Genome-wide association and genomic selection for resistance to Amoebic Gill Disease in Atlantic salmon

D. Robledo[1], O. Matika[1], A. Hamilton[2] & R.D. Houston[1]

[1] The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, EH25 9RG Midlothian, United Kingdom
ross.houston@roslin.ed.ac.uk (Corresponding Author)
[2] Landcatch Natural Selection Ltd., 15 Beta Centre, Stirling University Innovation Park, Stirling, United Kingdom

Summary

Amoebic gill disease (AGD) is one of the largest threats to salmon aquaculture. Although initially limited to Tasmania, AGD is now observed in salmon production worldwide and causes serious economic and animal welfare burden. Treatments are expensive and can be environmentally damaging, hence the need for alternative strategies. Breeding for disease resistance can contribute to control of AGD, providing long-term cumulative benefits in selected stocks. The use of genomic selection can expedite selection for disease resistance, since selection accuracy is typically improved compared to pedigree-based approaches. The aim of this work was to quantify and characterise genetic variation in AGD resistance in salmon, the architecture of the trait, and the potential of genomic selection as a contributor to disease control. An AGD challenge was performed in ~1,500 Atlantic salmon post-smolts, using gill damage (mean gill score) and amoebic load (qPCR) as indicator traits. Our findings show that both traits are heritable and are positively correlated, indicating they may be good measurements of host resistance to AGD. Heritability of resistance was moderate in magnitude (h2 ~0.25-0.30) and the trait was largely polygenic in nature. However, two regions on chromosome 18 were identified with a suggestive association with both AGD resistance traits, and merit independent verification and potentially functional studies to identify underlying genes. Accuracy of prediction for both traits using genetic markers was up to 18 % higher than using pedigree alone, and reduction in marker density to 2,000 SNPs was sufficient to obtain the accuracy advantage over pedigree records. This work shows that resistance to AGD is a suitable trait for genomic selection, and the addition of this trait to Atlantic salmon breeding programs can lead to more resistant stocks.

Keywords: aquaculture, amoebic gill disease, Salmo salar, disease resistance

Introduction

Salmonids are a high-value group of fish and represent 16.6% of global fish trade in 2013 (FAO 2016). Demand for salmonid fish continues to grow steadily, and Atlantic salmon is the salmonid species with the highest production globally (FAO 2016). However, infectious disease outbreaks are a major threat to sustainable production and expansion of salmon aquaculture. In particular, parasitic diseases and their treatments represent a huge financial burden on the industry, and a major animal welfare issue. Amoebic gill disease (AGD), caused by the parasite *Neoparamoeba perurans*, has been a perennial problem for salmon aquaculture in Australia, and outbreaks have become increasingly frequent in European

salmon farms. The disease can impair growth and cause severe mobidity and mortality if untreated. Current treatment strategies are crude, laborious and potentially environmentally damaging. Therefore, alternatives that help limit the impact of AGD are highly desirable. One such method is improving the resistance of farmed salmon stocks to this disease via selective breeding, the benefits of which can be cumulative and long-term. The overall aims of this study were to estimate genetic parameters of AGD resistance in farmed Atlantic salmon, the genetic architecture of the trait, and to test genomic prediction with a view to maximising the effectiveness of genomic selection within commercial breeding programs.

Material and methods

A total of 1,481 Atlantic salmon fish originated from a commercial breeding programme (Landcatch, UK) were challenged with AGD. The challenge was performed by cohabitation of seeder fish infected with AGD at a ratio of 15% seeder to naïve fish, allowing three separate cycles of infection with a treatment and recovery period after the first two. For the first two challenges, fresh water treatment was performed 21 days after challenge, followed by a week of recovery. The disease was allowed to progress until the terminal sampling point in the third challenge. Subjective gill lesion scores (1 to 5 in order of severity) were recorded for both gills, and one of the gills was stored in RNAlater for qPCR measurement of amoebic load using *Neoparamoeba perurans* specific primers. DNA was extracted from fin tissue and samples were genotyped with an Illumina SNP array containing approximately 17 K SNPs, designed based on a subset of SNPs from a higher density array (Houston et al. 2014). Genotypes were removed if they met any of the following criteria: SNP call-rate < 0.9, individual call-rate < 0.9, FDR rate for high individual heterozygosity < 0.05, identity-by-state > 0.95 (both individuals removed), Hardy-Weinberg equilibrium FDR p-value < 0.05, minor allele frequency < 0.05. After filtering, a total of 1,430 fish and 7,168 SNPs remained for further analysis.

