Genetic Investigation of Disease Resistance in Saint John River Atlantic Salmon

Dubois N.L.
Department of Clinical Neurology, Amiens Medical University, Amiens, France

Abstract

Infectious salmon anemia (ISA) is an economically costly viral disease to the global Atlantic salmon aquaculture industry due to the high virulence and high mortality rate. Selective breeding for disease resistance traits using genomic information has been suggested to increase the resistance to various diseases in the breeding population. The aim of this study was to perform a genome wide association (GWA) analysis to identify regions of the genome associated with ISA resistance in the commercial Saint John River population. A total of 2233 parr from 72 families from the 2015 year class were challenged with ISA using a cohabitation trial. The heritability of ISA resistance was 0.16 ± 0.04 on the observed scale for the binary trait of ISA survival at 50% mortality of all cohabitating fish. Binary phenotypes were recorded for all fish with those that survived the challenge being classified as resistant and those that died being classified as susceptible. Selective genotyping using a custom North American 50K SNP chip was performed by choosing, where possible, four resistant fish and four susceptible fish from each challenged family (n = 572). After quality control a total of 547 fish and 38,954 SNP markers remained for GWA analysis. Two survival traits were analyzed for the genotyped fish: ‘survival to day 37’ when mortality had reached 50% and ‘survival at trial termination’ on day 54. Two SNPs on *Ssa*03 and on *Ssa*07 were significant associated with ‘survival to day 37’ but only at the chromosome wide level. More notably nine significant SNPs on nine different chromosomes, were significantly associated with ‘survival at trial termination’ at the genome wide level. The most significant SNP, on *Ssa*13, explained 8.6% of the phenotypic variation, while SNPs on *Ssa*12 and *Ssa*11 explained 7.6% and 6.4% respectively. The remaining six SNPs explained between 4-6% of the phenotypic variation, indicating oligogenic architecture of the ‘survival at trial termination’ trait. A total of 33 genes known to be differentially expressed during infection with ISA were located downstream on the same chromosome arm of the nine significant SNPs. Identification of causal mutations is not necessary to implement genomic information into selective breeding programs but improves understanding of the genetic basis of ISA resistance.

1. Introduction

Infectious salmon anemia (ISA) is a serious viral disease that mainly affects Atlantic salmon but has been also observed in other salmonid and non-salmonid species (Raynard et al., 2001; Nylund et al., 1997; Nylund and Jakobsen, 1995; Nylund et al., 2002; MacLean et al., 2003). The causative agent of ISA is infectious salmon anemia virus (ISAV), a waterborne virus in the *Orthomyxoviridae* family and closely related to influenza viruses (Falk et al., 1997; Krossøy et al., 1999). The disease is characterized by severe anemia and hemorrhagic lesions (Falk et al., 1995). Gross pathological symptoms of ISA also include gill paleness, ascites, liver congestion, spleen inflammation, intestinal congestion, and petechial hemorrhaging of the viscera (Thorud and Djupvik, 1988; Evensen et al., 1991; Falk et al., 1995; Jones et al., 1999; Simko et al., 2000; Speilberg et al., 1995).

ISA is a global concern to the aquaculture industry and is a disease of high importance due to the high mortality and thus high economic impact. ISA was first detected in Norway in 1984 (Thorud and Djupvik, 1988) and has since been reported to occur globally where Atlantic salmon aquaculture is prevalent: Atlantic Canada in 1997 (Mullins et al., 1998), Scotland in 1998 (Rodger et al., 1998), the Faroe Islands in 1999 (Lyngøy, 2003), Maine (Atlantic USA) in 2000 (Bouchard et al., 2001), and Chile in 2007 (Godoy et al., 2008), although ISAV was detected in Coho salmon in Chile before detection in Atlantic salmon (Kibenge et al., 2001). Clinical outbreaks of ISA can have a high mortality rate, as cumulative mortality for natural and experimental infections can drastically range from 0 to 100% (Kibenge et al., 2004).

ISAV has been detected in both aquaculture and wild fish (Raynard et al., 2001). ISA outbreaks typically occur during seawater lifecycle stages of Atlantic salmon, but ISA has been detected in freshwater stages as well (Lyngstad et al., 2008). Transmission of ISAV is thought to be primarily though horizontal transfer, but notably there can be transmission of ISAV from wild fish to hatchery-reared smolts into commercial sea cage Atlantic salmon (Nylund et al., 1994). While the impact of ISAV on wild fish stocks is not currently well known, it is hypothesized that wild populations possess a non-virulent but still transmissible strain of ISAV (ISAV-HPR0 strain) (Christiansen et al., 2011). The ISAV-HPR0 strain, after transmission to aquaculture fish mutates into more virulent strains that are adapted to the high host densities typical of intensive aquaculture practices due to the high mutation rate of the ISAV genome (Mjaaland et al., 2002;

Markussen et al., 2008) and there is some evidence that mutation from a non-pathogenic to a low-pathogenic virus can occur (Christiansen et al., 2017; Cárdenas et al., 2019). Due to the high virulence of ISAV, ISA can spread quickly through a cage site through fish to fish transfer and can also be passively transferred to neighbouring sites by equipment sharing and further by shipping containers as the virus can survive for 20 hour in seawater (Nylund et al., 1994; Murray et al., 2002). Sea lice, a prevalent ectoparasite of farmed salmonids, could be a mechanical vector for transmission of ISAV (Oelckers et al., 2014). There is also evidence of vertical transmission of ISAV from mother to offspring through ovarian fluid and eggs (Marshall et al., 2014). Therefore, improved management practices have been crucial at reducing the risk of the disease, however outbreaks still occur in the Atlantic salmon industry.

ISA outbreaks globally have resulted in major economic losses. The Chilean aquaculture industry experienced a massive ISA outbreak shortly after initial detection in Atlantic salmon in 2007, the worst disease outbreak in salmon aquaculture to date (Asche et al., 2009). This Chilean outbreak reduced the production of Atlantic salmon by more than 60%, causing a devastating loss of profits (Asche et al., 2009). ISA in New Brunswick has not been as severe as the Chilean outbreak, but there are several instances of ISAV being detected in New Brunswick and other maritime provinces in Canada (Canadian Food Inspection Agency, 2018). In Canada as well as other Atlantic salmon producing countries such as Norway, Scotland, and Chile, the confirmation of ISA at a cage site requires the cull of all exposed fish, not just fish that are symptomatic, to prevent the spread of the virus (Canadian Food Inspection Agency, 2013).

In addition to developing management practices and vaccines to control ISAV, a genomic approach can be taken to breed Atlantic salmon that have increased genetic resistant to ISA so that ISAV is less likely to be transmitted to fish (Kjøglum et al., 2008). As ISA primarily affects fish while in the seawater life stages and because many aquaculture companies keep broodstock quarantined in freshwater hatcheries, there has rarely been direct selection on the phenotypes of the broodstock for ISA resistance. Additionally, deliberate challenges of broodstock with pathogenic ISAV is ill-advised due to biosecurity concerns of vertical transmission to potential future offspring but also the high virulence of the disease could cause large mortalities in the only breeding stock available. Therefore, the use of sibling testing is beneficial for disease traits such as ISA resistance, and combined with genomic data, accurate evaluations of both family and individual levels of resistance is possible.

Several studies have utilized sib-testing methodology in commercial Atlantic salmon populations to look for genotype-phenotype associations for a variety of disease resistance traits using QTL mapping and genome wide association (GWA) studies (Moen et al., 2007; Correa et al., 2015; Correa et al., 2017, Tsai et al., 2015; Rochus et al., 2018; Holborn et al., 2018; Holborn et al., 2019). Moen et al. (2007) performed quantitative trait loci (QTL) analysis using ISA survival data and found a QTL that explained 6% of the phenotypic variation in resistance to ISA on *Ssa*15, using a population of Atlantic salmon from European origin. However, QTL mapping needs to be verified in different populations, and there is a large difference between the European population in Moen et al. (2007) and this North American population, as European and North American Atlantic salmon are currently considered separate subspecies (Verspoor, 1997; King et al., 2007; Bourret et al., 2013). QTL mapping with SNP molecular markers have been performed on the commercial Saint John River population of Atlantic salmon. Using a SNP dataset of approximately 11,000 markers, three QTL were identified to be significantly associated with ISA survival in the 2009 and 2010 year classes (Dussault et al., revision in review).

Investigating ISA resistance in Atlantic salmon using a GWA study design has some benefits over the previously performed QTL mapping designs. With a GWA study design, rather than QTL linkage mapping within large families which has been done in previous studies, more families can be included as fewer individuals per family require genotyping. A GWA study design utilizes the within-family genetic variation much like a QTL mapping design, but additionally GWA studies also exploit the between-family genetic variation (Sonesson and Meuwissen, 2009). The increase in the number of families is more representative of the breeding population, increasing the sampling of genetic variance within the population.

In this study, a GWA analysis using SNP markers from across the whole Atlantic salmon genome was performed to detect SNPs in linkage disequilibrium (LD) with QTLs associated with ISA resistance. This research presents the first GWA study of ISA resistance in this commercial population of Atlantic salmon. The information from this study will assist breeders in developing a genetic improvement program that incorporates increased ISA resistance.

2. Methods

2.1 Study Population

The Atlantic salmon used in this study were ~5 month old parr from the Kelly Cove Salmon breeding nucleus from the 2015 year class. This commercial population has undergone six generations of artificial selection for economically important traits (Liu et al., 2017) after establishment from the North American Saint John River strain. Fish were raised in family tanks until a passive integrated transponder (PIT) tag could be inserted, therefore the parentage of each fish was known. Prior to exposure to ISAV, all fish in the challenge were fin clipped for future DNA extractions, and the fin clip was stored in ethanol for preservation.

2.2 Disease Challenge

A controlled cohabitation disease challenge was conducted at the Center for Aquaculture Technologies Canada's (CATC) biological containment laboratory in Souris, Prince Edward Island. This challenge design involved the intraperitoneal injection of a virulent strain of ISAV into shedder fish (=Trojans) which would then cohabitate and horizontally transfer ISAV to the fish in the family challenge (called cohabs) (Jones and Groman, 2011). Prior to a full scale family disease challenge, a pre-challenge was conducted to determine the dosage of ISAV. In addition to testing different dosages, the pre-challenge also tested different shedder fish pressures. Three shedder fish pressures were tested, 20%, 30%, and 40%, at three different proprietary $TCID_{50}$ doses. The aim of the pre-challenge was to select the dosage and the number of shedder fish relative to cohab fish that would yield an end mortality of approximately 50%, the point with the most variation in phenotypes.

The family disease challenge had a total of 2982 fish, 749 shedder fish and 2233 cohab fish, from 72 full-sib families. Each family had a minimum of 15 fish and up to 40 fish, with an average of 30.9 fish per family. After a 14 day acclimation period after transfer to the facility, fish were anesthetized for handling to prepare for the challenge. Cohab designated fish were weighed (average weight of 94.3 g (SD = 17.7 g)) and fish from each family were distributed into 18 tanks after a fin clip was taken. Each shedder fish was injected with 100 µl of inoculate, which contained ISAV in the dosage determined from the pre-challenge and phosphate buffered saline. Each tank had a shedder fish pressure of approximately 25%. The challenge was conducted in freshwater kept at 12 ± 1 °C.

