Aquaculture 578 (2024) 740078

Contents lists available at ScienceDirect

Aquaculture

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

Genome-assisted prediction of amoebic gill disease resistance in different populations of Atlantic salmon during field outbreak

Afees Abiola Ajasa [a,b,*], Solomon Antwi Boison [c], Hans Magnus Gjøen [b], Marie Lillehammer [a]

[a] Nofima (Norwegian institute of Food, Fisheries and Aquaculture research), P.O. Box 210, N-1431 Ås, Norway
[b] Department of Animal and Aquacultural Sciences, Norwegian University of Life Sciences, N-1430 Ås, Norway
[c] Mowi Genetics AS, Sandviksboder 77AB, Bergen, Norway

ARTICLE INFO

Keywords:
Genomic selection
Amoebic gill disease resistance
Linkage disequilibrium
Bayesian models

ABSTRACT

The Atlantic salmon industry in northern Europe is experiencing increasing losses due to the amoeba *Paramoeba perurans*, which is the causative agent of amoebic gill disease (AGD); a disease that has a debilitating impact on fish's health and welfare. Successful implementation of genomic selection (GS) for AGD can potentially increase selection response and help reduce outbreaks in the commercial farming of Atlantic salmon. However, successful implementation of GS requires the existence of linkage disequilibrium (LD) between markers and quantitative trait loci (QTL). In this study, we evaluated separately the extent of LD present in six Atlantic salmon breeding populations from Mowi. We also investigated the benefit of using genomic information for selection in these populations, comprising 4 year-classes from Mowi's Norwegian population and 2 year-classes from Mowi's Irish population that was recently introgressed into the Norwegian population. The average distance between markers was 43 kb and the average LD (measured by r^2) between adjacent markers was approximately 0.3 for each population. As expected, LD decreased as the physical distance between markers increased. In addition, we observed long-range LD (LD extending to several megabases) across all chromosomes and for all the populations studied. Both the heritability and the accuracy of the breeding value estimates for AGD resistance varied considerably among populations, ranging between 0.06 and 0.24, and 0.32 to 0.77, respectively. The GS models studied had overall better performance than the pedigree based best linear unbiased prediction (PBLUP) model with respect to the accuracy of breeding values prediction, whereas no significant difference was found between the linear and nonlinear GS models. We recommend the use of genomic best linear unbiased prediction (GBLUP) model for the genetic evaluation of AGD resistance due to the higher computing requirements of nonlinear GS models.

1. Introduction

Amoebic gill disease (AGD), first identified in 1985 (Munday, 1985), is an increasing concern to the salmon industry around the world. The disease, which was initially found only in Tasmania (Munday, 1985) and USA (Kent et al., 1988), has now been reported in most salmon producing countries of the world, including Norway, Chile and Ireland (Bustos et al., 2011; Oldham et al., 2016; Steinum et al., 2008).

AGD is a disease caused by the amoeba *Paramoeba perurans* (Young et al., 2008). Clinical manifestation of the disease includes anorexia, cardiac dysfunction, lethargy, respiratory distress, and convergence of infected fish near the water surface (Kent et al., 1988; Kube et al., 2012; Munday et al., 1990; Powell et al., 2002). If left untreated, it can result in mortalities of up to 50% (Munday, 1985).

AGD presents itself with the appearance of white multifocal mucoid patches on the gills (Kube et al., 2012). These patches are usually also used to score or assess the severity of AGD in what is often termed a *gross gill (pathology) score*. Presently, histopathology and polymerase chain reaction (PCR) analysis are the only two ways of confirming the presence of the pathogen.

The only economically feasible treatment of the disease is fresh water or hydrogen peroxide bath for about 0.5–3 h, depending on the treatment type (Parsons et al., 2001; Rodger, 2014). Baths are scheduled when the average gross gill score is above 2 (Taylor et al., 2009b). This bath merely removes gill mucus and gill associated amoebas (Taylor et al., 2009b), and usually fish will become reinfected in a matter of days

* Corresponding author at: Nofima (Norwegian institute of Food, Fisheries and Aquaculture research), P.O. Box 210, N-1431 Ås, Norway.
E-mail address: afees.ajasa@nofima.no (A.A. Ajasa).

https://doi.org/10.1016/j.aquaculture.2023.740078
Received 23 April 2023; Received in revised form 8 August 2023; Accepted 7 September 2023
Available online 9 September 2023

after bath. The number of required baths within a production cycle can thus be considerable; for instance, in Tasmania, between 12 and 15 baths may be needed in a 15–18 months production cycle. This can result in a substantial increase in cost of production (Taylor et al., 2009b) and there are of course also severe welfare issues related both to the disease itself and the treatment of it (Lillehammer et al., 2019).

A supplementary strategy that can help to manage this disease is selective breeding for AGD resistance. In Tasmania, which has the longest history of AGD in salmon farming, selective breeding has proven to be a successful approach in mitigating the negative impacts of the disease, and the number of baths required in each production cycle has been significantly reduced (Evans et al., 2015).

Genetic evaluation of disease resistance of aquacultural species can be done as a challenge trial in experimental facilities or during natural disease outbreak in the sea (field testing). In a field test, the natural environment will encompass all factors and/or their interactions that may contribute to pathogenicity (Norris et al., 2008). However, it might be difficult to ascertain the actual cause of disease or mortality. Also, the randomness of an outbreak may make field testing less suitable for systemic selective breeding. In experimental challenge trials, group of fish are exposed to specific pathogens (Odegard et al., 2006), thus, the cause of disease/death can be assumed known, although clinical examination of challenged fish is usually also performed. Challenge trials are thus arguably the 'gold standard' for evaluating disease resistance in aquatic species. Some studies have also reported favorable genetic correlation between challenge trials and field testing (Gjøen et al., 1997; Odegard et al., 2006; Storset et al., 2007). However, recent results of Gjerde et al. (2019) and Lillehammer et al. (2019) show low genetic correlation between challenge trial and field testing for AGD resistance, which implies that challenge trials are not necessarily reflective of AGD resistance in the commercial production environment.

The use of genome wide dense SNP-markers (single nucleotide polymorphism) to predict the genetic merit of an individual, which can be done for instance by GS (genomic selection), has revolutionized plant and animal breeding. In GS schemes, phenotypes are regressed on all markers simultaneously in an informant population (also termed training/reference population) to obtain the marker effects/allele substitution effects, which are then used to predict the breeding values of genotyped selection candidates (Meuwissen et al., 2001). Genome wide markers with sufficient density are used in GS so that all the quantitative trait loci (QTLs) are in linkage disequilibrium (LD) with at least one marker (Goddard and Hayes, 2007). Few studies (Barria et al., 2018; Kijas et al., 2017) exist on the amount of LD present in Atlantic salmon populations and moreover, there is dearth of information on LD in the Atlantic salmon breeding populations in Norway.

In aquaculture breeding programs, where most traits are not measured directly on selection candidates, GS has the advantage that it can accurately capture the mendelian sampling terms, allowing for better utilization of within and between family genetic variation, and thereby improving prediction accuracy compared to using only pedigree information. GS has recently been implemented in Atlantic salmon populations for several growth related and disease traits (Tsai et al., 2016; Tsai et al., 2015; Verbyla et al., 2022) and increase in prediction accuracy have been reported to range from 10 to 27%, compared to pedigree-based models. For AGD, GS has been shown to increase accuracy of prediction (Aslam et al., 2020; Verbyla et al., 2022) and most importantly, higher realized genetic gains has also been reported for the Tasmania Atlantic salmon breeding population (Verbyla et al., 2022). The above studies reported increases in prediction accuracy when using linear GS models, however, Bayesian models have been reported to perform better than linear GS models for traits that are controlled by a small number of QTLs or genes (Joshi et al., 2021; Kemper et al., 2015; Neves et al., 2014; Zhu et al., 2019). For AGD, Aslam et al. (2020) reported putative QTLs explaining about 37% of total genetic variance. No previous studies on AGD resistance in Atlantic salmon have compared linear and nonlinear (bayesian) GS models.

In this study, we will: (i) explore the extent of linkage disequilibrium present in different breeding populations of Atlantic salmon (ii) estimate the genetic parameters for AGD resistance using pedigree and genomic information and (iii) evaluate the prediction accuracy of pedigree based and genomic based (linear and nonlinear) evaluation models.

