Aquaculture 557 (2022) 738332

Contents lists available at ScienceDirect

Aquaculture

journal homepage: www.elsevier.com/locate/aquaculture

Genome-wide association and genomic prediction of resistance to *Flavobacterium columnare* in a farmed rainbow trout population

Clémence Fraslin [a,*], Heikki Koskinen [b], Antti Nousianen [b], Ross D. Houston [a,1], Antti Kause [c]

[a] The Roslin Institute and Royal (Dick) School of Veterinary Studies, The University of Edinburgh, Midlothian EH25 9RG, United Kingdom
[b] Natural Resources Institute Finland (Luke), Aquaculture Solutions, FI-70211 Kuopio, Finland
[c] Natural Resources Institute Finland (Luke), Breeding and Genomics, FI-31600 Jokioinen, Finland

ARTICLE INFO

Keywords:
GWAS
Genomic selection
Disease resistance
Aquaculture
Flavobacterium columnare

ABSTRACT

Columnaris disease is an emerging disease affecting farmed rainbow trout (*Oncorhynchus mykiss*) globally. In aquaculture breeding, genomic selection has been increasingly used to improve traits that are difficult to measure on candidate fish (such as disease resistance traits). Following a natural outbreak of columnaris disease, 3054 exposed fish and their 81 parents (33 dams and 48 sires) were genotyped with the 57 K SNP Axiom™ trout genotyping array. Genetic parameters of host resistance (measured as a binary survival trait) were estimated, a genome wide association study was performed, and the accuracy of pedigree-based and genomic prediction was estimated. After quality controls, 2874 challenged fish (1403 dead fish and 1471 alive fish) and 78 parents genotyped for 27,907 SNPs remained. Pedigree based heritability was estimated to be 0.18 and 0.35 on the observed and underlying scale, respectively. Genomic heritability was estimated to be 0.21 and 0.43 on the observed and underlying scale, respectively. A quantitative trait loci (QTL) was detected on chromosome Omy3, significant at the genome-wide level, along with several suggestive QTLs on two other chromosomes. The additive effect on mortality proportion of the peak SNP from Omy3 was estimated to be 0.11 (0.018; se), implying approximately 22% difference in survival between alternate homozygous fish at the QTL. Pedigree-based prediction accuracy was 0.59, and the use of genomic evaluation increased the prediction accuracy by at least 13.6%. Using the second iteration of a weighted genomic-based evaluation increased the prediction accuracy by 18.6% compared to the pedigree-based model. These results confirm that resistance to columnaris disease is a suitable target trait for genetic improvement by selective breeding, that a natural outbreak of columnaris disease in a farm environment can be used to select for fish that are more resistant and that genomic selection is a useful approach to speed up this process.

1. Introduction

Rainbow trout (*Oncorhynchus mykiss*) is an important aquaculture species globally, and is produced both in freshwater and sea water. Infectious diseases are a major threat to aquaculture production worldwide, with major impacts on fish welfare, the environment, and the sustainability of aquaculture (Houston, 2017; Yáñez et al., 2014). Water temperatures are rising due to global warming, and fish farms could be subject to more frequent and longer periods of warm water, conditions often more conducive to fish pathogens (Karvonen et al., 2010). One of the pathogens that take advantage of warmer temperature is *F. columnare*, a gram-negative bacterium, responsible for columnaris disease (CD) that affects various fish species of importance for aquaculture worldwide, including rainbow trout (Declercq et al., 2013). Outbreaks of CD occur in rainbow trout farms mainly during the summer, when the water temperature rises above 18 °C, with up to 100% mortality in the absence of antibiotic treatment (Pulkkinen et al., 2010; Suomalainen et al., 2005a). *F. columnare* causes both acute and chronic infections with necrosis of tissues resulting in skin lesions, fin erosions, mouth rot and gill necrosis often leading to the death of the fish (Declercq et al., 2013). A modified live *F. columnare* vaccine for channel catfish (*Ictalurus punctatus*) and largemouth bass (*Micropterus salmoides*) has been developed and tested (Shoemaker et al., 2011) and is now available commercially AQUAVAC-COL™. However, for rainbow trout

* Corresponding author.
E-mail address: clemence.fraslin@roslin.ed.ac.uk (C. Fraslin).
[1] Present address: Benchmark Genetics, 1 Pioneer Building, Edinburgh Technopole, Milton Bridge, Penicuik, EH26 0GB.

https://doi.org/10.1016/j.aquaculture.2022.738332
Received 3 March 2022; Received in revised form 2 May 2022; Accepted 6 May 2022
Available online 10 May 2022

there is no licensed commercial vaccine available against *F. columnare*. The treatments used in rainbow trout consist of adding salted water to increase the salinity (Declercq et al., 2013; Suomalainen et al., 2005b) or using antibiotics, either in a bath treatment or in the feed (Bullock, 1986). The use of antibiotics and antimicrobial agents to treat CD is not a sustainable solution as it can contribute to antimicrobial resistance, a major concern for human and animal health (Serrano, 2005). Thus, other means to control the disease are needed.

Increasing innate genetic resistance of rainbow trout through selective breeding could be a sustainable solution to this major problem. The Finnish national breeding programme for rainbow trout was established in the late 1980's, targeting production traits such as growth, age of maturity, external appearance, fish welfare, visceral percentage, and survival recorded on fish reared in brackish sea and fresh water (Kause et al., 2003; Kuukka-Anttila et al., 2010; Vehviläinen et al., 2012). To date, estimation of breeding values has been based on pedigree information that is obtained by rearing families initially separated in a large number of family tanks followed by individual tagging and pooling of all the fish. The development and availability of new genomics tools such as genotyping-by-sequencing or single nucleotide polymorphism (SNP) arrays (Palti et al., 2015; Robledo et al., 2017) have facilitated studies of the genetic architecture of valuable production traits. The same technologies have also supported the testing and implementation of genomic selection (GS) in aquaculture breeding programmes for several major species for aquaculture (Houston et al., 2020; You et al., 2020). Genomic selection relies on the use of thousands of genetic markers (such as SNPs) spread over the entire genome to estimate the breeding values of selection candidates. A reference population with both phenotypes and genotype data is used to train a prediction model that is then used to estimate breeding values of the candidates that are typically genotyped but not phenotyped (Meuwissen et al., 2001). Although genomic selection has been routinely implemented commercially in the major aquaculture species such as Atlantic salmon (Norris, 2017), it is still in its early days for other species. Recent studies show that in rainbow trout, resistance against *F. columnare* is indeed heritable and that genomic selection could be a potential way to improve the resistance (Evenhuis et al., 2015; Silva et al., 2019a, 2019b).

The objective of this study was to investigate the genetic architecture of resistance to *F. columnare* and quantify the potential of genomic selection to improve resistance in a rainbow trout breeding programme. Specifically, the heritability of resistance to *F. columnare* was estimated in a rainbow trout population from Finland, then a Genome Wide Association Study (GWAS) was performed to investigate the genetic architecture of resistance, and finally, the accuracy of breeding value predictions was compared between genomic evaluation and traditional pedigree-based evaluation approaches. The results contribute to the cumulating evidence of the benefits and suitable ways of implementing genomic selection in aquaculture breeding.

2. Material and methods

2.1. Ethical statement

The establishment of progeny families at Luke's research facilities followed the protocols approved by the Luke's Animal Care Committee, Helsinki, Finland. Hanka-Taimen Oy, a fish farming company, has authorisation for fish rearing and experiments, and both parties comply with the EU Directive 2010/63/EU for animal experiments.

2.2. Fish rearing and phenotyping

On the 15th of May 2019, 105 rainbow trout families (from 33 dams and 48 sires) were produced from the Finnish national breeding programme maintained by Luke at Enonkoski research station in east Finland. This breeding programme was established in early 90's. Annually, around 250–400 family tanks were used for initial rearing of the families, however for the current study, no family tanks were needed. The parents for each generation were selected based on their estimated breeding values (EBV) for growth (since 1992), maturity age (since 2001), external appearance (since 2001), skeletal deformations (since 2002), fillet colour (2003−2012), cataract caused by *Diplostomum* parasite (since 2003), visceral percentage (2005) and survival (since 2010). No selection for resistance to diseases caused by *F. columnare* has been practiced. To control the rate of inbreeding, the optimal genetic contributions method was used since 2002 to select the individuals with the highest selection index but which are not highly related, and to assign the number of mattings and mating pairs of the selected individuals (Kause et al., 2005).

