Aquaculture 519 (2020) 734767

Contents lists available at ScienceDirect

Aquaculture

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

Accuracy of selection for omega-3 fatty acid content in Atlantic salmon fillets

Siri S. Horn[a,b,*], Theo H.E. Meuwissen[b], Hooman Moghadam[c], Borghild Hillestad[c], Anna K. Sonesson[a]

[a] Nofima (Norwegian institute of Food, Fisheries and Aquaculture research), PO Box 210, N-1432 Ås, Norway
[b] Department of Animal and Aquaculture Sciences, Norwegian University of Life Sciences, N-1430 Ås, Norway
[c] SalmoBreed AS, Sandviksboder 3A, N-5035 Bergen, Norway

ARTICLE INFO

Keywords:
Selection accuracy
Omega-3
Atlantic salmon
Genomic prediction

ABSTRACT

The main aim of the current study was to compare the accuracy of selection for muscle content of fat and the health-promoting omega-3 fatty acids EPA and DHA in Atlantic salmon, by varying the sources of genetic information used in the estimation of breeding values. The following genetic information sources were compared: pedigree, SNP-chip markers and allele-specific expression markers.

The results showed that differences between information sources were in general small, and different genetic information performed best for different traits. SNP-chip performed best for DHA, and pedigree performed best for EPA.

Knowledge from gene expression analysis of a few individuals can be utilized to select a small panel of markers that perform relatively well. Genetic markers of allele-specific expression were able to capture a lot of genetic variation for DHA, but did not give significantly higher accuracies when combined with SNP-chip or pedigree information.

The cross-validation accuracies for selection for DHA and EPA were moderate and offer possibilities for selection for these traits, especially if one extends the reference data set to a much bigger population, with more sibs per selection candidate.

1. Background

Atlantic salmon (*Salmo salar L.*) is an important farmed fish species, known for its high content of the health-promoting long-chain polyunsaturated omega-3 fatty acids (n-3 LC-PUFA) eicosapentaenoic (EPA; 20:5n-3) and docosahexaenoic acid (DHA; 22:6n-3). The replacement of fish oil and fishmeal with plant ingredients in the diet of farmed Atlantic salmon the last few decades has resulted in reduced levels of EPA and DHA in the fish fillet (Sprague et al., 2016; Ytrestoyl et al., 2015). Quantitative genetic analyses have demonstrated the potential of selective breeding to increase n-3 LC-PUFA levels in salmon muscle (Horn et al., 2018; Leaver et al., 2011). Muscle content of EPA and DHA appear to be genetically different traits, with heritability of 0.09 and 0.26, respectively (Horn et al., 2018).

Selective breeding programs have historically been based on best linear unbiased prediction (BLUP) of individual breeding values that use pedigree-based relationship matrices. Genomic selection (GS) is a method that uses DNA marker-based relationship matrices as the genetic information to predict the breeding value of all genotyped individuals (Meuwissen et al., 2001). The GS methodology is of particular relevance for traits that cannot be measured directly on selection candidates, such as muscle EPA and DHA content, because it allows prediction of individual breeding values for non-phenotyped individuals. GS has been shown to increase the accuracy of breeding values for several traits in salmonids compared to conventional selection based on pedigree data (Bangera et al., 2017; Tsai et al., 2016; Yoshida et al., 2018). GS may be performed using several different sources of genetic information, where dense SNP-chip genotypes are the most common.

An alternative source of genetic information is markers of variation in gene expression. Variation in gene expression has the potential to contribute significantly to phenotypic variation (Pritchard et al., 2006; Wray et al., 2003). One technique to identify this variation is to screen for allele-specific expression (ASE) – unequal expression of the two alleles of a gene, caused by cis-regulatory elements (Yan et al., 2002). ASE markers are more closely linked to causative loci affecting traits than random markers, and can potentially be applied in selection

* Corresponding author at: Nofima (Norwegian institute of Food, Fisheries and Aquaculture research), PO Box 210, N-1432 Ås, Norway.
E-mail address: siri.storteig.horn@nofima.no (S.S. Horn).

https://doi.org/10.1016/j.aquaculture.2019.734767
Received 23 April 2019; Received in revised form 21 November 2019; Accepted 22 November 2019
Available online 23 November 2019

programs. ASE is widespread within a large number of species, and ASE genes have been found related to complex, economically important traits in chicken, pigs and cattle (Cheng et al., 2015; Muráni et al., 2009; Olbromski et al., 2013). Selection using ASE SNPs reduced disease incidence after one generation of selection in chicken (Cheng et al., 2015). ASE markers could therefore potentially be useful in selection for EPA and DHA in salmon.