Post challenge start, fish were monitored daily for mortalities and fed up to 1.5% body weight per day of commercial fish feed. Cohab mortalities had the day of death and external

disease symptoms recorded and a necropsy was performed to record sex and gross pathology. Fish that showed signs of early maturity were removed from the dataset as these fish are physiologically different from non-mature and may have different immune responses to disease. Heart samples from the first 100 cohab mortalities were collected and stored in RNAlater for subsequent qPCR analysis to confirm ISAV presence and estimate viral load. The challenge was terminated on day 54 post challenge once mortalities plateaued at approximately 1% per day for five days. All fish alive the day of challenge termination were classified as survivors, and survivors were sampled similarly to the cohab mortalities, however sex was not recorded.

The heritability of ISA resistance in this population was estimated with ASReml using data from 2223 cohabs that were not sexually mature or possessed deformities that may affect their performance in the ISA challenge. An animal model, with tank as a fixed effect, was used to estimate the heritability of ISA resistance, measured as survival at 34 days, the point in the challenge where 50% of the 2223 cohabs had died. The recorded pedigree used to estimate the heritability included up to the parent generation (n = 2413), and therefore provided a rough estimation of the heritability of the trait. The heritability estimate was based on 40,000 Gibbs samples after an 8000 burn-in. The model in equation analyzed ISA resistance as a binary trait:

$$y_{i,j} = \mu + t_i + a_j + e_{ij} \quad (1)$$

where, y_{ij} is the survival phenotype ('death before 34 days' at 50% mortality coded as 0, and 'survival to after 34 days' coded as 1) which is analyzed as a quantitative trait, μ is the overall mean of the population, t_i is the tank effect (a fixed effect), a_j is the animal effect from the pedigree (a random effect), and e_{ij} is the residual.

2.3 Genotyping

Genomic DNA was extracted from 770 cohabs, representing the performers based on residual ranking of fish within each of the 72 families using a DNeasy Blood & Tissue Kit (Qiagen). As there were differences in mortality levels between tanks, the binary survival phenotype was corrected for tank effects and using the residuals each fish within each family was ranked similar to an estimated breeding value. Selective genotyping was performed due to budget limitations and the high cost of genotyping and to ensure sampling of fish with variability in survival so the

aim was to genotype the four top performing fish and the four most poorly performing fish in each family, based on the residual ranking. In several families this represented four mortalities and four survivors, but in families with 100% mortality this was not possible. Due to the small size of the fish and the fact that fin tissue was taken antemortem before exposure to ISAV, many of the fin clip samples did not yield sufficient concentrations of DNA for genotyping and all sample was consumed in the initial extraction. Therefore, a replacement for that fish from the same family with a similar phenotype was selected. High quality and high molecular weight DNA from 572 fish samples was sent to CIGENE (Ås, Norway) for genotyping on the custom NA *Ssa*50K SNP chip (Holborn et al., 2018; Boulding et al., 2019). SNPs were exported from Axiom Analysis Suite (version 1.1.1.66) if they presented as 2 (NoMinorHom) or 3 (PolyHighResolution) well-separated clusters (FLD ≥ 5) indicating that they were in a diploid region of the ancestrally tetraploid Atlantic salmon genome (Gidskehaug et al., 2011).

As sex was not recorded for the surviving fish, the genotypic sex was determined for each fish with an available genotype using six previously determined sex SNPs that have an accuracy of prediction of ~95% (Boulding, unpublished data). The phenotypic sex recorded based on the presence of testes or ovaries, where applicable, of the challenged fish were compared with the genotypic sex and some differences in sex were noted. These differences between phenotypic and genotypic sex could likely be due to the small size of the fish and the fact that they were not sexually mature. Therefore, the genotypic sex was used for all fish in the analysis regardless of the recorded phenotypic sex.

2.4 Association Analysis

Additional quality control was performed on the returned genotypes to filter out poor SNPs and fish with low call rates. PLINK (version 1.9; www.cog-genomics.org/plink/1.9/) was used to filter based on a SNP call rate of > 0.95, a sample call rate of > 0.9, for a SNP minor allele frequency (MAF > 0.05), and for SNPs in Hardy-Weinberg Equilibrium (p value < 0.001). The fish samples that passed quality control were analyzed for continent of origin using STRUCTURE (version 2.3.4) using the methods in Liu et al. (2017). Previous disease resistance studies used a cut-off threshold for continent of origin of 90% or greater North America ancestry (Holborn et al., 2018) however this criterion would result in using only 325 fish genotypes which greatly lowered the power to detect any significant associations due to small sample size. Within

this dataset, the lowest percent of North American ancestry was 85%, and given the potential margin of error with the STRUCTURE analysis all individuals in the dataset were kept by using a threshold of 85% North American ancestry.

A multi-step mixed model approach was used for the GWA analysis for the survival trait using the package GenABEL (version 1.8-0) (Aulchenko et al., 2007) in R statistical software (version 3.3.2). This model incorporates a genomic kinship matrix to account for structure due to familial relatedness, which is important with a full-sib design, as well as structure due to historical population admixture. The genomic kinship matrix was estimated in GenABEL using the *ibs* function using all available SNPs, which estimates the relationships among individuals using identity by state. The *polygenic* and *mmscore* functions were used to perform a multi-step association analysis, called the Family-based Test for Association (FASTA) which is suitable for large family-based datasets (Chen and Abecasis, 2007). Using ANOVA, tank effects were significant on survival (p value <0.0001), genotypic sex approached significance (p value = 0.0686), and weight was not significant (p value = 0.887), therefore both tank and sex were included in the polygenic model.

The *polygenic* function is first used to estimate the residual, after accounting for other fixed and random factors (equation 2), and the residual is then fitted into a model using the *mmscore* function that tests the effect of each SNP on the trait (equation 3).

$$y_{ijkl} = \mu + G_i + t_j + s_k + e_{ijkl} \qquad (2)$$

where, y_{ijkl} is the survival phenotype (survivors coded as 0, mortalities as 1) which is analyzed as a quantitative trait, μ is the overall mean of the population, G_i is the additive polygenic effect estimated from the genomic kinship matrix (a random effect), t_j is the tank effect (a fixed effect), s_k is the sex effect (a fixed effect), and e_{ijkl} is the residual, that will be used in the next equation.

$$\hat{e} = \beta(SNP) + e_i \qquad (3)$$

where, $\hat{e}$ is the estimate of the residual from the polygenic model (equation 1), $\beta(SNP)$ is the SNP effect on the survival trait, and e_i is the residual.

Two survival traits were analyzed using this multi-step mixed model approach on the 547 fish that passed quality control. The first trait was 'survival at trial termination' on day 54, the second trait was 'survival to day 37'. The trait 'survival on day 37' was analyzed as this is the point in the trial where there were 50% mortalities in the subset of fish genotyped. This point in the trial represents the time with the most variation in phenotypes.

The mixed model approach involved the individual testing of all SNPs, and therefore a Bonferroni correction is commonly used to account for multiple testing. Bonferroni corrections are a conservative method where the threshold of significance is determined by α/M, where α is the significance level and M is the number of SNPs tested. Traditionally Bonferroni corrected critical values uses $\alpha = 0.05$, however Bonferroni correction is rather strict and can increases the type II error rate (Hochberg, 1988). Therefore, a value of $\alpha = 0.1$ was used, which is a suggestive level of significance and is useful when doing discovery work where thresholds can be relaxed. SNPs can also be significant at the chromosome wide level, which is calculated as α/C, where C is the number of SNPs tested on the chromosome of interest.

The percentage of phenotypic variation was also estimated using GenABEL. The number of samples genotyped for a given SNP divided by the chi-squared statistic value from the *mmscore* test function provides the proportion of phenotypic variation. This value was calculated for all SNPs that pass the suggestive Bonferroni threshold (0.1/M).

Although this SNP chip contains SNPs from across the genome, not all regions have significant LD. PLINK was used to estimate the LD (r^2) between any significant SNPs and the flanking regions within 0.1 Mb upstream and downstream.

The area surrounding any significant SNPs was searched for potential genes that could account for differences in resistance to ISA. SalmoBase was used to scan the area within ± 0.1 Mb of the estimated SNP location, and the gene product was recorded where applicable. The locations of significant SNPs were compared against genes that were differentially expressed during an ISA challenge of North American farmed salmon using the NCBI database (LeBlanc et al., 2010). All genes that were >2-fold differentially expressed between ISAV injected and sham injected fish in LeBlanc et al. (2010), either upregulated or downregulated, at different time points of exposure to ISAV were blasted using the NCBI database against the *Salmo salar* genome to identify the chromosome location of the gene. We note that not all differentially expressed genes between infected and naïve fish will be related to ISA resistance, however the

aim was to generate a list of genes for potential further study, specifically genes with known immune function. Phillips et al. (2009) and Brenna-Hansen et al. (2012) were used to determine the chromosome type for each chromosome with a significant SNP (acrocentric or metacentric).

3. Results

The ISA challenge had a higher than anticipated cumulative mortality with differences in mortality rates between tanks (Supplementary Material Fig. S2). The trial was terminated on day 54 and no mortalities occurred the day of trial termination. On day 53, the last day of mortalities, the trial had a cumulative mortality of 83% (Fig. 1), which was higher than expected based on the pre-challenge. There was a difference in tank mortalities across the 18 tanks, ranging from 74% mortality to 97% mortality, therefore the inclusion of a tank effect into the association analysis model was required. Most importantly though, there was a range in the family level mortality, which was expected if there was family level genetic variation in ISA resistance. The family level mortality ranged from 42% to 100% (Fig. 2).

The narrow sense heritability of ISA resistance was estimated to be 0.16 ± 0.04 (mean ± SE) on the observed scale for the binary trait of ISA survival at 50% mortality for all cohabs in the challenge including those that were not genotyped. The phenotypic mean was 0.495 ± 0.50 (mean ± SE). This was based on 2413 animals in pedigree file and 2223 animals with phenotypic observations for the ISA resistance trait.

The viral load for surviving fish was lower than the viral load of ISA mortalities within the first 6 days of the onset of cohab mortality, suggesting that surviving fish were in the process of clearing the virus. All early mortalities were symptomatic for ISA, and the 100 heart samples collected from the first mortalities were all positive for ISAV and had high viral loads (number of ISAV copies/µg RNA). This provides evidence that mortalities were caused by the viral infection and likely not other causes. The mean viral load for the 100 early cohab mortalities was 2.32×10^5 copies/µg RNA (SD = 1.29×10^5). The heart samples taken from the 378 fish that survived the challenge also had viral load estimated by qPCR. Of the survivors, 302 were positive for ISAV, 65 fish were near the limit of detection for the qPCR method, 9 fish were negative, and 2 fish had inconclusive results after several attempts of retesting. Of the 302 fish heart samples positive for ISAV, the mean viral load was 59.31 copies/µg RNA (SD = 117.68),

which is significantly ($p < 0.0001$) lower than the viral load from the early mortalities. This could indicate an ability to fight an infection, restricting the virus from entering the body, or another mechanism of ISAV resistance. However, we acknowledge that the viral load does not tell us whether the fish would have eventually died if the challenge had been permitted to continue.