After stripping and fertilisation, a sample of eggs from all families were pooled and sent to the multiplier farm of Hanka-Taimen Oy (Hankasalmi, Finland). The eggs were incubated and fingerlings reared in bulk. In June 2019, around 30,000 fry were distributed into three fingerling tanks, about 100 fish per family per tank and the fish were reared following commercial practices.

The multiplier farm uses water from a nearby stream, and natural *F. columnare* outbreaks occur frequently. From the day the fish were in the three tanks (day 0 of the study), signs of any disease and mortality were monitored twice a day. On days 11 fish in all three tanks started to show signs of CD (saddleback lesions), from day 14 to 16, seven dead or sick fish were collected by tank and the presence of the pathogen was confirmed by PCR. From day 20 to 24 a piece of tail was taken from around 510 fish per tank, randomly chosen amongst the dead fish with clear CD signs, for later genotyping. These 1538 fish were considered as susceptible fish in the analysis. The three tanks were treated, from day 26 to day 32, for *F. columnare* with approved treatments of salt, chloramine, and medical feeds. CD signs and increased mortality were observed also on days 56 and 71 and fish were treated again with an approved treatment. On the last day of the study (day 99, October 2019), a piece of tail was collected, for later genotyping, on around 506 randomly sampled surviving fish per tank. These 1519 fish were considered as resistant fish. By design the mortality rate in the genotyped samples was 50% since about 1500 dead fish and 1500 alive fish were sampled. During the three months of the trial, water temperature was recorded daily.

2.3. Genotyping, quality controls, parentage and sex assignment

Samples from 3057 fish from the trial and 567 fish from the parental generation including the 81 parents were all sent to IdentiGEN Ltd. (Dublin, Ireland) for DNA extraction and genotyping using the 57 K SNP Axiom ™ Trout Genotyping Array (Palti et al., 2015). Prior to the calling of genotypes, quality controls on the 57,501 SNPs from the SNP array were performed following D'Ambrosio et al. (2019): the SNP probes were aligned, using a BLASTn® procedure to the Omyk_1.0 genome assembly available at the time of the analysis (Gao et al., 2018; Pearse et al., 2019) and only SNPs mapping to a single position on the genome (50,820 SNPs) were retained for the next stage of quality control. Genotypes of all 3624 sampled fish were called together in a single run using Axiom Analysis Suite software (v. 4.0.3.3) with standard SNP quality control (QC) thresholds and customized sample QC thresholds (DQC $\geq$0.82, QC call rate $\geq$90, percent of passing samples $\geq$80, average call rate for passing samples $\geq$95). SNPs that were classified by Axiom Analysis Suite software as "off-target variant" or "other", monomorphic SNPs and SNPs for which no homozygous individual was observed for the minor allele were discarded and hence 36,020 polymorphic SNPs were kept for further analysis. Using plink software (v.1.9, Chang et al., 2015), further quality controls were performed on both SNPs and individuals. A total of 26 duplicated individuals were detected using the –genome option from plink, two individuals with an identity-by-descent value over 0.90 were considered as duplicated and both individuals were removed from the analysis. Only the individuals with a call rate over 0.90 and the SNPs with a minor allele frequency (MAF) higher than

0.05, call rate higher than 0.95 and that passed the Hardy-Weinberg equilibrium test (p-value $<10^{-6}$) were retained. The final dataset used for downstream genetic analyses comprised 3435 fish (2874 challenged fish, and 561 fish from the parental generation including 78 parents) genotyped for 27,907 high-quality SNPs.

Since three parents (two dams and one sire) were missing from the final dataset, parentage assignment was performed in two steps. First using a subset of 200 SNPs with a 100% call rate in both generations, the pedigree was reconstructed for fish with no missing parents using the R package APIS (Griot et al., 2020) with a mismatch assignment set to 1%. In the second step, the genomic relationship values derived from a genomic relationship matrix (GRM) built with GCTA software (Eq.1, Yang et al., 2011a) were used to infer the family where one parent was missing from the genotyped dataset. The full pedigree was inferred for 97.6% of the fish while the remaining 68 fish had one parent unassigned / missing.

2.4. Genomic relationship matrix

The GCTA software was used to compute the GRM, and the genetic relationship between individuals j and k (g_{jk}) was estimated following:

$$g_{jk} = \frac{1}{N}\sum_{i=1}^{N}\frac{(x_{ij}-2p_i)(x_{ik}-2p_i)}{2p_i(1-p_i)} \quad (1)$$

in which N is the total number of SNPs (27,907), $x_{ij}$and x_{ik} are the number of copies of the reference allele for the ith SNP for both the jth and kth fish, respectively, and p_i is the frequency of the reference allele estimated from the markers. Two GRM were constructed, a first one with all fish from both generation which was used to recover the pedigree. A second GRM, built only with the fish with a phenotype value (i.e. the offspring), was used for the GWAS.

2.5. Estimation of genetic parameters

Variance components and heritability were estimated based either on pedigree-based relationships or GRM using ASReml (v.4.1, Gilmour et al., 2015) with two different approaches, a linear mixed model (Eq.2) and a logistic regression model (Eq. 3) to assess the trait on the observed and underlying scales, respectively:

$$y_i = \mu + T_i + u_i + e_i \quad (2)$$

$$P(y_i = 1) = \frac{e^{(\mu+T_i+u_i+e_i)}}{1+e^{(\mu+T_i+u_i+e_i)}} \quad (3)$$

in which, for the ith fish, y_i is the phenotype recorded as binary survival (0 for alive and 1 for dead fish), μ is the population mean, T_i is the fixed effect of tank (3 levels), u_i is the random additive genetic value of individual i, following a normal distribution $\boldsymbol{u}\sim N(0,\boldsymbol{G}\sigma_g^2)$ or $\boldsymbol{u}\sim N(0,\boldsymbol{A}\sigma_g^2)$ with σ_g^2 the estimated genetic variance, where **G** is the GRM built with GCTA (Eq.1) and **A** is the pedigree-based relationship matrix. Finally, e_i is the residual effect following a normal and independent distribution $\boldsymbol{e}\sim N(0,\boldsymbol{I}\sigma_e^2)$ with σ_e^2 being the residual variance.

2.6. Genome wide association study

To identify SNPs associated with resistance to *F. columnare*, a GWAS was performed using a mixed linear model association (Eq.4) with the leave-one-chromosome-out (loco) option implemented in GCTA:

$$y_i = \mu + T_i + a_j g_{ij} + u_i + e_i \quad (4)$$

in which y_i is the observed phenotype of the ith individual (0 for alive and 1 for dead), μ the overall mean in the population, T_i the fixed effect of the tank (3 levels), a_j is the additive genetic effect of the reference allele for the jth SNP with its genotype for individual i (g_{ij}) coded as 0, 1 or 2. And e_i is the residual effect following a normal and independent distribution $\boldsymbol{e}\sim N(0,\boldsymbol{I}\sigma_e^2)$ with σ_e^2 the residual variance. Finally, u_i the random vector of polygenic effects followed a normal distribution $\boldsymbol{u}\sim N(0,\boldsymbol{G}\sigma_g^2)$ with σ_g^2 the estimated genetic variance and **G** a partial GRM constructed with 28 chromosomes after removing the chromosome containing the jth SNP since the analysis was performed using the leave-one-chromosome-out (mlma-loco) approach.

2.7. QTL characterisation

For the GWAS, a Bonferroni correction with $\alpha = 5\%$ was used to determine the genome-wide significance threshold $[-\log_{10}(\alpha/n)]$ and the chromosome-wide suggestive threshold $[-\log_{10}(\alpha/[n/29])]$, with n the number of SNPs in the analysis. Only the SNPs with a -$\log_{10}$(p-value) over the chromosome wide threshold were considered to detect QTL associated with the resistance. For each QTL, the additive effect (a) of the top SNP was used to estimate the proportion of genetic variance explained by this peak SNP using:

$$\%V_g = \frac{2p(1-p)a^2}{\sigma_g^2} * 100 \quad (5)$$

with σ_g^2 the total genetic variance estimated using the linear model (Eq.2) with ASReml and p the minor allele frequency of the target SNP.

The concordance between observed and predicted p-values was estimated with the λ value (Yang et al., 2011b).

Candidate genes located within a 2 Mb window around the peak SNP (1 Mb each side) for each QTL were investigated using the NCBI *O. mykiss* Annotation Release 100 (GCF_002163495.1).

