90: 3009-3013.
474 Liu, S., R. L. Vallejo, Y. Palti, G. Gao, D. P. Marancik *et al.*, 2015 Identification of single
475 nucleotide polymorphism markers associated with bacterial cold water disease
476 resistance and spleen size in rainbow trout. Front. Genet. 6: 1–10.
477 Macqueen, D. J., C. R. Primmer, R. D. Houston, B. F. Nowak, L. Bernatchez *et al.*, 2017
478 Functional Annotation of All Salmonid Genomes (FAASG): an international
479 initiative supporting future salmonid research , conservation and aquaculture. BMC
480 Genomics 18: 1–9.
481 Manon, R., N. Abed, F. Namdari, I. VIrlogeux-Payant, P. Velge *et al.*, 2012 The different
482 strategies used by Salmonella to invade host cells, pp. 339–364 in *Salmonella –*
483 *Distribution, Adaptation, Control Measures and Molecular Technologies*,.
484 Misztal, I., S. Tsuruta, D. Lourenco, Y. Masuda, I. Aguilar *et al.*, 2016 Manual for

BLUPF90 family of programs. 125.

Morel, F., and C. Aninat, 2011 The glutathione transferase kappa family. Drug Metab. Rev. 43: 281–291.

Neto, R., G. M. Yoshida, J. Paul, L. José, and M. Yáñez, 2019 Genome-wide association analysis for body weight identifies candidate genes related to development and metabolism in rainbow trout (Oncorhynchus mykiss). Mol. Genet. Genomics.

Ødegård, J., T. Moen, N. Santi, S. A. Korsvoll, S. Kjøglum *et al.*, 2014 Genomic prediction in an admixed population of Atlantic salmon (Salmo salar). Front. Genet. 5: 1–8.

Ødegård, J., I. Olesen, B. Gjerde, and G. Klemetsdal, 2006 Evaluation of statistical models for genetic analysis of challenge test data on furunculosis resistance in Atlantic salmon (Salmo salar): Prediction of field survival. Aquaculture 259: 116–123.

Oruc, E.O., Y. Sevgiler, and N. Uner, 2004 Tissue-specific oxidative stress response in fish exposed to 2,4-D and azinphosmethyl. Comparative Biochem. Physiol. Part C, 137: 43-51.

Palaiokostas, C., S. Ferarreso, R. Franch, R. D. Houston, and L. Bargelloni, 2016 Genomic prediction of resistance to pasteurellosis in gilthead sea bream (Sparus aurata) using 2b-RAD sequencing. G3 Genes Genomes Genet. X: 1–8.

Palti, Y., 2011 Toll-like receptors in bony fish : From genomics to function. Dev. Comp. Immunol. 35: 1263–1272.

Palti, Y., G. Gao, S. Liu, M. P. Kent, S. Lien *et al.*, 2015b The development and characterization of a 57K single nucleotide polymorphism array for rainbow trout. Mol. Ecol. Resour. 15: 662–672.

Palti, Y., R. L. Vallejo, G. Gao, S. Liu, A. G. Hernandez *et al.*, 2015a Detection and Validation of QTL Affecting Bacterial Cold Water Disease Resistance in Rainbow

509 Trout Using Restriction-Site Associated DNA Sequencing. PLoS One 10: e0138435.

510 Piscatelli, H. L., M. Li, and D. Zhou, 2016 Dual 4- and 5-phosphatase activities regulate

511 SopB-dependent phosphoinositide dynamics to promote bacterial entry. Cell.

512 Microbiol. 18: 705–719.

513 Procko, E., G. Raghuraman, D. Wiley, M. Raghavan, and R. Gaudet, 2005 Identification of

514 domain boundaries within the N-termini of TAP1 and TAP2 and their importance in

515 tapasin binding and tapasin-mediated increase in peptide loading of MHC class I.

516 Immunol. Cell Biol. 83: 475–482.

517 Purcell, S., B. Neale, K. Todd-brown, L. Thomas, M. A. R. Ferreira *et al.*, 2007 PLINK : A

518 Tool Set for Whole-Genome Association and Population-Based Linkage Analyses.

519 Am. J. Hum. Genet. 81: 559–575.

520 Ramírez, R., F. A. Ǵomez, and S. H. Marshall, 2015 The infection process of Piscirickettsia

521 salmonis in fish macrophages is dependent upon interaction with host-cell clathrin and

522 actin. FEMS Microbiol. Lett. 362: 1–8.