A total of 38,954 SNP markers and 547 fish, comprising of 195 survivors on day 54 and 352 mortalities, passed all quality control parameters. The genotyped fish represent all 72 families, and between 5 and 9 fish per family passed quality control, with an average of 7.6 fish per family. Not accounting for population stratification can cause over-inflation of association, however there was no evidence of over-inflation based on the QQplots (Supplementary Material Fig. S1). Rather, the association analyses showed under-inflation with lambda values of 0.95 for 'survival at day 37' and 0.95 for 'survival at trial termination', where the null hypothesis, indicating no over- or under-inflation, has a lambda value of 1.

The association analysis identified nine SNPs significantly associated with 'survival at trial termination' at the suggestive ($\alpha = 0.1$) genome wide level (see SNP details in Table 1). These SNPs were from nine different chromosomes (Fig. 3), indicating that ISA survival is a polygenic trait. For all nine of the significant SNPs the minor allele (A_2) was associated with ISA resistance. The most significant SNP located on *Ssa*13 accounted for 8.61% of the phenotypic variation of the trait, while the remaining SNPs ranged from 7.56% to 4.42% in this population (Table 1). In addition, SNPs on *Ssa*09, *Ssa*14, and *Ssa*18 were significant at their respective suggestive ($\alpha = 0.1$) chromosome wide thresholds.

The association analysis of 'survival to day 37' did not yield any suggestive genome wide significant SNPs, however two chromosome wide significant SNPs ($\alpha = 0.05$) were detected. The most significant SNP for this trait was located on *Ssa*03 while the second most significant SNP was located on *Ssa*07 (Fig. 4).

Although nine different SNPs on nine chromosomes were significantly associated with ISA resistance, there was no evidence that flanking SNPs (within ± 0.1Mb) were also associated with ISA resistance (Fig. 3). We hypothesized that this was because the pairwise LD between the significant SNP and the flanking SNPs was low. This hypothesis was supported in that the majority of the flanking SNPs had a r^2 value less than 0.2 with the significant SNP. An exception to this was *Ssa*05, *Ssa*12, and *Ssa*24 which had one or more SNPs with a r^2 value above 0.2 and *Ssa*14 had the highest amount of LD with two SNPs with r^2 values of 0.450 and 0.418

(Supplementary Material Fig. S3). All of the SNPs with a r^2 value above 0.2 had low MAF, with a range between 0.054 and 0.127.

In total 41 candidate genes were found within ± 0.1 Mb of the estimated SNP location of the nine significant SNPs (Table 2). From the LeBlanc et al. (2010) study, 324 unique gene sequences from the cGRASP microarray and 8 from pre-selected candidate genes for qRT-PCR were blasted. Ninety gene sequences from the cGRASP microarray were co-located on chromosomes with a significant SNP, with 30 genes located downstream from a significant SNP on the same chromosome arm (Table 3). The majority of the chromosomes of interest were acrocentric, meaning there is a single chromosome arm, however *Ssa*05 is metacentric, so the location of the centromere was estimated based on approximate location in Philips et al. (2009) and the total chromosome length from SalmoBase. Four genes from the candidate genes co-located with a significant SNP, of which three genes were located downstream on the same chromosome arm (Table 4).

4. Discussion

The results of this ISA challenge indicate genetic variation for ISA resistance in this North American Atlantic salmon population. This was expected as several other studies have shown ISA resistance to be genetically variable in Atlantic salmon of European origin (Moen et al., 2007; Kjøglum et al., 2008). The range in family level mortality from 42% to 100% and significant effect of family on survival indicated the existence of family level genetic variation for ISA resistance in this commercial Atlantic salmon population.

We demonstrated that ISA resistance in this population of Atlantic salmon is heritable, with a heritability estimate of 0.16. This estimate is lower than published estimates calculated based on European Atlantic salmon populations, which have estimates ranging from 0.24 to 0.40 (Ødegard et al., 2007; Olesen et al., 2007; Kjølum et al., 2008; Gjerde et al., 2009). However, this lower estimate is likely due to the large amount of environmental variation introduced through the use of 18 tanks. Regardless, this heritability estimate indicates that ISA resistance can be selected for in this population.

The STRUCTURE analysis indicated that this subset of fish from the 2015 year class had lower amounts of European ancestry within the genome (Holborn et al., 2019). Additionally, the

inclusion of fish with 85% or greater North American ancestry did not cause over-inflation of the association as lambda values were below a value of 1.

The two resistance traits analyzed, 'survival at trial termination' and 'survival to day 37', were associated with different SNPs confirming that they are not the same trait and that 'survival at trial termination' is a more accurate measure of ISA resistance. 'Survival at trial termination', when mortality naturally levels off, is likely to represent more fish that were truly able to resist infection, tolerate infection, or clear a viral infection from their system. In contrast, the 'survival to day 37' trait was defined as survival to a fixed point where 50% of the fish are alive, which resulted in fish classified as survived being a mixture of fish that had died and had survived at the termination of the trial. This method can introduce biases (Ødegård et al., 2011) as the classification of fish that would ultimately die throughout the challenge as survivors may conflict with the association analysis by reducing the power to detect SNPs associated with ISA resistance because these fish may possess the genotypes associated with mortality. This could explain why the 'survival to day 37' trait had no associated SNPs at the genome wide level (Fig. 4) but the trait 'survival at trial termination' did (Fig. 3).

The nine SNPs significantly associated with ISA resistance as 'survival at trial termination' were located on nine different chromosomes. A previous study on the commercial Saint John River population of Atlantic salmon was also performed using QTL linkage mapping and detected two QTL significant at the chromosome wide level on *Ssa*03 and *Ssa*04, and one QTL at the experiment wide level on *Ssa*25 (Dussault et al., revision in review). Our GWA analysis for 'survival at trial termination' did not detect any significant SNPs on these chromosomes, however 'survival to day 37' did have a chromosome wide significant SNP on *Ssa*03 but was not located within the same genome region as the QTL in Dussault et al., (revision in review). The QTL mapping study utilized fewer families but had more individuals genotyped from those families due to the experimental design. However, in the Dussault et al. (revision in review) study only QTL that happen to be segregating within the families selected for genotyping would be detected, as specific family crosses were not created for the study. Additionally, the year classes used differed between this GWA study and Dussault et al. (revision in review), which could also lead to differences in the number and location of significant SNPs associated with ISA resistance in the Saint John River strain.

No large, broad peaks of significant SNPs were detected in this GWA analysis as the majority of the flanking SNPs were not in LD with the significant SNP and therefore can be considered independent of each other. Several of the significant SNPs had flanking SNPs with a low level of LD while chromosome *Ssa*14 had moderate levels of LD. Therefore, it is possible that the significant SNP on *Ssa*14 is a false positive as these moderately linked SNPs did not show association with ISA resistance. However, the MAF of these two SNPs are low (0.091 and 0.127) so this may hamper the ability to detect associations. Further analysis with more individuals would be necessary to determine if the significant SNP on *Ssa*14 is a false positive or a true association by increasing the sample size of fish possessing the minor allele.

A study by Moen et al. (2007) using QTL mapping with microsatellites identified a QTL associated with ISA resistance on *Ssa*15 that accounted for 6% of the phenotypic variation of the trait in a European population of Atlantic salmon. The 'survival at trial termination' or 'survival to day 37' analyses did not indicate any region passing significance thresholds on *Ssa*15, however, there is a peak of several SNPs on *Ssa*15 for the 'survival to day 37' trait at approximately 33 Mb along the chromosome. This region, although not significant, is of interest as this is the region of the genome where the QTL was detected in Moen et al. (2007). The QTL found on *Ssa*15 in Moen et al. (2007) was nearby the gene HIV-EP2/MBP-2, which is located around 34.8 Mb according to SalmoBase. It is not uncommon that areas of the genome associated with traits differ between the North American and European subspecies of Atlantic salmon (Boulding et al., 2019), so the homolog of this region of *Ssa*15 was identified to be on *Ssa*06 using figure 2 in Lien et al. (2016). This region on *Ssa*06 did not have any significant SNPs or show any trend towards significance. Although it cannot be claimed that this region on *Ssa*15 is associated with ISA resistance in North American Atlantic salmon as it is in the European subspecies as significance thresholds have not been passed, it was of interest to note as future datasets may have more power to detect associations in this region of the genome.

A candidate gene on chromosome *Ssa*24, neu4, is of particular interest because 4-*O*-acetylated sialic acids (Neu4,5Ac), is the preferred receptor of ISAV. Neu4 is located at a 6.77 Mb in the Atlantic salmon genome and facilitates the binding of the virus to endothelial cells (Aamelfot et al., 2012). The most significant SNP on *Ssa*24 at the genome wide level was located upstream of this position at 3.60 Mb.

A total of 41 protein coding genes were detected within a 0.1 Mb distance from the nine significant SNPs that could be associated with ISA resistance in this population. There were also 30 differentially expressed genes from the cGRASP microarray and three from a candidate gene approach that were located downstream, on the same chromosome arm, from statistically significant SNPs. The SNPs located upstream from these genes, many of the genes involved in innate and adaptive immunity (LeBlanc et al., 2010), could be cis-regulating these genes, but further testing into the biological pathways of these genes is needed.

The underlying genetic cause is not required to implement genomic information into selective breeding programs, but it is important for fully understanding the genetic basis of ISA resistance. Li et al. (2011) were able to identify several candidate genes within the region of the QTL for ISA resistance identified by Moen et al. (2007). Li et al. (2011) identified one particular gene of interest, HIV-EP2/MBP-2, which is thought to influence the expression of other genes that have been linked to having a response to ISAV infection, likely though changes in the structure of the protein. Major histocompatibility complex class I and II genes have also been associated with resistance to ISA (Grimholt et al., 2003), but since ISA resistance is a polygenic trait there are likely more genes that contribute to resistance as well as epistasis, so a candidate gene approach would not be ideal (Yáñez et al., 2014).

As this is the first and only year class to undergo an ISAV challenge and subsequent GWA analysis, further investigation into the genetic architecture of ISA resistance in this population of Atlantic salmon is encouraged. Year classes are usually not crossed with other years, and these commercial salmon have a 4-year generation interval so each of the resulting four breeding lines show Fst values supporting their status as historically different populations (Liu et al., 2017). The inclusion of more year classes into the dataset would help estimate the effect of year class on the trait, so year class can be accounted for like sex or tank within the association analysis. Accounting for year class effects can potentially remove any significant SNPs associated with a year class effect that were initially attributed to ISA survival. This will allow researchers to further pinpoint regions of the genome truly associated with ISA resistance. Additionally, the continuation of ISA disease challenges will allow for the utilization of phenotypes and genotypes of the siblings of the broodstock to be applied to the broodstock using genomic estimated breeding values to make informed breeding decisions. As ISA resistance is a polygenic trait, genomic selection using all available SNPs or a subset of SNPs, including the nine significant

SNPs detected in this study, from all chromosomes would be the best course of action to breed increased resistance into the commercial strains of Atlantic salmon.

