1 **Genome-wide association study and genomic predictions for resistance against**

2 ***Piscirickettsia salmonis* in coho salmon (*Oncorhynchus kisutch*) using ddRAD**

3 **sequencing**

4

5 Agustín Barría[*,†], Kris A. Christensen[‡,1], Katharina Correa[*,§], Ana Jedlicki[*], Jean P.

6 Lhorente[§], William S. Davidson[‡], José M. Yáñez[*,§]

7

8 [*] Facultad de Ciencias Veterinarias y Pecuarias, Universidad de Chile, 8820808, La

9 Pintana, Santiago, Chile

10 [†] Doctorado en Acuicultura, Programa Cooperativo Universidad de Chile, Universidad

11 Católica del Norte, Pontificia Universidad Católica de Valparaíso, Chile.

12 [‡] Department of Molecular Biology and Biochemistry, Simon Fraser University, V5A

13 1S6, Burnaby, BC, Canada

14 [§] Aquainnovo S.A., 5503032, Puerto Montt, Chile

15 ORCID IDs: 0000-0002-2813-4559 (A.B.); 0000-0003-0021-0326 (Kr.C.); 0000-0003-

16 1386-8522 (K.C.); 0000-0003-0626-1247 (A.J.); 0000-0002-9157-4231 (JP.L.); 0000-

17 0002-3002-3225 (W.D.); 0000-0002-6612-4087 (JM.Y.)

18

19 [1] Present Address: Marine Ecosystems and Aquaculture Division, Science Branch

20 Fisheries and Oceans Canada / Government of Canada 4160 Marine Drive, West

21 Vancouver, BC, Canada V7V 1N6

22

23

1

24 **Genomic study for *Piscirickettsia salmonis* in coho salmon**

25

28 Corresponding author: *Facultad de Ciencias Agronómicas, Universidad de Chile,*
29 Avenida Santa Rosa 11735*, 8820808, La Pintana, Santiago, Chile*

30 E-mail: jmayanez@uchile.cl

31

32

33

34

35

36

37

38

39

40

41

42

43

2

44   ABSTRACT

45   *Piscirickettsia salmonis* is one of the main infectious diseases affecting coho salmon

46   (*Oncorhynchus kisutch*) farming. Current treatments have been ineffective for the control

47   of the disease. Genetic improvement for *P. salmonis* resistance has been proposed as a

48   feasible alternative for the control of this infectious disease in farmed fish. Genotyping

49   by sequencing (GBS) strategies allow genotyping hundreds of individuals with thousands

50   of single nucleotide polymorphisms (SNPs), which can be used to perform genome wide

51   association studies (GWAS) and predict genetic values using genome-wide information.

52   We used double-digest restriction-site associated DNA (ddRAD) sequencing to dissect

53   the genetic architecture of resistance against *P. salmonis* in a farmed coho salmon

54   population and identify molecular markers associated with the trait. We also evaluated

55   genomic selection (GS) models in order to determine the potential to accelerate the

56   genetic improvement of this trait by means of using genome-wide molecular information.

57   764 individuals from 33 full-sib families (17 highly resistant and 16 highly susceptible)

58   which were experimentally challenged against *P. salmonis* were sequenced using ddRAD

59   sequencing. A total of 4,174 SNP markers were identified in the population. These

60   markers were used to perform a GWAS and testing genomic selection models. One SNP

61   related with iron availability was genome-wide significantly associated with resistance to

62   *P. salmonis* defined as day of death. Genomic selection models showed similar accuracies

63   and predictive abilities than traditional pedigree-based best linear unbiased prediction

64   (PBLUP) method.

65

66

67

68

3

69   INTRODUCTION

70   Chile is the largest producer of coho salmon (*Oncorhynchus kisutch*) globally,

71   reaching about 160,000 tons in 2014, representing more than 90% of total production

72   (FAO 2016).   However, the success and sustainability of this industry is constantly

73   threatened by infectious diseases, including Salmon Rickettsial Syndrome (SRS). This

74   disease is caused by *Piscirickettsia salmonis*, a gram-negative and facultative intracellular

75   bacteria which was isolated for the first time in Chile in coho salmon (Cvitanich *et al.*

76   1991). Data from the Chilean National Fisheries and Aquaculture Service (Sernapesca)

77   estimates that 59.3% of the morality rates ascribed to infectious disease were associated

78   with SRS (Sernapesca 2016a). To date, control measures and treatments for SRS are

79   based on antibiotics and vaccines. However, both strategies have not had the expected

80   effectiveness under field conditions. Because of this, it is necessary to develop alternative

81   strategies for the control of this disease (Yáñez *et al.* 2014a). In this regard, breeding for

82   enhanced disease resistance is a feasible and sustainable option to improve animal health,

83   welfare and productivity (Stear *et al.* 2001). A primary requisite for including disease

84   resistance into a breeding program is the presence of significant additive genetic variation

85   for the trait (Falconer and Mackay 1996). Commonly, data to evaluate resistance comes

86   from experimental challenges carried out using siblings of the selection candidates

87   (Ødegård *et al.* 2011;Yañez and Martinez 2010; Yáñez *et al.* 2014a). Quantitative studies

88   have estimated significant genetic variation for resistance against different pathogens in

89   salmonid species (Ødegård *et al.* 2011; Yáñez *et al.* 2014a). For instance, low to moderate

90   heritabilities for resistance against *P. salmonis* in Atlantic salmon (*Salmo salar*) ($h^2 =$

91   0.11 to 0.41) (Yáñez *et al.* 2013; Yáñez *et al.* 2014b) and coho salmon ($h^2 = 0.16$) (Yañez

92   *et al.* 2016a) have been estimated.

       4

93   Marker assisted selection (MAS) can improve production traits in cases where the

94   phenotypes are difficult to measure in the selection candidates (e.g. disease resistance

95   traits) and the total additive genetic variance explained by genetic markers is high (Hayes

96   and Goddard 2010). This methodology has been successfully applied for the improvement

97   of resistance against infectious pancreatic necrosis in Atlantic salmon, which is controlled

98   by a major quantitative trait locus (QTL) (Houston *et al.* 2012; Moen *et al.* 2015). In the

99   case of polygenic traits, genomic selection (GS) (Meuwissen *et al.* 2001) can significantly

100   improve selection accuracy of breeding values compared to traditional selection, and

101   therefore enhance the response to selection for disease resistance in salmonid species

102   (Tsai *et al.* 2016; Vallejo *et al.* 2016, 2017; Bangera *et al.* 2017; Correa *et al.* 2017).

103   Genotyping by sequencing (GBS) is an alternative for genotyping in cases when

104   SNP panels are not available. This approach reduces the complexity of the genome, and

105   can be used to identify thousands of markers without prior marker discovery efforts or a

106   reference genome. Currently, several approaches of GBS have been developed,

107   significantly reducing the cost and labor (Baird *et al.* 2008; Elshire *et al.* 2011; Peterson

108   *et al.* 2012). These methodologies have been widely used in salmonid species, either to

109   generate high density linkage maps (Brieuc *et al.* 2014; Gonen *et al.* 2014) or to perform

110   association studies to identify genomic regions involved in the resistance against

111   pathogens (Campbell *et al.* 2014; Palti *et al.* 2015b).

112   Double-digest restriction-site associated DNA (ddRAD) reduces DNA

113   complexity by digesting DNA with two restriction enzymes (REs) simultaneously, and

114   avoiding random shearing (Peterson *et al.* 2012). This approach has been widely used in

115   genetic studies in aquaculture species (Robledo *et al.* 2017). In the present study, we used

116   ddRAD sequencing to dissect the genetic architecture of resistance against *P. salmonis* in

117   a farmed coho salmon population and identify molecular markers associated with the trait.