The main aim of the current study was to compare the accuracy of selection for muscle content of EPA, DHA and fat in Atlantic salmon, by varying the sources of genetic information used in the estimation of breeding values.

2. Materials and methods

2.1. Fish population and recordings

The 563 fish studied consisted of 174 full-sib families from 92 sires and 174 dams. All fish originated from year-class 2014 of the Atlantic salmon breeding population of SalmoBreed AS. Four generations of pedigree information on direct ancestors of the fish were available. Pedigree was tracked by PIT-tagging. The fish were transferred to sea at a mean weight of 0.1 kg, and slaughtered approximately 12 months later, at a mean weight of 3.6 kg. The fish were fed a commercial broodstock feed from Skretting (https://www.skretting.com/en/products/atlantic-salmon/?lifephase=474980) with a high fish oil content, and were fasted 13–14 days prior to slaughter. All fish were reared under the same conditions.

At slaughter, sex was determined by visual inspection of the gonads, body weight was recorded, and skeletal muscle samples for lipid and fatty acid analysis taken from the Norwegian Quality Cut were collected, frozen and stored at −20 °C.

2.2. Muscle fat and fatty acid analysis

The traits analyzed in this study were muscle EPA and DHA content, and total muscle fat (MFAT). EPA and DHA content was expressed as a percentage of the total amount of fatty acids in the sampled muscle tissue, and muscle fat (MFAT) expressed as total lipid percentage in the sampled muscle tissue.

Muscle fat content was measured by extracting total lipids from homogenized skeletal muscle samples of individual fish, according to the Folch method (Folch et al., 1957). Using one milliliter from the chloroform-methanol phase, fatty acid composition of total lipids was analyzed following the method described by Mason and Waller (1964). The extract was dried briefly under nitrogen gas and residual lipid extract was *trans*-methylated overnight with 2′,2′-dimethoxypropane, methanolic-HCl, and benzene at room temperature. The methyl esters formed were separated in a gas chromatograph (Hewlett Packard 6890; HP, Wilmington, DE, USA) with a split injector, using an SGE BPX70 capillary column (length 60 m, internal diameter 0.25 mm, and film thickness 0.25 μm; SGE Analytical Science, Milton Keynes, UK) and a flame ionization detector. The results were analyzed using HP Chem Station software. The carrier gas was helium, and the injector and detector temperatures were both 270 °C. The oven temperature was raised from 50 to 170 °C at the rate of 4 °C / min, and then raised to 200 °C at a rate of 0.5 °C / min and finally to 240 °C at 10 °C / min. Individual fatty acid methyl esters were identified by reference to well-characterized standards.

2.3. Genotyping

The 563 fish studied were genotyped using a customized 57 K axiom Affymetrix SNP Genotyping Array (NOFSAL02). From the initial 57 K SNPs, we retained those with call rate > 0.8, minor allele frequencies > 0.05, and Hardy-Weinberg equilibrium correlation *p*-value > .001. A total of 49,726 SNPs passed quality control (QC) filtering and were used to construct the genomic relationship matrix (G-matrix) $G_{SNPCHIP}$ (see section Genomic relationship matrices).

2.4. Allele specific expression

ASE is unequal expression of the two alleles of a gene and indicates the presence of one or more variants that act in cis to affect the expression level of the gene (Yan et al., 2002). We identified SNPs on the NOFSAL02 SNP-chip located within genes showing ASE by using RNA-seq data on 59 of the 563 fish. The methods and fish for the RNA-seq data is described in Horn et al. (2019). First, genes exhibiting ASE for the traits EPA and DHA were identified by selecting all genes whose expression was identified as associated with EPA and/or DHA content in muscle and/or liver in Horn et al. (2019). This resulted in 6194 genes detected as trait-associated (TA) at a nominal probability-value (*p*-value) of 0.10. Second, heterozygote individuals for the TA genes were identified among the 59 fish with RNA-seq data using the Freebayes software (https://github.com/ekg/freebayes). Third, the relative expression levels of the reference and the alternative allele were compared by counting how many times the heterozygous individuals had the alternative allele more often expressed than the reference allele. This number was compared to the Binomial distribution with $p = .5$ as a null-hypothesis distribution to determine statistical significance in a one-sided-test for overexpression of the alternative allele at a p-value of 0.001. A one-sided statistical test was performed because we observed a general tendency for the reference allele to be over-expressed (possibly because the reads with the reference allele are more likely to align with the reference genome than reads with alternative alleles). This statistical test resulted in a total of 537 genes with significant allele specific expressions. Within the SNP-chip we found 395 SNPs within these 537 genes, which were used to construct a genomic relationship matrix, G_{ASE}.