Conclusion

This was the first GWA analysis for ISA resistance using a high density SNP array in this commercial population of Atlantic salmon. This GWA analysis on the 2015 year class detected nine SNPs from nine different chromosomes significantly associated with ISA resistance at the suggestive genome wide level. For all nine SNPs the minor allele was the resistant allele, so increasing the proportion of these alleles in the population is possible. Although it is entirely possible that other selection pressures are keeping these allele variants at a lower frequency. We identified 41 genes within the 0.1 Mb surrounding each of the nine SNPs that could be associated with ISA resistance. However, understanding the genes underlying ISA resistance is not necessary for genomics selection but could be of future interest in understanding the genetic basis of ISA resistance. Continued ISA disease challenges would benefit the breeding program to increase ISA resistance in the population to mitigate the costly effects of the disease at sea cage sites due to the highly infectious and fatal nature of ISAV.

Acknowledgements

This project was supported by funding provided by Genome Canada (Genome Atlantic and Ontario Genomics) through the Genomic Applications Partnership Program awarded to EGB, KPA, JAKE, and FP. Additional funding for the challenges was provided by IRAP awarded to Cooke Aquaculture. A portion of MKH's salary was from a NSERC Industrial Postgraduate Scholarship and from University of Guelph teaching assistantships. We also acknowledge D. Plouffe and the Center for Aquaculture Technologies for the execution of the ISA disease challenge at their facility in Prince Edward Island. L.R. Schaeffer calculated the heritability estimates. We are grateful to J. Ødegård and L.R. Schaeffer for guidance on data analysis and for providing comments on the manuscript. Technical assistance was provided by Kamini Rajakumar and Camden Moir for DNA extractions, and S. Karoliussen of CIGENE for SNP chip processing.

Author Contributions

MKH performed the genome wide association analysis, prepared the manuscript and led the team that performed the DNA extractions. FP reared pedigreed fish. KPA, JAKE, and FP provided genetic material from their pedigreed year-classes for DNA extractions. EGB contributed with the filtering of SNPs, genetic sex determination, advice on how to remove fish with European ancestry, statistical analysis discussion, manuscript discussion and editing, and overall supervision of the work of MKH. All authors have read and provided comments on the manuscript.

References

Aamelfot, M., Dale, O.B., Weli, S.C., Koppang, E.O., Falk, K., 2012. Expression of the infectious salmon anemia virus receptor on Atlantic salmon endothelial cells correlated with the cell tropism of the virus. J. Virol. 86(19), 10571-10578. https://doi.org/10.1128/JVI.00047-12

Asche, F., Hansen, H., Tveteras, R., Tveterås, S., 2009. The salmon disease crisis in Chile. Mar. Resourc. Econ. 24, 405-411. https://doi.org/10.1086/mre.24.4.42629664

Aulchenko, Y.S., Ripke, S., Isaacs, A., van Dujin, C.M., 2007. GenABEL: an R library for genome-wide association analysis. Bioinformatics 23, 1294-1296. https://doi.org/10.1093/bioinformatics/btm108

Bouchard, D.A., Brockway, K., Giray, C., Keleher, W., Merrill, P.L., 2001. First report of infectious salmon anaemia (ISA) in the United States. Bull. Eur. Assoc. Fish Pathol. 21(2), 86-88.

Boulding, E.G., Ang, K.P., Elliott, J.A.K., Powell, F., Schaeffer, L.R., 2019. Difference in genetic architecture between continents at a major locus previously associated with sea age at sexual maturity in European Atlantic salmon. Aquaculture 500, 670-678. https://doi.org/10.1016/j.aquaculture.2018.09.025

Bourret, V., Kent, M.P., Primmer, C.R., Vasemägi, A., Karlsson, S., Hindar, K., McGinnity, P., Verspoor, E., Bernatchez, L., Lien, S., 2013. SNP-array reveals genome-wide patterns of geographical and potential adaptive divergence across the natural range of Atlantic salmon (*Salmo salar*). Mol. Ecol. 22(3), 532-551. https://doi.org/10.1111/mec.12003

Brenna-Hansen, S., Li, J., Kent, M.P., Boulding, E.G., Dominik, S., Davidson, W.S., Lien, S., 2012. Chromosomal differences between European and North American Atlantic salmon discovered by linkage mapping and supported by fluorescence *in situ* hybridization analysis. BMC Genetics 13, 432. https://doi.org/10.1186/1471-2164-13-432

Canadian Food Inspection Agency., 2013. Infectious Salmon Anemia – Fact Sheet. (22 December 2013). Canadian Food Inspection Agency – Government of Canada. http://www.inspection.gc.ca/animals/aquatic-animals/diseases/reportable/isa/fact-sheet/eng/1327198930863/1327199219511

Canadian Food Inspection Agency., 2018. Locations infected with infectious salmon anaemia in 2018. (6 December 2018). Canadian Food Inspection Agency – Government of Canada. http://www.inspection.gc.ca/animals/aquatic-animals/diseases/reportable/2018/infectious-salmon-anaemia-2018-/eng/1520361142560/1520361212232

Cárdenas, C., Ojeda, N., Labra, Á., Marshall, S.H., 2018. Molecular features associated with the adaptive evolution of Infectious Salmon Anemia Virus (ISAV) in Chile. Infect. Genet. Evol. 68, 203-211. https://doi.org/10.1016/j.meegid.2018.12.028

Chen, W.M., Abecasis, G.R., 2007. Family-based association tests for genome wide association scans. Am. J. Hum. Genet. 81, 913-926. https://doi.org/10.1086/521580

Christiansen, D.H., McBeath, A.J., Aamelfot, M., Matejusova, I., Fourrier, M., White, P., Petersen, P.E., Falk, K., 2017. First field evidence of the evolution from a non-virulent HPR0 to a virulent HPR-deleted infectious salmon anaemia virus. J. Gen. Virol. 98(4), 595-606. https://doi.org/10.1099/jgv.0.000741

Christiansen, D.H., Østergaard, P.S., Snow, M., Dale, O.B., Falk, K., 2011. A low-pathogenic variation of infectious salmon anemia virus (ISAV-HPR0) is highly prevalent and causes a non-clinical transient infection in farmed Atlantic salmon (*Salmo salar* L.) in the Faroe Islands. J. Gen. Virol. 92: 909-918. https://doi.org/10.1099/vir.0.027094-0

Correa, K., Lhorente, J.P., Bassini, L., López, M.E., Di Genova, A., Maass, A., Davidson, W.S., Yáñez, J.M., 2017. Genome wide association study for resistance to *Caligus rogercresseyi* in Atlantic salmon (*Salmo salar* L.) using a 50K SNP genotyping array. Aquaculture 472, 61-65. https://doi.org/10.1016/j.aquaculture.2016.04.008

Correa, K., Lhorente, J.P, López, M.E., Bassini, L., Naswa, S., Deeb, N., Di Genova, A., Maass, A., Davidson, W.S., Yáñez, J.M., 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. https://doi.org/10.1186/s12864-015-2038-7

Dussault, F.M., Ang, K.P., Elliott, J.A.K., Glebe, B.D., Leadbeater, S., Manning, A.J., Powell, F., Boulding, E.G., (revision in review) Aquaculture Research ARE-OA-19-Jan-083.R1 submitted April 3, 2019). Mapping quantitative trait loci for infectious salmon anemia resistance in a North American strain of Atlantic salmon. Aquac. Res.

Evensen, Ø., Thorud, K.E., Olsen, Y.A., 1991. Morphological study of the gross and light microscopic lesion of infectious anaemia in Altantic salmon (*Salmo salar*). Res. Vet. Sci. 51(2), 215-222. https://doi.org/10.1016/0034-5288(91)90017-I

Falk, K., Namork, E., Rimstad, E., Mjaaland, S., Dannevig, B.H., 1997. Characterization of infectious salmon anemia virus, an Orthomyxo-like virus isolated from Atlantic salmon (*Salmo salar* L.). J. Virol. 71(12), 9016-9023. https://doi.org/10.1128/JVI.79.19.12544-12553.2005

Falk, K., Press, C.M., Landsverk, T., Dannevig, B.H., 1995. Spleen and kidney of Atlantic salmon (*Salmo salar* L.) show histochemical changes early in the course of experimentally induced infectious salmon anaemia (ISA). Vet. Immunol. Immunop. 49, 115-126. https://doi.org/10.1016/0165-2427(95)05427-8

Gidskehaug, L., Kent, M., Hayes, B.J., Lien, S., 2011. Genotype calling and mapping of multisite variations using an Altantic salmon iSelect SNP array. Bioinformatics 27(3), 303-310. https://doi.org/10.1093/bioinformatics/btq673

Gjerde, B., Evensen, Ø., Bentsen, H.B., Storset, A., 2009. Genetic (co)variation of vaccine injuries and innate resistance to furunculosis (Aeromonas salmonicida) and infectious salmon anaemia (ISA) in Atlantic salmon (Salmo salar). Aquaculture 287, 52-58. https://doi.org/10.1016/j.aquaculture.2008.10.028

Godoy, M.G., Aedo, A., Kibenge, M.J.T., Broman, D.B., Yason, C.V., Grothusen, H., Lisperguer, A., Calbucura, M., Avendaño, F., Imilán, M., Jarpa, M., Kibenge, F.S.B., 2008. First detection, isolation and molecular characterization of infectious salmon anaemia virus associated with clinical disease in farmed Atlantic salmon (*Salmo salar*) in Chile. BMC Vet. Res. 4, 28. https://doi.org/10.1186/1746-6148-4-28

Grimholt, U., Larsen, S., Nordmo, R., Midtlyng, P., Kjoeglum, S., Storset, A., Saebø, S., Stet, R.J., 2003. MHC polymorphism and disease resistance in Atlantic salmon (*Salmo salar*); facing pathogens with single expressed major histocompatibility class I and class II loci. Immunogenetics 55(4), 210–219. https://doi.org/10.1007/s00251-003-0567-8

Hochberg, Y., 1988. A sharper Bonferroni procedure for multiple tests of significance. Biometrika 75, 800-802. https://doi.org/10.1093/biomet/75.4.800

Holborn, M.K., Ang, K.P., Elliott, J.A.K., Powell, F., Boulding, E.G., 2018. Genome wide association analysis for bacterial kidney disease resistance in a commercial North American Atlantic salmon (*Salmo salar*) population using a 50K SNP panel. Aquaculture 495, 465-471. https://doi.org/10.1016/j.aquaculture.2018.06.014

Holborn, M.K., Rochus, C.M., Ang, K.P., Elliott, J.A.K., Leadbeater, S. Powell, F., Boulding, E.G., (in review Aquaculture AQUA_2019_227). Family-based genome wide association analysis for salmon lice (*Lepeophtheirus salmonis*) resistance in North American Atlantic salmon using a 50K SNP array. Aquaculture 511, XXX-XXX. https://doi.org/10.1016/j.aquaculture.2019.734215

Jones, S.R.M., Groman, D.B., 2011. Cohabitation transmission of infectious salmon anemia virus among freshwater-reared Atlantic salmon. J. Aquat. Anim. Health. 13(4), 340-346. https://doi.org/10.1577/1548-8667(2001)013<0340:CTOISA>2.0.CO;2

