

Genomic Predictions For Fillet Yield And Firmness In Rainbow Trout Using Reduced-Density SNP Panels

Rafet Al-Tobasei

Middle Tennessee State University

Ali R. Ali

University of Maryland, College Park

Andre L. S. Garcia

University of Georgia

Daniela Lourenco

University of Georgia

Tim Leeds

USDA-ARS National Center for Cool and Cold Water Aquaculture

Mohamed Salem (✉ mosalem@umd.edu)

Department of Animal and Avian Sciences, University of Maryland, college park <https://orcid.org/0000-0003-2142-6716>

Research article

Keywords: genomic selection, GEBV, EBV, LD pruning, predictive ability

Posted Date: December 23rd, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-36925/v3>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published on January 30th, 2021. See the published version at <https://doi.org/10.1186/s12864-021-07404-9>.

Genomic Predictions For Fillet Yield And Firmness In Rainbow Trout Using Reduced-Density SNP Panels

Rafet Al-Tobasei¹, Ali Ali², Andre L. S. Garcia³, Daniela Lourenco³, Tim Leeds⁴, and Mohamed Salem²

¹Computational Science Program, Middle Tennessee State University, Murfreesboro, TN 37132, U.S.

²Department of Animal and Avian Sciences, University of Maryland, College Park, Maryland 20742 U.S.

³Department of Animal and Dairy Science, University of Georgia, Athens, GA 30602, U.S.

⁴National Center for Cool and Cold Water Aquaculture, Agricultural Research Service, United States Department of Agriculture Kearneysville, WV, U.S.

Corresponding author: Mohamed Salem

E-mail: mosalem@umd.edu

ABSTRACT

Background

One of the most important goals for the rainbow trout aquaculture industry is to improve fillet yield and fillet quality. Previously, we showed that a 50K transcribed-SNP chip can be used to detect quantitative trait loci (QTL) associated with fillet yield and fillet firmness. In this study, data from 1,568 fish genotyped for the 50K transcribed-SNP chip and ~774 fish phenotyped for fillet yield and fillet firmness were used in a single-step genomic BLUP (ssGBLUP) model to compute the genomic estimated breeding values (GEBV). In addition, pedigree-based best linear unbiased prediction (PBLUP) was used to calculate traditional, family-based estimated breeding values (EBV).

Results

The genomic predictions outperformed the traditional EBV by 35% for fillet yield and 42% for fillet firmness. The predictive ability for fillet yield and fillet firmness was 0.19 - 0.20 with PBLUP, and 0.27 with ssGBLUP. Additionally, reducing SNP panel densities indicated that using 500 – 800 SNPs in genomic predictions still provides predictive abilities higher than PBLUP.

Conclusion

These results suggest that genomic evaluation is a feasible strategy to identify and select fish with superior genetic merit within rainbow trout families, even with low-density SNP panels.

Keywords: genomic selection, GEBV, EBV, LD pruning, predictive ability

BACKGROUND

Aquaculture has great potential to enhance food security and meet the increasing consumer demand for seafood [1, 2]. However, one of the challenges is the lack of genetically improved strains of fish for aquaculture [3, 4]. Selective breeding programs can produce fish of improved genetics for heritable traits that positively impact aquaculture production [5, 6]. Breeding programs in rainbow trout have focused on the growth rate, disease resistance, and fat content [7-9].

Fillet yield has been ranked among the top traits impacting the industry returns [10]. Also, quality attributes can affect industry profitability and determine consumer's attitudes towards the product. For instance, loss of fillet firmness contributes to fillet downgrading and economic losses to the industry [11]. Even though important for the industry, fillet yield and firmness have not received much attention because they cannot be measured directly on breeding candidates, which makes genetic selection for these traits hard to implement [12, 13]. Moderate levels of heritability estimates for fillet yield [12] and fillet firmness [13] have been reported in rainbow trout, allowing potential genetic improvement through selective breeding programs [14].

Statistical models based on phenotypes and pedigree information have been widely used in traditional genetic improvement programs to estimate EBV and identify the best selection candidates in animal populations [15]. However, applying genomic approaches has the potential to enhance and expedite genetic gains in breeding programs [16]. Although the implementation of genomic selection (GS) in livestock animals started in 2008 [17], it took more time to have this technology adopted in aquaculture species. The delayed incorporation of genomic information in rainbow trout breeding programs was mainly due to the lack of dense SNP arrays [18]. A recently developed 50K SNP chip revealed the complex polygenic nature of fillet yield [19] and fillet firmness [13], suggesting GS as a practical strategy for rainbow trout breeding.[20]

Selection based on GEBV can be performed at an early age with high accuracy, once a DNA sample for selection candidates is obtained [21]. The use of genomic information has been shown to be effective in increasing the gains in accuracy of GEBV for multiple traits in aquaculture species, including Atlantic salmon [22, 23], catfish [24], tilapia [20, 25], and rainbow trout [12, 26-29]. For rainbow trout, the accuracy of GEBV was assessed for body weight, carcass weight, fillet weight, fillet yield [12], and resistance to diseases such as bacterial cold water disease

(BCWD) [26, 27], columnaris disease [28], and infectious pancreatic necrosis virus [29]. GS will allow within-family selection and hence increase the accuracy of genomic predictions and selection response, especially for lethally measured traits where within-family selection relies on traits measured in sibs of breeding candidates [25]. Additionally, GS has the potential to decrease the rates of inbreeding by selecting non-sib candidates from more families [26]. Genomic prediction can lead to higher genetic gains relative to pedigree-based selective breeding methods, which may, partially, cover the extra cost of genotyping [30]. Furthermore, GS is particularly advantageous in aquaculture species because the high fecundities of these species allow for the rapid amplification of elite genetics.

Although the use of genomic information generated from high-density SNP arrays has been demonstrated to expedite the rate of genetic gain in breeding programs, the genotyping cost is still high and alternative strategies are needed to reduce the cost of identifying elite breeding candidates [25]. Cost-effective strategies were previously assessed, yielding higher genomic prediction accuracies than those estimated using the pedigree-based PBLUP model [23, 25, 26]. The cost-reducing methods include using reduced-density SNP panels [26, 31, 32] and genotype imputation [21, 25, 33, 34]. However, imputations are prone to errors, which leads to less reliable genomic predictions [33]. Recently, the impact of low-density SNP panels on the accuracy of genomic predictions has been increasingly studied [26, 31, 32, 35, 36]. In rainbow trout, a 500 SNP panel showed higher genomic prediction accuracies for BCWD resistance (0.50 – 0.56) than traditional EBV (0.36) [26]. The study showed that the ssGBLUP model outperformed other genomic models, yielding high accuracy of genomic predictions with the least bias when reducing SNP panel density [26].

Recently, the accuracy of genomic prediction was assessed for fillet yield in rainbow trout using a 57K genomic SNP panel [12]; however, the impact of reducing the SNP panel density on the gains in accuracy for fillet yield relative to the traditional PBLUP has not been investigated. Additionally, there are no reports on the benefits of using GS for quality traits such as fillet firmness in rainbow trout. Therefore, the objectives of this study were to utilize a recently developed 50K transcribed SNP chip [19] to 1) evaluate the predictive ability of GS for fillet yield and fillet firmness in rainbow trout; and 2) investigate the impact of reducing the SNP panel density on the ability to predict fillet yield and fillet firmness.