5

118   Furthermore, GS models were used to evaluate the potential to accelerate the genetic

119   improvement of resistance against *P. salmonis* in this coho salmon breeding population

120   by means of using genome-wide molecular information.

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

6

138   MATERIALS AND METHODS

139   **Coho Salmon Sampling**

140   The coho salmon used in the present study were from the 2012 year-class of a

141   genetic improvement program established in 1997 (owned by Pesquera Antares) and

142   managed by Aquainnovo (Puerto Montt, Chile). Further details about this breeding

143   population, in terms of reproductive management, rearing conditions, fish tagging, and

144   breeding objectives are described by Yáñez *et al.* (2014c; 2016) and Dufflocq *et al.*

145   (2016).

146

147   **Experimental Challenge**

148   The experimental challenge is described in details by (Correa *et al.* 2015a) and

149   Yáñez et al. (2016). Briefly, 2,607 individuals belonging to 108 maternal full-sib families

150   (60 paternal half-sib families), were challenged against *Piscirickettsia salmonis*. Prior the

151   challenge experiment, each fish was marked with a passive integrated transponder (PIT-

152   tag), placed in the abdominal cavity for genealogy traceability during the challenge test.

153   The experimental challenge test was performed at Aquainnovo´s Research Station,

154   located in Lenca River, Xth Region, Chile. For the lethal dose 50 (LD_{50}) calculation, a

155   random sample of 80 fish were selected from the population. Four different dilutions from

156   the *P. salmonis* inoculum were evaluated (1/10, 1/100, 1/1000 and 1/10000). Twenty fish

157   per dilution were intraperitoneally (IP) injected with 0.2ml/fish. Daily mortality was

158   recorded. This preliminary test spanned 26 days and a dilution of 1:680 was estimated as

159   the LD_{50}.

160   For the definitive challenge, fish were distributed into three tanks ($7m^3$) with a salt

161   water concentration of 31 ppt. An average of eight individuals (ranging from 1 to 18)

162   from each of the 108 families were distributed into each tank.

163   The experimental challenge was performed through an intraperitoneal (IP)

164   injection with 0.2ml/fish of the LD_{50} inoculum. The average weight of the fish at the

165   inoculation was 279g (SD = 138g). Additionally, qRT-PCR was previously performed

166   in order to control for the presence of Infectious Salmon Anemia Virus (ISAV), Infectious

167   Pancreatic Necrosis Virus (IPNV) and *Flavobacterium spp.*

168   The challenge test was continued up to 50 days post injection. Throughout the

169   challenge, environmental parameters (pH, temperature, salinity and oxygen) were

170   measured and controlled. Daily mortality was removed from each tank, and a sample from

171   anterior kidney was taken and stored at -80ºC in RNALater. A necropsy assay was

172   performed in conjunction with qRT-PCR to confirm the cause of death and the presence

173   of *P. salmonis.* This was also done to control for the presence of other pathogens, such as

174   *Vibrio ordalii*, *Renibacterium salmoninarum* and IPNV.

175

176   **Trait definitions and quantitative analysis**

177   Resistance was defined as the day of death (DD) with values ranging from 1 to 50

178   depending on the time of death. Additionally, resistance was evaluated as a binary (BIN)

179   trait, either dead or alive at the end of the challenge. Values for this trait were 1 in cases

180   where the fish died during the challenge, or 0 if the fish survived until the end of the

181   challenge. Initial Body Weight (IW) for each fish, was measured prior to the IP injection.

182   A linear bivariate animal model was used in order to estimate covariance components for

183   DD and BIN. Sex and tank were used as fixed effects, and IW was included as a covariate

8

184   in the model. The animal model was fitted using ASREML (Gilmour *et al.* 2009).

185   Heritability estimation was computed as follow:

186
$$h_i^2 = \frac{\sigma_{Gi}^2}{\sigma_{Gi}^2 + \sigma_{Ei}^2}$$

187   Where σ_{Gi}^2 and σ_{Ei}^2 are the additive genetic and residual variances respectively.

188   The genetic correlation (r_{xy}) between both traits was defined as Falconer and Mackay

189   (1996):

190
$$r_{xy} = \frac{\sigma_{ax,ay}}{\sqrt{\sigma_{ax}^2 \sigma_{ay}^2}}$$

191   where $\sigma_{ax,ay}$ is the additive genetic covariance between traits, σ_{ax}^2 is the additive genetic

192   variance of DD, and σ_{ay}^2 is the additive genetic variance of BIN.

193

194   **ddRAD library preparation and sequencing**

195   Ten ddRAD libraries were produced by multiplexing 828 individuals following

196   the protocol described by Peterson et al. (2012). For this, 64 parents (males and females)

197   and 764 offspring representing the 17 most resistant and 16 most susceptible families

198   were selected. An average of 23 (ranging from 11 to 43 individuals) offspring per family

199   were chosen. Briefly, total DNA was extracted using the commercial kit Wizard SV

200   Genomic DNA purification System (Promega) according to manufacturer's protocol.

201   Between 80 and 200 ng of DNA, from each individual was digested with two restriction

202   enzymes (New England Biolabs, UK; NEB); 10 U of S*bf*I (specific for the CCTGCA|GG

203   motif)) and *Mse*I (specific for the T|TAA motif) in 12 µl reaction volume, including 1 µl

204   of S*bf*I and *Mse*I adapter (8.3 pM), for 90 min at 37°. The ligation reaction was carried

9

205   out by adding 1 µl of T4 ligase (NEB) diluted 1:100 in T4 buffer and incubating for 150

206   min at 37° and subsequently at 16° overnight.

207   Each ligation mix was diluted with 189 µl of dilute TE buffer (1:10). Kodak DNA

208   Polymerase (ABM), a high fidelity polymerase, was used to amplify DNA fragments with

209   the correct adapters. PCR reactions (20 µl) were prepared containing 10 µl of PCR mix

210   2x, 1 µl of primer mix (10 µM each), 6 µl of diluted ligation mix and 3 µl of nuclease-

211   free water. Each sample was PCR amplified using the following conditions: 95° for 2

212   min, followed by 17 cycles of 95° for 20s, 66° for 30s and 68° for 40s. After PCR,

213   amplicon quality was checked by loading 5 µl on a 2% agarose gel. Subsequently,

214   samples were pooled, so that the final concentration was similar among them within each

215   library. Each library was concentrated through an evaporation step for 80 min in a

216   Centrivap Mobile Console Centrifugal Evaporator (Labconco). This step was conducted

217   until 300 µl of the generated library was obtained. Final volume of each library was

218   loaded on a 1% agarose gel. Size of the bands selected for sequencing ranged from 750

219   and 1500 bp and between 1,800 and 2,500 bp. DNA was purified through the QIAquick

220   gel extraction kit (Qiagen) following manufacturer's instructions. Finally, libraries were

221   sequenced on an Illumina Hiseq2500 platform, using 150 base single-end sequencing.

222

SNP identification

223   **SNP identification**

224   Raw sequence reads obtained from sequencing Illumina were analyzed using STACKS

225   v. 1.41 (Catchen *et al.* 2011, 2013). This software was specifically developed to analyze

226   short-read data generated through next generation sequencing (NGS) (Davey *et al.* 2013).