2. Material and methods

2.1. Fish strains

The year-classes used in this study came from two strains: Mowi Fanad and Mowi Norway strains. These strains comprised of four parallel sub-populations each, often referred to as year-classes to indicate the year the sub-populations were taken to sea. The reason for having four parallel year-classes in a salmon breeding program is due to the generation interval of Atlantic salmon which is usually four years (longer in the wild). Selection is usually done yearly to make seeds available for the salmon industry each year (Gjedrem, 2010).

The origin of the Mowi Norway strain dates as far back as 1969 when wild salmons from rivers Vosso and Åroy were used to form the base population (Våge, 1995). Between 1982 and 1986, ova from this strain were exported to Ireland to form the Mowi Fanad strain (Norris et al., 1999). Hence, these strains originated from a common base population.

The year-classes in the Mowi strain were crossed systematically in the years 1969 to 2002 in order to avoid rapid increase in inbreeding. This mixing was considerably reduced in the generations following 2002, when the use of DNA based assignment of parents and BLUP selection was implemented. Since then, the year-classes have undergone about six generations of selection in parallel with the occasional use of males from the previous year-class.

Two year-classes from the Mowi Fanad strain were recently introgressed into the year-class 2015 (YC2015) and year-class 2016 (YC2016) of the Norwegian strain. Thus, for these year-classes (YC2015 and YC2016), we appended A and B to indicate their origin. A and B refer to the Norwegian and Irish strain, respectively. We have a total of 6 year-classes in this study (YC2015A, YC2015B, YC2016A, YC2016B, YC2017 and YC2018). Population bottleneck and genetic drift are likely the major factors responsible for the differentiation of these year-classes (Fig. 1). As stated earlier, there is usually the presence of four year-classes in an Atlantic salmon breeding population, whereas this study has 6 year-classes, hence, to avoid any ambiguity, we shall refer to the year-classes in our study as populations.

2.2. Fish management

The populations in this study were AGD phenotyped in a field trial. For all populations, the fish were eyed-egg and start-fed in January and April the year before. All fish were individually tagged using passive integrated transponder (PIT) tag obtained from Biomark (https://www.biomark.com/product/apt12-pit-tag/). They were stocked into net cages between April and May when they were approximately 14–15 months old and have completed smoltification. For example, the 2015 populations were eyed-eggs and start-fed in January and April 2014, respectively. They were raised in freshwater flow-through system until they smoltified and were ponded/transferred into net cages in May 2015. The average size of the fish at stocking in sea were 101 g, 122 g, 120 g and 198 g for the populations in the year 2015, 2016, 2017 and 2018, respectively.

2.3. Field test

For each field trial, we monitored the occurrences of AGD in the net cages by regular gill scoring of a random sample of fish until the average score approached 2 at which point all fish were gill scored (major gill scoring). This monitoring started in the month of July in each year based on historical occurrence of AGD in Norway. All populations were gill

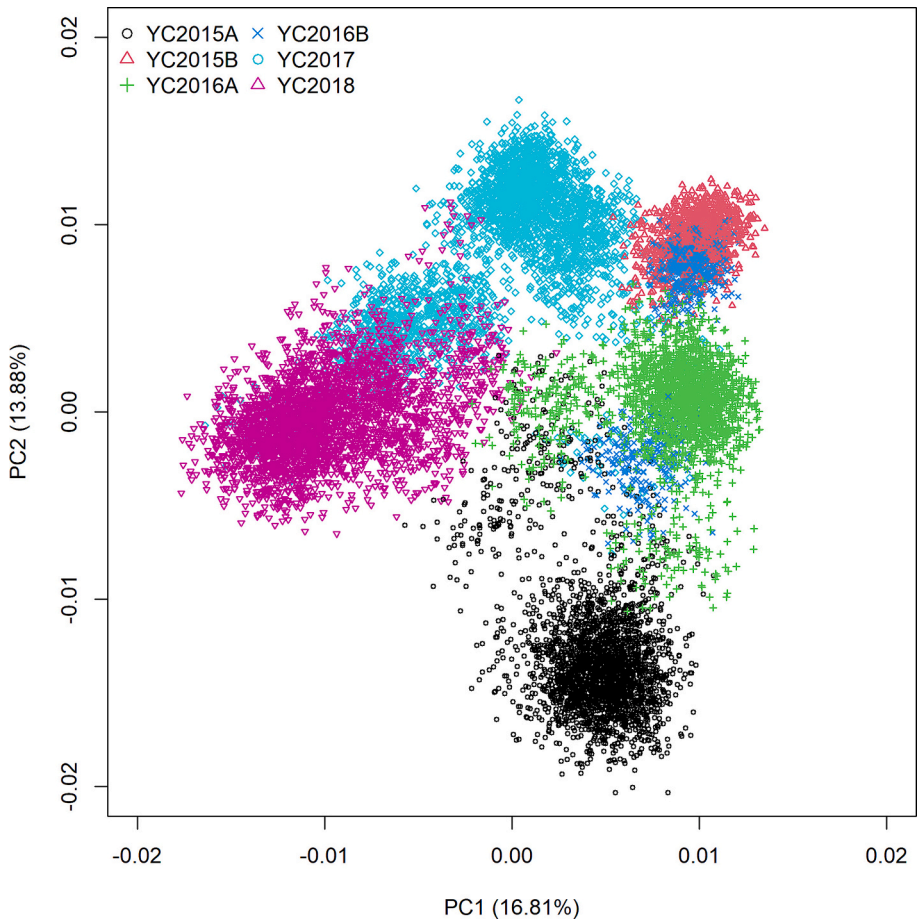

Fig. 1. Relationship between populations obtained from the first two PC of the genomic relationship matrix.

scored using the method of Taylor et al. (2009b). Briefly, each fish was scored between 0 and 5, where the numbers are on an ordinal scale with 0 meaning no infection and 5 meaning severe coverage of the gills with >50% of the gills covered with white spots. For the 2015 populations, two persons (A and B) scored all the fish, while only one person (B) gill scored all the other populations from 2016 to 2018. Two major gill scorings were done for populations in the year 2015, occurring in the months of September (first infection) and October (second infection). For other populations (YC2016A - YC2018), major gill scoring only happened in the month of November and this was assumed to be the first infection. After each major gill scoring, the fish were treated with fresh water. For populations in the year 2015, I and II is appended to the population's name to indicate first and second AGD infections, respectively.

2.4. Genotyping

All fish were genotyped with a non-commercial Affymetrix 55 K SNP array developed by Nofima in 2016 in collaboration with Benchmark and Mowi Genetics AS. Quality control was performed on the SNP data with Plink (Purcell et al., 2007). Markers and samples with call rate < 95% were removed. Also, SNPs with Hardy Weinberg *p* value (Fisher's exact test) of $<10^{-25}$ as well as those with minor allele frequency of <1% were discarded. Finally, to limit the impact of poor sample quality, only samples with heterozygosity rate between 0.25 and 0.45 were used for the analysis.

2.5. Genotype imputation

Sporadic missing genotypes were imputed with Beagle version 5.4 (Browning and Browning, 2016) using the following parameters: each chromosome was split into segments of size 20 Mb (*window = 20*) and overlaps between windows 5 Mb (*overlap = 5*) was used; 10 burn-in (*burnin = 10*) and 50 iterations was used for the phasing and imputation (*iterations = 50*), the rest of the parameters were set to default. A total of 50,456 SNPs were retained after quality control and imputation. The imputation step was necessary because missing genotypes are problematic for the genetic evaluation software (Wombat (Meyer, 2007)) used in this study.

2.6. Linkage disequilibrium between adjacent markers

Linkage disequilibrium (measured by r^2 (Hill and Robertson, 1968)) between adjacent markers for each population was computed using Plink (Purcell et al., 2007), with the following parameter setting; –-r2, –-ld-window-kb 1000, zero –ld-window-r2, and –-ld-window of 2 to obtain LD values between adjacent SNPs. The mean r^2 were then computed using R (Team, 2021).