2.8. Genomic prediction of breeding values

The efficiency of genomic prediction of breeding values was assessed using 20 replicates of Monte-Carlo "leave-one-group-out" cross-validation tests. For each replicate, fish were randomly assigned to five groups, four-fifths of the fish (n = 2300) with known phenotypes and genotypes were used as the training set and one fifth of the fish (n = 575) with known genotypes and masked phenotypes were used as the validation set. Mixed linear BLUP animal model (Eq.2) and logistic regression model (Eq.3) were used to estimate pedigree-based (EBV) and genomic breeding values (GEBV) of fish in the validation set using phenotypic values of fish in the training set and the relationship matrix, based on pedigree or genomic information, using two different software, BLUPF90 (Misztal et al., 2002) and ASReml.

For the Genomic BLUP (GBLUP), two approaches were implemented using the BLUPF90 software, a standard GBLUP and a weighted GBLUP (wGBLUP) approach. The wGBLUP is an iterative approach in which, at a given kth iteration, a weight is determined for each SNP based on the SNP variance (derived from the SNP additive effect) estimated at the (k-1) iteration (Wang et al., 2012). For the first iteration, weights are fixed to 1 which is equivalent to a standard GBLUP and we performed three iterations (labelled w2GBLUP and w3GBLUP).

The accuracy (r) of genomic prediction was computed as the Pearson correlation coefficient between the (G)EBV and the true phenotype (y) of the fish in the validation set divided by the square root of the genomic based heritability obtained on the observed scale (h^2_{obs}), following Legarra et al. (2008) using the formula:

$$r = \frac{[(G)EBV, y]}{\sqrt{h^2_{obs}}} \quad (6)$$

Since resistance was measured as binary survival, receiver operator characteristic (ROC) curves were also used to assess the accuracy of classifying animals as susceptible or resistant using the different models. The area under the ROC curve (AUC) metric was used to assess the performance of the classifier: an AUC value of 1 represents a perfect

classifier and an AUC value of 0.5 representing a random classifier (Hanley and McNeil, 1982). According to Wray et al. (2010) this AUC value should be considered relatively to a theoretical maximum AUC value (AUC_{max}) that would be obtained for a disease when the test classifier is a perfect predictor of the genetic risk. AUC_{max} depends on the disease prevalence and the heritability of the trait obtained on the underlying scale and was estimated using the formula from Wray et al. (2010) computed using a homemade R function (Supplementary File S1).

3. Results

3.1. Sample and data collection from farm outbreak

A small number of mortalities occurred in tanks 1 and 2 in the first few days after the transfer, with a peak at day 3 with 44 and 58 mortalities recorded in tank 1 and 2, respectively. Higher mortality was recorded after day 1 in tank 3, with a peak on day 4 (119 dead fish) after which mortality decreased to a base level of 30 mortality/day on average for a week. No specific causes were observed for these mortalities.

Mortalities accompanied by clear signs of *F. columnare* started on day 16 for fish in tanks 1 and 2 and on day 18 for fish in tank 3 (Fig. 1). On day zero at the end of June, the average daily water temperature was above 17 °C and started to rise above 18 °C from day 2 onwards with a peak temperature at 25 °C on days 33 and 34. At the end of the recording period, the final mortality was 39.4% (Fig. 1).

3.2. Genetic parameter estimates

The estimates of genetic parameters using the linear and the logistic regression models are summarised in Table 1. Estimates of the pedigree-based heritability of binary survival were 0.18 (± 0.038; se) and 0.35 (± 0.046; se) on the observed and underlying scale, respectively. Genomic heritability of binary survival to *F. columnare* were slightly higher than the pedigree-based estimates, at 0.21 (± 0.030; se) and 0.43 (± 0.042; se) on the observed and underlying scale, respectively.

3.3. Genome wide association study

A highly significant QTL affecting the binary trait of resistance was detected on chromosome 3 (Fig. 2). There were a total of 28 SNPs that were significantly associated with resistance to *F. columnare*, with a -$\log_{10}$(*p*-value) that was over the 5% chromosome-wide Bonferroni threshold (−$\log_{10}$(*p*-value) = 4.28). Those SNPs were located within three chromosomes with 23 SNPs on Omy3, one SNP on Omy12 and four SNPs on Omy15 (Fig. 2).

On chromosome 3, 23 SNPs (one SNP at 38.165 Mb and the rest spanning from 55.715 Mb to 79.557 Mb) had a -$\log_{10}$(*p*-value) at least over the 5% chromosome-wide threshold level (Supplementary Table S1), including 8 SNPs with a -$\log_{10}$(p-value) that surpassed the 5% genome-wide threshold with the peak SNP located at 64.39 Mb (Table 2). The peak SNP explained an estimated 11.2% of the additive genetic variation and the additive effect of the peak SNP from Omy3 on mortality proportion (the binary resistance trait) was estimated to be 0.11 (0.0018; se), implying approximately 22% difference in survival between alternate homozygous fish at the QTL.

On chromosome 12, the only SNP that exceeded the 5% suggestive threshold was located at 5.316 Mb and explained 4.40% of the total genetic variance of resistance to *F. columnare* (Table 2 and Supplementary Table S1). On chromosome 15, the first suggestive SNP was located at 13.36 Mb and the remaining three significant SNPs spanned from

Table 1
Estimates of variance components for resistance to *Flavobacterium columnare*.

Model	Relationship matrix	σ_a^2 (± se)	σ_p^2 (± se)	σ_e^2 (± se)	h^2 (± se)
Linear	A	0.044 ± 0.010	0.25 ± 0.008	0.21 ± 0.008	0.18 ± 0.038 (0.34)
Linear	G	0.054 ± 0.009	0.25 ± 0.008	0.20 ± 0.007	0.21 ± 0.030 (0.40)
Logit	A	0.54 ± 0.108	1.54 ± 0.108	1	0.35 ± 0.046
Logit	G	0.76 ± 0.132	1.76 ± 0.132	1	0.43 ± 0.042

A: the pedigree-based relationship matrix.
G: the genomic relationship matrix.
Logit: logistic regression.
σ_a^2: genetic variance.
σ_p^2: phenotypic variance.
σ_e^2: residual variance.
h^2: For the linear models, heritability is estimated on the observe scale and transformed on the underlying scale (value in brackets) using the formula from Dempster and Lerner (1950). For the logistic regression models, heritability is estimated on the underlying scale with the residual variance fixed to 1 by convention.

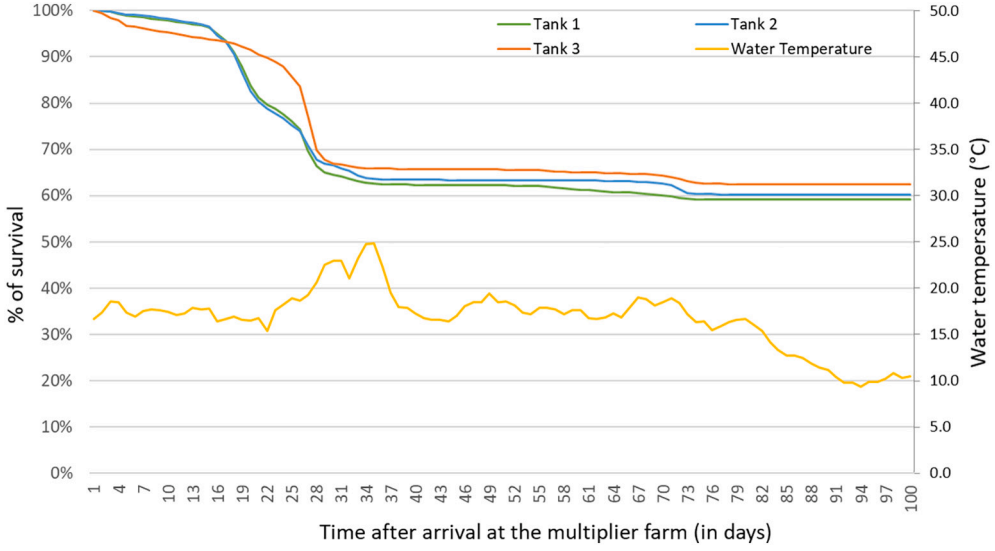

Fig. 1. Survival curves and water temperature for the three tanks for the duration of the study at the multiplier farm.