RESULTS AND DISCUSSION

Phenotypes and heritability estimates

The total numbers of phenotyped fish for fillet yield and shear force were 775 and 772, respectively, and varied per year-class (YC) with 471 fish in YC 2012, and 304 fish in YC 2010. Descriptive statistics of the data are provided in Table 1. A slightly higher fillet yield and shear force were observed for fish from the YC 2012 compared to that from YC 2010. A small phenotypic correlation ($R = 0.14$) was observed between fillet yield and shear force in this study.

A total of 1,572 genotyped fish passed the quality check (QC) and were available for this study (four samples were removed due to fish ID conflicts). The total number of genotyped individuals from YC 2010 and 2012 were 400 and 558, respectively (Table 1). In addition, 380 fish were genotyped before YC 2010, whereas 234 were genotyped after YC 2012. No phenotypic data were available for fish produced before YC 2010 or after YC 2012 (Table 1). All fish had complete pedigree information dating back to YC 2002. Figure S1 depicts a heatmap of the pedigree (A22) and the genomic (G) relationship matrices for data generated from the population used in this study.

Table 1. Descriptive statistics and the number of genotyped and phenotyped fish per year class.

Year class	Fillet yield (%)		Shear force (g/g)	
	2010	2012	2010	2012
Mean	48.21	49.59	448.92	491.91
Median	48.45	49.68	450.91	487.01
Max	55.30	55.62	724.39	822.67
Min	34.77	39.22	166.98	234.24
SD	2.64	1.94	87.53	81.44
CV%	5.48	3.91	5.13	6.04
Phenotyped	304	471	301	471
Genotyped	400	558	400	558

In this study, the heritability estimates were calculated based on all data from YC 2010 and YC 2012 using PBLUP and ssGBLUP models and are shown in Table 2. Variance components are provided in Additional file 1 (Table S1). The estimated heritability values using both methods were higher for fillet firmness than for fillet yield (Table 2). Other studies reported a slightly higher heritability (0.34-0.36) of fillet yield in other rainbow trout populations [12, 37]. The

estimated heritability values indicate a moderate additive genetic component for these traits. Garcia-Ruiz et al. (2016) showed that genomic information can help to increase accuracy even for lowly heritable traits. Additionally, increased accuracies in a trait with $h^2 = 0.14$ was reported in Atlantic salmon using GBLUP [22], which provides evidence for the feasibility of GS in the current NCCCWA fish population.

Table 2. Heritability, predictive ability, and regression coefficient of adjusted phenotype on (G)EBV for fillet yield and fillet firmness using PBLUP and ssGBLUP.

	Fillet yield		Fillet firmness	
	PBLUP (SE)	ssGBLUP (SE)	PBLUP (SE)	ssGBLUP (SE)
Heritability	0.25 (0.073)	0.26 (0.073)	0.38 (0.082)	0.38 (0.082)
Predictive ability	0.20 (0.078)	0.27 (0.077)	0.19 (0.079)	0.27 (0.077)
Regression coefficient (b1)	0.96	0.97	0.79	0.88

Genomic predictions using 50K SNP panel

The predictive ability of the genomic model was evaluated using a five-fold cross-validation scheme. Using five replicates helps to minimize errors that could be generated due to a single-fold sampling [28]. Table 2 shows the average predictive ability for fillet yield and firmness under the five-fold cross-validation strategy. For both traits, using genomic information through ssGBLUP enhanced the ability to predict fish performance relative to PBLUP. Genomic information has a greater impact on the predictive ability of lethally-measured traits than traits that can be directly measured on fish [12, 24]. In this study, genomic information increased predictive ability by 35% and 42%, compared to PBLUP, for fillet yield and firmness, respectively. Gonzalez-Pena *et al.* [12] also reported higher accuracy of GEBV compared to EBV; 0.55 and 0.13, respectively. Similar gains through GS have been achieved in other aquaculture species. For instance, a 29% increase in predictive ability, relative to PBLUP, was reported for residual carcass weight in channel catfish [24].

Variable gains of accuracy using genomic information were reported in Atlantic salmon for lice resistance (52%), fillet color (22%) [22], and growth traits (20%) [23]. The considerable variation in the relative improvement of accuracy between the traits is consistent with the difference in predictive ability between fillet yield and firmness in the current study. Even though high predictive abilities have been achieved by applying genomic information in this study, using

a higher number of fish in the training population would help to achieve higher prediction accuracies. Low predictive abilities (0.46 – 0.49) for BCWD phenotypes in trout were observed using ssGBLUP compared to PBLUP (0.50) when a small training sample size ($n = 583$) was used [38]. Incorporating more genotyped and phenotyped fish in the analysis improved the accuracy of the GEBVs by ~80% [27].

The regression coefficients (b_1), representing inflation for fillet yield and firmness, are shown in Table 2. For fillet firmness, GEBV was less inflated than EBV. These results are consistent with results reported for harvest weight and residual carcass weight in catfish [24] and BCWD resistance in rainbow trout [27], showing that the genomic information provides not only more accurate but also less biased evaluations. In this study, variance components estimated from the whole dataset were used. Using the same variance components for a five-fold cross-validation strategy yielded less inflated GEBV for fillet yield, close to 1 ($b_1 = 0.97$), than that of the harvest weight ($b_1 = 0.92$) and residual carcass weight ($b_1 = 0.91$) in catfish [24], and BCWD survival status in rainbow trout ($b_1 = 0.86$) [27]. On the other hand, breeding values for fillet firmness were more inflated ($b_1 = 0.88$) than those for fillet yield. Updating the variance components for the training datasets in catfish have been suggested to reduce the inflation of the genomic evaluations [24].

Fish used in this study were sampled from a genetic line selected for a high growth rate. Previously, when fillet yield and shear force phenotypes were regressed on body weight, coefficient of determination (R^2) values of 0.56 and 0.01 were observed, respectively [39]. Selection for growth in this population might have less representation of fish with low-ranked phenotype for fillet yield. Fish with phenotypes used in the study were sampled in a manner that captures between- and within-family variation for growth performance although. In an admixed population of Atlantic salmon, fish families produced from selection lines showing extreme phenotypes for lice resistance were over-represented among the genotyped fish leading to inflation in the between-family variation and less reliability of the GS model [22].

Linkage disequilibrium and effective population size

Long-range linkage disequilibrium (LD) and small effective population size (N_e) provided evidence of the possibility to reduce the marker density needed for GS in catfish and rainbow trout selectively-bred for BCWD resistance [24, 26]. When the historical LD is weak, the accuracy of

genomic prediction decreases [40]. In addition, higher genomic prediction accuracies are associated with small N_e [24, 41, 42]. LD and N_e were estimated for the rainbow trout population used in this study (Table 3). The mean LD per chromosome ranged from 0.21 to 0.34, with an overall genome-wide average r^2 of 0.26. Twenty-two chromosomes showed a mean $r^2 \geq 0.25$. The LD ($r^2 = 0.26$) in the current fish population was consistent with LD ($r^2 = 0.27$) reported in a Troutlodge, Inc., May-spawning, odd-year line [26]. Chromosomes 5 and 23 had the highest mean LD estimate ($r^2 = 0.34$) followed by chromosome 1 ($r^2 = 0.32$). Conversely, chromosomes 6 ($r^2 = 0.21$) and 19 ($r^2 = 0.22$) had the lowest mean LD estimates. LD decay with distance was estimated for all the rainbow trout chromosomes using only high-quality anchored SNPs (Figure 1). The LD decay plots show a long-range LD in all chromosomes. Interestingly, the strongest LD appeared on chromosome 5 and extended over 20Mb. In agreement with our results, Vallejo et al. [26] identified long-range LD spanning over 1Mb in all chromosomes and over 25Mb on chromosome 5 in a commercial rainbow trout population. This strong LD on chromosome 5 was likely due to a large chromosomal double inversion, which prevents recombination in heterozygous fish [26, 43]. Also, recent admixture events [22] and recombination interference on the male chromosome were shown to contribute to long-range LD in rainbow trout and other salmonids [12]. The population structure analysis suggested nine genetically different groups in the current population (data not shown), thus supporting population admixture as a contributor to the long-range LD observed in this study. In contrast, a Tasmanian Atlantic salmon population that was originated from a single founder strain showed short-range LD [44].