In order to compare this method to a random set of markers, 400 random markers were sampled from the SNP-chip, and a genomic relationship matrix was constructed based on these; G_{RANDOM}.

2.5. Statistical analyses

The different models compared in the current study differed solely with respect to their specification of the relationship matrix that was fitted:

- PED: Classical pedigree-based analysis with a numerator relationship matrix (A).
- SNPCHIP: G-matrix based on ~50,000 QC genome-wide SNP markers ($G_{SNPCHIP}$).
- ASE: G-matrix based on 395 markers identified in the allele specific expression analysis (G_{ASE}).
- RANDOM: G-matrix based on 400 randomly selected markers (G_{RANDOM}).
- ASE + PED: Fitting both the A and the G_{ASE} relationship matrix.
- ASE + SNPCHIP: Fitting both the G_{ASE} and $G_{SNPCHIP}$ matrices.

2.5.1. Genomic relationship matrices

The G-matrices were constructed using the GCTA software, following the method by Yang et al. (2011), using the following equation to estimate the genetic relationship between individuals j and k:

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

where x_{ij} is the number of copies of the reference allele for the i^{th} SNP of the j^{th} individual, x_{ik} is the number of copies of the reference allele for the i^{th} SNP of the k^{th} individual, and p_i is the frequency of the reference allele, estimated from the observed allele frequencies in the 563 samples.

In order to stabilize the relationship matrix (make it invertable), a small value (0.05) was added to the diagonal, in accordance with Forni et al. (2011).

The following G-matrices were constructed: $G_{SNPCHIP}$, G_{ASE} and G_{RANDOM}.

2.5.2. Breeding value estimation

The following linear mixed model was applied for the estimation of breeding values:

$$y = \mu + X_1 b_1 + X_2 b_2 + Zu_1 + (\ Zu_2 +)e$$

where y is a vector of the phenotype (EPA, DHA or MFAT content), μ is the overall mean, b_1 is the fixed effect of body weight as a covariable, b_2 is a vector of fixed effect of sex (Horn et al., 2018), X_1 and X_2 are incidence matrices for the effects contained in b_1 and b_2, u_1 is a vector of additive genetic effects distributed as $u_1 \sim N(0, G\sigma_{u_1}{}^2)$ (or $u_1 \sim N(0, A\sigma_{u_1}{}^2)$ in case of PED), where $\sigma_{u_1}{}^2$ is the additive genetic variance, G is the genomic relationship matrix, and A is the pedigree relationship matrix; Z is the corresponding incidence matrix to additive genetic effects, and e is the vector of random residual effects with $e \sim N(0, I\sigma_e{}^2)$. In models where two relationship matrices are fitted, the additive genetic effect of the second relationship matrix is included (Zu_2) with distribution $u_2 \sim N(0, G_2\sigma_{u_2}{}^2)$, where G_2 is the second genomic relationship matrix, and $\sigma_{u_2}{}^2$is its associated variance component. Variance components and breeding values were estimated by ASReml 4.0 (Gilmour et al., 2015). Estimated breeding values (EBV) were obtained as $EBV = \hat{u_1}(+\hat{u_2})$, where ˆ denotes estimates of the effects.

2.5.3. Selection accuracy

Selection accuracies of the different models were assessed through predictive ability, using a cross-validation scheme by randomly masking the phenotype of one of the siblings in every full-sib family consisting of more than three full-sibs. As 40 of the 174 families were represented by one or two individuals only, this resulted in 87 validation and 476 training individuals.

The mean selection accuracy (Acc) of 100 replicates was computed as:

$$\mathrm{Acc} = \mathrm{r}(\mathrm{EBV_v}, \hat{y})/\sqrt{\mathrm{h}^2}$$

where r represents the correlation between EBV_v and $\hat{y}$, EBV_v represents breeding values for validation individuals estimated using the reference data, $\hat{y}$ is y adjusted for fixed effects by calculating $\hat{y} = EBV_a +$ residual, with EBV_a denoting the pedigree based estimates of breeding values using all data, and h^2 is heritability estimates calculated as the ratio of additive genetic (σ_A^2) to total phenotypic (σ_P^2) variance ($h^2 = \sigma_A^2/\sigma_P^2$) using the pedigree based estimates of the variance components using all data.