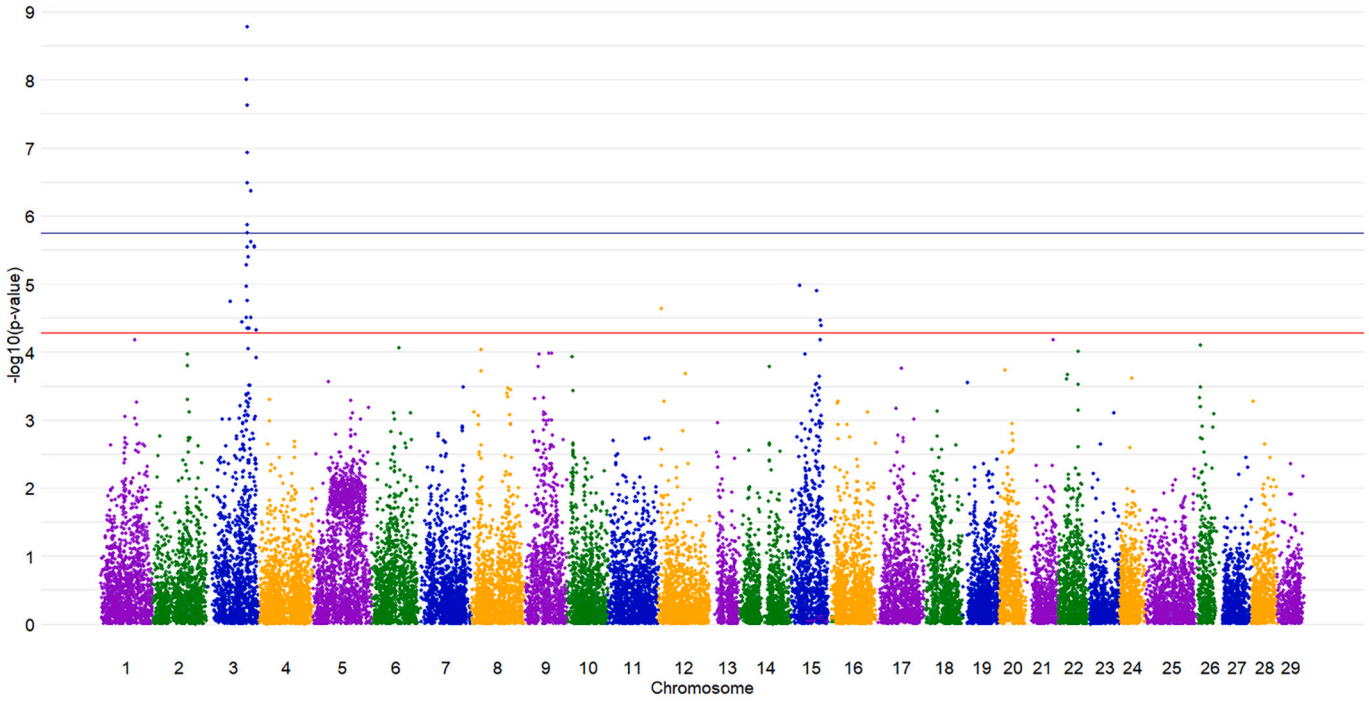

Fig. 2. Manhattan plot of QTL associated with resistance to *Flavobacterium columnare* detected using a GWAS under a MLMA-LOCO model.
The dark blue line is the 5% significance threshold at the genome-wide level, the red line is the 5% suggestive threshold at the chromosome-wide level. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Detection of QTLs associated with resistance to *Flavobacterium columnare* in rainbow trout.

Chr	Peak SNP name	Peak SNP position (bp)	Peak SNP additive effect (± se)	Peak SNP *p*-value	Peak SNP-$\log_{10}$(*p*-value)	*%Vg* explained by peak SNP
3	Affx-89944857	64,390,419	−0.11 ± 0.0018	1.69E-09	8.77	11.20%
12	Affx-88908225	5,316,265	−0.073 ± 0.0172	2.30E-05	4.64	4.40%
15	Affx-88949366	41,172,606	0.093 ± 0.0213	1.25E-05	4.90	5.70%

Chr: chromosome.
$\%Vg = 2p(1\text{-}p)a^2/0.054$ (Eq. 5), the proportion of total genetic variance explained by the peak SNP, with p the SNP minor allele frequency and a the SNP additive effect.

41.173 to 47.018 Mb with the peak SNP explaining 5.70% of the total genetic variance (Table 2).

The observed *p*-values were inflated with a λ value of 1.362 as expected for a population with large full and half-sib families (a λ of ~1.1 indicates a relatively good concordance between observed and predicted *p*-values, Yang et al., 2011b).

3.4. Genomic prediction of breeding values

For both linear and logistic models, genomic-based predictions of breeding values resulted in a higher accuracy than pedigree-based prediction, with an increase of 15.3% between the w3GBLUP and the pedigree-based BLUP (PBLUP) models and of 18.6% between the w2GBLUP and the PBLUP models (Table 3). Correlations and accuracies obtained under the linear or logistic regression models were very similar. The increase in accuracy between the pedigree-based and the genomic-based logistic regression (+13.6%) was slightly lower than between the pedigree-based and genomic-based linear model (+15.3%).

Table 3
Efficiency of genomic evaluation for resistance to *Flavobacterium columnare* in rainbow trout.

Analysis	Model	Relationship matrix	Correlation	Accuracy	AUC
PBLUP	BLUP	A	0.27 ± 0.037	0.59 ± 0.080	0.66 ± 0.021
GBLUP	BLUP	G	0.31 ± 0.035	0.68 ± 0.075	0.68 ± 0.020
w2GBLUP	BLUP	A	0.32 ± 0.032	0.70 ± 0.071	0.68 ± 0.019
w3GBLUP	BLUP	G	0.31 ± 0.031	0.68 ± 0.068	0.68 ± 0.019
PLOGIT	Logistic regression	A	0.27 ± 0.037	0.59 ± 0.080	0.66 ± 0.021
GLOGIT	Logistic regression	G	0.31 ± 0.035	0.67 ± 0.076	0.68 ± 0.020

Mean +/− sd are presented.
A: the pedigree-based relationship matrix.
G: the genomic relationship matrix.
PBLUP: Pedigree-based BLUP.
GBLUP: Genomic BLUP.
w2GBLUP and w3GBLUP are the second and third iteration of the weighted GBLUP, respectively.
PLOGIT: Pedigree-based logistic regression.
GLOGIT: Genomic-based logistic regression.
Correlation: Pearson correlation coefficient between phenotype and estimated breeding values [(G)EBVs].
Accuracy = cor(phenotype, (G)EBVs)/sqrt(h^2_{gen}). With h^2_{gen} the genomic heritability estimated on the observed scale (0.21).
AUC: area under the receiver operator characteristic (ROC) curve.

Based on the AUC values of the ROC curves, the genomic-based relationship matrix classified the fish in the validation population with a success rate of 68% for all models, 2% better than pedigree-based relationship matrix (Table 3). No differences were observed between the linear or logistic regression models on the classifier. To estimate the AUC_{max} of columnaris disease in this population the heritability on the underlying scale was estimated to be 0.40 (Table 1) and the disease prevalence in the sampled population was 0.5, which resulted in an AUC_{max} of 0.80.

4. Discussion

Rainbow trout, an aquaculture species of great importance worldwide, faces major threats due to infectious disease outbreaks in hatcheries and farms. Columnaris disease has become increasingly important over the past 20 years in rainbow trout production and could continue to increase as summer water temperatures rise because of global warming (Karvonen et al., 2010; Pulkkinen et al., 2010). The results presented herein show that selective breeding is a promising approach to enhance the natural resistance of broodstock, and that genomic selection may be an effective approach to increase genetic gain.

4.1. Resistance to F. columnare is moderately heritable

Resistance to *F. columnare* in this rainbow trout population was moderately heritable, with estimates ranging from of 0.18 to 0.43, pedigree-based heritability being slightly lower than genomic-based heritability. This implies that this trait can be improved by selective breeding. The heritability values estimated in the current study were in the range of previously estimated heritability (0.17–0.51) for resistance to *F. columnare* in rainbow trout populations from the USA infected in an experimental challenge (Evenhuis et al., 2015; Silva et al., 2019a, 2019b). As expected, the heritability estimated on the underlying scale using a logistic regression model were significantly higher than the heritability estimated on the observed scale. When the linear estimates obtained on the observed scale were corrected as proposed by Dempster and Lerner (1950), estimates were close to the one obtained on the underlying scale with the logistic regression.