523 Rise, M. L., S. R. M. Jones, G. D. Brown, K. R. von Schalburg, W. S. Davidson *et al.*, 2004

524 Microarray analyses identify molecular biomarkers of Atlantic salmon macrophage

525 and hematopoietic kidney response to Piscirickettsia salmonis infection. Physiol.

526 Genomics 20: 21–35.

527 Ritter, B., S. Murphy, H. Dokainish, M. Girard, M. V Gudheti *et al.*, 2013 NECAP 1

528 Regulates AP-2 Interactions to Control Vesicle Size , Number , and Cargo During

529 Clathrin-Mediated Endocytosis. PLoS One 11: e1001670.

530 Rodríguez, F. H., G. Cáceres, J. P. Lhorente, S. Newman, R. Bangera *et al.*, 2018 Genetic (

531 co) variation in skin pigmentation patterns and growth in rainbow trout. Animals 1–8.

532 Roher, N., A. Callol, J. V. Planas, F. W. Goetz and S. A. Mackenzie, 2011 Endotoxin

533 recognition in fish results in inflammatory cytokine secretion not gene expression.
534 Innate Immunity 17(1): 16-28.
535 Rozas, M., and R. Enríquez, 2014 Piscirickettsiosis and Piscirickettsia salmonis in fish: a
536 review. J. Fish Dis. 37: 163–188.
537 Rozas-Serri M., A. Peña, L. Maldonado, 2017 Transcriptomic profiles of post-smolt
538 Atlantic salmon challenged with Piscirickettsia salmonis reveal a strategy to evade the
539 adaptive immune response and modify cell-autonomous immunity. Developmental
540 and Comparative Immunology 81: 348-362.
541 Sernapesca, 2018 Informe Sanitario de Salmonicultura en Centros Marinos.
542 Siegel, R. M., F. K. Chan, H. J. Chun, and M. J. Lenardo, 2000 The multifaceted role of
543 Fas signaling in immune cell homeostasis and autoimmunity. Nat. Immunol. 1: 469–
544 474.
545 Tacchi, L., J. E. Bron, J. B. Taggart, C. J. Secombes, R. Bickerdike *et al.*, 2011 Multiple
546 tissue transcriptomic responses to Piscirickettsia salmonis in Atlantic salmon (Salmo
547 salar). Physiol. Genomics 43: 1241–1254.
548 Teitell, M. A., 2003 OCA-B regulation of B-cell development and function. Trends
549 Immunol. 24: 546–553.
550 Thornbrough, J. M., A. Gopinath, T. Hundley, and M. J. Worley, 2016 Human Genome-
551 Wide RNAi Screen for Host Factors That Facilitate Salmonella Invasion Reveals a
552 Role for Potassium Secretion in Promoting Internalization. PlosOne 11: e0166916.
553 Tsai, H.-Y., A. Hamilton, A. E. Tinch, D. R. Guy, J. E. Bron *et al.*, 2016 Genomic
554 prediction of host resistance to sea lice in farmed Atlantic salmon populations. Genet.
555 Sel. Evol. 48: 47.
556 Valencia-gallardo, C. M., N. Carayol, and G. T. Van Nhieu, 2015 Microreview