Jones, S.R.M., MacKinnon, A.M., Groman, D.B., 1999. Virulence and pathogenicity of infectious salmon anemia virus isolated from farmed salmon in Atlantic Canada. J. Aquat. Anim. Health. 11, 400-405. https://doi.org/10.1577/1548-8667(1999)011<0400:VAPOIS>2.0.CO;2

Kibenge, F.S.B., Garate, O.N., Johnson, G., Arriagada, R., Kibenge, M.J.T., Wadawska, D., 2001. Isolation and identification of infectious salmon anemia virus (ISAV) from Coho salmon in Chile. Dis. Aquat. Org. 45, 9-18. https://doi.org/10.3354/dao045009

Kibenge, F.S.B., Munir, K., Kibenge, M.J.T., Joseph, T., Moneke, E., 2004. Infectious salmon anemia virus: causative agent, pathogenesis and immunity. Anim. Health Res. Rev. 5(1), 65-78. https://doi.org/10.1079/AHRR200461

King, T.L., Verspoor, E., Spidle, A.P., Gross, R., Phillips, R.B., Koljonen, M.-L., Sanches, J.A., and Morrison, C.L., 2007. Biodiversity and population structure. In E. Verspoor, L. Stradmeyer, and J. Nielsen (Eds.), The Atlantic Salmon: Genetics, conservation and management. 117–166. Oxford: Blackwell. https://doi.org/10.1002/9780470995846.ch5

Kjøglum, S., Henryon, M., Aasmundstad, T., Korsgaard, I., 2008. Selective breeding can increase resistance of Atlantic salmon to furunculosis, infectious salmon anemia and infectious pancreatic necrosis. Aquac. Res. 39(5), 498-505. https://doi.org/10.1111/j.1365-2109.2008.01904.x

Krossøy, B., Hordvik, I., Nilsen, F., Nylund, A., Endresen, C., 1999. The putative polymerase sequence of infectious salmon anemia virus suggests a new genus within the *Orthomyxoviridae*. J. Virol. 73, 2136-2142.

LeBlanc, F., Laflamme, M., and Gangné, N., 2010. Genetic markers of the immune response of Atlantic salmon (*Salmo salar*) to infectious salmon anemia virus (ISAV). Fish Shellfish Immun. 29(2), 217-232. https://doi.org/10.1016/j.fsi.2010.03.007

Li, J., Boroevich, K.A., Davidson, W.S., 2011. Comparative genomics identifies candidate genes for infectious salmon anemia (ISA) resistance in Atlantic salmon (*Salmo salar*). Mar. Biotechnol. 13(2), 232-241. https://doi.org/10.1007/s10126-010-9284-0

Lien, S., Koop, B.F., Sandve, S.R., Miller, J.R., Kent, M.P., Nome, T., Hvidsten, T.R., Leong, J.S., Minkley, D.R., Zimin, A., Grammes, F., Grove, H., Gjuvsland, A., Walenz, B., Hermansen, R.A., von Schalburg, K., Rondeau, E.B., Di Genova, A., Samy, J.K.A., Vik, J.O., Vigeland, M.D., Caler, L., Grimholt, U., Jentoft, S., Våge, D.I., de Jong, P., Moen, T., Baranski, M., Palti, Y., Smith, D.R., Yorke, J.A., Nederbragt, A.J., Tooming-Klunderud, A., Jakobsen, K.S., Jiang, X., Fan, D., Hu, Y., Liberles, D.A., Vidal, R., Iturra, P., Jones, S.J.M., Jonassen, I., Maass, A., Omholt, S.W., Davidson, W.S., 2016.

The Atlantic salmon genome provides insights into rediploidization. Nature 533, 200-205. https://doi.org/10.1038/nature17164

Liu, L., Ang, K.P., Elliott, J.A.K., Kent, M.P., Lien, S., MacDonald, D., Boulding, E.G., 2017. A genome scan for selection signatures comparing farmed Atlantic salmon with two wild populations: Testing colocalization among outlier markers, candidate genes, and quantitative trait loci for production traits. Evol. Appl. 10, 276-296. http://dx.doi.org/10.1111/eva.12450.

Lyngøy, C., 2003. Infectious salmon anemia in Norway and the Faroe Islands: an industrial approach. In: Miller, O., Cipriano, R.C. (Eds.), International response to infectious salmon anemia: prevention, control, and eradication: proceedings of a symposium, New Orleans, LA, 3–4 September, 2002. Tech. Bull. No. 1902, Washington, DC: U.S. Department of Agriculture, Animal and Plant Health Inspection Service; U.S. Department of the Interior, U.S. Geological Survey; U.S. Department of Commerce, National Marine Fisheries Service, 97–109.

Lyngstad, T.M., Jansen, P.A., Sindre, H., Jonassen, C.M., Hjortaas, M.J., Johnsen, S., Brun, E., 2008. Epidemiological investigation of infectious salmon anaemia (ISA) outbreaks in Norway 2003-2005. Prev. Vet. Med. 84(3-4), 213-227. https://doi.org/10.1016/j.prevetmed.2007.12.008

MacLean, S., Bouchard, D., Ellis S., 2003. Survey of non-salmonid marine fishes for detection of infectious salmon anaemia virus and other salmonid pathogens. In: Miller, O., and Cipriano, R.C., technical coordinators. International response to infectious salmon anemia: prevention, control, and eradication: Proceedings of a symposium, 3-4 September 2002, New Orleans, LA. Technical Bulletin 1902 Washington, D.C.: Department of Agriculture, Animal and Plant Health Inspection Service pp. 135-143.

Markussen, T., Jonassen, C.M., Numanovic, S., Braaen, S., Hjortaas, M., Nilsen, H., Mjaaland, S., 2008. Evolutionary mechanisms involved in the virulence of infectious salmon anaemia virus (ISAV), a piscine orthomyxovirus. Virology 374, 515-527. https://doi.org/10.1016/j.virol.2008.01.019

Marshall, S.H., Ramírez, R., Labra, A., Carmona, M., Muñoz, C., 2014. Bona fide evidence for natural vertical transmission of infectious salmon anemia virus (ISAV) in freshwater brood stocks of farmed Atlantic salmon (*Salmo salar*) in southern Chile. J. Virol. 88(11), 6012-6018. https://doi.org/10.1128/JVI.03670-13

Mjaaland, S., Hungnes, O., Teig, A., Dannevig, B.H., Thorud, K., Rimstad, E., 2002. Polymorphism in the infectious salmon anemia virus hemagglutinin gene: importance and possible implications for evolution and ecology of infectious salmon anemia disease. Virology 304, 379-391. https://doi.org/10.1006/viro.2002.1658

Moen, T., Sonesson, A.K., Hayes, B., Lien, S., Munck, H., Meuwissen, T.H.E., 2007. Mapping of a quantitative trait locus for resistance against infectious salmon anaemia in Atlantic

salmon (*Salmo salar*): a comparing survival analysis with analysis on affected/resistant data. BMC Genet. 8, 53. https://doi.org/10.1186/1471-2156-8-53

Mullins, J.E., Groman, D., Wadowska, D., 1998. Infectious salmon anaemia in salt water Atlantic salmon (*Salmo salar* L.) in New Brunswick, Canada. Bull. Eur. Assoc. Fish Pathol. 18(4), 110-114.

Murray, A.G., Smith, R.J., Stagg, R.M., 2002. Shipping and the spread of infectious salmon anemia in Scottish aquaculture. Emerg. Infect. Dis. 8, 1-5. https://doi.org/10.3201/eid0801.010144

Nylund, A., Devold, M., Mullins, J., Plarre, H., 2002. Herring (*Clupea harengus*): a host for infectious salmon anemia virus (ISAV). Bull. Eur. Ass. Fish Pathol. 22, 311-318.

Nylund, A., Hovland, T., Hodneland, K., Nilsen, F., Levik, P., 1994. Mechanisms for transmission of infectious salmon anaemia (ISA). Dis. Aquat. Organ. 19, 95-100.

Nylund, A., Jakobsen, P., 1995. Sea-trout as a carrier of infectious salmon anemia virus. J. Fish Biol. 47, 174-176. https://doi.org/10.1111/j.1095-8649.1995.tb01885.x

Nylund, A., Kvenseth, A.M., Krossoy, B., Hodneland, K., 1997. Replication of the infectious salmon anaemia virus (ISAV) in rainbow trout, *Oncorhynchus mykiss* (Walbaum). J. Fish Dis. 20, 275-279. https://doi.org/10.1046/j.1365-2761.1997.00300.x

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(s1), 103-114. https://doi.org/10.1111/j.1365-2109.2010.02669.x

Ødegård, J., Olesen, I., Gjerde, B. Klemetsdal, G., 2007. Positive genetic correlation between resistance to bacterial (furunculosis) and viral (infectious salmon anaemia) diseases in farmed Atlantic salmon (Salmo salar). Aquaculture 273, 173-177. https://doi.org/10.1016/j.aquaculture.2007.06.006

Oelckers, K., Vike, S., Duesund, H., Gonzalez, J., Wadsworth, S., Nylund, A., 2014. *Caligus rogercresseyi* as a potential vector for transmission of Infectious Salmon Anaemia (ISA) virus in Chile. Aquaculture 420-421, 125-132. https://doi.org/10.1016/j.aquaculture.2013.10.016

Olesen, I., Hung, D., Ødegård, J., 2007. Genetic analysis of survival in challenge tests of furunculosis and ISA in Atlantic salmon. Genetic parameter estimates and model comparisons. Aquaculture 272, S297–S298. https://doi.org/10.1016/j.aquaculture.2007.07.155

Philips, R.B., Keatley, K.A., Morasch, M.R., Ventura, A.B., Lubieniecki, K.P., Koop, B.F., Danzmann, R.G., Davidson, W.S., 2009. Assignment of Atlantic salmon (*Salmo salar*) linkage groups to specific chromosomes: Conservation of large syntenic blocks

corresponding to whole chromosome arms in rainbow trout (*Oncorhynchus mykiss*). BMC Genetics 10, 46. https://doi.org/10.1186/1471-2156-10-46

Raynard, R.S., Murray, A.G., Gregory, A., 2001. Infectious salmon anaemia virus in wild fish from Scotland. Dis. Aquat. Organ. 46(2), 93-100. https://doi.org/10.3354/dao046093

Rochus, C.M., Holborn, M.K., Ang, K.P., Elliott, J.A.K., Glebe, B.D., Leadbeater, S., Tosh, J.J., and Boulding, E.G., 2018. Genome-wide association analysis of salmon lice (*Lepeophtheirus salmonis*) resistance in a North American Atlantic salmon population. Aquac. Res. 49(3), 1329-1338. https://doi.org/10.1111/are.13592

Rodger, H.D., Turnbull, T., Muir, F., Millar, S., Richards, R.H., 1998. Infectious salmon anaemia (ISA) in the United Kingdom. Bull. Eur. Assoc. Fish Pathol. 18(4), 115-116.