For each of the 100 replicates, a data subset was created. The same 100 data subsets were used for estimating accuracy for all methods. Thus, we obtained accuracy-estimates 1 to 100 for each method (for data subset 1 to 100). To test if the accuracy-estimates for two methods were significantly different from each other we compared the correlations between EBV and masked phenotypes for each of the 100 replicated sub data sets. We counted the number of replicates in which one method yielded a higher correlation. Next, this number was compared to a binomial distribution with null hypothesis H_0 of both methods being equally good (i.e. expected frequency was 0.5). This yielded a one-tail p-value. The two-tail significance value was obtained by doubling the one-tail p-value, because the expected proportion was 50%. Differences between the two models were considered significant if the two-tail p-value was < 0.01.

3. Results and discussion

The variance components showed that the genetic variance (and thus heritability) estimates with genomic information were lower than the estimates using pedigree, except for DHA where PED and SNPCHIP were very similar (Table 1). Several authors have previously reported a reduction in heritability estimates with genetic marker information (Boison et al., 2019; Erbe et al., 2013; Robledo et al., 2018). This may be explained by factors such as lack of markers that are in linkage disequilibrium with the causative mutations, and large numbers of markers that are in linkage equilibrium with the causative mutations (de Los Campos et al., 2015). It should also be noted that the standard errors of the estimates of the variance components were high (Table 1).

Table 1
Estimates of variance components for EPA, DHA and muscle fat (MFAT) using different genetic information.

		DHA	EPA	MFAT
PED				
	h^2	0.21 (0.09)	0.06 (0.05)	0.36 (0.10)
	V_a	0.038 (0.017)	0.061 (0.057)	2.216 (0.665)
	V_e	0.144 (0.016)	0.988 (0.078)	3.877 (0.542)
SNPCHIP				
	h^2	0.20 (0.07)	0.04 (0.04)	0.25 (0.07)
	V_a	0.036 (0.013)	0.044 (0.043)	1.505 (0.486)
	V_e	0.144 (0.014)	1.001 (0.072)	4.501 (0.455)
ASE				
	h^2	0.14 (0.05)	0.03 (0.03)	0.16 (0.05)
	V_a	0.026 (0.010)	0.033 (0.031)	0.960 (0.316)
	V_e	0.154 (0.012)	1.013 (0.067)	5.012 (0.372)
RANDOM				
	h^2	0.02 (0.04)	0.02 (0.03)	0.15 (0.05)
	V_a	0.003 (0.007)	0.026 (0.030)	0.872 (0.304)
	V_e	0.177 (0.120)	1.021 (0.068)	5.106 (0.382)
ASE + PED				
	h^2	0.21 (0.08)	0.05 (0.05)	0.36 (0.10)
G_{ASE}	V_a	0.020 (0.011)	0.020 (0.041)	0.324 (0.321)
A	V_a	0.018 (0.017)	0.031 (0.073)	1.861 (0.710)
	V_e	0.143 (0.015)	0.996 (0.079)	3.877 (0.541)
ASE + SNPCHIP				
	h^2	0.20 (0.07)	0.04 (0.04)	0.24 (0.07)
G_{ASE}	V_a	0.022 (0.016)	0.025 (0.073)	0.970 (0.599)
$G_{SNPCHIP}$	V_a	0.015 (0.012)	0.016 (0.056)	0.463 (0.399)
	V_e	0.143 (0.014)	1.004 (0.073)	4.553 (0.456)

V_a: Additive genetic variance. V_e: Residual variance. h^2: Estimated heritability. Standard errors in brackets.

The accuracies were generally low (0.27–0.61; Table 2), which is similar to a study on fatty acid traits in cattle ($N = 1366$), where the accuracy of genomic prediction was < 0.40 for the majority of the fatty acids (Chen et al., 2015). We would expect to get higher accuracies compared to Chen et al. (2015) as we used a different cross-validation approach which included siblings of validation animals in the reference population. The trait with the highest heritability (MFAT) had highest accuracy. Higher heritability may be expected to result in increased accuracy of genomic prediction (e.g., Sonesson and Meuwissen 2009). Moreover, several studies have shown that low heritability can result in low prediction accuracy (Nirea et al., 2012; Vela-Avitúa et al., 2015; Daetwyler et al., 2008).