557 Cytoskeletal mechanics during Shigella invasion and dissemination in epithelial cells.
558 Cell. Microbiol. 17: 174–182.
559 Vallejo, R. L., T. D. Leeds, B. O. Fragomeni, G. Gao, A. G. Hernandez *et al.*, 2016
560 Evaluation of genome-enabled selection for bacterial cold water disease resistance
561 using progeny performance data in rainbow trout: Insights on genotyping methods and
562 genomic prediction models. Front. Genet. 7: 1–13.
563 Vallejo, R. L., T. D. Leeds, G. Gao, J. E. Parsons, K. E. Martin *et al.*, 2017a Genomic
564 selection models double the accuracy of predicted breeding values for bacterial cold
565 water disease resistance compared to a traditional pedigree-based model in rainbow
566 trout aquaculture. Genet. Sel. Evol. 49: 17.
567 Vallejo, R., S. Liu, G. Gao, B. O. Fragomeni, A. G. Hernandez *et al.*, 2017b Similar genetic
568 architecture with shared and unique quantitative trait loci for bacterial cold water
569 disease resistance in two rainbow trout breeding populations. Front. Genet. 8: 1–15.
570 Wang, T., J. W. Holland, N. Bols, and C. J. Secombes, 2005 Cloning and expression of the
571 first nonmammalian interleukin-11 gene in rainbow trout Oncorhynchus mykiss.
572 FEBS J. 272(5): 1136-47.
573 Wang, H., I. Misztal, I. Aguilar, A. Legarra, and W. M. Muir, 2012 Genome-wide
574 association mapping including phenotypes from relatives without genotypes. Genet.
575 Res. (Camb). 94: 73–83.
576 Yañez, J. M., R. Bangera, J. P. Lhorente, A. Barria, M. Oyarzun *et al.*, 2016 Negative
577 genetic correlation between resistance against Piscirickettsia salmonis and harvest
578 weight in coho salmon (Oncorhynchus kisutch). Aquaculture 459: 8–13.
579 Yañez, J. M., R. Bangera, J. P. Lhorente, M. Oyarzún, and R. Neira, 2013 Quantitative
580 genetic variation of resistance against Piscirickettsia salmonis in Atlantic salmon

(Salmo salar). Aquaculture 414–415: 155–159.

Yañez, J. M., R. D. Houston, and S. Newman, 2014 Genetics and genomics of disease resistance in salmonid species. Front. Genet. 5: 1–13.

Yañez, J. M., J. P. Lhorente, L. N. Bassini, M. Oyarzún, R. Neira *et al.*, 2014 Genetic co-variation between resistance against both Caligus rogercresseyi and Piscirickettsia salmonis, and body weight in Atlantic salmon (Salmo salar). Aquaculture 433: 295–298.

Yañez, J. M., and V. Martinez, 2010 Genetic factors involved in resistance to infectious diseases in salmonids and their application in breeding programmes. Arch. Med. Vet. 42: 1–13.

Yañez, J. M., S. Naswa, M. E. Lopez, L. Bassini, K. Correa *et al.*, 2016 Genomewide single nucleotide polymorphism discovery in Atlantic salmon (Salmo salar): validation in wild and farmed American and European populations. Mol. Ecol. Resour. 16: 1002–1011.

Yáñez, J., S. Newman, and R. D. Houston, 2015 Genomics in aquaculture to better understand species biology and accelerate genetic progress. Front. Genet. 6: 1–3.

Yoshida, G. M., R. Bangera, R. Carvalheiro, K. Correa, R. Figueroa *et al.*, 2018a Genomic prediction accuracy for resistance against Piscirickettsia salmonis in farmed rainbow trout. G3 Genes Genomes Genet. 8: 719–726.

Yoshida, G. M., R. Carvalheiro, F. H. Rodríguez, and J. P. Lhorente, 2018b Genomics Single-step genomic evaluation improves accuracy of breeding value predictions for resistance to infectious pancreatic necrosis virus in rainbow trout. Genomics 1–6.

Zheng, W., Y. Hu, M. Zhang, L. Sun, 2010 Analysis of the expression and antioxidative property of a peroxiredoxin 6 from *Scophthalmus maximus*. Fish & Shellfish

Immunology 29: 305-311.

Zuo, W., and P. T. K. Woo, 1997 Natural anti-proteases in rainbow trout, Oncorhynchus mykiss and brook charr, Salvelinus fontinalis and the in vitro neutralization of fish alpha 2-macroglobulin by the metalloprotease from the pathogenic haemoflagellate, Cryptobia salmositica. Parasitology 114: 375–81.

630 **Table 1.** Summary statistics for time to death (TD), binary survival (BS) and final weight
631 (FW) measured in 2,130 rainbow trout individuals.

Trait	Mean	SD	CV(%)	Min	Max
TD	23.26	7.86	33.27	10	32
BS	0.59	0.49	0.83	0	1
FW	173.80	52.27	30.07	46.10	448

632

633

634 **Table 2.** Genetic parameters and heritabilities for resistance to *Piscirickettsia salmonis* as
635 time to death (TD) and binary survival (BS).