The mean estimated effective population size based on LD was $N_e = 113$ (Table 3). An average N_e of 145 was reported in the early generations of the NCCCWA rainbow trout selective breeding program and was expected to decline with selection [45]. The estimated N_e in this population is comparable to that estimated in other livestock species; Jersey cattle ($N_e = 101$), Angus cattle ($N_e = 113$), Holstein cattle ($N_e = 149$) [46] and a commercial rainbow trout population ($N_e = 155$) [26], and it is considerably larger than the N_e estimated for catfish ($N_e = 27$) [24], swine ($N_e = 32$), and chicken ($N_e = 44$) [46]. Chromosome 22 had the largest effective population size ($N_e = 180$) followed by chromosome 24 ($N_e = 171$). Chromosomes 5 ($N_e = 18$) and 4 ($N_e = 80$) had the smallest effective population size. In the early generations produced at NCCCWA, an N_e range of 75.51 – 203.35 was reported, and the authors suggested the presence of suites of genes on each chromosome that are disproportionately under selection pressure [45]. The N_e in this population

may have limited the predictive ability because only a small number of genotyped fish were available for computing genomic predictions. Based on the average effective population size ($N_e = 113$) and trout genome length ($L = 32.3$ M), the number of independent chromosome segments (Me) for this population was $\sim 14,600$, according to the equation $Me = 4N_eL$ [47]. In a commercial rainbow trout population, an $Me = 20,026$ was reported [26]. Overall, our results suggest the feasibility of reducing the SNP panel size, which can reduce the genotyping cost. In the next section, further investigation was conducted to determine the impact of reducing the SNP panel size on the predictive ability of GEBV in the current rainbow trout population.

Table 3: Linkage disequilibrium and effective population size by chromosome. Chromosome 5 had the highest mean LD estimate (r^2) and the smallest effective population size.

CHR	Number of SNPs	Size (bp)	Mean (r^2)	N_e
1	1,298	80,480,703	0.32	81
2	1,580	85,221,383	0.26	92
3	1,409	84,915,088	0.25	104
4	1,380	84,297,858	0.27	80
5	2,060	91,890,047	0.34	18
6	1,189	77,071,016	0.21	105
7	1,278	79,714,361	0.24	91
8	1,655	83,683,676	0.27	94
9	1,339	68,350,947	0.30	110
10	1,169	64,390,894	0.26	114
11	1,119	79,489,160	0.26	87
12	1,711	81,222,362	0.24	98
13	1,069	47,531,856	0.25	128
14	1,284	80,268,421	0.26	94
15	913	59,307,523	0.25	100
16	1,572	70,790,504	0.27	113
17	1,379	76,497,411	0.26	104
18	1,133	61,193,526	0.25	125
19	872	58,039,566	0.22	147
20	677	40,981,438	0.26	105
21	783	51,606,644	0.23	165
22	792	48,370,041	0.25	180
23	602	48,805,734	0.34	97
24	633	40,198,439	0.25	171
25	1,330	82,450,733	0.26	97
26	453	30,795,589	0.23	166
27	657	45,210,144	0.27	127
28	843	40,702,581	0.23	153
29	731	42,409,034	0.26	141
Average	1,135	65,030,575	0.26	113
SD	389	17941345	0.031	34

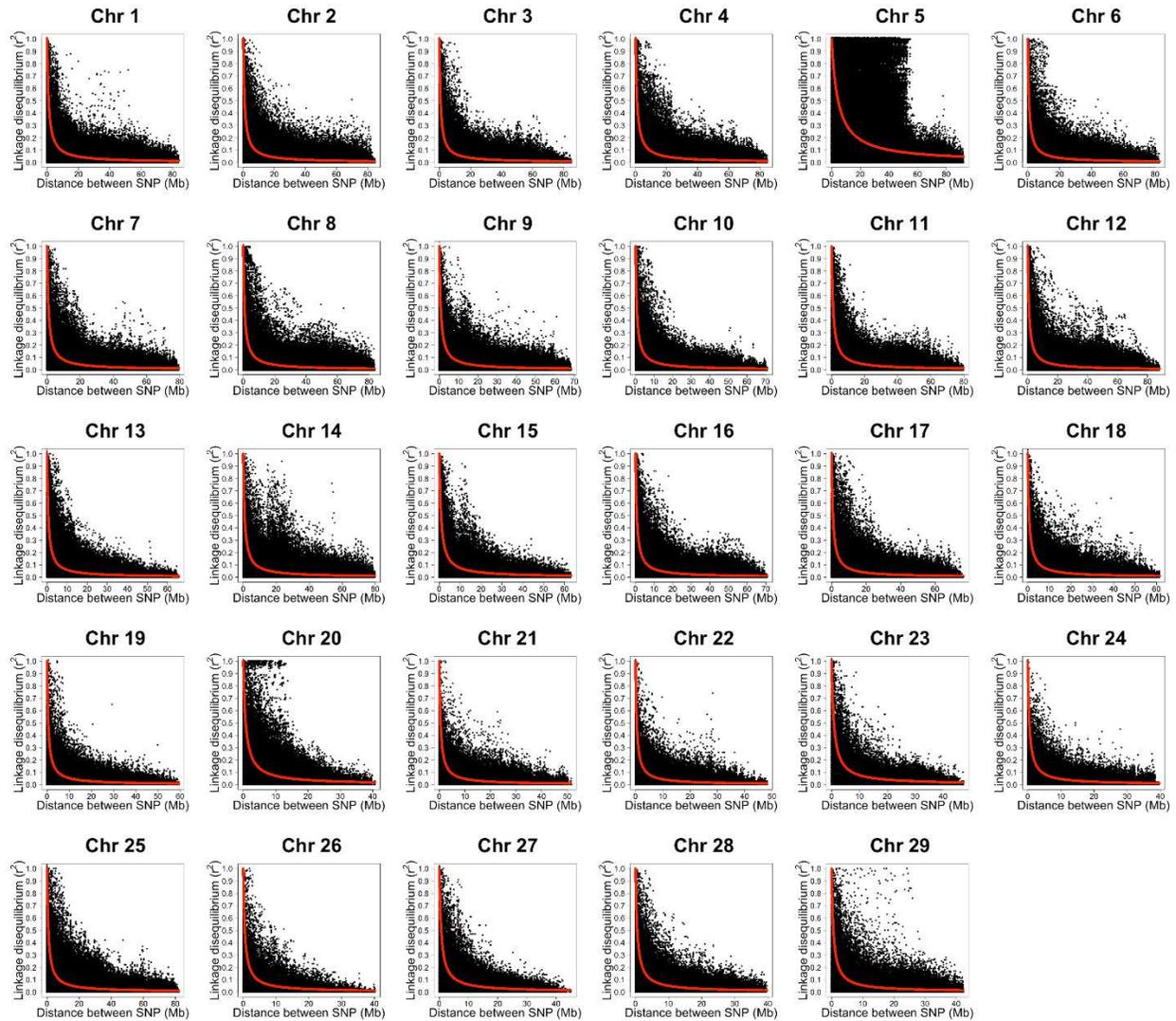


Figure 1. Plot of LD decay estimated for all chromosomes using 50K SNP panel in rainbow trout from the NCCCWA genetic line. Chromosome 5 showed a long-range LD extended over 20Mb.