Inaccurate estimates of variance components may have reduced cross-validation prediction accuracies in Table 2. In fact, for EPA none of the heritability estimates were significantly different from zero, which may explain the low cross validation accuracies. In a quantitative genetic analysis using a bigger dataset (668 fish), the heritability estimate was also not significantly different from zero (0.09 ± 0.06) (Horn et al., 2018), suggesting that a larger dataset is required to detect significant additive genetic variation for EPA. This could be explained by that EPA muscle content is highly variable over time because EPA serves many metabolic roles in the body, such as energy production and conversion to bioactive components (Glencross et al., 2014; Sanden et al., 2011). DHA content, on the other hand is more stable as it is mainly incorporated into phospholipids in cell membranes (Ruiz-Lopez

Table 2
Estimated accuracies by cross-validation for DHA, EPA and muscle fat (MFAT) for six different genetic information sources.

Trait	PED	SNPCHIP	ASE	ASE + PED	ASE + SNPCHIP
DHA	0.328 (0.018)	0.413 (0.019)	0.359 (0.02)	0.342 (0.019)	0.369 (0.019)
EPA	0.374 (0.046)	0.321 (0.042)	0.331 (0.039)	0.268 (0.041)	0.267 (0.04)
MFAT	0.603 (0.013)	0.565 (0.014)	0.564 (0.013)	0.606 (0.013)	0.566 (0.014)

Standard errors in brackets.

Table 3
P-values from binomial test indicating significant differences in accuracies between sources of genetic information for DHA.

DHA	PED	SNPCHIP	ASE	ASE + PED	ASE + SNPCHIP
PED	–	*	ns	ns	ns
SNPCHIP		–	ns	*	ns
ASE			–	ns	*
ASE + PED				–	*
ASE + SNPCHIP					–

ns = not significant ($p > .01$).
* $p < .01$.

Table 5
P-values from binomial test indicating significant differences in accuracies between sources of genetic information for muscle fat (MFAT).

MFAT	PED	SNPCHIP	ASE	ASE + PED	ASE + SNPCHIP
PED	–	*	*	*	*
SNPCHIP		–	ns	*	*
ASE			–	*	*
ASE + PED				–	*
ASE + SNPCHIP					–

ns = not significant (p > .01).
* p < .01.

et al., 2015). The differences in biological roles of the two fatty acids are reflected in the higher heritability of DHA compared to EPA (Table 1).

Based on the results of this study, it cannot be concluded which genetic information source is best because different genetic information performed best for different traits (Table 2). Differences between information sources were in general small. SNPCHIP performed significantly better than PED for DHA, where the use of SNPCHIP genotypes resulted in 15.5% (on average) higher selection accuracy than all other information sources (Tables 2 & 3). However, for the two other traits, PED performed significantly better than SNPCHIP (Tables 4 & 5). For EPA, PED gave significantly higher accuracy than the other sources, except for ASE, which was similar to PED (Tables 2 & 4). Thus, in this study GS did not result in higher accuracy than pedigree for a low heritability trait (EPA). This is contrary to a previous study by Calus et al. (2008), and may be due to the very low (non-significant) heritability of EPA combined with the small dataset of the current study. Moreover, the performance of SNPCHIP is dependent on the presence of markers in genes influencing the trait. The NOFSAL02 SNP-chip does not contain markers in the genes of fatty acid synthase (*LOC106610271*), peroxisome proliferator-activated receptors *pparb2b* and *pparb2a,* carnitine palmitoyltransferases *cpt1b* and *cpt2*, hormone-sensitive lipase, and acetyl-CoA carboxylase (*LOC106603271*), which are all known to influence lipid metabolism in mammals and/or fish (Sul and Smith, 2008; Tocher, 2003; Varga et al., 2011). Therefore, it is possible that including markers in these genes could increase the performance of SNPCHIP.

The G-matrix based on ASE markers seemed able to capture a relatively high portion of genetic variation for DHA (72% of $G_{SNPCHIP}$; although standard errors of the components were high), indicating that markers selected through the ASE method were relevant for this trait.

Table 4
P-values from binomial test indicating significant differences in accuracies between sources of genetic information for EPA.