The heritability of host resistance to AGD was estimated using both the genomic kinship matrix and the pedigree-based numerator relationship matrix. Linear mixed models were used to estimate heritability, and genetic / phenotypic correlations between traits were estimated using a bivariate analysis using ASReml 3.0. The GWAS was performed using the mmscore method in the GenABEL R package. Significance thresholds were calculated using a Bonferroni correction according to the number of independently segregating SNPs calculated using the variance inflation factor (VIF) in Plink v.1.9. A VIF threshold of 10 (R^2 of 0.9), a sliding window of 50 SNPs and a step size of 25 SNPs was used.

The accuracy of genomic selection was estimated by five replications of five-fold cross-validation analysis, selecting two sets of five random non-overlapping training (80% of data) and validation (20% of data) sets. The phenotype recorded in the validation population was then masked and breeding values were estimated using both pedigree-based BLUP and GBLUP. Accuracy was calculated as the correlation between the predicted EBVs of the validation set and the actual phenotypes divided by the square root of the heritability [$\sim r(y_1, y_2)/h$] using all individuals. Since high-density genotyping is expensive, we also evalutated the impact of reduced SNP density on prediction accuracy. To choose the SNPs for the low density panels, we progressively increased the minimum allele frequency threshold (maf, 0.05, 0.10, 0.15, …) resulting in genotype datasets with progresively lower SNP density and progressively higher MAF.

Results and Discussion

The mean and standard deviations for AGD resistance traits were 2.79 ± 0.85 for the subjective gill score and 31.36 ± 3.24 (amplification cycle) for amoebic load respectively. Moderate heritability estimates ranged between 0.25 and 0.36 for both phenotypes, and both the phenotypic and genetic correlaitons between the two traits were high and positive. A previous study on AGD disease resistance within the Tasmanian Atlantic salmon population found similar heritabilities, ranging from 0.16 for gross gill score to 0.35 for digital image gill score (Taylor et al. 2007); and similar to those obtained for sea lice (~0.2-0.3; Gjerde et al., 2011; Lhorente et al. , 2012; Yáñez et al., 2014; Tsai et al., 2016), another problematic parasite for salmon aquaculture.

The GWAS results pointed to a polygenic architecture for both gill score and amoebic load (Figure 1), with no markers surpassing the genome-wide significance threshold. However, both measurements identified two suggestive regions on chromosome 18, seemingly located in two different regions around 9-12 Mb and 54-61 Mb respectively. While a few major disease resistance loci have been described in aquaculture species, such as for infectious pancreatic necrosis virus in Atlantic salmon (Houston et al., 2008, Moen et al. 2009), the majority of disease resistance traits for aquaculture species are controlled by many loci of minor or moderate effect. Other examples include sea lice (Tsai et al., 2016) and *Piscirickettsia salmonis* in Atlantic salmon (Correa et al., 2015), and pasteurellosis in gilthead sea bream (Palaiokostas et al., 2016). Nonetheless, the suggestive QTL for AGD resistance merits verification in independent populations, and potentially functional studies to investigate the underlying causative gene(s). Experiments are underway to compare the gene expression response to infection in AGD-resistant and AGD-susceptible salmon using RNA-Seq to interrogate the mechanisms underpinning host resistance to AGD.

Accuracy of prediction using the genomic relationship matrix was ~18% higher than using the pedigree for both mean gill score and amoebic load (Table 1). Further, accuracies were also ~20% higher for amoebic load than for mean gill score. Amoebic load is measured using qPCR and has a wider range than the subjective mean gill score, which could be considered as a categorical trait. Taylor et al. (2007) found that gill damage scores obtained using image analysis or histopathology showed high positive genetic correlation, but correlation between these traits and subjective gill score was lower. Our results suggest that both subjective gill score and qPCR measure of amoebic load are useful traits for selection for AGD resistance.

Since genotyping with SNP chips is relatively expensive, and aquaculture species tend to have closely related animals in training and validation populations (e.g. in sib testing schemes), well designed low density genotyping panels may be useful in genomic selection. When SNP density was reduced via progressive increase in MAF, accuracy remained stable until 1,808 SNPs (MAF threshold 0.40) where a gradual drop off was observed. Even at very low SNP density (435 SNPs), the prediction accuracy was higher than using pedigree data (Figure 2). Genotype imputation approaches may further reduce the density of genotyping required in many individuals to improve cost-effectiveness (Tsai et al. 2017).

Conclusions

AGD resistance in Atlantic salmon is moderately heritable (h^2 ~ 0.25-0.30) and can be measured using the correlated traits of subjective gill score or amoebic load measured by qPCR. Suggestive QTL were identified on chromosome 18 which should be tested in independent populations, and may form the basis for identification of underlying causative

genes. Genomic prediction accuracy was found to be substantially higher (~18%) when using genomic relationships rather than pedigree-based relationships, even when marker density was substantially reduced. Since AGD is a large threat for salmon aquaculture in most major salmon production countries, genomic selection for improved resistance is likely to play an important role in breeding programmes.