2.7. LD decay

LD between markers at distances of up to 1000 kb was estimated using Plink (Purcell et al., 2007) with the parameters –r2, –-ld-window-r2 0, –-ld-window 100, and –ld-window-kb 1000. The last two parameters indicate that markers of up to 100 within 1000 kb distance should be used in the computation. Based on the physical distance of each SNP pair, SNPs pair with the same physical distance were grouped together (bin) and the mean r^2 of each bin was computed and visualized using R (Team, 2021).

2.8. Model for analysis

For each population, the following models were used to analyse the gill score records.

2.8.1. PBLUP

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{W}\mathbf{u} + \mathbf{e}$$

where **y** is a vector of phenotypes (gill score), **μ** is the overall mean, **1** is a vector of ones, **W** is an incidence matrix relating the phenotype to the additive genetic effects **u** and **e** is a vector of residual effects. $\mathbf{e} \sim N(0, \mathbf{I}\sigma_e^2)$, $\mathbf{u} \sim N(0, \mathbf{A}\sigma_u^2)$, where **I** is an identity matrix, σ_e^2 is the residual variance, **A** is the numerator relationship matrix, σ_u^2 is the additive genetic variance.

Heritability (h^2) was estimated as: $h^2 = \frac{\sigma_u^2}{\sigma_u^2+\sigma_e^2}$

2.8.2. GBLUP

SNP BLUP is equivalent to Genomic best linear unbiased prediction (GBLUP) (Goddard, 2009; Strandén and Garrick, 2009). However, when the number of individuals is less than the number of markers (as is the case for the datasets in this study), GBLUP is desirable because it has less computational requirement (Meuwissen et al., 2013), hence our choice of GBLUP. The GBLUP model used was similar to the PBLUP model but with **u** being replaced with **g** and **A** being replaced with **G**. **g** is the vector of additive genomic effects and **G** is the genomic relationship matrix. **G** was computed as described by VanRaden (2008).

$$\mathbf{G} = \frac{\mathbf{Z}\mathbf{Z}'}{\sum_{i}^{m} 2p_i(1-p_i)}$$

where **Z** is a matrix of centered genotypes (0 -2p = homozygous, 1 -2p = heterozygous, 2 -2p = homozygous), p_i is the frequency of the reference allele for the i^{th} marker, and *m* is the total number of markers.

Genomic heritability was estimated as: $h^2 = \frac{\sigma_g^2}{\sigma_g^2+\sigma_e^2}$

2.8.3. Bayes B

Bayes B was proposed by Meuwissen et al. (2001). In BGLR (Pérez and de Los Campos, 2014), this model has a mixture density prior with some SNPs having no effect with probability 1-π, while the remaining SNPs have an effect with probability π. The marker effect is sampled from a scaled t-distribution and π is sampled from a beta distribution. The genetic and residual variance are sampled from an inverted chi-square distribution with scale parameter (S_u and S_e) and degree of freedom ($d.f_U$ and $d.f_e$), respectively.

2.8.4. Bayes R

Bayes R was proposed by Erbe et al. (2012) to address the deficiency of Bayes B and to improve computational efficiency. It assumes the SNPs effect comes from a series of normal distributions with variance ranging from zero to 1% of the genetic variance. The residual variance is sampled from an inverted chi-square distribution as described above.

2.9. Genetic correlation between first and second infection

For populations in the year 2015, a bivariate model was used for the estimation of the genetic correlation between the first and second gill-score record.

$$\begin{bmatrix}\mathbf{y_1}\\ \mathbf{y_2}\end{bmatrix} = \begin{bmatrix}\mathbf{1} & 0\\ 0 & \mathbf{1}\end{bmatrix}\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_1 & 0\\ 0 & \mathbf{Z}_2\end{bmatrix}\begin{bmatrix}\mathbf{u_1}\\ \mathbf{u_2}\end{bmatrix} + \begin{bmatrix}\mathbf{e_1}\\ \mathbf{e_2}\end{bmatrix}$$

$$var\begin{bmatrix}\mathbf{u_1}\\ \mathbf{u_2}\end{bmatrix} = \begin{bmatrix}\mathbf{A}\sigma_{u1}^2 & \mathbf{A}\sigma_{u1,2}\\ \mathbf{A}\sigma_{u2,1} & \mathbf{A}\sigma_{u2}^2\end{bmatrix}$$

$$var\begin{bmatrix}\boldsymbol{e_1}\\ \boldsymbol{e_2}\end{bmatrix} = \begin{bmatrix}\sigma_{e1}^2 & \sigma_{e1,2}\\ \sigma_{e2,1} & \sigma_{e2}^2\end{bmatrix}$$

$\mathbf{y_1}$ and $\mathbf{y_2}$ are vectors of phenotypes (gill score) for infections 1 and 2, μ_1 and μ_2 are the overall means of gill scores for infections 1 and 2, $\mathbf{Z_1}$ and $\mathbf{Z_2}$ are incidence matrices relating the phenotypes to the additive genetic effects $\mathbf{u_1}$ and $\mathbf{u_2}$ for infections 1 and 2, respectively, σ_{e1}^2 and σ_{e2}^2 are vectors of residual effects for infections 1 and 2.

σ_{u1}^2 and σ_{u2}^2 are the additive genetic variances for infections 1 and 2, and $\sigma_{u1,2} = \sigma_{u2,1}$ is the additive genetic covariance between both infections, σ_{e1}^2 and σ_{e2}^2 are the residual variances for infections 1 and 2, and $\sigma_{e1,2} = \sigma_{e2,1}$ is the residual covariance between both infections.

2.10. Genomic correlation between first and second infection

This model is similar to the bivariate model described above but with $\mathbf{u_i}$ being replaced with $\mathbf{g_i}$ and **A** being replaced with **G**. **g** is the vector of additive genomic effects and **G** is the genomic relationship matrix.

The 'parameter expanded' expectation maximisation (PX-EM) followed by the "average information" restricted maximum likelihood (AI REML) algorithm of the Wombat software was used to estimate the genetic parameters for both the PBLUP and GBLUP models (Meyer, 2007). To aid convergence of the GBLUP model in Wombat the option –dense was specified at run time. Bayes B and Bayes R model was analyzed with the BGLR package (Pérez and de Los Campos, 2014) and GCTB (Zeng et al., 2018), respectively.

To ensure that the Bayesian models converge to their stationary posterior distribution, trace plots and auto correlation plots were visually inspected. Also, geweke diagnostics was performed using R package coda (Plummer et al., 2006). The number of iterations of the Bayesian models ranged from 100,000 to 800,000 depending on the population and model. 50% of the iterations were discarded as burn-in with the thinning interval ranging from 1 to 5%. These values were chosen based on the convergence diagnostics tests.

2.11. Accuracy and bias

For each of the model (linear and nonlinear) in a population, accuracy was assessed by cross validation, mimicking sib-based GS performed for aquacultural species. A random sample of one fish was selected from each family (families with <5 individuals were excluded), these fish were selected as the selection candidates and thus their phenotype was masked. The remaining fish in the dataset were used to train the model. The process was repeated 50 times. For each run, the accuracy (r) of prediction was estimated by:

$$Accuracy\left(r\right) = \frac{cor(gill\ score(g)ebv)}{\sqrt{h_{pblup}^2}}$$

These estimates were then averaged to give the estimated mean accuracy and the standard deviation of these estimates was taken as the standard error (SE) even though the sample size in each family varied. The variance components derived from analyzing each population's gill score record were utilized in predicting the (g)ebv for the linear models (PBLUP, GBLUP).

Bias was obtained from the regression coefficient of gill score on the (g)ebv. A bias value higher than 1 indicates that the (g)ebv estimates are deflated while a value lower than 1 indicates the (g)ebv is inflated.

$$Bias = \frac{cov\ (gill\ score(g)ebv)}{var((g)ebv)}$$

3. Results

3.1. Descriptive statistics

Table 1 shows the descriptive statistics of gill score records for each population. The number of phenotyped and genotyped fish for each population were more or less the same. The average gill score spanned from 1.52 to 2.69. The distribution of gill score record for each population is shown in Fig. 2. The number of phenotyped fish ranged from 640 to 3114, with YC2016B having the lowest number of fish while YC2017 had the highest number of fish.