4.2. A major QTL on chromosome Omy3

One main genome-wide significant QTL was located on chromosome Omy3 (peak at 64.390 Mb), several less significant QTLs located on two other chromosomes (Omy12 and Omy15) as well as a polygenic background contribution. To the best of our knowledge, there is only one published study that investigated genetic resistance to *F. columnare* in rainbow trout after an experimental bath challenge in a controlled environment with a water temperature of 16 °C (Silva et al., 2019a). They detected 40 QTLs across 14 chromosomes associated with resistance in two populations, however none of the QTLs detected in the present study were reported in their study (see Supplementary Table S3). Chromosome 12 was the only common chromosome reported in both studies with two QTLs (between 45.5 and 46.5 Mb and second one between 49.2 and 50.1 cM) reported by Silva et al. (2019a) and one QTL detected in the current study (peak SNP at 5.3 Mb) but the distance between those QTLs suggest they are different ones. In the current study, several SNPs located on chromosome Omy3 were associated with resistance. The first significant SNP, with a *p*-value over the 5% chromosome-wide threshold, was located at 38.165 Mb and the next significant SNP located at 55.715 Mb. Such a long distance between the two successive significant SNPs (17.55 Mb) may potentially reflect the presence of two QTLs (Supplementary Fig. S1 and Table S1). This first SNP association may also reflect linkage disequilibrium (LD) with the peak SNP from the most significant QTL at 64.390 Mb, given that long range LD is common in rainbow trout. LD values between two markers at about 1 Mb distance ranging between 0.13 and 0.25 were reported previously by D'Ambrosio et al. (2019) and Vallejo et al. (2018). In the current study, the LD between two SNPs about 1 Mb apart was 0.11 on average (± 16; SD) (data not shown). The detection of one main significant QTL on Omy3 with a 22% difference in survival between alternate homozygous fish at the peak SNP that explained 11.2% of the total genetic variance along with other suggestive minor-effect QTLs and a polygenic background contribution show that resistance to *F. columnare* in rainbow trout is oligogenic in this population.

In the *Flavobacterium* genus there are two bacteria species which are responsible for diseases with similar signs and that target similar fish species. *F. columnare*, the focus of the current study, is responsible for CD in warm waters, and *Flavobacterium psychrophilum* is responsible for cold water disease (BCWD) (Bernardet and Bowman, 2006). Those two bacteria are closely related (Kumru et al., 2017), thus it is plausible that a certain proportion of genetic resistance mechanisms in the fish might be common between the two diseases. In the two studies on rainbow trout populations from the USA Evenhuis et al. (2015) and Silva et al. (2019b) estimated a moderate positive genetic correlation (ranging between 0.35 and 0.40) between the resistance to *F. columnare* and to *F. psychrophilum*. In addition, some of the QTLs detected in our study as associated with resistance *to F. columnare* may co-localised with previously published QTLs associated with resistance to *F. psychrophilum*. For instance, the main QTL on chromosome Omy3 (peak at 64 Mb) has been identified also in two previous studies on resistance to *F. psychrophilum*, after a natural outbreak in a French rainbow trout population (Fraslin et al., 2019) as well as after an experimental challenge in isogenic lines of rainbow trout (Fraslin et al., 2018) or rainbow trout population from the USA (Vallejo et al., 2017). QTLs associated with resistance to *F. psychrophilum* in different rainbow trout populations have also been previously detected on chromosomes Omy12 (Liu et al., 2015; Palti et al., 2015; Vallejo et al., 2014) with QTL located between 18.722 and 78.020 Mb (see supplementary Table S3) and Omy15 with QTLs located between 5.701 and 48.678 Mb in Fraslin et al. (2018) or at 38.446–39.557 Mb in Vallejo et al. (2017). Even if those QTL were detected at different positions on the chromosome, they were located within wide confidence intervals and thus might be identical between the two diseases. The favourable genetic correlation for resistance to both diseases as well as the potentially common QTLs associated with resistance to *F. columnare* and *F. psychrophilum* on chromosomes Omy3, 12 and 15, although estimated in different rainbow trout populations with different genetic backgrounds, are encouraging for breeders. They suggest that improving the resistance to one pathogen could improve the resistance to its cold/warm counterpart.

We investigated the putative candidate genes located within a 2 Mb window around the peak SNPs using the NCBI *O. mykiss* Annotation Release 100 (GCF_002163495.1). Overall eleven genes involved in immune response, through the pro-inflammatory response of cytokine or the receptor-mediated endocytosis by macrophage or dendritic cells in response to bacteria activity, were located around the peak SNPs on the three chromosomes with QTLs (supplementary Table S2). Focusing on the peak SNP with the lowest *p*-value on chromosome Omy3 (located at 64,390 Mb), two genes were identified as being involved in the pro-inflammatory response of cytokine, *transforming growth factor beta receptor type-2-like (TGF-beta 2)* located between 63,826,317 and 63,853,315 bp, and an *interleukin-1 receptor type 1* (*il-1r1*) located between 65,103,069 and 65,123,204 bp. Since the QTLs detected in our study covered a large region of the chromosome, this list of putative candidate genes has to be confirmed by more studies. One approach could be to refine the QTL position using whole-genome-sequence and imputation (Fraslin et al., 2020a; Yoshida and Yáñez, 2021) to refine the list of positional candidate genes. Other approaches such as RNAseq (Marancik et al., 2015; Robledo et al., 2019; Robledo et al., 2018; Zwollo et al., 2017) or knout-out by CRISPR/Cas9 approaches that have been recently used successfully in other fish species (Gratacap et al., 2020; Luo et al., 2022; Pavelin et al., 2021) could be used to validate their implication in the immune response to *F. columnare*.

4.3. Genomic evaluation increases the accuracy of breeding values

Using genomic information significantly increased the accuracy of prediction, by at least 13.6%, compared to breeding values estimated with only pedigree-based information. Both linear and logistic regression models gave similar results. The highest accuracy and highest AUC values were obtained with the weighted GBLUP approach, using the second iteration (w2GBLUP). The small decrease of the accuracy of genomic prediction after the third iteration (w3GBLUP) compare to w2GBLUP has been reported previously (Irano et al., 2016; Melo et al., 2016; Vallejo et al., 2017; Vallejo et al., 2016). Wang et al. (2012) explained this decrease of accuracy by an over weighting of SNPs in large effect QTL and an underweighting of SNPs in small effect QTLs. The better performance of the weighted model compared to the standard GBLUP approach reflects the fact that, resistance to *F. columnare* in this population is oligogenic, controlled by a main QTL, several smaller effect QTLs together with a polygenic background contribution. The utility of using a weighted approach in the presence of a main QTL have been demonstrated in various fish species for other traits of interest (Lu et al., 2020; Song and Hu, 2021; Vallejo et al., 2018). The AUC metrics obtained from the ROC curves are a way of accounting for both sensitivity (true positive rate) and specificity (true negative rate) of a test, and are a complementary approach to accuracy of genomic prediction to estimate the predictability of a model for disease resistance traits (Wray et al., 2010). In the current study, AUC values were already high for the PBLUP model (0.66) the models using genomic information ((w)GBLUP and GLOGIT) better predicted the outcome of the disease with a 0.2 increase in AUC. Although 0.2 represent a small increase in the predictability of the model, this value should be consider relative to the AUC_{max} (Wray et al., 2010) that represent an upper limit to the predictability of the model taking in account the disease prevalence and the heritability on the underlying scale and was estimated to be 0.8 in the current study. The AUC value of 0.68 obtained were in the range of what was obtained by Palaiokostas et al. (2018) for viral nervous necrosis resistance in European seabass (*Dicentrarchus labrax*), a trait with similar disease prevalence and heritability as resistance to *F. columnare* in our study. However, the 0.2 increase of AUC in our study was lower than the increase in Palaiokostas et al. (2018) that reported an increase between 0.8 and 0.13 but the predictability of the pedigree-based model in the current study (0.66) was better than in their study (0.62).

Prediction accuracy using genomic information in the w2GBLUP model was 18.6% higher than the pedigree-based prediction accuracy. This result confirmed that genomic selection is a useful approach to increase resistance to *F. columnare* in this rainbow trout population. A recently published study (Silva et al., 2019a) also concluded that genomic selection was a more promising approach than pedigree-based selection or marker assisted selection to increase rainbow trout resistance to *F. columnare* due to the polygenic architecture of the trait. They observed an improved prediction accuracy of about 40% when using genomic models compared to pedigree-based models. Those results along with the one of the current study confirm the major benefit of using genomic selection to improve resistance to *F. columnare* in different rainbow trout populations, with different genetic architecture of resistance. It should be noted that diseases like *F. columnare* infect small fish well before they can be individually tagged, and hence genotyping is needed to establish relationships between individuals. Therefore, the cost-efficiency of genomic selection may be reasonable, since genotyping is routinely performed anyway, even if at lower density than the SNP array.