List of References

Correa K, Lhorente JP, López ME, Bassini L, Naswa S, Deeb N, Di Genova A, Maass A, Davidson WS & Yáñez JM, 2015. Genome-wide association analysis reveals loci associated with resistance against *Piscirickettsia salmonis* in two Atlantic salmon (*Salmo salar* L.) chromosomes. BMC Genomics 16: 854.

FAO, 2016. The State of World Fisheries and Aquaculture 2016. Contributing to food security and nutrition for all. Rome, Italy, 200 pp.

Gjerde B, Ødegård J & Thorland I, 2011. Estimates of genetic variation in the susceptibility of Atlantic salmon (*Salmo salar*) to the salmon louse *Lepeophtheirus salmonis*. Aquaculture 314: 66–72.

Houston RD, Haley CS, Hamilton A, Guy DR, Tinch AE, Taggart JB, McAndrew BJ & Bishop SC, 2008. Major quantitative trait loci affect resistance to infectious pancreatic necrosis in Atlantic salmon (*Salmo salar*). Genetics 178: 1109-1115.

Houston RD, Taggart JB, Cézard T, Bekaert M, Lowe NR, Downing A, Talbot R, Bishop SC, Archibald AL, Bron JE, Penman DJ, Davassi A, Brew F, Tinch AE, Gharbi K & Hamilton A, 2014. Development and validation of a high density SNP genotyping array for Atlantic salmon (*Salmo salar*). BMC Genomics 15: 90.

Lhorente JP, Gallardo JA, Villanueva B, Araya AM, Torrealba DA, Toledo XE & Neira R, 2012. Quantitative genetic basis for resistance to *Caligus rogercresseyi* sea lice in a breeding population of Atlantic salmon (*Salmo salar*). Aquaculture 324–325: 55–59.

Moen T, Baranski M, Sonesson AK &, Kjøglum S, 2009. Confirmation and fine-mapping of a major QTL for resistance to infectious pancreatic necrosis in Atlantic salmon (*Salmo salar*): population-level associations between markers and trait. BMC Genomics 10: 368.

Palaiokostas C, Ferraresso S, Franch R, Houston RD & Bargelloni L, 2016. Genomic prediction of resistance to pasteurellosis in gilthead sea bream (*Sparus aurata*) using 2b-RAD sequencing. G3 (Bethesda). 6: 3693-3700.

Taylor RS, Wynne JW, Kube PD & Elliott NG, 2007. Genetic variation of resistance to amoebic gill disease in Atlantic salmon (*Salmo salar*) assessed in a challenge system. Aquaculture 272S1: S94-S99.

Tsai HY, Hamilton A, Tinch AE, Guy DR, Bron JE, Taggart JB, Gharbi K, Stear M, Matika O, Pong-Wong R, Bishop SC & Houston RD, 2016. Genomic prediction of host resistance to sea lice in farmed Atlantic salmon populations. Gen. Sel. Evol. 48: 47.

Tsai HY, Matika O, Edwards SM, Antolín-Sánchez R, Hamilton A, Guy DR, Tinch AE, Gharbi K, Stear MJ, Taggart JB, Bron JE, Hickey JM & Houston RD, 2017. Genotype imputation to improve the cost-efficiency of genomic selection in farmed Atlantic salmon. G3 (Bethesda) 7: 1377-1383.

Yáñez JM, Lhorente JP, Bassini LN, Oyarzún M, Neira R & Newman S, 2014. Genetic co-variation between resistance against both *Caligus rogercresseyi* and *Piscirickettsia salmonis*, and body weight in Atlantic salmon (*Salmo salar*) Aquaculture 433: 295–298.

Table 1. Accuracy of genomic selection.