3.2. Linkage disequilibrium

LD between marker pairs at various distances is shown in Supplementary file 1: Fig. S1 and Table 2. All populations had similar r^2 value at the various distances examined. LD between marker pairs decreased as distance increased (Supplementary file 1: Fig. S1 and Table 2). The mean LD between adjacent SNPs is also shown in Table 2. All populations had similar r^2 value between adjacent markers. Table 3 has the average r^2 by chromosome for the various populations. The average LD between markers across chromosomes and populations ranged from 0.21 to 0.37. Chromosome 8 had the least amount of LD while chromosome 19 had the highest amount of LD across populations. LD between adjacent markers (by chromosome) was similar across populations.

3.3. Variance component and heritability

Table 4 and Supplementary file 2: Table S1 report on variance components and heritability estimates obtained from the different linear and nonlinear models investigated. The heritability estimates ranged from 0.05 to 0.24, with the lowest estimates (0.06 to 0.09) recorded in the 2015 populations and for the first infections, except for the first infection of YC2015B. The highest heritability estimates (0.17–0.24) were found in the 2016, 2017, and 2018 populations.

The results of the genomic selection (GS) models yielded similar variance components and heritability estimates within each population. The estimates obtained with the pedigree based best linear unbiased prediction (PBLUP) model were also similar to those of the GS models, except for YC2015B-I, where the GS model estimates were lower. For the first infection, the heritability estimates varied from 0.08 to 0.24 with YC2015B and YC2016A having the highest heritability estimate for the PBLUP model. However, for the GS models, YC2016A had the highest heritability estimate. For the second infection, the heritability estimates ranged from 0.06 to 0.09, with YC2015A having the highest value.

Overall, the study found that heritability estimates varied depending on the population, with higher estimates observed in more recent populations. The results also showed that GS and PBLUP models provided similar estimates of variance components and heritability.

3.4. Genetic correlation between first and second infection

Genetic correlation between first and second infection for YC2015A and YC2015B is shown in Table 5. The genetic correlation between the first and second gill score using pedigree information is 0.913 and 0.723 for YC2015A and YC2015B, respectively. The genetic correlation estimates using genomic information were slightly lower (0.81 and 0.714 for YC2015A and YC2015B, respectively) than using pedigree information.

3.5. Accuracy and bias of prediction

The within population prediction accuracy for the various models and AGD incidence are shown in Table 6. The within population accuracy ranged from 0.32 to 0.77. For each population, the accuracies with GS models were higher than the PBLUP estimates. GS models did not differ significantly from each other. In general, all bias estimates were close to 1.

4. Discussion

To facilitate selective breeding, it is important to understand the genetic basis of AGD resistance and develop accurate genetic evaluation models. This study uses both pedigree and genome assisted models to estimate the genetic parameters for AGD resistance and evaluate the accuracy of different evaluation models, including genomic-based linear and nonlinear models.

4.1. Linkage disequilibrium and decay

We have explored for the first time, the extent of linkage disequilibrium present in different populations of Atlantic salmon in Norway, which is important for successful implementation of genomic selection. The linkage disequilibrium observed between adjacent markers for each population studied was about 0.3 (Table 2). Meuwissen et al. (2001), using haplotypes to predict the additive genetic value in the concept paper on GS, achieved an accuracy of 0.85 with microsatellite markers 1 centimorgan apart, which is approximately the same as having SNP markers with LD (measured by r^2) of 0.2 (Calus et al., 2008). This indicates that the density of markers applied in this study should be sufficient for accurate prediction of additive genetic values. The average LD values between adjacent markers observed in our study population were generally higher than what was previously reported by Barria et al. (2018), where LD estimate ranged from 0.07 to 0.26 in different Atlantic salmon populations of North American and European origin. In contrast, Kijas et al. (2017) reported LD for markers at different physical distances in Tasmanian and Finnish Atlantic salmon populations. Comparing LD values from different studies is however not straight forward because marker types, measures of LD, manners of obtaining haplotypes, ways of reporting (with some studies reporting LD at certain physical distance) and so on may not be the same (Pritchard and Przeworski, 2001). In addition, the density of markers will affect the LD values obtained, with higher LD values between adjacent markers observed at higher densities

Table 1
Descriptive statistics of the data set.

	Phenotyped fish	No of families	Mean (SD)	Median	Range (min-max)	Genotyped fish
YC2015A-I	2464	273	2.06 (1.02)	2	0–4	2464
YC2015A-II	2316	269	2.69 (1.00)	3	0–5	2316
YC2015B-I	1049	116	1.87 (1.04)	2	0–5	1049
YC2015B-II	948	117	2.48 (1.05)	3	0–5	948
YC2016A	2006	180	2.35 (1.35)	2	0–5	2006
YC2016B	640	70	1.91 (1.29)	2	0–5	640
YC2017	3114	275	1.52 (1.31)	1	0–5	2911
YC2018	3033	139	1.52 (1.18)	1	0–5	2949

SD = Standard Deviation, min = minimum, max = maximum.

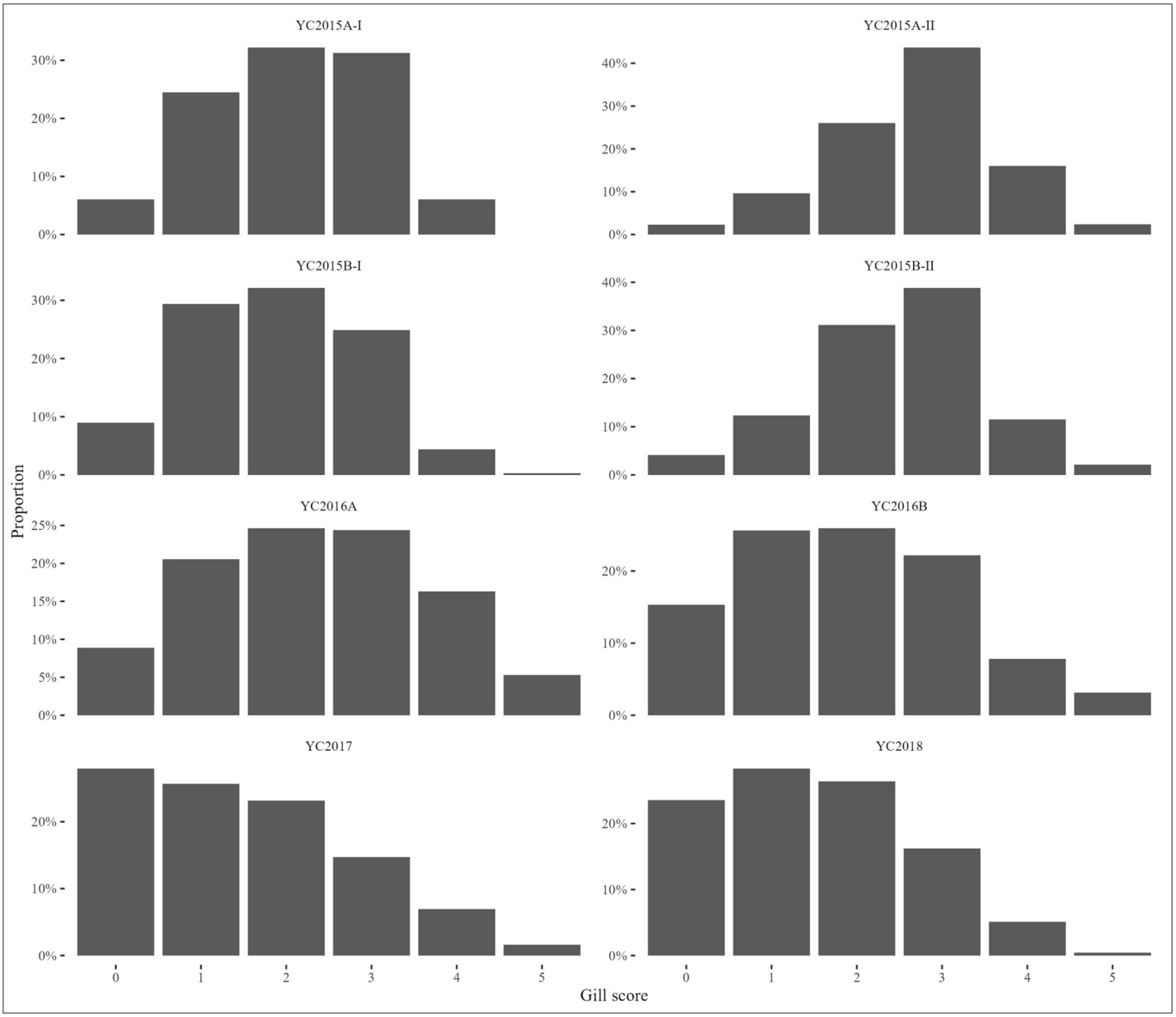

Fig. 2. Distribution of gill scores of the various populations.