Impact of reducing the SNP panel density

There is an interest from the aquaculture industry to implement GS in breeding programs. However, genome-wide genotyping using high-density SNP chips may be cost-prohibitive, especially for small and mid-size hatcheries and companies; therefore, a cost-efficient genotyping is necessary.

In the current study, two approaches have been adopted to evaluate the effect of reducing the density of the SNP panel on the genomic predictive ability for fillet yield and firmness.

Reducing the SNP panel density based on LD

In the first approach, LD-based SNP pruning was used to reduce the density of SNP panel throughout the genome [48]. Five LD values were used to prune SNPs ($r^2 > 0.7$, > 0.4 , > 0.1 , > 0.05 , and > 0.01). The predictive ability for fillet yield, based on LD, was reasonably well maintained with the reduction of the SNP panel density down to ~11K SNPs (Table 4). However, further reduction of the marker density, below 11K SNPs, led to lower predictive abilities than those estimated using the PBLUP model (0.20). Similar results have been reported in other aquaculture populations/species. For instance, a moderate decrease in genomic accuracy was reported when reducing the SNP density down to 10K SNPs in a trout population selected for BCWD resistance [26]. Also, increasing SNP density above ~10K SNPs in Atlantic salmon showed no improvement in genomic accuracy for growth traits [23]. This can be related to the theory of dimensionality of genomic information, where the number of animals and SNP needed to achieve a reasonable level of accuracy approaches $4NeL$ [49, 50], which was 14,600 in our study.

In the case of fillet firmness, a reduction of the SNP density down to 1,153K SNPs yielded the same predictive ability (0.27) as the full 50K SNP panel; however, GEBV were more inflated. Interestingly, further reduction of the SNP density down to 497 SNPs yielded predictive abilities (0.24) that, although were lower than the 50K panel, still outperformed the traditional PBLUP model (0.19) (Table 4). Recently, gains of accuracy relative to traditional PBLUP, using 500 SNP panel, were reported for BCWD in rainbow trout. Such gains in accuracy with a significantly reduced SNP panel was attributed to the long-range LD in the studied population [26].

Table 4: Predictive ability and regression coefficient of adjusted phenotype on (G)EBV for fillet yield and firmness with reduced-SNP panel density based on LD.

LD (r^2)	N of SNPs	Predictive ability \pm SE		Regression coefficient (b1)	
		GEBV		GEBV	
		Fillet yield	Fillet firmness	Fillet yield	Fillet firmness
50K chip	35,303	0.27 \pm 0.077	0.27 \pm 0.077	0.97	0.88
> 0.7	20,042	0.27 \pm 0.077	0.27 \pm 0.077	0.99	0.90
> 0.4	11,433	0.26 \pm 0.077	0.26 \pm 0.078	0.98	0.88
> 0.1	1,153	0.19 \pm 0.079	0.27 \pm 0.077	0.71	0.79
> 0.05	497	0.17 \pm 0.079	0.24 \pm 0.078	0.57	0.67
> 0.01	128	0.07 \pm 0.080*	0.02 \pm 0.080*	0.21	0.04

* indicates a significant difference between the low-density SNP panel and the 50K SNP panel in predicting the future performance of fish (t-test p-value < 0.05).

Reducing SNP density based on the percentage of additive genetic variance explained by SNPs

The second approach involved reducing the density of the panel based on the percentage of additive genetic variance explained by SNPs for each trait, which were already determined in our previous publications using weighted single-step genomic best linear unbiased prediction (wssGBLUP) analysis [13, 19]. Reduced SNP panels with the percentage of additive genetic variance between > 0.05% and > 1.8% were used to evaluate the predictive ability using the five-fold cross-validation strategy. Tables 5 and 6 show the predictive ability for fillet yield and firmness using those reduced SNP panels. For fillet yield, reduction of the SNP marker density down to ~9K SNPs yielded the same predictive ability (0.27) that was obtained using the 50K panel. Vallejo et al. [26] reported that SNP panel densities less than 20K SNPs are suitable for GS in rainbow trout. Simulation studies in plants and livestock showed that marker densities higher than ~10K SNPs had little or no improvement in genomic accuracy [51, 52]. Interestingly, further reduction of the number of SNPs down to ~800 SNPs caused a minimal decrease in the predictive ability for fillet yield (0.26). For fillet firmness, reducing the number of SNPs down to 10K SNPs had a small increase in predictive ability (0.30) compared to the 50K SNP panel (0.27). Interestingly, the inflation of GEBV greatly improved for firmness without a significant change in fillet yield. For both traits, the predictive ability of GEBV for all SNP panels down to ~800 SNPs was higher than those estimated using the pedigree-based model. Overall, prioritizing SNPs based on their percentage of additive genetic variance explained allowed a high reduction of the SNP density, down to ~800 SNPs, while maintaining the potential to enhance the accuracy of genomic predictions in this population, but with a trade-off of increasing inflation.

Table 5: Predictive ability and regression coefficient of adjusted phenotype on (G)EBV for fillet yield with reduced-SNP panel density based on the proportion of additive genetic variance explained by SNPs.

% Additive Genetic Variance explained by SNPs	N of SNPs	Predictive ability	Regression coefficient (b1)
		\pm SE	
		GEBV	GEBV
50K	35,303	0.27 \pm 0.077	0.97
> 0.05	16,381	0.25 \pm 0.078	0.87
> 0.1	9,493	0.27 \pm 0.077	0.96
> 0.5	857	0.26 \pm 0.077	0.85
> 1.0	232	0.19 \pm 0.079	0.59
> 1.5	102	0.10 \pm 0.080*	0.34
> 1.8	68	0.12 \pm 0.080*	0.44

* indicates a significant difference between the low-density SNP panel and the 50K SNP panel in predicting the future performance of fish (t-test p-value < 0.05).

Table 6: Predictive ability and regression coefficient of adjusted phenotype on (G)EBV for fillet firmness with reduced-SNP panel density based on the proportion of additive genetic variance explained by SNPs.

% Additive Genetic Variance explained	N of SNPs	Predictive ability \pm	Regression coefficient (b1)
		SE	
		GEBV	GEBV
50K	35,303	0.27 \pm 0.077	0.88
> 0.05	17,592	0.27 \pm 0.077	0.91
> 0.1	10,533	0.30 \pm 0.077	0.99
> 0.5	808	0.23 \pm 0.078	0.70
> 0.8	273	0.17 \pm 0.079	0.47
> 1.0	139	0.12 \pm 0.080*	0.33

* indicates a significant difference between the low-density SNP panel and the 50K SNP panel in predicting the future performance of fish (t-test p-value < 0.05).

Using the proportion of additive genetic variance explained by SNPs for pruning was more effective than the LD approach for fillet yield. Conversely, the LD approach for SNP pruning was more effective in predicting fillet firmness with low-density SNP panels. This could be attributed to the different genetic architecture between traits combined with the number of SNPs retained for analysis.

Altogether, reducing the SNP density to about 10,000 SNPs would likely help to reduce the genotyping cost of implementing GS in this population with no loss in predictive ability and no bias. A further reduction to around 1000 SNP would have a small impact on predictive ability, but GEBV may have an increased level of bias depending on which method was used to reduce SNP density.