	Pedigree	Genomic
Mean gill score	0.51	0.62
Amoebic load	0.60	0.70

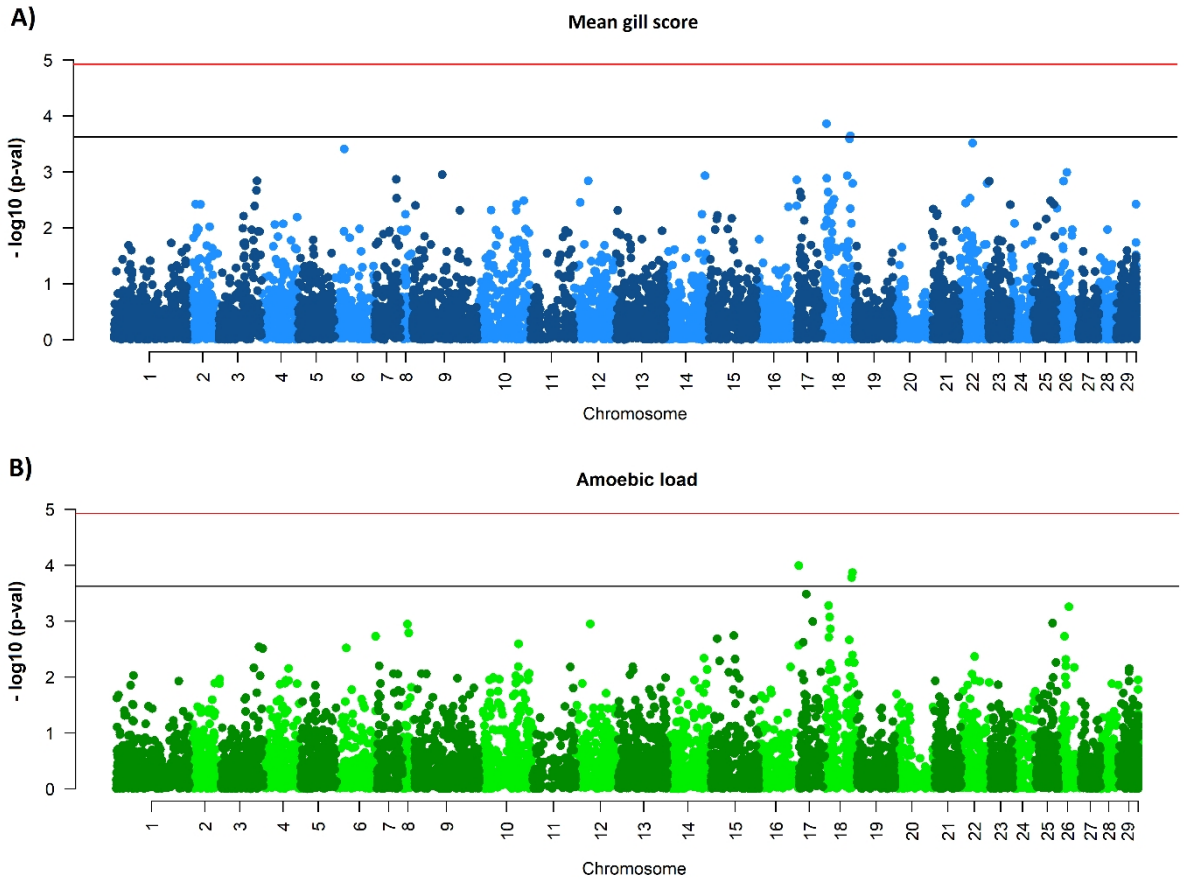

Figure 1. Genome-wide association study for resistance to AGD.

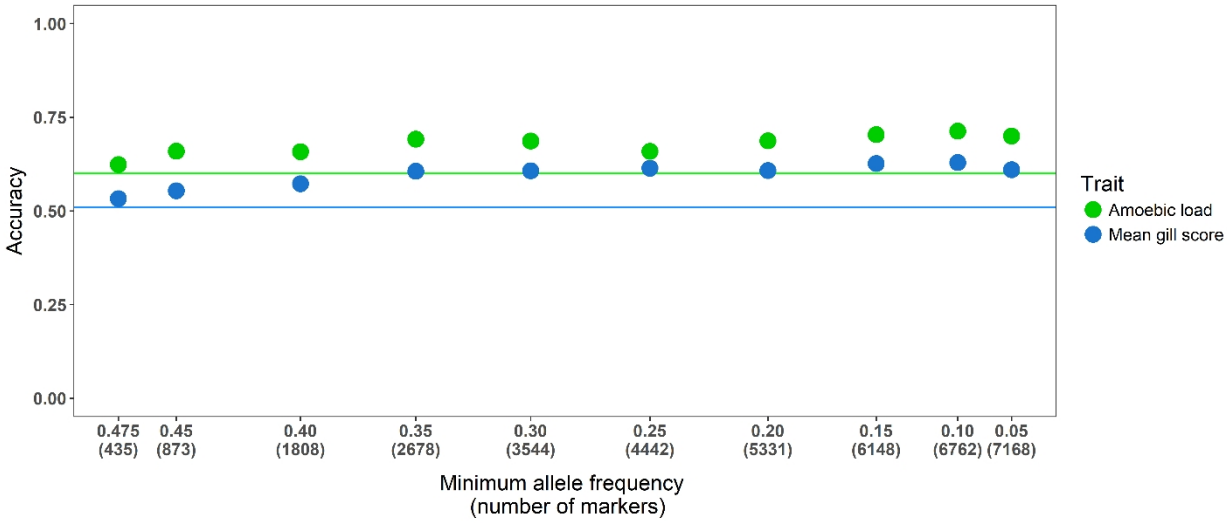

Figure 2. Prediction accuracy for different SNP densities. Solid lines represent prediction accuracy when using the pedigree.