227   Sample reads were trimmed to 134 bp for all subsequent analyses. Additionally, these

228   reads were demultiplexed and filtered using *process_radtags*. Rad-tags which passed the

10

229 quality filter were aligned to the *Oncorhynchus kisutch* reference genome (GenBank:

230 MPKV00000000.1) using BWA v. 0.7.12 (Li and Durbin 2009). The reference genome

231 was indexed and alignments were performed using the *mem* algorithm, all other

232 parameters were set as default. Loci were then built using *pstacks* allowing a minimum

233 depth of coverage of three to build a locus (-m 3). A catalog of loci was constructed

234 through *cstacks* program using only the parents' reads. To build the catalog, the maximum

235 number of mismatches allowed between sample tags was three (-n 3), and the matching

236 was based on genomic location (g). After this, the *sstacks* program was used in order to

237 match rad-tags against the catalog based again on genomic location (g), followed by

238 *populations* software, using defaults parameters. Loci were considered as valid if they are

239 present, in at least, 75% of the individuals of the population.

240

241 **Genomic Association Study**

242 In order to associate the molecular markers to *P. salmonis* resistance, either as DD or

243 BIN, a GWAS was performed using the GenABEL library (Aulchenko *et al.* 2007)

244 implemented in R (http://www.r-project.org). Prior to analyses, identified SNPs were

245 submitted to a Quality Control (QC), with the following parameters: Minor Allele

246 Frequency (MAF) ≥ 0.01, Hardy-Weinberg Equilibrium (HWE) = $p < 1\mathrm{x}10^{-6}$, individual

247 call rate > 0.70 and SNP call rate > 0.90. Additionally, all markers were used in order to

248 estimate identity by state (IBS). Association analysis was performed through the

249 Grammar-Gamma method (Svishcheva *et al.* 2012) including a genomic kinship matrix

250 estimated using filtered SNPs. The polygenic function (Thompson and Shaw 1990) was

251 used to fit two different univariate additive polygenic models. The linear regression

252 model was defined as follows:

11

253 $$y_i = \beta_0 + \beta_1 SNP + \beta_2 Sex + \beta_3 Tank + \beta_4 IW + e_i$$

254 Where y_i is the vector of phenotypes (DD), β_0 is the mean, β_1 is the effect of each SNP,

255 β_2 is the fixed effect of the sex, β_3 is the fixed effect of the tank, β_4 is the effect of the

256 initial weight as covariate and e_i is the random residual.

257 The logistic regression models for BIN was:

258 $$P(y = 1) = \frac{e^{\beta_0 + \beta_1 SNP + \beta_2 Sex + \beta_3 Tank + \beta_4 IW}}{1 + e^{\beta_0 + \beta_1 SNP + \beta_2 Sex + \beta_3 Tank + \beta_4 IW}}$$

259 Where P is the probability of the random variable to be one, β_0 is the mean, β_1 is the

260 effect of each SNP, β_2 is the fixed effect of the sex, β_3 is the fixed effect of the tank, β_4

261 is the effect of the initial weight as covariate and e_i is the random residual.

262 Genomic-wide significance was assessed by False Discovery Rate (FDR) (Benjamini and

263 Hochberg 1995). Briefly, p values were ordered from $p_1 \leq p_2 \leq \ldots \leq p_k$ where k is the

264 number of SNPs tested. Beginning from the largest p value, the first p value (p_i) that

265 satisfied: $p_i \leq /k*0.05$, where i = ith observation. Finally, the p_i that satisfied the condition

266 became the significant value.

267 Finally, *BLASTn* was used to align the genome-wide associated region, against Atlantic

268 salmon genome (Genbank: GCA_000233375.4) and identify candidate genes associated

269 with *P. salmonis* resistance.

270

271 **Genomic Prediction**

272 The pedigree-based approach, PBLUP was used as the control for the genomic

273 evaluations, and EBV for each individual were estimated using a linear mixed model

12

274   implemented in BLUPF90 (Misztal *et al*. 2016). The model used was as follows:

$$y = X\beta + Tg + e$$

276   where y is a vector of phenotypes (BIN or DD), β is a vector of fixed effects (Sex, Tank

277   and initial weight effects), g is a vector of random additive polygenic genetic effects that

278   follows a normal distribution $\sim N(0, A\sigma_g^2)$, X and T are incidence matrices, A is the

279   additive relationship matrix, and e is the residual (Lynch and Walsh 1998).

280   The genomic EBV (GEBV) for each individual were estimated using GBLUP (*Genomic*

281   *BLUP*) as implemented in the BLUPF90 software. GBLUP is a modification of the

282   PBLUP method, where the numerator relationship matrix A is replaced by a genomic

283   relationship matrix G, as described by VanRaden (2008). Pedigree (PBLUP) and genomic

284   (GBLUP) heritabilities were calculated using the AIREMLF90 software (Misztal *et al.*

285   2016) as previously described.

286   The different models were compared using a five-fold cross validation scheme. Briefly,

287   all genotyped and phenotyped animals were randomly separated into five validations sets,

288   which were predicted one at a time by masking their phenotypes and using the remaining

289   animals as a training set to estimate the marker effects. Thus, for each validation run, the

290   dataset was split into a training set (80%) and a validation set (20%). To reduce the

291   stochastic effects, this cross-validation analysis was replicated 10 times. Accuracy was

292   used to assess the performance of each model and was estimated as:

$$R_{EBV,BV} = \frac{R_{EBV,y}}{h},$$

294   where $R_{EBV,y}$ is the correlation between the EBV of a given model (predicted for the

295   validation set using information from the training set) and the actual phenotype, while h

13

296   is the square root of the pedigree-based estimate of heritability (Legarra *et al.* 2008;

297   Ødegård *et al.* 2014). To test prediction accuracies obtained using different SNP densities,

298   1 K, 2 K, 3 K and 4 K SNPs were randomly selected from the SNPs that passed the quality

299   control. Finally, accuracies were calculated for each model and SNP density, and

300   compared to those obtained with the PBLUP model.

301

302   **Data availability**

303   Table S1 contains genotypic data. Data of pedigree and phenotypes that supports the

304   findings in this study are available from Aquainnovo and AquaChile but restrictions apply

305   to the availability of these data, which were used under license for the current study, and

306   so are not publicly available. However, data are available from the corresponding author

307   upon reasonable request and with permission of Aquainnovo and AquaChile.

308

309

310

311

312

313

314

315

316

317

318

319

14

320   RESULTS

321   Challenge test

322   Mortality began on day 10 post challenge, with evident clinical signs and pathological

323   lesions typical of the Salmon Rickettsial Syndrome. These signs include swollen kidney,

324   splenomegaly and liver with a yellowish tone (Rozas and Enríquez 2014). Challenged

325   families showed considerable phenotypic variation for *P. salmonis* resistance. Average

326   mortality of all families reached 41% during the 50-day challenge. Average cumulative

327   mortality rate among the 17 best and 16 worst families selected for genotyping, reached

328   19% and 63%, respectively (Figure 1).

329

330   Heritabilities and correlations between traits

331   Moderate significant additive genetic variation was estimated for both DD and BIN.

332   Using a pedigree relationship matrix, h^2 was estimated as 0.26 ($\pm$ 0.01) and 0.42 ($\pm$ 0.01)

333   for DD and BIN, respectively. In the case of using a G matrix, the h^2 was estimated to be

334   0.32 and 0.38 for DD and BIN, respectively. Genetic correlation between traits was very

335   high and significantly different from zero (-0.95 $\pm$ 0.03). Additionally, high significant

336   phenotypic correlation was estimated, being -0.77 ($\pm$ 0.01).

337

338   ddRAD sequencing

339   Prior to quality control (QC), per base quality (Phred score) was evaluated. The average

340   quality score ranged from 36 to 38 among libraries, indicating high quality of data.