Table 2
Average r^2 at various distances in six populations.

	0 - 100kp	100 - 200 kb	200 -300 kb	300 -400 kb	400 -500 kb	adjacent SNPs	Total mean r^2
YC2015A	0.22	0.14	0.12	0.12	0.11	0.31	0.13
YC2015B	0.21	0.12	0.11	0.09	0.09	0.29	0.11
YC2016A	0.21	0.12	0.11	0.09	0.09	0.29	0.11
YC2016B	0.19	0.11	0.09	0.08	0.07	0.28	0.09
YC2017	0.19	0.11	0.09	0.09	0.08	0.27	0.09
YC2018	0.21	0.13	0.11	0.11	0.10	0.29	0.12

of markers (Larmer et al., 2014). Our study and that of Kijas et al. (2017) also differ in marker density used and thus difficult to compare. Moreso, LD values are population specific and factors such as selection, mutation, genetic drift, admixture/migration and recombination shape the LD observed in a population (Ardlie et al., 2002). As explained earlier, the Atlantic salmon populations studied have a history of admixture (see materials section). Admixture leads to linkage disequilibrium between linked and unlinked loci (spurious LD), but the spurious LD dissipates after a few generations. However, LD induced by admixture between linked loci can persist for several generations (Stephens et al., 1994), and recent admixture history can result in LD extending over several megabases (Al-Tobasei et al., 2021; Vallejo et al., 2018). We indeed observed strong LD at distances spanning several megabases (Supplementary file 1: Figs. S2 - S7), indicating that accurate genomic prediction can be performed with even low-density markers for these populations. While dense markers are not necessary for significant association with QTL, long-range LD makes fine mapping of the causal variant challenging, as markers far from the QTL can also be associated with it.

The similar pattern of LD observed across populations (Tables 2 & 3, Supplementary file 1: Figs. S2 - S7) probably reflect similar population history or similar historical LD generating events and/or similar forces of selection acting on the populations. The effect of selection on LD depends on the intensity of selection, the direction, duration and consistency (Du et al., 2007).

Variable recombination rates exist across the genome (Ardlie et al., 2002; Barria et al., 2018; Yu et al., 2001), and this might affect the LD seen on different chromosomes as LD is inversely related to

Table 3
Summary statistics for the evaluated SNPs and average linkage disequilibrium values (by chromosome) of different Atlantic salmon populations.

Chromosome	Average[1] number of SNPs	Average[1] SNP distance (kbp)	Average linkage disequilibrium values					
			YC2015A	YC2015B	YC2016A	YC2016B	YC2017	YC2018
1	3712	41	0.31	0.30	0.29	0.27	0.28	0.30
2	1394	47	0.28	0.24	0.25	0.22	0.22	0.26
3	2383	37	0.29	0.3	0.28	0.27	0.27	0.29
4	2029	39	0.31	0.28	0.28	0.26	0.26	0.29
5	1983	38	0.28	0.27	0.28	0.25	0.25	0.28
6	2016	41	0.27	0.25	0.25	0.24	0.23	0.25
7	1378	39	0.26	0.25	0.24	0.24	0.23	0.24
8	454	55	0.23	0.24	0.21	0.21	0.21	0.22
9	2695	50	0.32	0.32	0.30	0.29	0.30	0.34
10	2665	42	0.32	0.32	0.32	0.3	0.28	0.32
11	1842	49	0.29	0.30	0.28	0.28	0.26	0.29
12	2054	44	0.27	0.28	0.27	0.26	0.26	0.28
13	2583	41	0.36	0.32	0.34	0.32	0.32	0.34
14	2264	41	0.35	0.33	0.32	0.31	0.31	0.34
15	2000	50	0.32	0.31	0.29	0.30	0.29	0.31
16	1658	48	0.32	0.30	0.30	0.29	0.28	0.30
17	1177	45	0.28	0.26	0.25	0.26	0.24	0.27
18	1434	47	0.33	0.29	0.29	0.28	0.28	0.32
19	1618	49	0.37	0.35	0.35	0.35	0.33	0.36
20	1875	45	0.35	0.32	0.34	0.31	0.31	0.32
21	1230	45	0.29	0.27	0.27	0.26	0.26	0.27
22	1465	41	0.32	0.30	0.29	0.30	0.28	0.32
23	1450	34	0.27	0.27	0.26	0.25	0.25	0.26
24	1200	37	0.33	0.33	0.29	0.30	0.27	0.29
25	1154	43	0.33	0.29	0.31	0.28	0.27	0.30
26	1016	45	0.31	0.27	0.29	0.25	0.25	0.29
27	1222	35	0.31	0.28	0.29	0.27	0.27	0.30
28	991	39	0.29	0.27	0.27	0.26	0.26	0.26
29	892	46	0.27	0.27	0.25	0.25	0.23	0.26

[1] average number per chromosome was averaged across populations.

Table 4
Genetic parameters of gill scores using pedigree or genomic information.

	PBLUP			GBLUP		
	Va ± SE	Ve ± SE	h^2 ± SE	Va ± SE	Ve ± SE	h^2 ± SE
YC2015A-I	0.08 ± 0.03	0.96 ± 0.04	0.08 ± 0.03	0.07 ± 0.02	0.97 ± 0.03	0.07 ± 0.02
YC2015A-II	0.09 ± 0.03	0.92 ± 0.03	0.09 ± 0.03	0.07 ± 0.02	0.93 ± 0.03	0.07 ± 0.02
YC2015B-I	0.26 ± 0.07	0.83 ± 0.06	0.24 ± 0.06	0.13 ± 0.04	0.93 ± 0.05	0.12 ± 0.03
YC2015B-II	0.06 ± 0.04	1.04 ± 0.06	0.06 ± 0.04	0.06 ± 0.03	1.04 ± 0.05	0.05 ± 0.03
YC2016A	0.43 ± 0.09	1.40 ± 0.08	0.24 ± 0.05	0.43 ± 0.07	1.42 ± 0.06	0.23 ± 0.03
YC2016B	0.29 ± 0.14	1.39 ± 0.14	0.17 ± 0.08	0.33 ± 0.11	1.36 ± 0.11	0.19 ± 0.06
YC2017	0.36 ± 0.07	1.35 ± 0.06	0.21 ± 0.04	0.36 ± 0.05	1.34 ± 0.43	0.21 ± 0.03
YC2018	0.28 ± 0.06	1.14 ± 0.05	0.19 ± 0.04	0.30 ± 0.04	1.12 ± 0.03	0.21 ± 0.03

Va = additive genetic variance, Ve = residual variance, h^2 = heritability.

Table 5
Genetic and phenotypic correlation between repeated gill scores, first and second infection, on the same individuals for the 2015 populations.

		rp ± SE		rg ± SE	
		Pedigree	Genomic	Pedigree	Genomic
Populations	Incidence	II		II	
YC2015A	I	0.227 ± 0.021	0.220 ± 0.020	0.913 ± 0.129	0.810 ± 0.130
YC2015B	I	0.225 ± 0.033	0.229 ± 0.033	0.723 ± 0.262	0.714 ± 0.255

rg = genetic correlation, rp = phenotypic correlation, I = first infection, II = second infection.

recombination rate, regions with high recombination rate are usually referred to as "recombination jungle" whereas regions with low recombination rate are commonly referred to as "recombination desert" (Yu et al., 2001). To investigate this, we computed the average LD between adjacent markers on each chromosome for the different populations (Table 3). We observed variable LD across the genome (chromosomes), with chromosome 19 consistently having the highest LD across populations, making chromosome 19 a recombination desert. In contrast, chromosome 8 had the lowest LD across populations, indicating that it is a recombination jungle, but this could also be a result of low density of SNP on that chromosome (Table 3). Nonetheless, since the amount of LD on chromosome 8 is still higher than 0.2, we rather refer to it as a "recombination grassland".