4.4. Phenotyping for resistance after a natural disease outbreak

Natural outbreaks and controlled challenge tests can be both used to obtain information to improve disease resistance using breeding programmes. In our study, we took advantage of a natural outbreak of columnaris disease to detect QTL associated with resistance to the pathogen agent responsible of the disease and to estimate the potential of selective breeding to increase the resistance of this rainbow trout population. Experimental challenges are usually used for disease resistance studies since they allow a more controlled experiment, knowing the time when the fish was infected and when it died or sometimes showed signs of infection (Fraslin et al., 2020b; Ødegård et al., 2011; Robinson et al., 2017; Saura et al., 2019). However, experimental challenge requires advanced knowledge of the bacteria in order to isolate and replicate it while still keeping the infectivity high enough to induce mortality in an infectious challenge. Experimental pathogen challenge usually needs to be performed in different facilities, under strict controls, and thus could be expensive to set up within a breeding programme. Furthermore the results obtained in a controlled environment would potentially need to be validated in a field setting before being implemented in a commercial breeding programme (Wiens et al., 2018). Field or natural outbreak data can provide very valuable phenotypes for disease resistance, as the fish are exposed to realistic commercial conditions. Opportunistic sampling of fish during a natural outbreak could be used in selection for increased resistance and to benchmark the experimental challenge trial with the disease resistance measured in a farm environment. In various fish species resistance measured after an natural disease outbreak has been successfully used to estimate genetic parameters (Bangera et al., 2014; Barría et al., 2020; Lillehammer et al., 2013) and detect QTL associated with resistance (Barría et al., 2021; Fraslin et al., 2018; Houston et al., 2008). However, they also have disadvantages such as being unpredictable and the uncertainty of which specific pathogens are responsible for the observed disease or mortality. Furthermore, in the current study, the fish had to be treated for the disease. Thus, the fish that were considered as resistant fish (still alive at the end of the 3 months rearing periods) might not all be truly resistant fish but fish that were very slow to develop disease symptoms, or by chance were not in contact with the bacteria before the treatment, or fish that were infected but cured by the treatment they received. Thus, the resistance phenotype measured in our study may not be the true resistance to the pathogen as usually defined in disease resistance studies (Fraslin et al., 2020b; Robinson et al., 2017). However, the co-location between the QTL detected in our study and previously published QTLs associated with resistance to *F. psychrophilum* (mainly on Omy3) as well as the concordance of heritability estimated in the current study and in previous study after a controlled challenges (Evenhuis et al., 2015; Silva et al., 2019b) suggests that resistance measured in the current study is indeed an appropriate measure of genetic resistance. Furthermore, the results of the current study suggest that phenotyping of resistance after a natural outbreak can be performed as part of the normal rearing process with little effort (collecting the dead fish prior to the treatment) which is potentially less time consuming and less expensive than designing an experimental challenge on siblings of the breeding candidates.

5. Conclusion

In the current study, a moderate heritability of resistance to *F. columnare* was estimated in a rainbow trout population after a natural disease outbreak. Resistance was controlled by a major QTL, located on chromosome Omy3, together with various smaller QTLs and a polygenic background contribution. Finally, genomic selection was shown to be an efficient solution to improve genetic resistance, giving approximately 14% higher breeding values accuracy than pedigree approaches.

Data availability

The data that support the findings of this study are not publicly available. The data can be made available for reproduction of the results from LUKE (antti.kause@luke.fi) upon reasonable request via a material transfer agreement."

Funding

This work is part of the AquaIMPACT project and was supported by the European Union's Horizon 2020 research and innovation programme under the grant agreement No 818367. The Roslin Institute received BBSRC Institute Strategic Program funding (BB/P013732/1, BB/P013740/1, BB/P013759/1).

CRediT authorship contribution statement

Clémence Fraslin: Formal analysis, Data curation, Visualization, Writing – original draft, Writing – review & editing. **Heikki Koskinen:** Conceptualization, Resources, Writing – review & editing. **Antti Nousianen:** Conceptualization, Resources, Writing – review & editing. **Ross D. Houston:** Conceptualization, Resources, Supervision, Project administration, Funding acquisition, Writing – original draft, Writing – review & editing. **Antti Kause:** Conceptualization, Resources, Supervision, Project administration, Funding acquisition, Investigation, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The skilled staff of Savon Taimen Oy and Hanka Taimen Oy are thanked for their expertise in data collection and fish rearing.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aquaculture.2022.738332.

References

Bangera, R., Ødegård, J., Mikkelsen, H., Nielsen, H.M., Seppola, M., Puvanendran, V., Mortensen, A., 2014. Genetic analysis of francisellosis field outbreak in Atlantic cod (Gadus morhua L.) using an ordinal threshold model. In: Aquaculture, International Symposium on Genetics in Aquaculture - XI, 420–421, pp. S50–S56. https://doi.org/10.1016/j.aquaculture.2013.08.029.

Barría, A., Trinh, T.Q., Mahmuddin, M., Benzie, J.A.H., Chadag, V.M., Houston, R.D., 2020. Genetic parameters for resistance to Tilapia Lake virus (TiLV) in Nile tilapia (Oreochromis niloticus). Aquaculture 522, 735126. https://doi.org/10.1016/j.aquaculture.2020.735126.

Barría, A., Trịnh, T.Q., Mahmuddin, M., Peñaloza, C., Papadopoulou, A., Gervais, O., Houston, R.D., 2021. A major quantitative trait locus affecting resistance to Tilapia lake virus in farmed Nile tilapia (Oreochromis niloticus). Heredity 127, 334–343. https://doi.org/10.1038/s41437-021-00447-4.

Bernardet, J.-F., Bowman, J.P., 2006. The genus flavobacterium. In: Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K.-H., Stackebrandt, E. (Eds.), The Prokaryotes, Proteobacteria: Delta, Epsilon Subclass, vol. 7. Springer New York, New York, NY, pp. 481–531. https://doi.org/10.1007/0-387-30747-8_17.

Bullock, G.L., 1986. Columnaris Disease of Fishes. U.S. Department of the Interior, Fish and Wildlife Service, Division of Fisheries and Wetlands Research.

Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., Lee, J.J., 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience 4, 7. https://doi.org/10.1186/s13742-015-0047-8.

D'Ambrosio, J., Phocas, F., Haffray, P., Bestin, A., Brard-Fudulea, S., Poncet, C., Dupont-Nivet, M., 2019. Genome-wide estimates of genetic diversity, inbreeding and effective size of experimental and commercial rainbow trout lines undergoing selective breeding. Genet. Sel. Evol. 51, 26. https://doi.org/10.1186/s12711-019-0468-4.

Declercq, A.M., Haesebrouck, F., Van den Broeck, W., Bossier, P., Decostere, A., 2013. Columnaris disease in fish: a review with emphasis on bacterium-host interactions. Vet. Res. 44, 27. https://doi.org/10.1186/1297-9716-44-27.

Dempster, E.R., Lerner, I.M., 1950. Heritability of threshold characters. Genetics 35, 212–236.

Evenhuis, J.P., Leeds, T.D., Marancik, D.P., LaPatra, S.E., Wiens, G.D., 2015. Rainbow trout (Oncorhynchus mykiss) resistance to columnaris disease is heritable and favorably correlated with bacterial cold water disease resistance1,2. J. Anim. Sci. 93, 1546–1554. https://doi.org/10.2527/jas.2014-8566.

Fraslin, C., Dechamp, N., Bernard, M., Krieg, F., Hervet, C., Guyomard, R., Quillet, E., 2018. Quantitative trait loci for resistance to Flavobacterium psychrophilum in rainbow trout: effect of the mode of infection and evidence of epistatic interactions. Genet. Sel. Evol. 50, 60. https://doi.org/10.1186/s12711-018-0431-9.

Fraslin, C., Brard-Fudulea, S., D'Ambrosio, J., Bestin, A., Charles, M., Haffray, P., Phocas, F., 2019. Rainbow trout resistance to bacterial cold water disease: two new quantitative trait loci identified after a natural disease outbreak on a French farm. Anim. Genet. 50, 293–297. https://doi.org/10.1111/age.12777.

Fraslin, C., Phocas, F., Bestin, A., Charles, M., Bernard, M., Krieg, F., Quillet, E., 2020a. Genetic determinism of spontaneous masculinisation in XX female rainbow trout: new insights using medium throughput genotyping and whole-genome sequencing. Sci. Rep. 10, 17693. https://doi.org/10.1038/s41598-020-74757-8.

Fraslin, C., Quillet, E., Rochat, T., Dechamp, N., Bernardet, J.-F., Collet, B., Boudinot, P., 2020b. Combining multiple approaches and models to dissect the genetic architecture of resistance to infections in fish. Front. Genet. 11 https://doi.org/10.3389/fgene.2020.00677.