EPA	PED	SNPCHIP	ASE	ASE + PED	ASE + SNPCHIP
PED	–	*	ns	*	*
SNPCHIP		–	ns	ns	*
ASE			–	*	*
ASE + PED				–	ns
ASE + SNPCHIP					–

ns = not significant (p > .01).
* p < .01.

This was further supported by testing the G_{RANDOM} matrix, which for DHA resulted in a very low estimate of genetic variance (Table 1). This confirms that the 395 markers selected based on ASE explain substantially more variance for DHA than 400 random markers. The results for selection accuracies showed that ASE alone performed similarly to PED for DHA (Tables 2 & 3). However, the ASE markers did not capture a higher portion of genetic variance than RANDOM markers for the traits EPA and MFAT. For EPA, this could be explained by the very low, non-significant heritability of EPA in this dataset. The reason why ASE markers did not perform better than random markers for MFAT, is likely because the ASE markers were identified and selected based on them being located in genes significantly associated with EPA and DHA phenotypes, not MFAT. Although muscle fat is highly interconnected with the muscle's EPA and DHA content (Horn et al., 2018), it is likely that MFAT is regulated by different genes than those regulating the proportional content of EPA and DHA. Thus, it is not expected that ASE markers should perform as well for MFAT.

When combining ASE with either PED or SNPCHIP, ASE surprisingly explained the greatest portion of the genetic variation of the two information sources (Table 1). However, the total genetic variation captured for the traits did not increase compared to PED or SNPCHIP alone, thus ASE did not significantly improve prediction accuracies when combined with SNP-chip or pedigree information (Tables 2–5). Except for MFAT, where ASE + PED gave slightly higher accuracy than PED. ASE + SNPCHIP surprisingly resulted in lower accuracy than SNPCHIP alone (11% lower) for DHA, but this difference was not significant. Moreover, ASE + PED did not give significantly higher accuracy than PED alone. A possible reason why the combined models did not perform better could be that there were twice as many effects to estimate ($\hat{u}_1$ and $\hat{u}_2$), but the size of the reference data is too small to accurately estimate that many effects. In addition, the variance due to ASE may be overestimated due to sampling, which may explain why the accuracy is not improved.

The main limitation of this study is the size of the reference population. The cross-validation method is unfavorable for this dataset, because we are masking one individual from each family whilst we have small family sizes. The latter hampers the accuracy of GS, as there are insufficient siblings to assign which chromosome segment is best within a family. Yet, the choice of this validation scenario is realistic because for most traits in aquaculture (and especially for carcass quality traits), selection of breeding candidates is based on full- and half-sib phenotypic information. However, to obtain high accuracies of GS we need high numbers of validation animals per family (Odegård et al., 2014).

The model used for estimating breeding values does not include a term to account for dominance effects. Although dominance may influence variance components, as estimates of the additive genetic variance may be overestimated when dominance variance is ignored. However, it is not expected to affect accuracies of breeding value estimates substantially (Gallardo et al., 2010).

The accuracies of selection for DHA and EPA were moderate and offer possibilities for selection for these traits, especially if one uses a much larger reference data set with more sibs per selection candidate. The latter requires cost-effective recording of DHA and EPA in large reference populations, which is currently unavailable. When using the 395 ASE markers alone, the reference population does not need to be so large since there are not as many marker effects to estimate. In the current case we cannot recommend using ASE markers, but the results show that knowledge from gene expression analysis of a few individuals (59 in this case) can be utilized to select a small panel of markers that perform relatively well.

4. Conclusions

The results of this study show that different genetic information sources performed best for different traits, and that differences between information sources were in general small. SNPCHIP genotype information performed relatively poor, possibly due to the small size of the reference population and the small number of animals per family.

Knowledge from gene expression analysis of a few individuals can be utilized to select a small panel of markers that perform relatively well. Markers based on allele-specific expression were able to capture a substantial fraction of the genetic variation for DHA (72%), but did not give significantly higher accuracies when combined with SNP-chip or pedigree information in this dataset.

The cross-validation accuracies for selection for DHA and EPA were moderate and offer possibilities for selection for these traits, especially if one extends the reference data set to a much bigger population containing more sibs per selection candidate.

Declaration of Competing Interest

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

Acknowledgements

This work was supported by the Norwegian research council [grant number NFR 244200].