Trait	$\sigma_a^{2\,a}$	$\sigma_e^{2\,b}$	h^2(SD)[c]
TD	25.95	28.92	0.48(0.04)
BS	6.27×10^{-2}	1.21×10^{-1}	0.34(0.04)

636 [a] Additive genetic variance

637 [b] Residual variance

638 [c] Heritability and standard deviation

639

640

641

642

643

644

645

646 **Table 3.** Top markers associated with *Piscirickettsia salmonis* resistance defined as TD and

647 BS in rainbow trout, using ssGWAS,

Ranking	Name	Chr[a]	Pos (Bp)	PGV[b]	Genes[c]
Time to death					
1	Affx-88923370	27	9998276	2.43	*usp2, nlrc3, tap, pitpna*
2	Affx-88916453	3	14818380	1.41	*stl2, aicda, il11, gstk1*
3	Affx-88922612	14	10975036	1.21	*tlr4, tax1bp1, satb1*
4	Affx-88927397	24	11828385	1.02	*a2m, pou2af1, nfkbiz*
Binary survival					
1	Affx-88923370	27	9998276	1.50	*usp2, nlrc3, tap, pitpna*
2	Affx-88951679	5	68055053	1.12	*faslg, prdx6, plpp6*
3	Affx-88908715	29	32519588	1.01	*mapk12, gls*

648 [a] Chromosome.

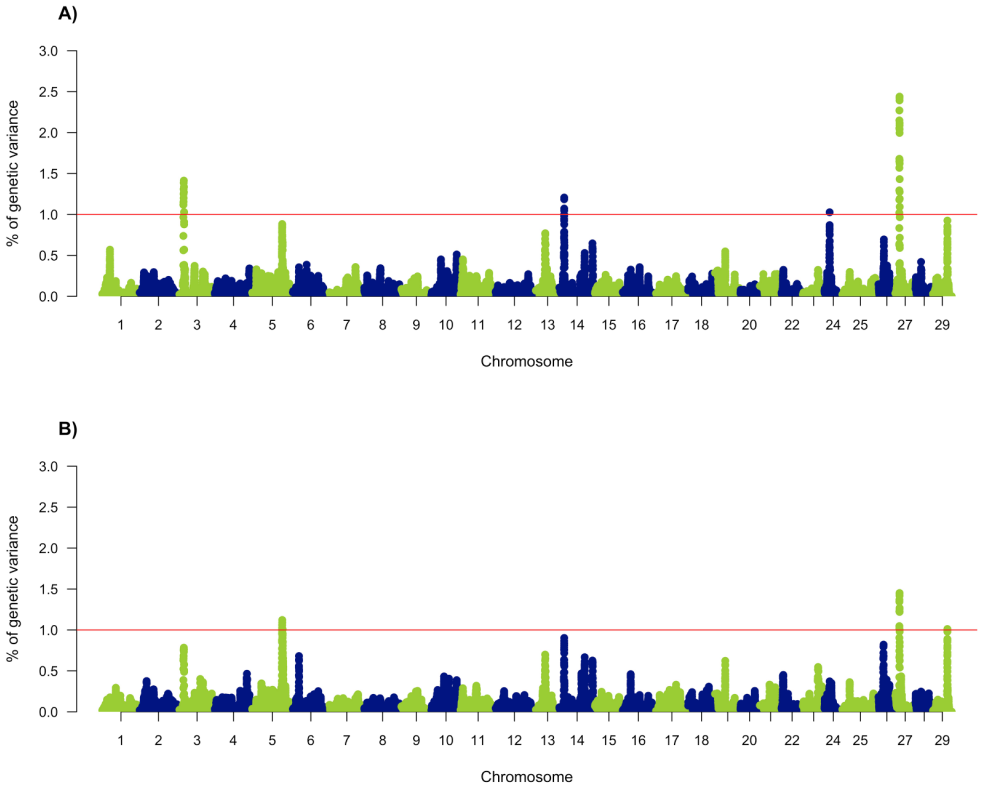
649 [b] Percentage of genetic variance.

650 [c] Summary of the genes located within 1Mb window.

651

652

653



654

655 **Figure 1.** Genomic association analysis for resistance to *Piscirickettsia salmonis* in

656 rainbow trout (*Oncorhynchus mykiss*). Resistance was defined as time to death (A) and as

657 binary survival (B).