Simko, E., Brown, L.L., MacKinnon, A.M., Byrne, P.J., Ostland, V.E., Ferguson, H.W., 2000. Experimental infection of Atlantic salmon, *Salmo salar* L., with infectious salmon anaemia virus: a histopathological study. J. Fish Dis. 23, 27-32. https://doi.org/10.1046/j.1365-2761.2000.00203.x

Sonesson, A.K., Meuwissen, T.H.E., 2009. Testing strategies for genomic selection in aquaculture breeding programs. Gen. Sel. Evol. 41, 37. https://doi.org/10.1186/1297-9686-41-37

Speilberg, L., Evensen, Ø., Dannevig, B.H., 1995. A sequential study of the light and electron microscopic liver lesion of infectious anemia in Atlantic salmon (*Salmo salar* L.). Vet. Pathol. 32(5), 466-478. https://doi.org/10.1177/030098589503200503

Thorud, K.E., Djupvik, H.O., 1988. Infectious salmon anaemia in Atlantic salmon (*Salmo salar* L.). Bull. Eur. Assoc. Fish Pathol. 8, 109-111.

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 trait in juvenile farmed Atlantic salmon using a high density SNP array. BMC Genomics 16, 969. https://doi.org/10.1186/s12864-015-2117-9

Verspoor, E., 1997. Genetic diversity among Atlantic salmon (*Salmo salar* L.) populations. ICES J. Mar. Sci. 54, 965-973. https://doi.org/10.1016/S1054-3139(97)80001-X

Yáñez, J.M., Houston, R.D., Newman, S., 2014. Genetics and genomics of disease resistance in salmonid species. Front. Genet. 5, 415. https://doi.org/10.3389/fgene.2014.00415

Table 1. Genome wide significant SNP markers for ISA 'survival at trial termination' and corresponding SNP parameters. The variable n represents the number of fish successfully genotyped at this SNP. The A_2 allele represents the minor allele.

SNP	Chromosome	Position[a]	p value	n	A_1	A_2	MAF	Effect A_2 (SE)	Percent Phenotypic Variation
AX-87024549	*Ssa*13	55007327	1.53×10^{-11}	533	G	A	0.106	-0.222 (0.0329)	8.61
AX-87248478	*Ssa*12	34981886	1.08×10^{-9}	523	C	T	0.162	-0.164 (0.0269)	7.56
AX-87806676	*Ssa*11	9190503	1.21×10^{-8}	532	G	A	0.0987	-0.197 (0.0345)	6.42
AX-96490180	*Ssa*05	8566307	3.55×10^{-8}	532	C	T	0.140	-0.161 (0.0291)	5.83
AX-87545217	*Ssa*10	103636938	3.42×10^{-7}	535	G	A	0.124	-0.153 (0.0230)	5.22
AX-87710310	*Ssa*01	142039673	1.68×10^{-6}	532	A	C	0.195	-0.116 (0.0243)	4.59
AX-87453983	*Ssa*24	3601575	1.98×10^{-6}	520	C	A	0.135	-0.141 (0.0296)	4.52
AX-86955267	*Ssa*14	32136776	1.48×10^{-6}	527	C	T	0.189	-0.121 (0.0252)	4.38
AX-96136557	*Ssa*16	27381873	1.65×10^{-6}	520	T	C	0.133	-0.144 (0.0300)	4.42

[a] The position on the chromosome is an approximate location.

Table 2. Protein coding genes within a 0.1Mb distance from nine genome wide significant SNPs associated with the trait 'survival at trial termination' at the suggestive ($\alpha = 0.05$) genome-wide level.

Chromosome	Gene Name	Position	Gene Product
*Ssa*13	LOC106567521	ssa13:54801869..54945875 (+ strand)	rap guanine nucleotide exchange factor 6-like, transcript variant X1
*Ssa*13	LOC106567522	ssa13:55079115..55297652 (- strand)	follistatin-related protein 4-like
*Ssa*12	LOC106565061	ssa12:34898127..34980877 (- strand)	dystroglycan-like, transcript variant X3
*Ssa*12	LOC106565062	ssa12:34943220..34945025 (+ strand)	uncharacterized LOC106565062, transcript variant X2
*Ssa*11	LOC106561937	ssa11:8949290..9126433 (- strand)	leucine zipper protein 2-like, transcript variant X2
*Ssa*11	LOC106561938	ssa11:9144268..9150312 (- strand)	39S ribosomal protein L21, mitochondrial-like, transcript variant X1
*Ssa*11	LOC106561939	ssa11:9159006..9215406 (- strand)	CUGBP Elav-like family member 1, transcript variant X3
*Ssa*11	LOC106561940	ssa11:9215641..9219193 (+ strand)	uncharacterized LOC106561940, transcript variant X5
*Ssa*11	LOC106561941	ssa11:9226163..9238858 (- strand)	complement C1q tumor necrosis factor-related protein 4-like

*Ssa*11	LOC106561942	ssa11:9221145..9225781 (+ strand)	NADH dehydrogenase [ubiquinone] iron-sulfur protein 3, mitochondrial-like
*Ssa*05	LOC106604223	ssa05:8462248..8469211 (- strand)	uncharacterized LOC106604223
*Ssa*05	LOC106604222	ssa05:8470224..8498096 (+ strand)	BTB/POZ domain-containing protein KCTD16-like
*Ssa*05	LOC106604221	ssa05:8541796..8581300 (+ strand)	SH3 domain-containing RING finger protein 3-like, transcript variant X3
*Ssa*05	kif3a	ssa05:8582691..8608263 (+ strand)	kinesin family member 3A, transcript variant X5
*Ssa*10	LOC106561554	ssa10:103520912..103530846 (- strand)	caveolin-1
*Ssa*10	LOC106561555	ssa10:103532194..103539415 (- strand)	caveolin-2-like
*Ssa*10	LOC106561556	ssa10:103633788..103649768 (- strand)	testin-like, transcript variant X2
*Ssa*10	LOC106561558	ssa10:103652046..103669422 (+ strand)	carboxypeptidase A1-like, transcript variant X3
*Ssa*10	LOC106561560	ssa10:103659509..103663532 (+ strand)	carboxypeptidase A1-like, transcript variant X1
*Ssa*10	cep41	ssa10:103669441..103677257 (- strand)	centrosomal protein 41kDa, transcript variant X1
*Ssa*10	poc1b	ssa10:103677333..103746798 (+ strand)	POC1 centriolar protein B, transcript variant X1
*Ssa*01	LOC106569485	ssa01:141997378..142008536 (- strand)	probable global transcription activator SNF2L2, transcript variant X2

*Ssa*01	LOC106569472	ssa01:142023415..142053554 (- strand)	KN motif and ankyrin repeat domain-containing protein 1-like, transcript variant X3
*Ssa*01	LOC106569451	ssa01:142072572..142087824 (+ strand)	urea transporter 2-like, transcript variant X1
*Ssa*01	LOC106569446	ssa01:142095352..142099584 (+ strand)	dnaJ homolog subfamily B member 5-like
*Ssa*01	LOC106569438	ssa01:142105919..142107960 (+ strand)	uncharacterized LOC106569438
*Ssa*01	LOC106569430	ssa01:142108057..142127112 (- strand)	transitional endoplasmic reticulum ATPase-like
*Ssa*01	LOC106569424	ssa01:142128054..142134348 (- strand)	ATP synthase subunit alpha, mitochondrial
*Ssa*01	haus1	ssa01:142134127..142136664 (+ strand)	HAUS augmin-like complex, subunit 1, transcript variant X2
*Ssa*24	LOC106585228	ssa24:3507761..3631854 (+ strand)	reticulon-4 receptor-like, transcript variant X1
*Ssa*24	CIGSSA_115407	ssa24:3652916..3653113 (- strand)	PiggyBac transposable element-derived protein 4|hypothetical; weakly similar(predicted)
*Ssa*14	LOC106569323	ssa14:32014784..32028932 (+ strand)	THO complex subunit 1-like, transcript variant X1
*Ssa*14	LOC106569325	ssa14:32029033..32036081 (- strand)	ubiquitin carboxyl-terminal hydrolase 14-like, transcript variant X1
*Ssa*14	LOC106569322	ssa14:32037088..32114484 (- strand)	rho-associated protein kinase 1-like, transcript variant X1

*Ssa*14	LOC106569321	ssa14:32122105..32214857 (+ strand)	GREB1-like protein, transcript variant X6
*Ssa*14	CIGSSA_076364	ssa14:32146699..32148356 (+ strand)	Uncharacterized protein|hypothetical(predicted)
*Ssa*14	CIGSSA_076368	ssa14:32210119..32210855 (+ strand)	Unknown(predicted)
*Ssa*14	LOC106569320	ssa14:32235160..32266752 (- strand)	phospholipase ABHD3-like
*Ssa*16	LOC106573504	ssa16:27155158..27295386 (- strand)	talin-2, transcript variant X1
*Ssa*16	LOC106573529	ssa16:27300078..27301788 (+ strand)	C2 calcium-dependent domain-containing protein 4C-like
*Ssa*16	LOC106573506	ssa16:27440279..27710925 (+ strand)	nuclear receptor ROR-alpha-like, transcript variant X1

Table 3. Atlantic salmon genes from the cGRASP microarray that showed differential expression after ISA infection from LeBlanc et al. (2010) that co-located with a statistically significant SNPs associated with ISA resistance measured as 'survival at trial termination' located upstream and on the same chromosome arm.

GB Accession number	Gene ID	Chromosome	Gene location on chromosome (Mb)	Up or Down Regulation
CB493346	Cyclophilin_ABH-like	*Ssa*01	142.8	up
CB505886	Peptidyl-protyl cis-trans isomerase A	*Ssa*01	142.8	up
CA054167	Lysozyme C II precursor	*Ssa*01	150.6	up
CA040401	Stonustoxin subunit beta (TRIM protein)	*Ssa*05	13.6	up
EG871962	Tripartite motif-containing protein 25	*Ssa*05	13.6	up
CA042241	Tripartite motif-containing protein 16	*Ssa*11	19.4	up

CA049433	Ferritin heavy subunit	*Ssa*11	24.2	up
CB501070	Cytochrome p450 1A3	*Ssa*11	26.1	up
CB511232	Peptidyl-protyl cis-trans isomerase A	*Ssa*11	49.6	up
EG936494	Glucose regulated protein precursor	*Ssa*11	65.7	up
CB497878	Proactivator polypeptide precursor	*Ssa*11	72.4	up
CA049405	Ras association domain-containing protein 1	*Ssa*12	44.4	up
CA055175	Translocator protein (benzodiazepine receptor)	*Ssa*12	50.1	down
CB512696	Cathepsin Z precursor	*Ssa*12	52.8	up
CA041370	Transcription factor MafB	*Ssa*12	70.8	up
DW565581	Sulfate transporter	*Ssa*13	90.1	up
CK990501	Guanine nucleotide-binding protein	*Ssa*14	32.7	down

CA056363	Peroxiredoxin	*Ssa*14	33.8	up
EG770376	CD9 tetraspannin	*Ssa*14	38.7	up
CB496983	Hematopoietically-expressed homeobox protein hex	*Ssa*14	45.5	down
CA044407	Class I histocompatibility antigen	*Ssa*14	51.0	up
CA044472	BOLA class I histocompatibility antigen, alpha chain	*Ssa*14	59.5	up
CA043257	Class I histocompatibility antigen	*Ssa*14	59.5	up
DY714754	Tripartite motif-containing protein 29	*Ssa*14	63.0	up
CA049981	Selenoprotein Pa precursor	*Ssa*24	5.8	down
BU965885	Myosin light chain	*Ssa*24	7.3	down
CA061671	Unknown	*Ssa*24	24.8	up
CB496964	Linker histone (Histone H1)	*Ssa*24	38.5	down
CK991241	Linker histone (Histone H1)	*Ssa*24	38.5	down
CA063753	Chemokine_CC_DCCL	*Ssa*24	41.0	up

Table 4. Atlantic salmon candidate genes differentially expressed after ISA infection from LeBlanc et al. (2010) that co-located with a statistically significant SNP associated with ISA resistance measured as 'survival at trial termination' located upstream and on the same chromosome arm.