CONCLUSIONS

This study reveals the impact of using genomic information on progressing the rainbow trout breeding programs for fillet yield and firmness. Using genomic information improves the ability to predict future performance and reduces the inflation of breeding values. The long-range LD detected in this study, partially due to population admixture, is likely contributing to the high genomic predictive ability in case of the reduced density SNP panels. Reducing the SNP panel density to approximately 10,000 SNPs is a feasible strategy to help to reduce the cost of implementing GS in rainbow trout.

METHODS

Fish population and phenotypes

The fish population and sample collection were previously described in detail [53]. The selective breeding program has been established in 2004 and underwent five generations of selection for body weight. Details on how the population was formed and selected for growth can be found in [8, 54]. Phenotypic records on fillet yield and shear force were collected from the third and fourth generations, comprising YC 2010 and 2012. A total of 775 fish representing 76 full-sib families from YC 2010 and 98 families from YC 2012, were used. To maintain unique pedigree information, each family was kept in a separate 200-L tank until they were tagged at ~5-month post-hatching. Subsequently, fish were reared together in 800-L tanks and fed a commercial fishmeal-based diet.

For phenotyping, fish were harvested over five consecutive weeks (one fish/family/week), yielding 5 harvest groups. The YC 2010 fish were 410 to 437 days old at harvest with a mean body weight of 985 ± 239 (g). Fish from the YC 2012 were 446 to 481 days old, with a mean body weight of $1,803 \pm 305$ (g). Head-on gutted carcasses were manually processed into skinless fillet.

The trimmed fillet was weighed, and fillet yield was calculated as a percent of whole-body weight. The fish (YC 2010 & YC 2012) had an average fillet yield of 48.91 ± 2.42 (%). The shear force of a cooked fillet section was used to measure the fillet firmness as fully described in a previous publication [55].

Genotyping data and quality control check

In the current study, we used a 50K gene-transcribed SNP chip that has recently been developed for rainbow trout [19]. A total of 1,728 fish were genotyped as previously described [13, 19]. Affymetrix SNPolisher software was used at the default parameters to perform quality control and filter out samples that did not meet the threshold filtration criteria. In addition, about 5K SNPs that were not anchored to the newly assembled rainbow trout genome were filtered out because LD calculation (explained in the next sections) requires physically mapped SNPs. Anchored SNPs were subjected to QC analysis using PREGSF90 software, which belongs to the BLUPF90 family [56]. After QC, 35,303 SNPs (70.6%) were retained, those SNPs had minor allele frequency (MAF) > 0.05, call rate for SNP > 0.90, and deviation from the Hardy–Weinberg equilibrium (HWE) < 0.15. The filtered SNPs were used both for GS and LD analysis.

Model and analysis

Two single-trait mixed models were used for fillet yield and shear force (fillet firmness) as follows:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_1\mathbf{a} + \mathbf{Z}_2\mathbf{c} + \mathbf{e}$$

where \mathbf{y} is a vector of phenotypes (fillet yield or fillet firmness), \mathbf{b} is a vector of fixed effects of hatch-year and harvest group, \mathbf{a} is a vector of additive direct genetic effect, \mathbf{c} is a vector of random of common environmental effect (i.e., family effect), and \mathbf{e} is the vector of residuals. Incidence matrices for effect contained in \mathbf{b} , \mathbf{a} , and \mathbf{c} are represented by \mathbf{X} , \mathbf{Z}_1 , and \mathbf{Z}_2 , respectively.

The BLUPF90 from the BLUPF90 family of programs was used to perform both traditional PBLUP and ssGBLUP analyses [56]. The ssGBLUP model uses both pedigree and genomic information to calculate GEBV. Those two sources of information are combined into a realized relationship matrix (\mathbf{H}), where the inverse (\mathbf{H}^{-1}) replaces the inverse of the pedigree relationship matrix (\mathbf{A}^{-1}) in the BLUP mixed model equations. The \mathbf{H}^{-1} was described in [57] as follows:

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix}$$

where \mathbf{A}_{22}^{-1} is the inverse of the pedigree relationship matrix for genotyped animals.

\mathbf{G}^{-1} is the inverse of the genomic relationship matrix, constructed as we previously described [24]

$$\mathbf{G} = \frac{\mathbf{MDM}'}{2 \sum p_j(1 - p_j)'}$$

where M is a matrix of genotypes centered by twice the current allele frequencies (p); j is the jth locus; D is a diagonal matrix of SNP weights with a dimension equal to the number of SNPs. All SNPs were assigned to have homogeneous weights, i.e., D was an identity matrix (I).

Variance components were estimated using single-trait models with all data, with and without genomic information for both traits, using AIREMLF90 [58]. However, pedigree-based variance components, estimated with all the data, were used in the validation study to have fair comparisons between PBLUP and ssGBLUP. Heritabilities were calculated as:

$$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_c^2 + \sigma_e^2}$$

where σ_a^2 is the additive genetic variance, σ_c^2 is the common environmental variance, and σ_e^2 is the residual variance.

Cross-validation

To evaluate predictive abilities of both traditional pedigree-based and genomic evaluations, as well as the impact of different SNP densities, five-fold cross-validation was conducted [24, 28, 59]. The genotyped animals with phenotypes (fillet yield N = 775; shear force N = 772) were randomly split into five mutually exclusive folds/groups. Then, phenotypes were removed from the data, one group at a time from the validation groups (fillet yield N = 155; shear force N = ~154). The remaining animals (i.e., training group) were used to predict the future performance of the validation group. Results were presented as the overall mean of the 5 replicates. This cross-validation approach was chosen given the small number of genotyped animals with phenotypes.

To calculate the predictive ability, phenotypes were adjusted for fixed effects (y^*) as we previously described in [54]. Predictive ability was defined as the Pearson's correlation between adjusted phenotypes (y^*) and (G)EBV.

$$\text{Predictive ability} = \text{cor} [(G)EBV, y^*]$$

Further, inflation was assessed as the regression coefficient (b_1) of adjusted phenotypes (y^*) on (G)EBV as follows:

$$y^* = b_0 + b_1 \times (G)EBV + e$$

where $b_1 < 1$ indicates inflation and $b_1 > 1$ indicates deflation of (G)EBV[24].

Linkage disequilibrium and effective population size

The LD was calculated using PREGSF90 [56] according to the following equation:

$$r^2 = \frac{D^2}{P_A P_a P_B P_b}$$

where P_A , P_a , P_B , and P_b represent the allele frequencies; $D = P_{AB} - P_A P_B$ where P_{AB} is the frequency of the genotype AB. The mean LD was estimated for each chromosome as the average estimate of r^2 from all pairwise SNPs [26].

For each chromosome, LD decay with the distance between SNP markers was calculated by fitting Sved's equation [60] as follows:

$$E[r_t^2] = \frac{1}{(1 + 4N_{et}d_{ij})}$$

where d_{ij} is the distance between the SNP-marker pair i and j ; N_{et} is the effective population size for chromosome t ; N_e was calculated using the equation proposed by Saura *et al.* [61]:

$$N_{et} = (4 d_t)^{-1} [(r_t^2 - n^{-1})^{-1} - \alpha]$$

where d_t is the average length of chromosome t in Morgan, r^2 is the average LD of chromosome t , $(n)^{-1}$ represents the adjustment term for the sample size, and α is a fixed parameter that equals to 1 in the absence of mutations or to 2 in the presence of mutations. In this study, $\alpha = 2$ was used.