341   Illumina sequencing, including parents, yielded an average of 156,058,078 ($\pm$ 16

342   millions) of raw sequences. After QC, which included the removal of low quality

15

343   sequences and reads with either missing or ambiguous barcodes, an average of

344   31,660,024 of the reads were removed. In parallel with QC, reads were trimmed to 134

345   bp. In this regard, 79% of the raw reads were retained for subsequent analysis. In order

346   to create a set of all possible alleles in the population, data sets of the parental samples

347   were used to create the STACKS catalogue. This catalogue consisted of 106,309 unique

348   ddRAD loci from which 55,770 markers from 757 individuals were obtained.

349

350   **Association Analysis**

351   A total of 4,174 markers and 592 animals passed all QC criteria (see Table S1).

352   Association analysis identified one marker showing genome-wide statistical significance

353   (p value = 5.50E-05) (Fig. 2) for DD. This marker was located on scaffold01025.

354   However, none of the markers were significant using the FDR significance threshold

355   when resistance was defined as a binary trait (Fig. 3). For this trait only one marker was

356   identified as suggestively associated (p value = 1.50E-05). This marker was located in

357   Chromosome 29. BLAST results determined that the genomic region surrounding the

358   significant SNP associated with DD, was located on chromosome Ssa03, proximate to a

359   *heme oxygenase 2-like* gen. Location, proportion of heritability and phenotypic variance

360   explained, and p-values for the significant and suggestive marker are summarized in table

361   1.

362

363

364

365

    16

Genomic Prediction

Prediction was assessed using phenotype and genotype information from 4,174 SNPs which passed all QC steps from 592 individuals. These individuals were randomly split into five different training (474) and validation (118) sets, either for DD or BIN. PBLUP accuracy reached 0.58 and 0.61 for DD and BIN, respectively. These values were similar to those obtained using genomic information, ranging from 0.57-0.60 for DD and 0.60-0.62 for BIN at different SNPs densities (Table 2). Moreover, using only phenotypic information, reliability was also higher for BIN, reaching 0.38 in contrast to 0.36 for DD. When genomic data was used, reliability values were similar to those obtained with PBLUP ranging from 0.34 to 0.39. Finally, predictive ability reached up to 38% in the case of DD, showing an increase around 15% in the case of BIN, independently of the use of genotype information.

17

388   DISCUSSION

389   *Piscirickettsia salmonis* is the etiological agent of SRS, the main disease affecting

390   salmon production in Chile. Currently, control strategies have not shown the expected

391   effectiveness in field conditions (Rozas and Enríquez 2014). Some strains of *P. salmonis*

392   have presented resistance to Quinolones and oxytetracycline and intermediate

393   susceptibility to florfenicol (Henríquez *et al.* 2016), which are some of the most used

394   antibiotics in the Chilean industry for the treatment of this infectious disease (nearly 90%

395   of the total amount) (Sernapesca 2016b).

396   Selective breeding has been successfully applied commercially for important

397   production traits of aquaculture species, including salmonids (Gjedrem 2012; Yáñez *et*

398   *al.* 2014a). Genetic improvement has included disease resistance, based on information

399   of relatives, but does not allow for the use of within-family genetic variation and

400   consequently affects the achievable genetic progress due to the lower accuracy on EBVs

401   (Falconer and Mackay 1996). GS has had a major impact on traits difficult to measure on

402   the selection candidates themselves, as is the case of disease resistance, and has improved

403   genetic gain and selection accuracy (Goddard and Hayes 2009). In this sense, breeding

404   schemes can incorporate genomic information, either through GS or via MAS, in order

405   to accelerate the genetic progress for disease resistance traits (Yáñez *et al.* 2015).

406   Significant genetic variation for *P. salmonis* resistance was detected in the present

407   study. Moderate heritabilities were estimated using different trait definitions, either for

408   DD or BIN using both A and G matrices. Estimated heritabilities were higher for

409   resistance as binary trait when compared to DD, independent of the matrix used in the

410   analysis. Additionally, high genetic correlation between traits was estimated (-0.95 ±

411   0.03), suggesting that DD and BIN, are likely measures of the same trait.

        18

412   Previously, Yáñez (2016), estimated a heritability of 0.16 for resistance against *P.*

413   *salmonis*, defined as day of death, in the same coho salmon population, through a linear

414   model. However, our estimations used a directed sample of the population, which

415   explains the difference between estimations. These results are in agreement with previous

416   findings which estimated heritabilities for resistance against *P. salmonis* in Atlantic

417   salmon ranging from 0.11 to 0.41 (Yáñez *et al.* 2013, 2014b; Correa *et al.* 2015b). The

418   genetic variation for *P. salmonis* resistance, and values of heritability, are in accordance

419   with different studies that also found significant genetic resistance against other bacterial

420   disease in salmonid species (Gjøen *et al.* 1997; Ødegård *et al.* 2006; Vallejo *et al.* 2016).

421   Association analysis identified one molecular marker significantly associated with

422   *P. salmonis* resistance defined as day of death. The reference genome of coho salmon is

423   not completely assembled into chromosomes (~75% anchored), and it was not possible

424   to elucidate to which chromosome this marker belongs. However, the availability of a

425   high quality *Salmo salar* genome reference (Lien *et al.* 2016) made it possible to identify

426   the marker on a genomic region of chromosome Ssa03 proximate to *Hem Oxygenase-2*

427   (HO2), an enzyme involved in the iron accessibility through the degradation of a heme

428   group (Kikuchi *et al.* 2005). Different studies have already demonstrated the ability of *P.*

429   *salmonis* to evade the immune system and infect, survive and replicate inside

430   macrophages (McCarthy *et al.* 2008; Rojas *et al.* 2009). Transcriptional studies performed

431   by Rise *et al.* (2004) showed that after *P. salmonis* infection, ferritin and transferrin, both

432   related with binding and the storing of iron, were up regulated in Atlantic salmon

433   macrophages. Similar results were obtained by Pulgar *et al.* (2015) who showed that after

434   an IP challenge against *P. salmonis*, genes related to iron homeostasis were differentially

435   up-regulated in head kidney of resistant families of Atlantic salmon compared with the

436   susceptible ones. Moreover, head-kidney iron content was significantly lower in resistant

19

437   families, which was positive correlated with *P. salmonis* load. This evidence makes it

438   possible to suggest that the deprivation of the intracellular iron content is a key immune

439   response mechanism against *P. salmonis* infection, and may reduce its replication rate. A

440   similar mechanism was previously demonstrated to reduce bacterial growth inside

441   murine macrophages for a variety of bacterial infections (Paradkar *et al.* 2011). *Hem*

442   *oxygenase 1* has been shown to be up-regulated 2 hr after an infection generated by the

443   gram negative bacteria, *Salmonella typhimurium* (Nairz *et al.* 2007). This gene was able

444   to reduce oxidative stress, has been involved in murine-macrophages iron export and a

445   subsequent intracellular iron deprivation (making iron inaccessible to bacteria) (Nairz *et*

446   *al.* 2007). These results support the potential role that *HO2* could be playing in the

447   resistance against *P. salmonis* in coho salmon (i.e. limiting bacteria replication rate

448   through iron deprivation at the intra-cellular level). However, further studies need to be

449   performed in order to have a better understanding of the host immune response against

450   the infection and confirm the role of *HO2* in the variation of *P. salmonis* resistance.

451   Previously, resistance against bacterial diseases has been suggested as a polygenic

452   trait in aquaculture species. For example, Palaiokostas *(*2016) suggested resistance

453   against *Photobacterium damselae subsp.* had a polygenic genetic architecture in Gilthead

454   Sea Bream (*Sparus aurata*). Using a 50K SNP genotyping array, it was possible to

455   elucidate a moderately polygenic architecture for resistance against *P. salmonis* in

456   Atlantic salmon (Correa *et al.* 2015b). In the current study, and using 4 K SNPs, a

457   moderate polygenic genetic architecture for resistance against *P. salmonis* in coho salmon

458   population is suggested. However, due to the low marker density, it might be possible

459   that this density is not enough to capture linkage disequilibrium (LD) between genetic

460   markers and all the important QTLs controlling *P. salmonis* resistance in coho salmon.