Barria et al. (2018) suggested that LD values may be inflated due to high relatedness within the studied population. We examined the impact of relatedness on LD in our study by randomly sampling one fish from each family and then estimating LD using these individuals. The results from this analysis (not shown) did not differ from using the whole dataset.

4.2. Variance component and heritability

Knowledge of genetic parameters is important for the prediction of genetic gain (Gianola and Rosa, 2015). Our pedigree based heritability estimate varied from 0.06 to 0.24. Lillehammer et al. (2019) and Gjerde et al. (2019) also reported estimates within this range for AGD resistance in different Norwegian Atlantic salmon populations. However, these estimates are in general lower than those reported from Tasmania (Kube et al., 2012; Taylor et al., 2009a), which could be as a result of different genetic makeup (different loci segregating in different populations). Moreso, in Tasmania, major gill scoring in the breeding program takes place when the level of infection exceeds the normal level used in commercial practise (Kube et al., 2012; Taylor et al., 2009a; Taylor et al., 2009b), whereas in Norway, major gill scoring takes place when

Table 6
Accuracy and bias of prediction.

	PBLUP		GBLUP		Bayes B		Bayes R	
	Mean Acc. ± SE	Mean Bias ± SE	Mean Acc. ± SE	Mean Bias ± SE	Mean Acc. ± SE	Mean Bias ± SE	Mean Acc. ± SE	Mean Bias ± SE
YC2015A-I	0.39 ± 0.23	1.08 ± 0.66	0.51 ± 0.22	0.98 ± 0.44	0.51 ± 0.25	0.97 ± 0.56	0.48 ± 0.20	0.95 ± 0.45
YC2015A-II	0.48 ± 0.17	1.07 ± 0.41	0.54 ± 0.19	1.03 ± 0.36	0.54 ± 0.19	1.02 ± 0.43	0.41 ± 0.19	0.75 ± 0.36
YC2015B-I	0.48 ± 0.15	0.92 ± 0.30	0.51 ± 0.16	0.96 ± 0.33	0.50 ± 0.16	0.94 ± 0.34	0.51 ± 0.18	0.94 ± 0.36
YC2015B-II	0.32 ± 0.39	1.05 ± 1.29	0.40 ± 0.35	0.99 ± 0.89	0.39 ± 0.35	0.93 ± 0.86	0.60 ± 0.39	1.15 ± 0.77
YC2016A	0.47 ± 0.12	0.95 ± 0.28	0.56 ± 0.12	0.91 ± 0.21	0.56 ± 0.12	0.90 ± 0.23	0.56 ± 0.12	0.91 ± 0.22
YC2016B	0.47 ± 0.38	1.13 ± 0.94	0.54 ± 0.38	1.05 ± 0.78	0.54 ± 0.38	1.05 ± 0.80	0.53 ± 0.38	0.99 ± 0.76
YC2017	0.59 ± 0.12	1.01 ± 0.24	0.73 ± 0.12	1.01 ± 0.19	0.74 ± 0.12	0.99 ± 0.19	0.73 ± 0.12	0.99 ± 0.19
YC2018	0.56 ± 0.15	0.99 ± 0.28	0.76 ± 0.14	1.02 ± 0.21	0.77 ± 0.14	1.02 ± 0.22	0.77 ± 0.14	1.01 ± 0.21

Acc. = accuracy.

there is a moderate level of infection in the field or challenge setting. Kube et al. (2012) indicated that higher heritabilities are usually found when there is a high incidence of AGD in the population. As mentioned earlier, in the 2015 populations in our study, gill scoring was recorded twice. For YC2015A, heritability estimates of the first infection was marginally lower than the second infection, whereas for YC2015B, estimated heritability of the second infection was significantly lower than the first infection. The latter is in close agreement with Lillehammer et al. (2019) who found a similar trend. Other authors have reported higher heritability estimates for the second infection when compared to the first infection (Kube et al., 2012; Taylor et al., 2009a). It is important to note that while AGD was recorded in the months of September and October for all populations in the year 2015, it was recorded in the month of November for the other populations (YC2016-YC2018), hence, the relatively high heritability we observed for these populations (YC2016-YC2018) may be indicative of these populations having a previous episode of AGD infection given that higher heritability is usually observed for subsequent infections (Kube et al., 2012).

4.3. Genetic correlation between first and second infection

Conventionally, the first and subsequent infections are considered as separate traits, probably because of the study by Kube et al. (2012) who found a generally low correlation between the first and subsequent infections. The first infection is thought to elicit an innate response while subsequent infections elicit an adaptive response (Kube et al., 2012). However, a high genetic correlation between the first and second infection have been found in most studies (Kube et al., 2012; Lillehammer et al., 2019; Taylor et al., 2009a), which also correspond to the findings in our study (YC2015A: 0.91 PBLUP, 0.81 GBLUP; YC2015B: 0.723 PBLUP, 0.714 GBLUP).

4.4. Within population accuracy

The prediction accuracy estimates in our study ranged from 0.32 to 0.77, corresponding to similar estimates by Aslam et al. (2020). Better prediction accuracy of GS models, compared to PBLUP, is generally observed (Robledo et al., 2018; Tsai et al., 2016; Vallejo et al., 2017), although in some cases, the better performance of the GS model is marginal (Aslam et al., 2020; Tsai et al., 2016), particularly when the informants population size is small. Other factors that influence prediction accuracy include choice of model, heritability of the trait, relationship between the validation set and informants, and LD between markers and QTL (Calus, 2010; Hayes et al., 2009).

Daetwyler et al. (2010) showed with simulated data that when a trait is controlled by a small number of QTLs (i.e not polygenic), nonlinear model performed better than GBLUP model. Joshi et al. (2021) observed better performance of nonlinear GS models with *Streptococcus agalactiae* challenge-test data in Nile Tilapia. However, in our study we observed in most cases similar performance for all GS models, which could indicate that AGD resistance is a polygenic trait. We have no obvious reason for the poorer performance of the Bayes R model for YC2015A-II and its exceptional performance for YC2015B-II, this however highlights the importance of validating the accuracy of GS models in independent data sets.

Since SNPs may not explain all the genetic variance, including a polygenic effect in the model has been suggested to account for background genes not explained by the SNPs (Meuwissen et al., 2013). The fact that GBLUP in general showed a similar genetic variance as PBLUP in our study indicates that the SNP markers managed to capture the genetic variance and hence there is no need to fit a polygenic effect in our model.

The heritability of a trait has been reported to have an effect on prediction accuracy (Hayes et al., 2009). Low accuracy has been observed for traits with low heritability (Hayes et al., 2009; Sukhavachana et al., 2020). In order to achieve a high accuracy for traits with low heritability, large informants population are required (Hayes et al., 2009). In addition, the size of the informants population affects how accurately SNPs effect can be estimated, and this is particularly important when a large number of QTLs with small effect controls the trait (Hayes et al., 2009). Accordingly, we observed that the population with a low sample size and heritability (YC2015B-II) had the lowest prediction accuracy with large standard errors (YC2016B included). Meuwissen et al. (2001) reported that sampling errors of the estimated marker effect will limit the accuracy of prediction and these sampling errors increases as heritability decreases. They posit this can be solved with higher sample size. We observed a reduced standard error when the sample size was high even with low heritability for some populations.

The relationship between our informants and validation set reflects what is commonly used in most aquacultural breeding programmes. This is relevant because the accuracy of prediction depends in part on the relationship between the informants and validation set (Fraslin et al., 2022; Habier et al., 2007), which is favorable for the sib-testing scheme used. Moreso, Daetwyler et al. (2013) recommended that when evaluating the prediction accuracy of genomic evaluation models, it should mimic how it is used in practice.

Please note that evaluating the performance of genetic evaluation models using accuracy is not without its drawbacks as the estimates will be affected by how accurately the variance components are estimated.

5. Conclusions

Our results suggest that accurate genomic prediction can be achieved even with relatively low marker density. We also confirm the existence of genetic variation for AGD resistance in the studied populations, which hence can be improved by selective breeding. In the majority of cases we studied, there was no significant difference between the GS models, indicating that resistance to AGD is likely polygenic. Nonlinear models have a higher computing requirement, hence in the genetic evaluation of AGD resistance, we recommend the use of the GBLUP model.