References

Bangera, R., Correa, K., Lhorente, J.P., Figueroa, R., Yáñez, J.M., 2017. Genomic predictions can accelerate selection for resistance against Piscirickettsia salmonis in Atlantic Salmon (Salmo salar). BMC Genomics 18 (1), 121. https://doi.org/10.1186/s12864-017-3487-y.

Boison, S.A., Gjerde, B., Hillestad, B., Makvandi-Nejad, S., Moghadam, H.K., 2019. Genomic and transcriptomic analysis of amoebic gill disease resistance in Atlantic Salmon (*Salmo salar L.*). Front. Genet. 10 (68).

Calus, M.P.L., Meuwissen, T.H.E., de Roos, A.P.W., Veerkamp, R.F., 2008. Accuracy of genomic selection using different methods to define haplotypes. Genetics 178 (1), 553. https://doi.org/10.1534/genetics.107.080838.

Chen, L.H., Ekine-Dzivenu, C., Vinsky, M., Basarab, J., Aalhus, J., Dugan, M.E.R., Fitzsimmons, C., Stothard, P., Li, C.X., 2015. Genome-wide association and genomic prediction of breeding values for fatty acid composition in subcutaneous adipose and longissimus lumborum muscle of beef cattle. BMC Genet. 16, 15.

Cheng, H.H., Perumbakkam, S., Pyrkosz, A.B., Dunn, J.R., Legarra, A., Muir, W.M., 2015. Fine mapping of QTL and genomic prediction using allele-specific expression SNPs demonstrates that the complex trait of genetic resistance to Marek's disease is predominantly determined by transcriptional regulation. BMC Genomics 16 (1), 816. https://doi.org/10.1186/s12864-015-2016-0.

Daetwyler, H.D., Villanueva, B., Woolliams, J.A., 2008. Accuracy of predicting the genetic risk of disease using a genome-wide approach. PLoS ONE 3 (10), e3395. https://doi.org/10.1371/journal.pone.0003395.

de Los Campos, G., Sorensen, D., Gianola, D., 2015. Genomic heritability: what is it? PLoS Genet. 11 (5), e1005048. https://doi.org/10.1371/journal.pgen.1005048.

Erbe, M., Gredler, B., Seefried, F.R., Bapst, B., Simianer, H., 2013. A function accounting for training set size and marker density to model the average accuracy of genomic prediction. PLoS ONE 8 (12), e81046. https://doi.org/10.1371/journal.pone.0081046.

Folch, J., Lees, M., Stanley, G.H.S., 1957. A simple method for the isolation and purification of total lipides from animal tissues. J. Biol. Chem. 226 (1), 497–509.

Forni, S., Aguilar, I., Misztal, I., 2011. Different genomic relationship matrices for single-step analysis using phenotypic, pedigree and genomic information. Genet. Sel. Evol. 43 (1) (1–1).

Gallardo, J.A., Lhorente, J.P., Neira, R., 2010. The consequences of including non-additive effects on the genetic evaluation of harvest body weight in Coho salmon (Oncorhynchus kisutch). Genet. Sel. Evol. 42 (1), 19. https://doi.org/10.1186/1297-9686-42-19.

Gilmour, A., Gogel, B., Cullis, B., Welham, S., Thompson, R., 2015. ASReml User Guide Release 4.1 Structural Specification. *Hemel hempstead: VSN international ltd*https://doi.org/10.1016/j.cobeha.2015.04.011.

Glencross, B.D., Tocher, D.R., Matthew, C., Bell, J.G., 2014. Interactions between dietary docosahexaenoic acid and other long-chain polyunsaturated fatty acids on performance and fatty acid retention in post-smolt Atlantic salmon (Salmo salar). Fish Physiol. Biochem. 40 (4), 1213–1227. https://doi.org/10.1007/s10695-014-9917-8.

Horn, S.S., Ruyter, B., Meuwissen, T.H.E., Hillestad, B., Sonesson, A.K., 2018. Genetic effects of fatty acid composition in muscle of Atlantic salmon. Genet. Sel. Evol. 50 (1), 23. https://doi.org/10.1186/s12711-018-0394-x.

Horn, S.S., Sonesson, A.K., Krasnov, A., Moghadam, H., Hillestad, B., Meuwissen, T.H.E., Ruyter, B., 2019. Individual differences in EPA and DHA content of Atlantic salmon are associated with gene expression of key metabolic processes. Sci. Rep. 9 (1), 3889. https://doi.org/10.1038/s41598-019-40391-2.