20

461   Currently, GS has been demonstrated to be a feasible alternative over traditional

462   breeding methods to increase response to selection due to higher accuracy breeding

463   values, in both, simulated and empirical studies in aquaculture species (Sonesson and

464   Meuwissen 2009; Tsai *et al.* 2015, 2016, Vallejo *et al.* 2016, 2017; Bangera *et al.* 2017;

465   Correa *et al.* 2017).

466   The current study, showed similar accuracies, reliabilities and predictive ability

467   (PA) when comparing PBLUP and GBLUP (even among different SNPs densities). In

468   this regard, estimated accuracies with PBLUP were 0.58 and 0.60, which are highly

469   similar to those obtained using genomic data, from 0.56 to 0.62 (independent of the

470   number of genetic markers used). Similar tendencies were obtained for reliability and PA,

471   where the average values were identical between PBLUP and GBLUP.

472   When resistance was defined as a binary trait, both accuracy and reliability, were

473   slightly higher in comparison with resistance as day of death. However, PA was almost

474   15% higher for BIN than DD, either with or without genomic data, reaching up to 0.44.

475   These results, are similar to those obtained in GS for Bacterial Cold Water Disease

476   (BCWD), in rainbow trout (*Oncorhynchus mykiss*), using SNP arrays and RAD seq

477   (Vallejo *et al.* 2016), obtaining similar PA values with or without genotypic information,

478   ranging from 0.37-0.50 and from 0.26-0.41 for DD or BIN, respectively. Vallejo *et al.*

479   (2016) attributed this results to the small training sample size of their study (n=583),

480   which is higher than that used in the current study (n=473). In this regard, PA increase

481   reached up to 108% relative to PBLUP when the number of phenotyped and genotyped

482   individuals increase, even when less SNPs were used (Vallejo *et al.* 2017). Interestingly,

483   estimated PA for BCWD resistance are similar using 10 K SNPs obtained by RAD as

484   those obtained using a 40 K SNP array, even if resistance is defined as day of death or as

21

485   binary. Authors attributed this to the long-range LD within the population (Vallejo *et al.*

486   2016).

487   In case of Atlantic salmon, improvement in accuracy raised 20% using 112 K

488   SNPs either for length and weight in juvenile individuals (Tsai *et al.* 2015). Moreover,

489   sea lice resistance has showed a relative improvement in accuracy of 22 and 27%, relative

490   to PBLUP, using 37 K and 33 K SNPs respectively. Additionally, improvement in

491   reliability reached up to 52% with 220 K (Ødegård *et al.* 2014; Tsai *et al.* 2016; Correa

492   *et al.* 2017). In case of *P. salmonis* resistance, relative reliability has been improved in 25

493   and 30% either for resistance, defined as day of death or as a binary trait respectively

494   (Bangera *et al.* 2017).

495   Using 2b-RAD methodology, Palaiokostas (2016) showed a reduction in the

496   improvement of accuracies, for resistance against a bacterial disease in *S. aurata*, when

497   lower SNPs densities were evaluated. In this regard, an improvement of 53% in the

498   estimated accuracy was observed using 12 K molecular markers, raising it up to 0.46 in

499   comparison with 0.30 obtained through PBLUP through a Bayesian model. However,

500   using 2 K SNPs, accuracy showed a maximum increase of 20%. However, this

501   improvement was completely abolished when the number of markers was reduced to 700.

502   We hypothesize that the similar results with or without genomic information in

503   the current study, could be due to the small training sample size and the low density of

504   SNPs markers used, making it difficult to capture the LD between markers and all of the

505   loci influencing this trait. Moreover, simulations performed by Perez Enciso (2015),

506   suggest that even using a higher number of molecular markers obtained by RAD seq,

507   accuracies are lower than using a medium density SNP array, likely due to that in the

508   RAD seq methodology; SNPs are tightly linked among them, whereas in arrays SNPs are

22

509 uniformly distributed along the genome. The availability of dense SNP arrays for coho

510 salmon, as it is already the case for Atlantic salmon (Houston *et al.* 2014; Yañez *et al.*

511 2016b) and rainbow trout (Palti *et al.* 2015a), may allow increase the accuracy for

512 predicting genomic breeding values and the power for the determination of the genetic

513 factors involved in economically-important traits, including *P. salmonis* resistance, in this

514 species.

515

516 **5. Conclusions**

517 Moderate significant genetic variation was estimated for resistance against *Piscirickettsia*

518 *salmonis* in coho salmon, using either pedigree or genomic information. These results

519 highlight the feasibility to include this character into genetic improvement programs. One

520 SNP was genome-wide significantly associated with resistance to *P. salmonis,* and may

521 be involved in the immune host response against this infection through iron homeostasis

522 mechanism. Genomic predictions using ddRAD genotypes including 4,174 loci showed

523 similar accuracies as PBLUP. To our knowledge, this is the first study aiming at

524 dissecting the genetic architecture of resistance against *P. salmonis*, in coho salmon

525 population.

526

527

528

529

530

23

531   Acknowledgements

540

541

542

543

544

545

546

547

548

549

550

551

24

552 Authors' contributions

553 AB performed DNA extraction, library construction, ddRAD analysis, GWAS analysis

554 and wrote the initial version of the manuscript. KrC performed library construction and

555 contributed on the data analysis. KC performed Genomic prediction analysis. AJ

556 performed DNA extraction. JPL contributed with study design. WD contributed with

557 analysis and discussion. JMY conceived and designed the study, supervised work of AB

558 and contributed to the analysis, discussion and writing. All authors have reviewed and

559 approved the manuscript.

560

561

562

563

564

565

566

567

568

569

570

571

572

25

573 Bibliography

574 Aulchenko, Y. S., S. Ripke, A. Isaacs, and C. M. van Duijn, 2007 GenABEL: An R

575   library for genome-wide association analysis. Bioinformatics 23: 1294–1296.

576 Baird, N. A., P. D. Etter, T. S. Atwood, M. C. Currey, A. L. Shiver *et al.*, 2008 Rapid

577   SNP discovery and genetic mapping using sequenced RAD markers. PLoS One 3:

578   1–7.

579 Bangera, R., K. Correa, J. P. Lhorente, R. Figueroa, and J. M. Yáñez, 2017 Genomic

580   predictions can accelerate selection for resistance against *Piscirickettsia salmonis*

581   in Atlantic salmon (*Salmo salar*). BMC Genomics 18: 121.

582 Benjamini, Y., and Y. Hochberg, 1995 Controlling the False Discovery Rate : A

583   Practical and Powerful Approach to Multiple Testing. J. R. Stat Soc. 57: 289–300.

584 Brieuc, M. S. O., C. D. Waters, J. E. Seeb, and K. a Naish, 2014 A dense linkage map

585   for Chinook salmon (*Oncorhynchus tshawytscha*) reveals variable chromosomal

586   divergence after an ancestral whole genome duplication event. G3 (Bethesda). 4:

587   447–460.

588 Campbell, N. R., S. E. LaPatra, K. Overturf, R. Towner, and S. R. Narum, 2014

589   Association mapping of disease resistance traits in rainbow trout using restriction

590   site associated DNA sequencing. G3 (Bethesda). 4: 2473–81.