Funding

This research has been performed in the scope of the EATFISH

project, which received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 956697.

CRediT authorship contribution statement

Afees Abiola Ajasa: Formal analysis, Methodology, Software, Writing – original draft. **Solomon Antwi Boison:** Conceptualization, Supervision, Software, Methodology, Writing – review & editing. **Hans Magnus Gjøen:** Supervision, Methodology, Writing – review & editing. **Marie Lillehammer:** Conceptualization, Supervision, Methodology, 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.

Data availability

The authors do not have permission to share data.

Acknowledgements

Thanks to Mowi Genetics AS for providing the data used in this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aquaculture.2023.740078.

References

Al-Tobasei, R., Ali, A., Garcia, A.L., Lourenco, D., Leeds, T., Salem, M., 2021. Genomic predictions for fillet yield and firmness in rainbow trout using reduced-density SNP panels. BMC Genomics 22, 1–11.

Ardlie, K.G., Kruglyak, L., Seielstad, M., 2002. Patterns of linkage disequilibrium in the human genome. Nat. Rev. Genet. 3, 299–309.

Aslam, M.L., Boison, S.A., Lillehammer, M., Norris, A., Gjerde, B., 2020. Genome-wide association mapping and accuracy of predictions for amoebic gill disease in Atlantic salmon (Salmo salar). Sci. Rep. 10, 6435. https://doi.org/10.1038/s41598-020-63423-8.

Barria, A., López, M.E., Yoshida, G., Carvalheiro, R., Lhorente, J.P., Yáñez, J.M., 2018. Population genomic structure and genome-wide linkage disequilibrium in farmed Atlantic salmon (Salmo salar L.) using dense SNP genotypes. Front. Genet. 9, 649.

Browning, B.L., Browning, S.R., 2016. Genotype imputation with millions of reference samples. Am. J. Hum. Genet. 98, 116–126.

Bustos, P.A., Young, N.D., Rozas, M.A., Bohle, H.M., Ildefonso, R.S., Morrison, R.N., Nowak, B.F., 2011. Amoebic gill disease (AGD) in Atlantic salmon (Salmo salar) farmed in Chile. Aquaculture 310, 281–288. https://doi.org/10.1016/j.aquaculture.2010.11.001.

Calus, M.P., 2010. Genomic breeding value prediction: methods and procedures. Animal 4, 157–164.

Calus, M., Meuwissen, T., De Roos, A., Veerkamp, R., 2008. Accuracy of genomic selection using different methods to define haplotypes. Genetics 178, 553–561.

Daetwyler, H.D., Pong-Wong, R., Villanueva, B., Woolliams, J.A., 2010. The impact of genetic architecture on genome-wide evaluation methods. Genetics 185, 1021–1031.

Daetwyler, H.D., Calus, M.P., Pong-Wong, R., de Los Campos, G., Hickey, J.M., 2013. Genomic prediction in animals and plants: simulation of data, validation, reporting, and benchmarking. Genetics 193, 347–365.

Du, F.-X., Clutter, A.C., Lohuis, M.M., 2007. Characterizing linkage disequilibrium in pig populations. Int. J. Biol. Sci. 3, 166.

Erbe, M., Hayes, B., Matukumalli, L., Goswami, S., Bowman, P., Reich, C., Mason, B., Goddard, M., 2012. Improving accuracy of genomic predictions within and between dairy cattle breeds with imputed high-density single nucleotide polymorphism panels. J. Dairy Sci. 95, 4114–4129.

Evans, B., Kube, P., Taylor, R., Elliott, N., 2015. Genetic gains achieved over 10 years of selective breeding for resistance to amoebic gill disease in Atlantic Salmon (*Salmo salar*). In: Proceedings of the International Symposium on Genetics in Aquaculture xii. Santiago de Compostela, p. 2015.

Fraslin, C., Yáñez, J.M., Robledo, D., Houston, R.D., 2022. The impact of genetic relationship between training and validation populations on genomic prediction accuracy in Atlantic salmon. Aquac. Rep. 23, 101033.

Gianola, D., Rosa, G.J., 2015. One hundred years of statistical developments in animal breeding. Annu. Rev. Anim. Biosci. 3, 19–56.

Gjedrem, T., 2010. The first family-based breeding program in aquaculture. Rev. Aquac. 2, 2–15.

Gjerde, B., Boison, S.A., Aslam, M.L., Løvoll, M., Bakke, H., Rey, S., Lillehammer, M., 2019. Estimates of genetic correlations between susceptibility of Atlantic salmon to amoebic gill disease in a bath challenge test and a field test. Aquaculture 511, 734265. https://doi.org/10.1016/j.aquaculture.2019.734265.

Gjøen, H.M., Refstie, T., Ulla, O., Gjerde, B., 1997. Genetic correlations between survival of Atlantic salmon in challenge and field tests. Aquaculture 158, 277–288.

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

Goddard, M.E., Hayes, B.J., 2007. Genomic selection. J. Anim. Breed. Genet. 124, 323–330. https://doi.org/10.1111/j.1439-0388.2007.00702.x.

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

Hayes, B.J., Bowman, P.J., Chamberlain, A.J., Goddard, M.E., 2009. Invited review: genomic selection in dairy cattle: Progress and challenges. J. Dairy Sci. 92, 433–443.

Hill, W., Robertson, A., 1968. Linkage disequilibrium in finite populations. Theor. Appl. Genet. 38, 226–231.

Joshi, R., Skaarud, A., Alvarez, A.T., Moen, T., Ødegård, J., 2021. Bayesian genomic models boost prediction accuracy for survival to Streptococcus agalactiae infection in Nile tilapia (Oreochromus nilioticus). Genet. Sel. Evol. 53, 1–10.

Kemper, K.E., Reich, C.M., Bowman, P.J., Vander Jagt, C.J., Chamberlain, A.J., Mason, B. A., Hayes, B.J., Goddard, M.E., 2015. Improved precision of QTL mapping using a nonlinear Bayesian method in a multi-breed population leads to greater accuracy of across-breed genomic predictions. Genet. Sel. Evol. 47, 1–17.

Kent, M.L., Sawyer, T., Hedrick, R., 1988. Paramoeba pemaquidensis (Sarcomastigophora: Paramoebidae) infestation of the gills of coho salmon Oncorhynchus kisutch reared in sea water. Dis. Aquat. Org. 5, 163–169.

Kijas, J., Elliot, N., Kube, P., Evans, B., Botwright, N., King, H., Primmer, C., Verbyla, K., 2017. Diversity and linkage disequilibrium in farmed Tasmanian Atlantic salmon. Anim. Genet. 48, 237–241.

Kube, P.D., Taylor, R.S., Elliott, N.G., 2012. Genetic variation in parasite resistance of Atlantic salmon to amoebic gill disease over multiple infections. Aquaculture 364, 165–172. https://doi.org/10.1016/j.aquaculture.2012.08.026.

Larmer, S., Sargolzaei, M., Schenkel, F., 2014. Extent of linkage disequilibrium, consistency of gametic phase, and imputation accuracy within and across Canadian dairy breeds. J. Dairy Sci. 97, 3128–3141.

Lillehammer, M., Boison, S.A., Norris, A., Løvoll, M., Bakke, H., Gjerde, B., 2019. Genetic parameters of resistance to amoebic gill disease in two Norwegian Atlantic salmon populations. Aquaculture 508, 83–89. https://doi.org/10.1016/j.aquaculture.2019.04.072.

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. https://doi.org/10.1093/genetics/157.4.1819.

Meuwissen, T., Hayes, B., Goddard, M., 2013. Accelerating improvement of livestock with genomic selection. Annu. Rev. Anim. Biosci. 1 (1), 221–237. https://doi.org/10.1146/annurev-animal-031412-103705.

Meyer, K., 2007. WOMBAT—A tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). J Zhejiang Univ Sci B 8, 815–821.

Munday, B., 1985. Diseases of salmonids. In: Proceedings of the Workshop on Diseases of Australian Fish and Shellfish. Benalla.