Gao, G., Nome, T., Pearse, D.E., Moen, T., Naish, K.A., Thorgaard, G.H., Palti, Y., 2018. A new single nucleotide polymorphism database for rainbow trout generated through whole genome resequencing. Front. Genet. 9 https://doi.org/10.3389/fgene.2018.00147.

Gilmour, A.R., Gogel, B.J., Cullis, B.R., Welham, S.J., Thompson, R., 2015. ASReml User Guide Release 4.1 Functional Specification. VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.

Gratacap, R.L., Regan, T., Dehler, C.E., Martin, S.A.M., Boudinot, P., Collet, B., Houston, R.D., 2020. Efficient CRISPR/Cas9 genome editing in a salmonid fish cell line using a lentivirus delivery system. BMC Biotechnol. 20, 35. https://doi.org/10.1186/s12896-020-00626-x.

Griot, R., Allal, F., Brard-Fudulea, S., Morvezen, R., Haffray, P., Phocas, F., Vandeputte, M., 2020. APIS: an auto-adaptive parentage inference software that tolerates missing parents. Mol. Ecol. Resour. 20, 579–590. https://doi.org/10.1111/1755-0998.13103.

Hanley, J.A., McNeil, B.J., 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143, 29–36. https://doi.org/10.1148/radiology.143.1.7063747.

Houston, R.D., 2017. Future directions in breeding for disease resistance in aquaculture species. Rev. Bras. Zootec. 46, 545–551. https://doi.org/10.1590/s1806-92902017000600010.

Houston, R.D., Haley, C.S., Hamilton, A., Guy, D.R., Tinch, A.E., Taggart, J.B., Bishop, S. C., 2008. Major quantitative trait loci affect resistance to infectious pancreatic necrosis in Atlantic salmon (Salmo salar). Genetics 178, 1109–1115. https://doi.org/10.1534/genetics.107.082974.

Houston, R.D., Bean, T.P., Macqueen, D.J., Gundappa, M.K., Jin, Y.H., Jenkins, T.L., Robledo, D., 2020. Harnessing genomics to fast-track genetic improvement in aquaculture. Nat. Rev. Genet. 1–21 https://doi.org/10.1038/s41576-020-0227-y.

Irano, N., de Camargo, G.M.F., Costa, R.B., Terakado, A.P.N., Magalhães, A.F.B., de Silva, R.M.O., de Albuquerque, L.G., 2016. Genome-wide association study for indicator traits of sexual precocity in Nellore cattle. PLoS One 11, e0159502. https://doi.org/10.1371/journal.pone.0159502.

Karvonen, A., Rintamäki, P., Jokela, J., Valtonen, E.T., 2010. Increasing water temperature and disease risks in aquatic systems: climate change increases the risk of some, but not all, diseases. Int. J. Parasitol. 40, 1483–1488. https://doi.org/10.1016/j.ijpara.2010.04.015.

Kause, A., Ritola, O., Paananen, T., Mäntysaari, E., Eskelinen, U., 2003. Selection against early maturity in large rainbow trout Oncorhynchus mykiss: the quantitative genetics of sexual dimorphism and genotype-by-environment interactions. Aquaculture 228, 53–68. https://doi.org/10.1016/S0044-8486(03)00244-8.

Kause, A., Ritola, O., Paananen, T., Wahlroos, H., Mäntysaari, E.A., 2005. Genetic trends in growth, sexual maturity and skeletal deformations, and rate of inbreeding in a breeding programme for rainbow trout (Oncorhynchus mykiss). Aquac. Genet. Aquac. VIII 247, 177–187. https://doi.org/10.1016/j.aquaculture.2005.02.023.

Kumru, S., Tekedar, H.C., Gulsoy, N., Waldbieser, G.C., Lawrence, M.L., Karsi, A., 2017. Comparative analysis of the *Flavobacterium columnare* genomovar I and II genomes. Front. Microbiol. 8.

Kuukka-Anttila, H., Peuhkuri, N., Kolari, I., Paananen, T., Kause, A., 2010. Quantitative genetic architecture of parasite-induced cataract in rainbow trout, Oncorhynchus mykiss. Heredity 104, 20–27. https://doi.org/10.1038/hdy.2009.123.

Legarra, A., Robert-Granié, C., Manfredi, E., Elsen, J.-M., 2008. Performance of genomic selection in mice. Genetics 180, 611–618. https://doi.org/10.1534/genetics.108.088575.

Lillehammer, M., Ødegård, J., Madsen, P., Gjerde, B., Refstie, T., Rye, M., 2013. Survival, growth and sexual maturation in Atlantic salmon exposed to infectious pancreatic necrosis: a multi-variate mixture model approach. Genet. Sel. Evol. 45, 8. https://doi.org/10.1186/1297-9686-45-8.

Liu, S., Vallejo, R.L., Palti, Y., Gao, G., Marancik, D.P., Hernandez, A.G., Wiens, G.D., 2015. Identification of single nucleotide polymorphism markers associated with bacterial cold water disease resistance and spleen size in rainbow trout. Front. Genet. 6 https://doi.org/10.3389/fgene.2015.00298.

Lu, S., Zhu, J., Du, X., Sun, S., Meng, L., Liu, S., Chen, S., 2020. Genomic selection for resistance to Streptococcus agalactiae in GIFT strain of Oreochromis niloticus by GBLUP, wGBLUP, and BayesCπ. Aquaculture 523, 735212. https://doi.org/10.1016/j.aquaculture.2020.735212.

Luo, M., Wang, J., Dong, Z., Wang, C., Lu, G., 2022. CRISPR-Cas9 sgRNA design and outcome assessment: bioinformatics tools and aquaculture applications. Aquac. Fish. SI 7, 121–130. https://doi.org/10.1016/j.aaf.2021.10.002.

Marancik, D., Gao, G., Paneru, B., Ma, H., Hernandez, A.G., Salem, M., Wiens, G.D., 2015. Whole-body transcriptome of selectively bred, resistant-, control-, and