591 Catchen, J. M., A. Amores, P. Hohenlohe, W. Cresko, J. H. Postlethwait *et al.*, 2011

592   Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences. G3

593   (Bethesda). 1: 171–182.

594 Catchen, J., P. A. Hohenlohe, S. Bassham, A. Amores, and W. A. Cresko, 2013 Stacks:

595   An analysis tool set for population genomics. Mol. Ecol. 22: 3124–3140.

26

596　Correa, K., R. Bangera, R. Figueroa, J. P. Lhorente, and J. M. Yáñez, 2017 The use of

597　　genomic information increases the accuracy of breeding value predictions for sea

598　　louse (*Caligus rogercresseyi*) resistance in Atlantic salmon (*Salmo salar*). Genet.

599　　Sel. Evol. 49: 15.

600　Correa, K., M. Filp, D. Cisterna, M. E. Cabrejos, C. Gallardo-Escárate *et al.*, 2015a

601　　Effect of triploidy in the expression of immune-related genes in coho salmon

602　　*Oncorhynchus kisutch* (Walbaum) infected with *Piscirickettsia salmonis*. Aquac.

603　　Res. 46: 59–63.

604　Correa, K., J. Lhorente, M. Lopez, L. Bassini, S. Naswa *et al.*, 2015b Genome-wide

605　　association analysis reveals loci associated with resistance against *Piscirickettsia*

606　　*salmonis* in two Atlantic salmon (*Salmo salar L.*) chromosomes. BMC Genomics

607　　16: 854.

608　Cvitanich, J., O. Garate, and C. E. Smith, 1991 The isolation of a rickettsia-like

609　　organism causing disease and mortality in Chilean salmonids and its confirmation

610　　by Koch's postulate. J. Fish Dis. 14: 121−146.

611　Davey, J. W., T. Cezard, P. Fuentes-Utrilla, C. Eland, K. Gharbi *et al.*, 2013 Special

612　　features of RAD Sequencing data: Implications for genotyping. Mol. Ecol. 22:

613　　3151−3164.

614　Dufflocq, P., J. P. Lhorente, R. Bangera, R. Neira, S. Newman *et al.*, 2016 Correlated

615　　response of flesh color to selection for harvest weight in coho salmon

616　　(*Oncorhynchus kisutch*). Aquaculture 6–11.

617　Elshire, R. J., J. C. Glaubitz, Q. Sun, J. A. Poland, K. Kawamoto *et al.*, 2011 A robust,

618　　simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS

27

619   One 6: 1–10.

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

621   FAO, 2016 Fisheries and Aquaculture Information and Statistical Branch.

622   Gilmour, A. R., B. J. Gogel, B. R. Cullis, and R. Thompson, 2009 ASReml User Guide

623   release 3.0. 372 p. VSN Int. Ltd, Hemel Hempstead, UK 372.

624   Gjedrem, T., 2012 Genetic improvement for the development of efficient global

625   aquaculture: A personal opinion review. Aquaculture 344–349: 12–22.

626   Gjøen, H. M., T. Refstie, O. Ulla, and B. Gjerde, 1997 Genetic correlations between

627   survival of Atlantic salmon in challenge and field tests. Aquaculture 158: 277–288.

628   Goddard, M., and B. Hayes, 2009 Mapping genes for complex traits in domestic

629   animals and their use in breeding programmes. Nat. Rev. Genet. 10: 381–391.

630   Gonen, S., N. R. Lowe, T. Cezard, K. Gharbi, S. C. Bishop *et al.*, 2014 Linkage maps of

631   the Atlantic salmon (*Salmo salar*) genome derived from RAD sequencing. BMC

632   Genomics 15: 166.

633   Hayes, B., and M. Goddard, 2010 Genome-wide association and genomic selection in

634   animal breeding. Genome 876–883.

635   Henríquez, P., M. Kaiser, H. Bohle, P. Bustos, and M. Mancilla, 2016 Comprehensive

636   antibiotic susceptibility profiling of Chilean *Piscirickettsia salmonis* field isolates.

637   J. Fish Dis. 39: 441–448.

638   Houston, R. D., J. W. Davey, S. C. Bishop, N. R. Lowe, J. C. Mota-Velasco *et al.*, 2012

639   Characterisation of QTL-linked and genome-wide restriction site-associated DNA

640   (RAD) markers in farmed Atlantic salmon. BMC Genomics 13: 244.

28

641 Houston, R. D., J. B. Taggart, T. Cézard, M. Bekaert, N. R. Lowe *et al.*, 2014

642   Development and validation of a high density SNP genotyping array for Atlantic

643   salmon (*Salmo salar*). BMC Genomics 15: 90.

644 Kikuchi, G., T. Yoshida, and M. Noguchi, 2005 Heme oxygenase and heme

645   degradation. Biochem. Biophys. Res. Commun. 338: 558–567.

646 Legarra, A., C. Robert-Granié, E. Manfredi, and J. M. Elsen, 2008 Performance of

647   genomic selection in mice. Genetics 180: 611–618.

648 Li, H., and R. Durbin, 2009 Fast and accurate short read alignment with Burrows-

649   Wheeler transform. Bioinformatics 25: 1754–1760.

650 Lien, S., B. F. Koop, S. R. Sandve, J. R. Miller, P. Matthew *et al.*, 2016 The Atlantic

651   salmon genome provides insights into rediploidization. Nature 533: 200–205.

652 Lynch, M., and B. Walsh, 1998 *Genetics and analysis of quantiative traits. Sunderland:*

653   *Sinauer Associates.*

654 McCarthy, Ú. M., J. E. Bron, L. Brown, F. Pourahmad, I. R. Bricknell *et al.*, 2008

655   Survival and replication of *Piscirickettsia salmonis* in rainbow trout head kidney

656   macrophages. Fish Shellfish Immunol. 25: 477–484.

657 Meuwissen, T. H., B. J. Hayes, and M. E. Goddard, 2001 Prediction of Total Genetic

658   Value Using Genome-Wide Dense Marker Maps. Genetics 157: 1819–1829.

659 Misztal, I., S. Tsuruta, D. Lourenco, Y. Masuda, I. Aguilar *et al.*, 2016 Manual for

660   BLUPF90 family of programs.

661 Moen, T., J. Torgersen, N. Santi, W. S. Davidson, M. Baranski *et al.*, 2015 Epithelial

662   Cadherin Determines Resistance to Infectious Pancreatic Necrosis Virus in

663   Atlantic Salmon. Genetics 200: 1313–26.

664   Nairz, M., I. Theurl, S. Ludwiczek, M. Theurl, S. M. Mair *et al.*, 2007 The co-ordinated

665   regulation of iron homeostasis in murine macrophages limits the availability of

666   iron for intracellular *Salmonella typhimurium*. Cell. Microbiol. 9: 2126–2140.

667   Ødegård, J., M. Baranski, B. Gjerde, and T. Gjedrem, 2011 Methodology for genetic

668   evaluation of disease resistance in aquaculture species: Challenges and future

669   prospects. Aquac. Res. 42: 103–114.

670   Ødegård, J., T. Moen, N. Santi, S. A. Korsvoll, S. Kjøglum *et al.*, 2014 Genomic

671   prediction in an admixed population of Atlantic salmon (*Salmo salar*). Front.

672   Genet. 5: 1–8.

673   Ødegård, J., I. Olesen, B. Gjerde, and G. Klemetsdal, 2006 Evaluation of statistical

674   models for genetic analysis of challenge test data on furunculosis resistance in

675   Atlantic salmon (*Salmo salar*): Prediction of field survival. Aquaculture 259: 116–

676   123.