Munday, B., Foster, C., Roubal, F., Lester, R., 1990. Paramoebic gill infection and associated pathology of Atlantic salmon, Salmo salar, and rainbow trout, Salmo gairdneri, in Tasmania. In: Perkins, F.O., C.T (Eds.), Pathology in Marine Science. Academic Press, pp. 215–222.

Neves, H.H., Carvalheiro, R., O'brien, A.M.P., Utsunomiya, Y.T., Do Carmo, A.S., Schenkel, F.S., Sölkner, J., McEwan, J.C., Van Tassell, C.P., Cole, J.B., 2014. Accuracy of genomic predictions in Bos indicus (Nellore) cattle. Genet. Sel. Evol. 46, 1–13.

Norris, A., Bradley, D., Cunningham, E., 1999. Microsatellite genetic variation between and within farmed and wild Atlantic salmon (Salmo salar) populations. Aquaculture 180, 247–264.

Norris, A., Foyle, L., Ratcliff, J., 2008. Heritability of mortality in response to a natural pancreas disease (SPDV) challenge in Atlantic salmon, Salmo salar L., post-smolts on a west of Ireland Sea site. J. Fish Dis. 31, 913–920. https://doi.org/10.1111/j.1365-2761.2008.00982.x.

Odegard, J., Olesen, I., Gjerde, B., Klemetsdal, G., 2006. Evaluation of statistical models for genetic analysis of challenge test data on furunculosis resistance in Atlantic salmon (Salmo salar): prediction of field survival. Aquaculture 259, 116–123. https://doi.org/10.1016/j.aquaculture.2006.05.034.

Oldham, T., Rodger, H., Nowak, B.F., 2016. Incidence and distribution of amoebic gill disease (AGD) — an epidemiological review. Aquaculture 457, 35–42. https://doi.org/10.1016/j.aquaculture.2016.02.013.

Parsons, H., Nowak, B., Fisk, D., Powell, M., 2001. Effectiveness of commercial freshwater bathing as a treatment against amoebic gill disease in Atlantic salmon. Aquaculture 195, 205–210.

Pérez, P., de Los Campos, G., 2014. Genome-wide regression and prediction with the BGLR statistical package. Genetics 198, 483–495.

Plummer, M., Best, N., Cowles, K., Vines, K., 2006. CODA: convergence diagnosis and output analysis for MCMC. R News 6, 7–11.

Powell, M., Nowak, B., Adams, M., 2002. Cardiac morphology in relation to amoebic gill disease history in Atlantic salmon, Salmo salar L. J. Fish Dis. 25, 209–215.

Pritchard, J.K., Przeworski, M., 2001. Linkage disequilibrium in humans: models and data. Am. J. Hum. Genet. 69, 1–14.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., De Bakker, P.I., Daly, M.J., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575.

Robledo, D., Matika, O., Hamilton, A., Houston, R.D., 2018. Genome-wide association and genomic selection for resistance to amoebic gill disease in Atlantic salmon. G3 (Bethesda) 8, 1195–1203. https://doi.org/10.1534/g3.118.200075.

Rodger, H., 2014. Amoebic gill disease (AGD) in farmed salmon (Salmo salar) in Europe. Fish. Vet. J. 14, 16–26.

Steinum, T., Kvellestad, A., Ronneberg, L.B., Nilsen, H., Asheim, A., Fjell, K., Nygard, S. M.R., Olsen, A.B., Dale, O.B., 2008. First cases of amoebic gill disease (AGD) in Norwegian seawater farmed Atlantic salmon, Salmo salar L., and phylogeny of the causative amoeba using 18S cDNA sequences. J. Fish Dis. 31, 205–214. https://doi.org/10.1111/j.1365-2761.2007.00893.x.

Stephens, J.C., Briscoe, D., O'Brien, S.J., 1994. Mapping by admixture linkage disequilibrium in human populations: limits and guidelines. Am. J. Hum. Genet. 55, 809.

Storset, A., Strand, C., Wetten, M., Sissel, K., Rarnstad, A., 2007. Response to selection for resistance against infectious pancreatic necrosis in Atlantic salmon (Salmo salar L.). Aquaculture 272, S62–S68. https://doi.org/10.1016/j.aquaculture.2007.08.011.

Strandén, I., Garrick, D., 2009. Derivation of equivalent computing algorithms for genomic predictions and reliabilities of animal merit. J. Dairy Sci. 92, 2971–2975.

Sukhavachana, S., Tongyoo, P., Massault, C., McMillan, N., Leungnaruemitchai, A., Poompuang, S., 2020. Genome-wide association study and genomic prediction for resistance against Streptococcus agalactiae in hybrid red tilapia (Oreochromis spp.). Aquaculture 525, 735297.

Taylor, R.S., Kube, P.D., Muller, W.J., Elliott, N.G., 2009a. Genetic variation of gross gill pathology and survival of Atlantic salmon (Salmo salar L.) during natural amoebic gill disease challenge. Aquaculture 294, 172–179. https://doi.org/10.1016/j.aquaculture.2009.06.007.

Taylor, R.S., Muller, W.J., Cook, M.T., Kube, P.D., Elliott, N.G., 2009b. Gill observations in Atlantic salmon (Salmo salar, L.) during repeated amoebic gill disease (AGD) field exposure and survival challenge. Aquaculture 290, 1–8. https://doi.org/10.1016/j.aquaculture.2009.01.030.

Team, R.C., 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Tsai, H.-Y., Hamilton, A., Tinch, A.E., Guy, D.R., Gharbi, K., Stear, M.J., Matika, O., Bishop, S.C., Houston, R.D., 2015. Genome wide association and genomic prediction for growth traits in juvenile farmed Atlantic salmon using a high density SNP array. BMC Genomics 16, 1–9.

Tsai, H.-Y., Hamilton, A., Tinch, A.E., Guy, D.R., Bron, J.E., Taggart, J.B., Gharbi, K., Stear, M., Matika, O., Pong-Wong, R., 2016. Genomic prediction of host resistance to sea lice in farmed Atlantic salmon populations. Genet. Sel. Evol. 48, 1–11.

Våge, R., 1995. Avlsarbeidet til NLA og A/S MOWI. In: Masteroppgave. Høgskolen i Molde, p. 43.

Vallejo, R.L., Leeds, T.D., Gao, G., Parsons, J.E., Martin, K.E., Evenhuis, J.P., Fragomeni, B.O., Wiens, G.D., Palti, Y., 2017. Genomic selection models double the accuracy of predicted breeding values for bacterial cold water disease resistance compared to a traditional pedigree-based model in rainbow trout aquaculture. Genet. Sel. Evol. 49, 1–13.

Vallejo, R.L., Silva, R.M., Evenhuis, J.P., Gao, G., Liu, S., Parsons, J.E., Martin, K.E., Wiens, G.D., Lourenco, D.A., Leeds, T.D., 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.

VanRaden, P.M., 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91, 4414–4423.

Verbyla, K.L., Kube, P.D., Evans, B.S., 2022. Commercial implementation of genomic selection in Tasmanian Atlantic salmon: scheme evolution and validation. Evol. Appl. 15, 631–644.

Young, N.D., Dyková, I., Snekvik, K., Nowak, B.F., Morrison, R.N., 2008. Neoparamoeba perurans is a cosmopolitan aetiological agent of amoebic gill disease. Dis. Aquat. Org. 78, 217–223. https://doi.org/10.3354/dao01869.

Yu, A., Zhao, C., Fan, Y., Jang, W., Mungall, A.J., Deloukas, P., Olsen, A., Doggett, N.A., Ghebranious, N., Broman, K.W., 2001. Comparison of human genetic and sequence-based physical maps. Nature 409, 951–953.

Zeng, J., De Vlaming, R., Wu, Y., Robinson, M.R., Lloyd-Jones, L.R., Yengo, L., Yap, C.X., Xue, A., Sidorenko, J., McRae, A.F., 2018. Signatures of negative selection in the genetic architecture of human complex traits. Nat. Genet. 50, 746–753.

Zhu, B., Guo, P., Wang, Z., Zhang, W., Chen, Y., Zhang, L., Gao, H., Wang, Z., Gao, X., Xu, L., 2019. Accuracies of genomic prediction for twenty economically important traits in Chinese Simmental beef cattle. Anim. Genet. 50, 634–643.