GB Accession number	Gene ID	Chromosome	Gene location on chromosome (Mb)
BT044710	Class I MHC antigen alpha chain	*Ssa*14	51.0
EG871962	Tripartite motif-containing protein 25	*Ssa*10	13.6
BT049801	CC chemokine 19	*Ssa*24	41.0

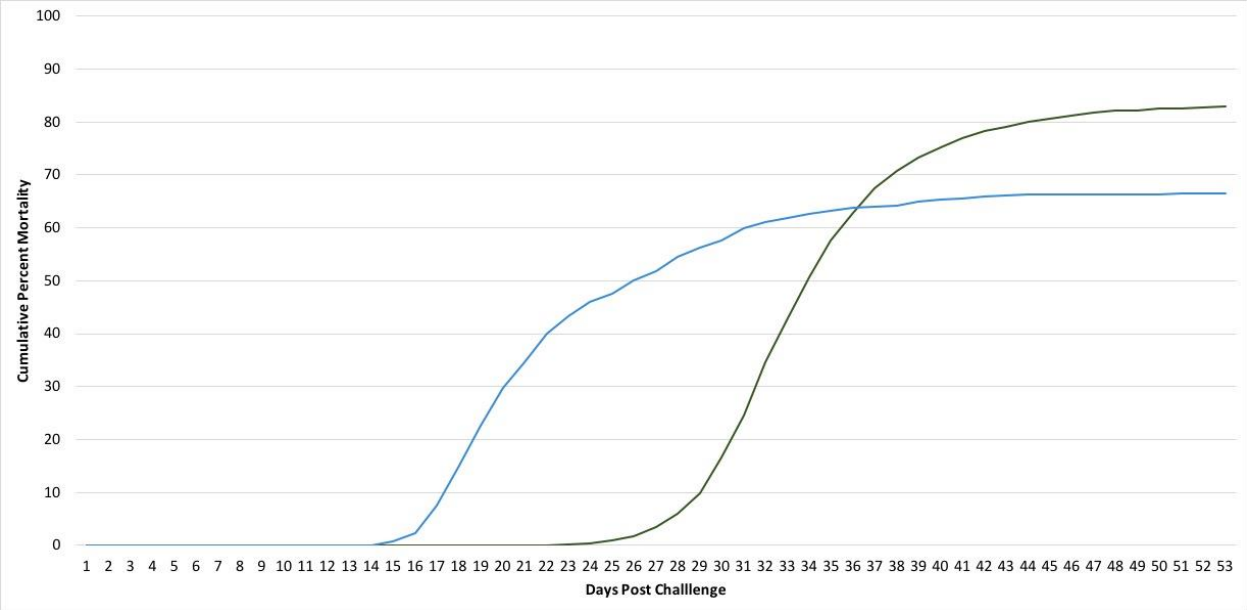


Figure 1. Cumulative percent mortality for the 53 days of the ISA challenge. The green line represents the mortality curve of the cohab fish, while the blue line represents the mortality of the shedder fish that were injected with the ISA virus. The trial was terminated on day 54 after mortalities leveled off.

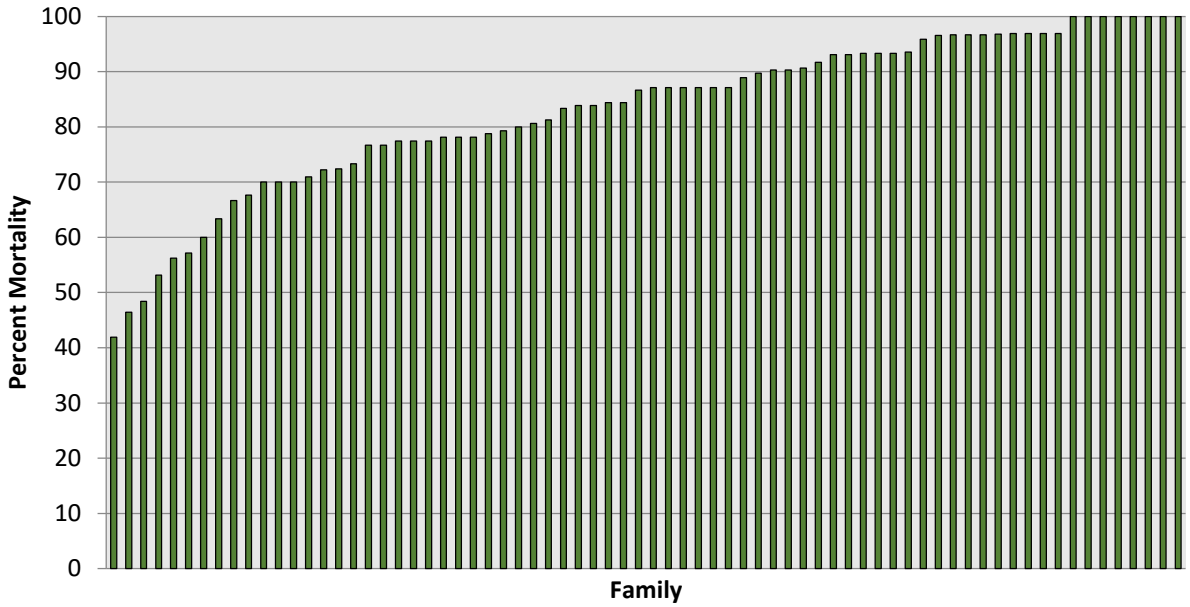


Figure 2. Overall percent mortality for all 72 families present in the ISA challenge. The average percent mortality at termination of the trial at 53 days.

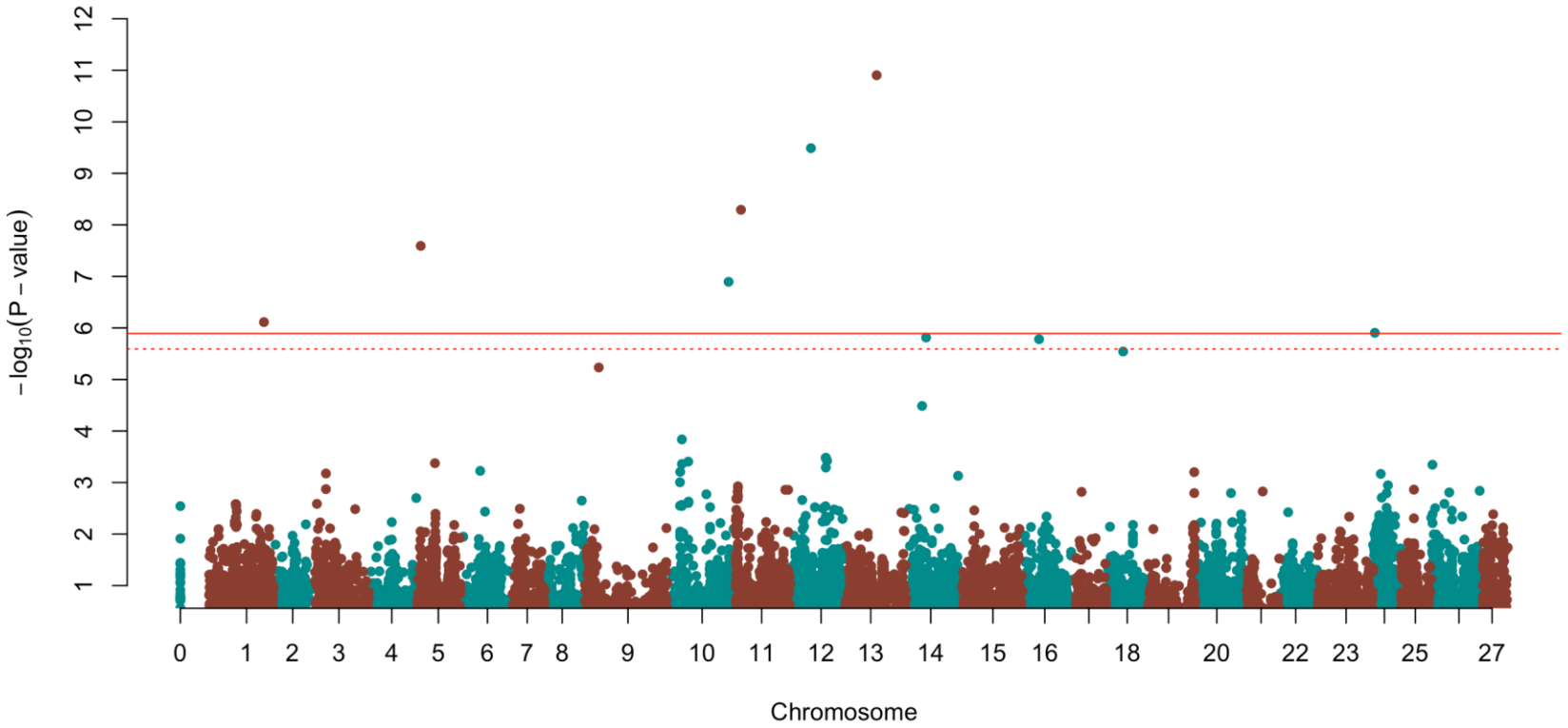


Figure 3. Genome wide association analysis for ISA survival as 'survival at trial termination'. The solid line represents genome wide Bonferroni significance while the dotted line represents the suggestive genome wide Bonferroni significance threshold.