677   Palaiokostas, C., S. Ferarreso, R. Franch, R. D. Houston, and L. Bargelloni, 2016

678   Genomic prediction of resistance to pasteurellosis in gilthead sea bream (*Sparus*

679   *aurata*) using 2b-RAD sequencing. G3 (Bethesda). X: 1–8.

680   Palti, Y., G. Gao, S. Liu, M. P. Kent, S. Lien *et al.*, 2015a The development and

681   characterization of a 57K single nucleotide polymorphism array for rainbow trout.

682   Mol. Ecol. Resour. 15: 662–672.

683   Palti, Y., R. L. Vallejo, G. Gao, S. Liu, A. G. Hernandez *et al.*, 2015b Detection and

684   Validation of QTL Affecting Bacterial Cold Water Disease Resistance in Rainbow

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

686   e0138435.

687   Paradkar, P. N., I. De Domenico, N. Durchfort, I. Zohn, J. Kaplan *et al.*, 2011 Iron

688   depletion limits intracellular bacterial growth in macrophages Iron depletion limits

689   intracellular bacterial growth in macrophages. Blood 112: 866–874.

690   Peterson, B. K., J. N. Weber, E. H. Kay, H. S. Fisher, and H. E. Hoekstra, 2012 Double

691   digest RADseq: An inexpensive method for de novo SNP discovery and

692   genotyping in model and non-model species. PLoS One 7:.

693   Pulgar, R., C. Hödar, D. Travisany, A. Zuñiga, C. Domínguez *et al.*, 2015

694   Transcriptional response of Atlantic salmon families to *Piscirickettsia salmonis*

695   infection highlights the relevance of the iron-deprivation defence system. BMC

696   Genomics 16: 495.

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

698   2004 Microarray analyses identify molecular biomarkers of Atlantic salmon

699   macrophage and hematopoietic kidney response to *Piscirickettsia salmonis*

700   infection. Physiol. Genomics 20: 21–35.

701   Robledo, D., C. Palaiokostas, L. Bargelloni, P. Martínez, and R. Houston, 2017

702   Applications of genotyping by sequencing in aquaculture breeding and genetics.

703   Rev. Aquac. 1–13.

704   Rojas, V., N. Galanti, N. C. Bols, and S. H. Marshall, 2009 Productive infection of

705   *Piscirickettsia salmonis* in macrophages and monocyte-like cells from rainbow

706   trout, a possible survival strategy. J. Cell. Biochem. 108: 631–637.

707   Rozas, M., and R. Enríquez, 2014 Piscirickettsiosis and *Piscirickettsia salmonis* in fish:

708   a review. J. Fish Dis. 37: 163–188.

31

709 Sernapesca, 2016a Informe Sanitario de Salmonicultura en Centros Marinos 2015.

710   Servicio Nacional de Pesca y Acuicultura.

711 Sernapesca, 2016b Informe sobre uso de antimicrobianos en la salmonicultura nacional.

712 Sonesson, A. K., and T. H. E. Meuwissen, 2009 Testing strategies for genomic selection

713   in aquaculture breeding programs. Genet. Sel. Evol. 41: 37.

714 Stear, M. J., S. C. Bishop, B. a Mallard, and H. Raadsma, 2001 The sustainability,

715   feasibility and desirability of breeding livestock for disease resistance. Res. Vet.

716   Sci. 71: 1–7.

717 Svishcheva, G. R., T. I. Axenovich, N. M. Belonogova, C. M. van Duijn, and Y. S.

718   Aulchenko, 2012 Rapid variance components–based method for whole-genome

719   association analysis. Nat. Genet. 44: 1166–1170.

720 Thompson, E. A., and R. G. Shaw, 1990 Pedigree Analysis for Quantitative Traits :

721   Variance Components without Matrix Inversion. Biometrics 46: 399–413.

722 Tsai, H.-Y., A. Hamilton, A. E. Tinch, D. R. Guy, J. E. Bron *et al.*, 2016 Genomic

723   prediction of host resistance to sea lice in farmed Atlantic salmon populations.

724   Genet. Sel. Evol. 48: 47.

725 Tsai, H. Y., A. Hamilton, A. E. Tinch, D. R. Guy, K. Gharbi *et al.*, 2015 Genome wide

726   association and genomic prediction for growth traits in juvenile farmed Atlantic

727   salmon using a high density SNP array. BMC Genomics 1–9.

728 Vallejo, R. L., T. D. Leeds, B. O. Fragomeni, G. Gao, A. G. Hernandez *et al.*, 2016

729   Evaluation of genome-enabled selection for bacterial cold water disease resistance

730   using progeny performance data in rainbow trout: Insights on genotyping methods

731   and genomic prediction models. Front. Genet. 7: 1–13.

32

732 Vallejo, R. L., T. D. Leeds, G. Gao, J. E. Parsons, K. E. Martin *et al.*, 2017 Genomic

733   selection models double the accuracy of predicted breeding values for bacterial

734   cold water disease resistance compared to a traditional pedigree-based model in

735   rainbow trout aquaculture. Genet. Sel. Evol. 49: 17.

736 VanRaden, P. M., 2008 Efficient methods to compute genomic predictions. J. Dairy Sci.

737   91: 4414–23.

738 Yañez, J. M., R. Bangera, J. P. Lhorente, A. Barria, M. Oyarzun *et al.*, 2016a Negative

739   genetic correlation between resistance against Piscirickettsia salmonis and harvest

740   weight in coho salmon (Oncorhynchus kisutch). Aquaculture 459: 8–13.

741 Yañez, J. M., S. Naswa, M. E. Lopez, L. Bassini, K. Correa *et al.*, 2016b Genomewide

742   single nucleotide polymorphism discovery in Atlantic salmon (Salmo salar):

743   validation in wild and farmed American and European populations. Mol. Ecol.

744   Resour. 16: 1002–1011.

745 Yáñez, J.M. S. Newman, and R. D. Houston, 2015 Genomics in aquaculture to better

746   understand species biology and accelerate genetic progress. Front. Genet. 6: 1–3.

747 Yáñez, J. M., R. D. Houston, and S. Newman, 2014a Genetics and genomics of disease

748   resistance in salmonid species. Front. Genet. 5: 1–13.

749 Yáñez, J. M., J. P. Lhorente, L. N. Bassini, M. Oyarzún, R. Neira *et al.*, 2014b Genetic

750   co-variation between resistance against both Caligus rogercresseyi and

751   Piscirickettsia salmonis, and body weight in Atlantic salmon (Salmo salar).

752 Yáñez, J. M., L. N. Bassini, M. Filp, J. P. Lhorente, R. W. Ponzoni *et al.*, 2014c

753   Inbreeding and effective population size in a coho salmon (Oncorhynchus kisutch)

754   breeding nucleus in Chile. Aquaculture 420–421: S15–S19.

33

755   Aquaculture 433: 295–298.

756   Yáñez, J. M., R. Bangera, J. P. Lhorente, M. Oyarzún, and R. Neira, 2013 Quantitative

757   genetic variation of resistance against Piscirickettsia salmonis in Atlantic salmon

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

759   Yañez, J. M., and V. Martinez, 2010 Genetic factors involved in resistance to infectious

760   diseases in salmonids and their application in breeding programmes. Arch. Med.

761   Vet. 42: 1–13.

762

763

764
765

766

767

768

769

770

771

772

773

774

34

775 **Table 1.** Location, p value, proportion of heritability and phenotypic variance explained

776 by the two most important markers associated with resistance against *P. salmonis*.