Reducing SNP density based on LD

PLINK 1.9 [48] was used to generate reduced/pruned SNP subsets based on variable LD values of 0.7, 0.4, 0.1, 0.05, and 0.01. SNP pruning was achieved using the command line (plink –file data –indep-pairwise 50 5 r^2). In brief, LD was calculated between each pair of SNPs within a genomic window of 50 SNPs. A pair of SNPs was removed if the LD between the two pairs of SNPs exceeded the LD value, e.g., 0.7. The window was shifted 5 SNPs forward, and the procedure was repeated. The command line was re-executed using the four other r^2 values. Each one of the five resulting pruned SNP subsets with 20042, 11433, 1153, 497, and 128 SNPs were used in genomic predictions.

Reducing SNP density based on the percentage of additive genetic variance

The second approach used to reduce the density of the SNP panel was based on the percentage of additive genetic variance explained by SNPs in each trait. The SNP variances were previously computed using wssGBLUP for fillet yield and fillet firmness [13, 19]. In this approach, SNPs were clustered into five to six groups based on the percentage of additive variance explained, with clusters ranging from > 0.05% to > 1.0%. Accordingly, the reduced number of SNPs of each group was used to evaluate the predictive ability using the five-fold cross-validation strategy.

Abbreviations

QTL: Quantitative Trait Loci

GEBV: Genomic Estimated Breeding Values

PBLUP: Pedigree-based Best Linear Unbiased Prediction

EBV: Estimated Breeding Values

ssGBLUP: single-step Genomic Best Linear Unbiased Prediction

GS: Genomic Selection

BCWD: Bacterial Cold Water Disease

YC: Year-Class

QC: Quality Check

NCCCWA: National Center of Cool and Cold Water Aquaculture

LD: Linkage Disequilibrium

HWE: Hardy–Weinberg Equilibrium

wssGBLUP: weighted single-step Genomic Best Linear Unbiased Prediction

DECLARATIONS

Ethics approval and consent to participate

The Institutional Animal Care and Use Committee of the United States Department of Agriculture, National Center for Cool and Cold Water Aquaculture (Leetown, WV, United States) specifically reviewed and approved all husbandry practices used in this study (IACUC approval #056).

Consent for publication

Not applicable.

Competing interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

This study was supported by a competitive grant No. 2014-67015-21602 from the United States Department of Agriculture, National Institute of Food and Agriculture (MS), and by the USDA, Agricultural Research Service CRIS Project 1930-31000-010 "Utilizing Genetics and Physiology for Enhancing Cool and Cold Water Aquaculture Production". The content is solely the responsibility of the authors and does not necessarily represent the official views of any of the funding agents.

Authors' Contributions

MS, and TL, conceived and designed the experiments. RA-T, AA, AG, MS, and TL performed the experiments and analyzed the data. DL helped to design the analyses. AA, RA-T and MS wrote the paper. All authors reviewed and approved the publication.

Acknowledgments

The authors acknowledge J. Everson, M. Hostuttler, K. Jenkins, J. Kretzer, J. McGowan, K. Melody, T. Moreland, and D. Payne for technical assistance. The use of trade, firm, or corporation names in this publication is for the information and convenience of the reader. Such use does not

constitute an official endorsement or approval by the USDA or the ARS of any product or service to the exclusion of others that may be suitable. USDA is an equal opportunity provider and employer.

Availability of data and materials

All datasets generated for this study are included in the manuscript and/or the additional file. The genotypes (ped and map files) and phenotypes are available in our previous publication [13].

REFERENCES

1. Burbridge, Hendrick, Roth, Rosenthal: **Social and economic policy issues relevant to marine aquaculture**. *Journal of Applied Ichthyology* 2001, **17**(4):194-206.
2. Fornshell G: **Rainbow Trout — Challenges and Solutions**. In., vol. 10; 2002: 545–557.
3. Gjedrem T: **Selection and Breeding Programs in Aquaculture**. New York: Springer; 2008.
4. WorldFish Center: **Climate Change: Research to Meet the Challenges Facing Fisheries and Aquaculture**. In: **Issues Brief 1915**. In.; 2009.
5. Gjedrem T: **Breeding plans for rainbow trout**. In: *The Rainbow Trout: Proceedings of the First Aquaculture-sponsored Symposium held at the Institute of Aquaculture, University of Sterling, Scotland*. Edited by Gall GAE, vol. 100; 1992: 73–83.
6. Gjedrem T: **Genetic variation in quantitative traits and selective breeding in fish and shellfish**. *Aquaculture* 1983, **33**(1-4):51-72.
7. Leeds TD, Silverstein JT, Weber GM, Vallejo RL, Palti Y, Rexroad CE, Evenhuis J, Hadidi S, Welch TJ, Wiens GD: **Response to selection for bacterial cold water disease resistance in rainbow trout**. *J Anim Sci* 2010, **88**(6):1936-1946.
8. Leeds TD, Vallejo RL, Weber GM, Pena DG, Silverstein JS: **Response to five generations of selection for growth performance traits in rainbow trout (*Oncorhynchus mykiss*)**. *Aquaculture* 2016, **465**:341-351.
9. Florence L, Mireille C, Jérôme B, Laurent L, Françoise M, Edwige Q: **Selection for muscle fat content and triploidy affect flesh quality in pan-size rainbow trout, *Oncorhynchus mykiss***. *Aquaculture* 2015, **448**:569-577.
10. Sae-Lim P, Komen H, Kause A, van Arendonk JA, Barfoot AJ, Martin KE, Parsons JE: **Defining desired genetic gains for rainbow trout breeding objective using analytic hierarchy process**. *J Anim Sci* 2012, **90**(6):1766-1776.
11. Torgersen JS, Koppang EO, Stien LH, Kohler A, Pedersen ME, Mørkøre T: **Soft texture of atlantic salmon fillets is associated with glycogen accumulation**. *PLoS One* 2014, **9**(1):e85551.
12. Gonzalez-Pena D, Gao GT, Baranski M, Moen T, Cleveland BM, Kenney PB, Vallejo RL, Palti Y, Leeds TD: **Genome-Wide Association Study for Identifying Loci that Affect Fillet Yield, Carcass, and Body Weight Traits in Rainbow Trout (*Oncorhynchus mykiss*)**. *Frontiers in Genetics* 2016, **7**.
13. Ali A, Al-Tobasei R, Lourenco D, Leeds T, Kenney B, Salem M: **Genome-Wide Association Study Identifies Genomic Loci Affecting Filet Firmness and Protein Content in Rainbow Trout**. *Front Genet* 2019, **10**(386):386.