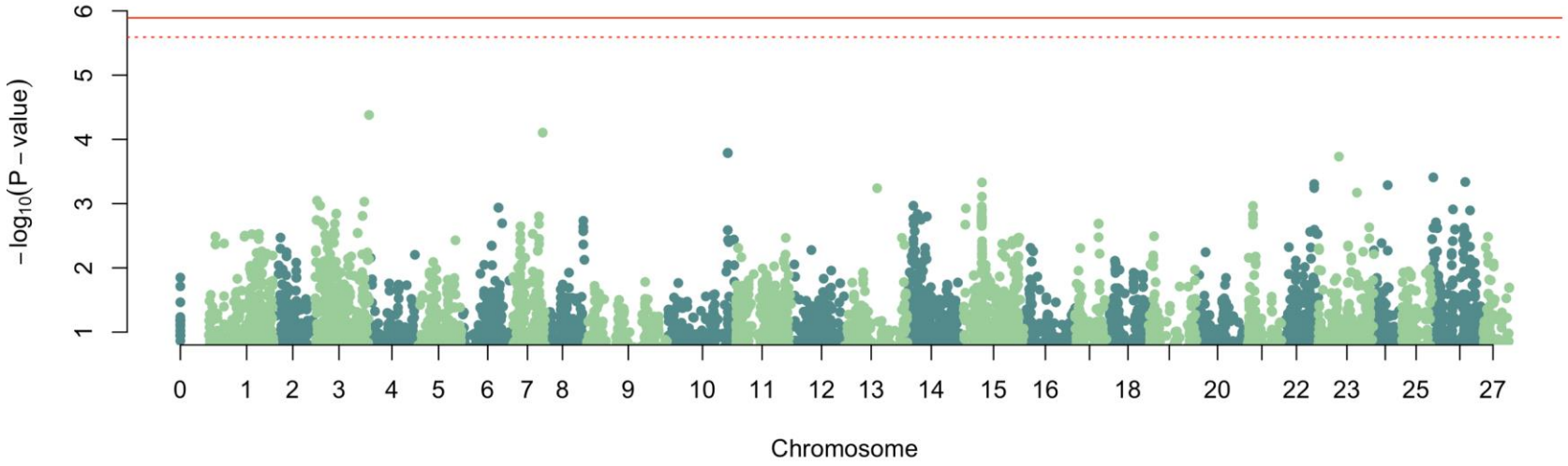
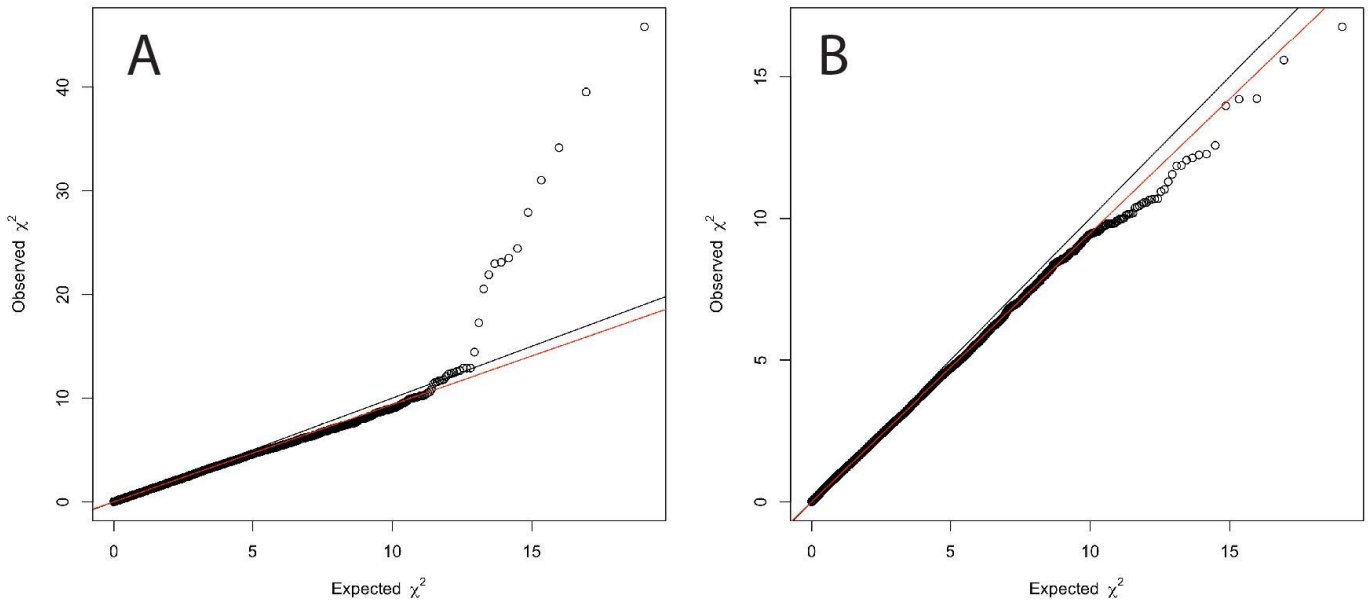
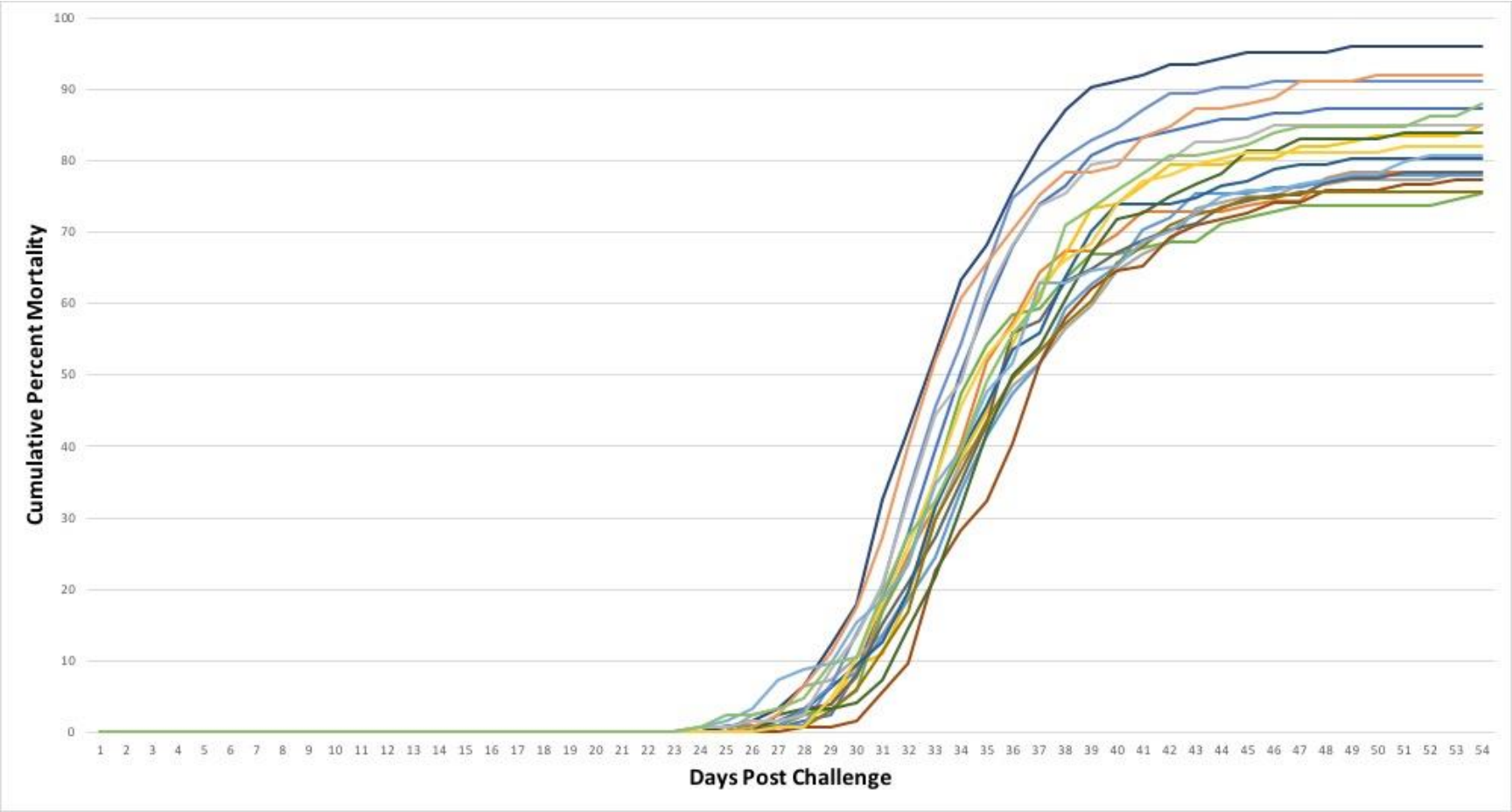


Figure 4. Genome wide association analysis for ISA 'survival to day 37', where there was 50% mortality in the genotyped dataset. The solid line represents genome wide Bonferroni significance while the dotted line represents the suggestive genome wide Bonferroni significance threshold.

Supplementary Material



Supplementary Material Figure S1. QQplots for genome wide association analyses for ISA resistance indicated that there was slight under-inflation that was corrected by adjusting the p values using the genomic control factor lambda, λ. Black line represents the theoretical χ^2 values expected under the null hypothesis. A. QQplot for 'survival at trial termination' where λ = 0.94. B. QQplot for 'survival to day 37' where λ = 0.95.



Supplementary Material Figure S2. The cumulative percent mortality across all 18 tanks in the ISA challenge until trial termination at day 54. Each line represents a different tank.

	Upstream SNPs										Downstream SNPs								
	9	8	7	6	5	4	3	2	1		1	2	3	4	5	6	7	8	9
Chromosome 1	0.008	0.000	0.006	0.006	0.013	0.014	0.004	0.013	0.014	Sig SNP	0.041	0.014	0.014	0.009	0.006	0.005	0.000	0.000	0.000
Chromosome 5	0.024	0.045	0.015	0.119	0.012	0.100	0.018	0.229	0.173	Sig SNP	0.211	0.034	0.010	0.007	0.007	0.006	0.006	0.005	0.001
Chromosome 10	0.018	0.029	0.000	0.034	0.031	0.000	0.000	0.029	0.031	Sig SNP	0.000	0.028	0.000	0.031	0.029	0.032	0.031	0.154	0.154
Chromosome 11	0.001	0.000	0.001	0.002	0.002	0.002	0.001	0.000	0.001	Sig SNP	0.001	0.003	0.000	0.017	0.002	0.002	0.002	0.002	0.002
Chromosome 12	0.007	0.005	0.004	0.005	0.001	0.001	0.002	0.002	0.002	Sig SNP	0.310	0.020	0.001	0.001	0.000	0.002	0.003	0.000	0.007
Chromosome 13	0.000	0.000	0.000	0.000	0.004	0.004	0.003	0.000	0.002	Sig SNP	0.004	0.040	0.035	0.001	0.009	0.001	0.000	0.002	0.008
Chromosome 14	0.072	0.002	0.003	0.088	0.088	0.005	0.083	0.450	0.090	Sig SNP	0.418	0.029	0.042	0.040	0.046	0.005	0.050	0.040	0.014
Chromosome 16	0.005	0.005	0.067	0.028	0.058	0.076	0.085	0.000	0.006	Sig SNP	0.085	0.004	0.005	0.123	0.044	0.112	0.088	0.020	0.043
Chromosome 24	0.001	0.001	0.001	0.001	0.005	0.005	0.018	0.029	0.027	Sig SNP	0.214	0.222	0.217	0.231	0.222	0.090	0.090	0.090	0.055

Supplementary Material Figure S3. The level of pairwise linkage disequilibrium (r^2) between the significant single nucleotide polymorphism (Sig SNP – see Table 1) and the nine upstream and nine downstream flanking SNPs, representing 0.1 Mb distance to the significant SNP. The grey coded SNPs represent SNPs with a r^2 value of 0, light blue a r^2 value lower than 0.2, medium blue a r^2 value between 0.2 and 0.4, and dark blue a r^2 value of above 0.4.