susceptible-line rainbow trout following experimental challenge with Flavobacterium psychrophilum. Front. Genet. 5 https://doi.org/10.3389/fgene.2014.00453.
Melo, T.P., Takada, L., Baldi, F., Oliveira, H.N., Dias, M.M., Neves, H.H.R., Carvalheiro, R., 2016. Assessing the value of phenotypic information from non-genotyped animals for QTL mapping of complex traits in real and simulated populations. BMC Genet. 17, 89. https://doi.org/10.1186/s12863-016-0394-1.
Meuwissen, T.H., Hayes, B.J., Goddard, M.E., 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157, 1819–1829.
Misztal, I., Tsuruta, S., Strabel, T., Auvray, B., Druet, T., Lee, D.H., 2002. BLUPF90 and related programs (BGF90). In: Proceedings of the 7th World Congress on Genetics Applied to Livestock Production. Presented at the 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France. Communication No.28–27.
Norris, A., 2017. Application of genomics in salmon aquaculture breeding programs by Ashie Norris: who knows where the genomic revolution will lead us?. In: Mar. Genomics, Navigating the Future: Cross-sector Marine Genomics, 36, pp. 13–15. https://doi.org/10.1016/j.margen.2017.11.013.
Ødegård, J., Baranski, M., Gjerde, B., Gjedrem, T., 2011. Methodology for genetic evaluation of disease resistance in aquaculture species: challenges and future prospects. Aquac. Res. 42, 103–114. https://doi.org/10.1111/j.1365-2109.2010.02669.x.
Palaiokostas, C., Cariou, S., Bestin, A., Bruant, J.-S., Haffray, P., Morin, T., Houston, R.D., 2018. Genome-wide association and genomic prediction of resistance to viral nervous necrosis in European sea bass (Dicentrarchus labrax) using RAD sequencing. Genet. Sel. Evol. GSE 50, 30. https://doi.org/10.1186/s12711-018-0401-2.
Palti, Y., Gao, G., Liu, S., Kent, M.P., Lien, S., Miller, M.R., Moen, T., 2015. The development and characterization of a 57K single nucleotide polymorphism array for rainbow trout. Mol. Ecol. Resour. 15, 662–672. https://doi.org/10.1111/1755-0998.12337.
Pavelin, J., Jin, Y.H., Gratacap, R.L., Taggart, J.B., Hamilton, A., Verner-Jeffreys, D.W., Houston, R.D., 2021. The nedd-8 activating enzyme gene underlies genetic resistance to infectious pancreatic necrosis virus in Atlantic salmon. Genomics 113, 3842–3850. https://doi.org/10.1016/j.ygeno.2021.09.012.
Pearse, D.E., Barson, N.J., Nome, T., Gao, G., Campbell, M.A., Abadía-Cardoso, A., Lien, S., 2019. Sex-dependent dominance maintains migration supergene in rainbow trout. Nat. Ecol. Evol. 3, 1731–1742. https://doi.org/10.1038/s41559-019-1044-6.
Pulkkinen, K., Suomalainen, L.-R., Read, A.F., Ebert, D., Rintamäki, P., Valtonen, E.T., 2010. Intensive fish farming and the evolution of pathogen virulence: the case of columnaris disease in Finland. Proc. R. Soc. B Biol. Sci. 277, 593–600. https://doi.org/10.1098/rspb.2009.1659.
Robinson, N.R., Gjedrem, T., Quillet, E., 2017. Improvement of Disease Resistance by Genetic Methods, in: Fish Diseases, Prevention and Control Strategies. Galina Jeney, UK.
Robledo, D., Palaiokostas, C., Bargelloni, L., Martínez, P., Houston, R., 2017. Applications of genotyping by sequencing in aquaculture breeding and genetics. Rev. Aquac. 0 https://doi.org/10.1111/raq.12193.
Robledo, D., Gutiérrez, A.P., Barría, A., Yáñez, J.M., Houston, R.D., 2018. Gene expression response to sea lice in Atlantic Salmon skin: RNA sequencing comparison between resistant and susceptible animals. Front. Genet. 9, 287. https://doi.org/10.3389/fgene.2018.00287.
Robledo, D., Gutiérrez, A.P., Barría, A., Lhorente, J.P., Houston, R.D., Yáñez, J.M., 2019. Discovery and functional annotation of quantitative trait loci affecting resistance to sea lice in Atlantic Salmon. Front. Genet. 10, 56. https://doi.org/10.3389/fgene.2019.00056.
Saura, M., Carabaño, M., Fernández, A., Cabaleiro, S., Doeschl-Wilson, A., Anacleto, O., Villanueva, B., 2019. Disentangling genetic variation for resistance and endurance to Scuticociliatosis in turbot using pedigree and genomic information. Front. Genet. 7 https://doi.org/10.3389/fgene.2019.00539.
Serrano, P.H., 2005. Responsible Use of Antibiotics in Aquaculture. FAO Fish. Tech. Papp, p. 469.
Shoemaker, C.A., Klesius, P.H., Drennan, J.D., Evans, J.J., 2011. Efficacy of a modified live Flavobacterium columnare vaccine in fish. Fish Shellfish Immunol. 30, 304–308. https://doi.org/10.1016/j.fsi.2010.11.001.
Silva, R.M.O., Evenhuis, J.P., Vallejo, R.L., Gao, G., Martin, K.E., Leeds, T.D., Lourenco, D.A.L., 2019a. Whole-genome mapping of quantitative trait loci and accuracy of genomic predictions for resistance to columnaris disease in two rainbow trout breeding populations. Genet. Sel. Evol. GSE 51, 42. https://doi.org/10.1186/s12711-019-0484-4.
Silva, R.M.O., Evenhuis, J.P., Vallejo, R.L., Tsuruta, S., Wiens, G.D., Martin, K.E., Leeds, T.D., 2019b. Variance and covariance estimates for resistance to bacterial cold water disease and columnaris disease in two rainbow trout breeding populations. J. Anim. Sci. https://doi.org/10.1093/jas/sky478.
Song, H., Hu, H., 2021. Strategies to improve the accuracy and reduce costs of genomic prediction in aquaculture species. Evol. Appl. 00, 1–13. https://doi.org/10.1111/eva.13262.
Suomalainen, L.R., Tiirola, M.A., Valtonen, E.T., 2005a. Influence of rearing conditions on Flavobacterium columnare infection of rainbow trout, Oncorhynchus mykiss (Walbaum). J. Fish Dis. 28, 271–277. https://doi.org/10.1111/j.1365-2761.2005.00631.x.
Suomalainen, L.R., Tiirola, M.A., Valtonen, E.T., 2005b. Treatment of columnaris disease of rainbow trout: low pH and salt as possible tools? Dis. Aquat. Org. 65, 115–120. https://doi.org/10.3354/dao065115.
Vallejo, R.L., Palti, Y., Liu, S., Evenhuis, J.P., Gao, G., Rexroad, C.E., Wiens, G.D., 2014. Detection of QTL in rainbow trout affecting survival when challenged with flavobacterium psychrophilum. Mar. Biotechnol. 16, 349–360. https://doi.org/10.1007/s10126-013-9553-9.
Vallejo, R.L., Leeds, T.D., Fragomeni, B.O., Gao, G., Hernandez, A.G., Misztal, I., Palti, Y., 2016. Evaluation of genome-enabled selection for bacterial cold water disease resistance using progeny performance data in rainbow trout: insights on genotyping methods and genomic prediction models. Front. Genet. 7 https://doi.org/10.3389/fgene.2016.00096.
Vallejo, R.L., Liu, S., Gao, G., Fragomeni, B.O., Hernandez, A.G., Leeds, T.D., Palti, Y., 2017. Similar genetic architecture with shared and unique quantitative trait loci for bacterial cold water disease resistance in two rainbow trout breeding populations. Front. Genet. 8 https://doi.org/10.3389/fgene.2017.00156.
Vallejo, R.L., Silva, R.M.O., Evenhuis, J.P., Gao, G., Liu, S., Parsons, J.E., Palti, Y., 2018. Accurate genomic predictions for BCWD resistance in rainbow trout are achieved using low-density SNP panels: evidence that long-range LD is a major contributing factor. J. Anim. Breed. Genet. 135, 263–274. https://doi.org/10.1111/jbg.12335.
Vehviläinen, H., Kause, A., Kuukka-Anttila, H., Koskinen, H., Paananen, T., 2012. Untangling the positive genetic correlation between rainbow trout growth and survival. Evol. Appl. 5, 732–745. https://doi.org/10.1111/j.1752-4571.2012.00251.x.
Wang, H., Misztal, I., Aguilar, I., Legarra, A., Muir, W.M., 2012. Genome-wide association mapping including phenotypes from relatives without genotypes. Genet. Res. 94, 73–83. https://doi.org/10.1017/S0016672312000274.
Wiens, G.D., Palti, Y., Leeds, T.D., 2018. Three generations of selective breeding improved rainbow trout (Oncorhynchus mykiss) disease resistance against natural challenge with Flavobacterium psychrophilum during early life-stage rearing. Aquaculture 497, 414–421. https://doi.org/10.1016/j.aquaculture.2018.07.064.
Wray, N.R., Yang, J., Goddard, M.E., Visscher, P.M., 2010. The genetic interpretation of area under the ROC curve in genomic profiling. PLoS Genet. 6, e1000864 https://doi.org/10.1371/journal.pgen.1000864.
Yáñez, J.M., Houston, R.D., Newman, S., 2014. Genetics and genomics of disease resistance in salmonid species. Front. Genet. 5 https://doi.org/10.3389/fgene.2014.00415.
Yang, J., Lee, S.H., Goddard, M.E., Visscher, P.M., 2011a. GCTA: a tool for genome-wide complex trait analysis. Am. J. Hum. Genet. 88, 76–82. https://doi.org/10.1016/j.ajhg.2010.11.011.
Yang, J., Weedon, M.N., Purcell, S., Lettre, G., Estrada, K., Willer, C.J., Visscher, P.M., 2011b. Genomic inflation factors under polygenic inheritance. Eur. J. Hum. Genet. 19, 807–812. https://doi.org/10.1038/ejhg.2011.39.
Yoshida, G.M., Yáñez, J.M., 2021. Multi-trait GWAS using imputed high-density genotypes from whole-genome sequencing identifies genes associated with body traits in Nile tilapia. BMC Genomics 22, 57. https://doi.org/10.1186/s12864-020-07341-z.
You, X., Shan, X., Shi, Q., 2020. Research advances in the genomics and applications for molecular breeding of aquaculture animals. Aquaculture 526, 735357. https://doi.org/10.1016/j.aquaculture.2020.735357.
Zwollo, P., Hennessey, E., Moore, C., Marancik, D.P., Wiens, G.D., Epp, L., 2017. A BCWD-resistant line of rainbow trout exhibits higher abundance of IgT+ B cells and heavy chain tau transcripts compared to a susceptible line following challenge with Flavobacterium psychrophilum. Dev. Comp. Immunol. 74, 190–199. https://doi.org/10.1016/j.dci.2017.04.019.