14. Gutierrez AP, Yanez JM, Fukui S, Swift B, Davidson WS: **Genome-wide association study (GWAS) for growth rate and age at sexual maturation in Atlantic salmon (*Salmo salar*)**. *PLoS One* 2015, **10**(3):e0119730.
15. Schaeffer LR: **C. R. Henderson: contributions to predicting genetic merit**. *J Dairy Sci* 1991, **74**(11):4052-4066.
16. Meuwissen TH, Hayes BJ, Goddard ME: **Prediction of total genetic value using genome-wide dense marker maps**. *Genetics* 2001, **157**(4):1819-1829.
17. VanRaden PM, Van Tassell CP, Wiggans GR, Sonstegard TS, Schnabel RD, Taylor JF, Schenkel FS: **Invited Review: Reliability of genomic predictions for North American Holstein bulls**. *Journal of Dairy Science* 2009, **92**(1):16-24.
18. Palti Y, Gao G, Liu S, Kent MP, Lien S, Miller MR, Rexroad CE, 3rd, Moen T: **The development and characterization of a 57K single nucleotide polymorphism array for rainbow trout**. *Mol Ecol Resour* 2015, **15**(3):662-672.
19. Salem M, Al-Tobasei R, Ali A, Lourenco D, Gao G, Palti Y, Kenney B, Leeds TD: **Genome-Wide Association Analysis With a 50K Transcribed Gene SNP-Chip Identifies QTL Affecting Muscle Yield in Rainbow Trout**. *Front Genet* 2018, **9**:387.
20. Joshi R, Skaarud A, de Vera M, Alvarez AT, Ødegård J: **Genomic prediction for commercial traits using univariate and multivariate approaches in Nile tilapia (*Oreochromis niloticus*)**. *Aquaculture* 2020, **516**:734641.
21. Cleveland MA, Hickey JM: **Practical implementation of cost-effective genomic selection in commercial pig breeding using imputation**. *J Anim Sci* 2013, **91**(8):3583-3592.
22. Odegard J, Moen T, Santi N, Korsvoll SA, Kjøglum S, Meuwissen TH: **Genomic prediction in an admixed population of Atlantic salmon (*Salmo salar*)**. *Front Genet* 2014, **5**:402.
23. Tsai HY, Hamilton A, Tinch AE, Guy DR, Gharbi K, Stear MJ, Matika O, Bishop SC, Houston RD: **Genome wide association and genomic prediction for growth traits in juvenile farmed Atlantic salmon using a high density SNP array**. *BMC Genomics* 2015, **16**:969.
24. Garcia ALS, Bosworth B, Waldbieser G, Misztal I, Tsuruta S, Lourenco DAL: **Development of genomic predictions for harvest and carcass weight in channel catfish**. *Genet Sel Evol* 2018, **50**(1):66.
25. Yoshida GM, Lhorente JP, Correa K, Soto J, Salas D, Yanez JM: **Genome-Wide Association Study and Cost-Efficient Genomic Predictions for Growth and Fillet Yield in Nile Tilapia (*Oreochromis niloticus*)**. *G3 (Bethesda)* 2019, **9**(8):2597-2607.
26. Vallejo RL, Silva RMO, Evenhuis JP, Gao G, Liu S, Parsons JE, Martin KE, Wiens GD, Lourenco DAL, Leeds TD *et al*: **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* 2018.
27. Vallejo RL, Leeds TD, Gao G, Parsons JE, Martin KE, Evenhuis JP, Fragomeni BO, Wiens GD, Palti Y: **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* 2017, **49**(1):17.
28. Silva RMO, Evenhuis JP, Vallejo RL, Gao G, Martin KE, Leeds TD, Palti Y, Lourenco DAL: **Whole-genome mapping of quantitative trait loci and accuracy of genomic predictions for resistance to columnaris disease in two rainbow trout breeding populations**. *Genet Sel Evol* 2019, **51**(1):42.
29. Yoshida GM, Carvalheiro R, Rodriguez FH, Lhorente JP, Yanez JM: **Single-step genomic evaluation improves accuracy of breeding value predictions for resistance to infectious pancreatic necrosis virus in rainbow trout**. *Genomics* 2019, **111**(2):127-132.

30. Sonesson AK, Meuwissen TH: **Testing strategies for genomic selection in aquaculture breeding programs.** *Genet Sel Evol* 2009, **41**:37.
31. Tsai HY, Hamilton A, Tinch AE, Guy DR, Bron JE, Taggart JB, Gharbi K, Stear M, Matika O, Pong-Wong R *et al*: **Genomic prediction of host resistance to sea lice in farmed Atlantic salmon populations.** *Genet Sel Evol* 2016, **48**(1):47.
32. Correa K, Bangera R, Figueroa R, Lhorente JP, Yanez JM: **The use of genomic information increases the accuracy of breeding value predictions for sea louse (*Caligus rogercresseyi*) resistance in Atlantic salmon (*Salmo salar*).** *Genet Sel Evol* 2017, **49**(1):15.
33. Chen L, Li C, Sargolzaei M, Schenkel F: **Impact of genotype imputation on the performance of GBLUP and Bayesian methods for genomic prediction.** *PLoS One* 2014, **9**(7):e101544.
34. Tsai HY, Matika O, Edwards SM, Antolin-Sanchez R, Hamilton A, Guy DR, Tinch AE, Gharbi K, Stear MJ, Taggart JB *et al*: **Genotype Imputation To Improve the Cost-Efficiency of Genomic Selection in Farmed Atlantic Salmon.** *G3 (Bethesda)* 2017, **7**(4):1377-1383.
35. Bangera R, Correa K, Lhorente JP, Figueroa R, Yanez JM: **Genomic predictions can accelerate selection for resistance against *Piscirickettsia salmonis* in Atlantic salmon (*Salmo salar*).** *BMC Genomics* 2017, **18**(1):121.
36. Yoshida GM, Bangera R, Carvalheiro R, Correa K, Figueroa R, Lhorente JP, Yanez JM: **Genomic Prediction Accuracy for Resistance Against *Piscirickettsia salmonis* in Farmed Rainbow Trout.** *G3 (Bethesda)* 2018, **8**(2):719-726.
37. Fraslin C, Dupont-Nivet M, Haffray P, Bestin A, Vandeputte M: **How to genetically increase fillet yield in fish: New insights from simulations based on field data.** *Aquaculture* 2018, **486**:175-183.
38. Vallejo RL, Leeds TD, Fragomeni BO, Gao G, Hernandez AG, Misztal I, Welch TJ, Wiens GD, Palti Y: **Evaluation of Genome-Enabled Selection for Bacterial Cold Water Disease Resistance Using Progeny Performance Data in Rainbow Trout: Insights on Genotyping Methods and Genomic Prediction Models.** *Front Genet* 2016, **7**:96.
39. Ali A, Al-Tobasei R, Kenney B, Leeds T, Salem M: **Integrated analysis of lncRNA and mRNA expression in rainbow trout families showing variation in muscle growth and fillet quality traits.** *Scientific Reports* 2018, **Under review**.
40. Sun X, Fernando R, Dekkers J: **Contributions of linkage disequilibrium and co-segregation information to the accuracy of genomic prediction.** *Genet Sel Evol* 2016, **48**(1):77.
41. Pocnic I, Lourenco DA, Masuda Y, Legarra A, Misztal I: **The Dimensionality of Genomic Information and Its Effect on Genomic Prediction.** *Genetics* 2016, **203**(1):573-581.
42. Muir WM: **Comparison of genomic and traditional BLUP-estimated breeding value accuracy and selection response under alternative trait and genomic parameters.** *J Anim Breed Genet* 2007, **124**(6):342-355.
43. Pearse DE, Barson NJ, Nome T, Gao G, Campbell MA, Abadia-Cardoso A, Anderson EC, Rundio DE, Williams TH, Naish KA *et al*: **Publisher Correction: Sex-dependent dominance maintains migration supergene in rainbow trout.** *Nat Ecol Evol* 2020, **4**(1):170.
44. Kijas J, Elliot N, Kube P, Evans B, Botwright N, King H, Primmer CR, Verbyla K: **Diversity and linkage disequilibrium in farmed Tasmanian Atlantic salmon.** *Anim Genet* 2017, **48**(2):237-241.
45. Rexroad CE, 3rd, Vallejo RL: **Estimates of linkage disequilibrium and effective population size in rainbow trout.** *BMC Genet* 2009, **10**:83.
46. Pocnic I, Lourenco DA, Masuda Y, Misztal I: **Dimensionality of genomic information and performance of the Algorithm for Proven and Young for different livestock species.** *Genet Sel Evol* 2016, **48**(1):82.
47. Stam P: **The distribution of the fraction of the genome identical by descent in finite random mating populations.** *Genetical research* 1980, **35**(2):131-155.

48. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, et al.: **Plink: A tool set for whole-genome association and population-based linkage analyses.** *Am J Hum Genet* 2007, **81**:559-575.
49. Pocnic I, Lourenco DAL, Masuda Y, Legarra A, Misztal I: **Limited dimensionality of genomic information and effective population size.** In: *Proceedings of the World Congress on Genetics Applied to Livestock Production: 2018*; 2018: 32.
50. Misztal I: **Inexpensive Computation of the Inverse of the Genomic Relationship Matrix in Populations with Small Effective Population Size.** *Genetics* 2016, **202**(2):401-409.
51. Hickey JM, Dreisigacker S, Crossa J, Hearne S, Babu R, Prasanna BM, Grondona M, Zambelli A, Windhausen VS, Mathews K et al: **Evaluation of Genomic Selection Training Population Designs and Genotyping Strategies in Plant Breeding Programs Using Simulation.** *Crop Sci* 2014, **54**:1476.
52. Gorjanc G, Cleveland MA, Houston RD, Hickey JM: **Potential of genotyping-by-sequencing for genomic selection in livestock populations.** *Genet Sel Evol* 2015, **47**:12.
53. Al-Tobasei R, Ali A, Leeds TD, Liu S, Palti Y, Kenney B, Salem M: **Identification of SNPs associated with muscle yield and quality traits using allelic-imbalance analyses of pooled RNA-Seq samples in rainbow trout.** *BMC Genomics* 2017, **18**(1):582.
54. Ali A, Al-Tobasei R, Lourenco D, Leeds T, Kenney B, Salem M: **Genome-wide identification of loci associated with growth in rainbow trout.** *BMC Genomics* 2020, **21**(1):209.
55. Manor ML, Cleveland BM, Kenney PB, Yao J, Leeds T: **Differences in growth, fillet quality, and fatty acid metabolism-related gene expression between juvenile male and female rainbow trout.** *Fish Physiol Biochem* 2015, **41**(2):533-547.
56. Misztal I, Tsuruta S, Lourenco D, Aguilar I, style="font-style: italic;">A, Vitezica Z: **Manual for BLUPF90 family of programs.** In: Univ. Georg., Athens, USA; 2014.
57. Aguilar I, Misztal I, Johnson DL, Legarra A, Tsuruta S, Lawlor TJ: **Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score.** *J Dairy Sci* 2010, **93**(2):743-752.
58. Misztal I, Tsuruta S, Lourenco D, Masuda Y, Aguilar I, Legarra A, Vitezica Z: **Manual for BLUPF90 family of programs.** Univ. Georg., Athens, USA; 2018.
59. Reverter A, Golden BL, Bourdon RM, Brinks JS: **Method R variance components procedure: application on the simple breeding value model.** *J Anim Sci* 1994, **72**(9):2247-2253.
60. Sved JA: **Linkage disequilibrium and homozygosity of chromosome segments in finite populations.** *Theor Popul Biol* 1971, **2**(2):125-141.
61. Saura M, Tenesa A, Woolliams JA, Fernandez A, Villanueva B: **Evaluation of the linkage-disequilibrium method for the estimation of effective population size when generations overlap: an empirical case.** *BMC Genomics* 2015, **16**:922.

Figures

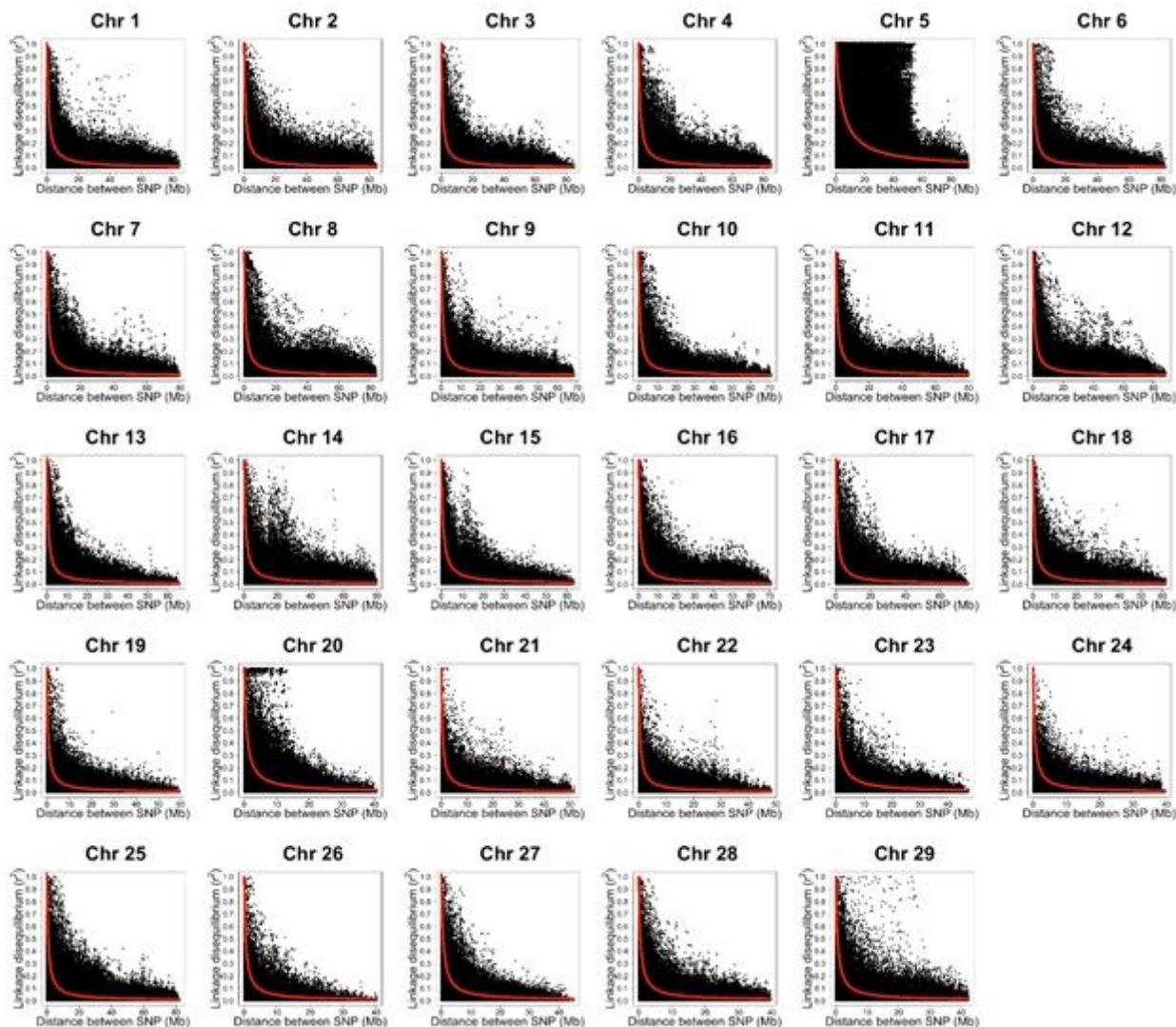


Figure 1

Plot of LD decay estimated for all chromosomes using 50K SNP panel in rainbow trout from the NCCCWA genetic line. Chromosome 5 showed a long-range LD extended over 20Mb.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile1R2.docx](#)