777

	SNP	Position	p value	Proportion of heritability	Proportion of phenotypic variance
BIN	58185_41	Okis 29	1.50E-05	0.056	0.021
	62738_67	Scaffold 01025	1.34E-04	0.044	0.016
DD	58185_41	Okis 29	3.67E-04	0.066	0.021
	62738_67	Scaffold 01025	5.50E-05*	0.123	0.039

* Significant Association; BIN: binary (dead or alive); DD: day of death

778

779

780

781

782

783

784

785

786

787

788

789

35

790   **Table 2.** Accuracy, reliability and predictive ability of GEBVs assessed by five-fold

791   cross validation for resistance against *Piscirickettsia salmonis* defined as day of

792   death or as binary trait in coho salmon population using four SNPs densities.

	DD			BIN		
	Accuracy	Reliability	Predictive Ability	Accuracy	Reliability	Predictive Ability
PBLUP	0.58	0.36	0.38	0.61	0.38	0.44
1000	0.60	0.38	0.38	0.60	0.36	0.44
2000	0.58	0.36	0.37	0.62	0.39	0.44
3000	0.56	0.34	0.37	0.60	0.37	0.44
4174	0.57	0.35	0.37	0.60	0.38	0.44

DD: day of death; BIN: binary (dead or alive)

793

794

795
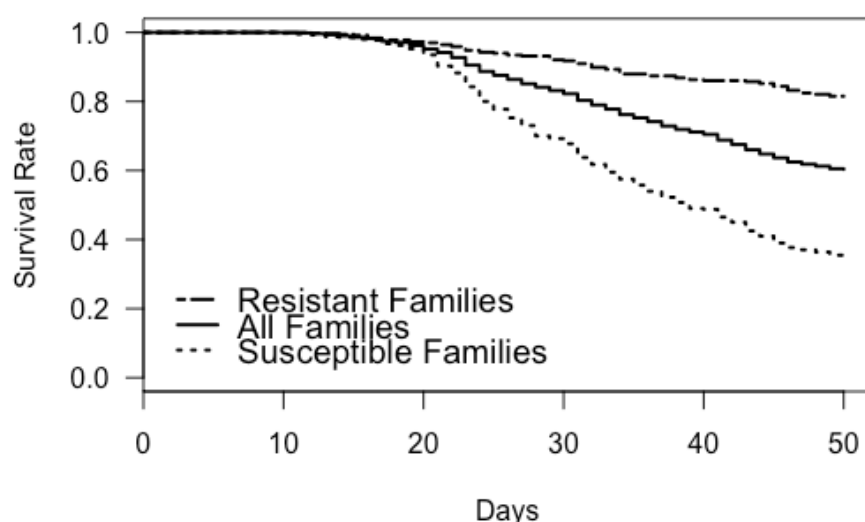
796

797

798

799

800

801

36

802

Figure 1. Kaplan-Meier curves for *Pisciricketssia salmonis* experimental challenge in coho salmon. Average mortality curves for the 108 full-sib families, and the 17 best and 16 worst families.
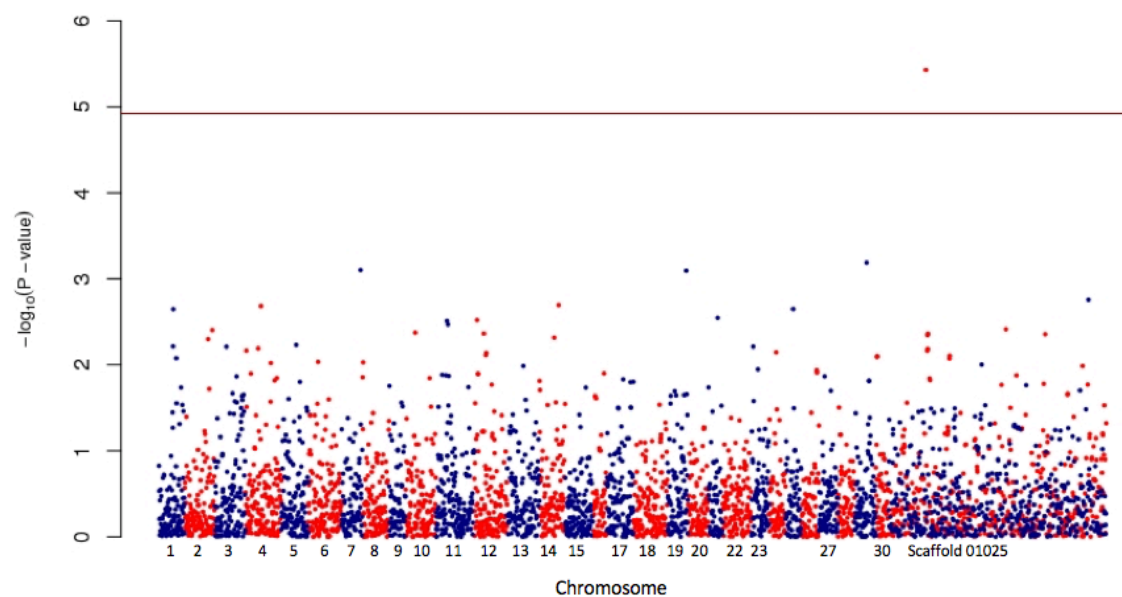
806

807

808

809

810

811

812

Figure 2. Genomic association analysis for resistance against *Pisciricketssia salmonis* in coho salmon population. Resistance was defined as day of death. The horizontal line indicates the FDR significance threshold.
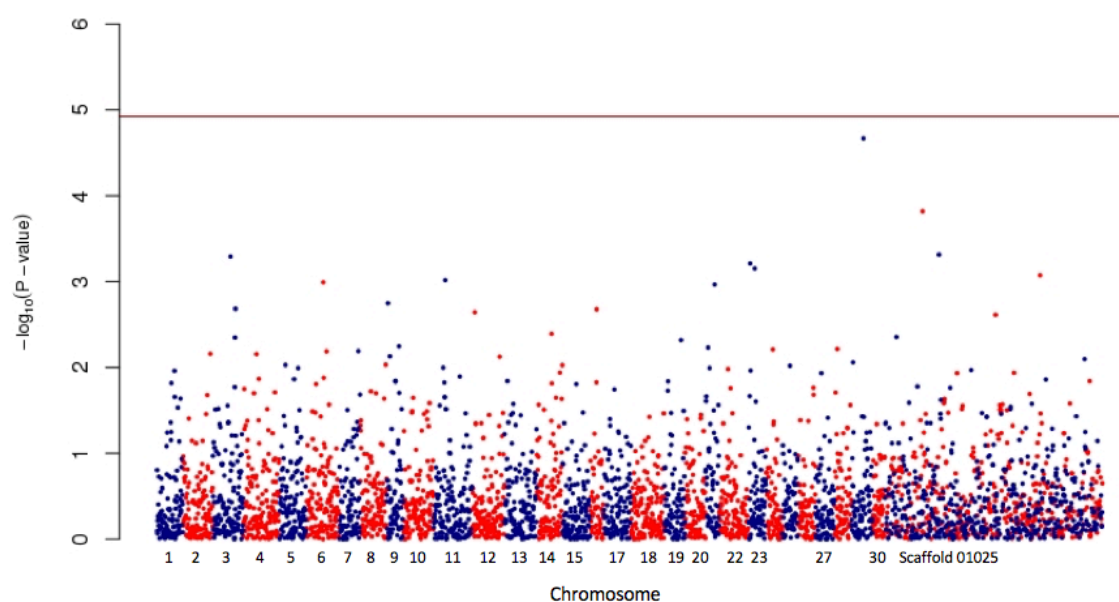
816

817

818

38

819

820 **Figure 3.** Genomic association analysis for resistance against *Piscirickettsia salmonis*

821 in coho salmon population. Resistance was defined as a binary trait. The horizontal line

822 indicates the FDR significance threshold.

823

824

825

826

827