Leaver, M.J., Taggart, J.B., Villeneuve, L., Bron, J.E., Guy, D.R., Bishop, S.C., Houston, R.D., Matika, O., Tocher, D.R., 2011. Heritability and mechanisms of n-3 long chain polyunsaturated fatty acid deposition in the flesh of Atlantic salmon. Comp. Biochem. Physiol. Part D Genomics Proteomics 6 (1), 62–69. https://doi.org/10.1016/j.cbd.2010.04.002.

Mason, M.E., Waller, G.R., 1964. Dimethoxypropane induced transesterification of fats and oils in preparation of methyl esters for gas chromatographic analysis. Anal. Chem. 36 (3), 583–586. https://doi.org/10.1021/ac60209a008.

Meuwissen, T.H.E., Hayes, B.J., Goddard, M.E., 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157.

Muráni, E., Ponsuksili, S., Srikanchai, T., Maak, S., Wimmers, K., 2009. Expression of the porcine adrenergic receptor beta 2 gene in longissimus dorsi muscle is affected by cis-regulatory DNA variation. Anim. Genet. 40 (1), 80–89. https://doi.org/10.1111/j.1365-2052.2008.01811.x.

Nirea, K.G., Sonesson, A.K., Woolliams, J.A., Meuwissen, T.H.E., 2012. Strategies for implementing genomic selection in family-based aquaculture breeding schemes: double haploid sib test populations. Genet. Sel. Evol. 44. https://doi.org/10.1186/1297-9686-44-30.

Odegård, J., Moen, T., Santi, N., Korsvoll, S.A., Kjøglum, S., Meuwissen, T.H.E., 2014. Genomic prediction in an admixed population of Atlantic Salmon (Salmo salar). Front. Genet. 5. https://doi.org/10.3389/fgene.2014.00402.

Olbromski, R., Siadkowska, E., Żelazowska, B., Zwierzchowski, L., 2013. Allelic gene expression imbalance of bovine IGF2, LEP and CCL2 genes in liver, kidney and pituitary. Mol. Biol. Rep. 40 (2), 1189–1200. https://doi.org/10.1007/s11033-012-2161-3.

Pritchard, C., Coil, D., Hawley, S., Hsu, L., Nelson, P.S., 2006. The contributions of normal variation and genetic background to mammalian gene expression. Genome Biol. 7 (3), R26. https://doi.org/10.1186/gb-2006-7-3-r26.

Robledo, D., Matika, O., Hamilton, A., Houston, R.D., 2018. Genome-wide association and genomic selection for resistance to amoebic gill disease in Atlantic Salmon. G3 (Bethesda) 8 (4), 1195–1203. https://doi.org/10.1534/g3.118.200075.

Ruiz-Lopez, N., Stubhaug, I., Ipharraguerre, I., Rimbach, G., Menoyo, D., 2015. Positional distribution of fatty acids in triacylglycerols and phospholipids from fillets of Atlantic Salmon (Salmo Salar) fed vegetable and fish oil blends. Marine Drugs 13 (7), 4255–4269. https://doi.org/10.3390/md13074255.

Sanden, M., Stubhaug, I., Berntssen, M.H.G., Lie, O., Torstensen, B.E., 2011. Atlantic Salmon (Salmo salar L.) as a net producer of long-chain marine omega-3 fatty acids. J. Agric. Food Chem. 59 (23), 12697–12706. https://doi.org/10.1021/jf203289s.

Sonesson, A.K., Meuwissen, T.H.E., 2009. Testing strategies for genomic selection in aquaculture breeding programs. Genetics Selection Evolution 41 (1).

Sprague, M., Dick, J.R., Tocher, D.R., 2016. Impact of sustainable feeds on omega-3 long-chain fatty acid levels in farmed Atlantic salmon, 2006-2015. Sci. Rep. 6, 9.

Sul, H.S., Smith, S., 2008. Fatty acid synthesis in eukaryotes. In: Vance, J.E., Vance, D.E. (Eds.), Biochemistry of Lipids, Lipoproteins and Membranes. Elsevier Science, pp. 156–188.

Tocher, D.R., 2003. Metabolism and functions of lipids and fatty acids in teleost fish. Rev. Fish. Sci. 11 (2), 107–184.

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

Varga, T., Czimmerer, Z., Nagy, L., 2011. PPARs are a unique set of fatty acid regulated transcription factors controlling both lipid metabolism and inflammation. Biochim. Biophys. Acta 1812 (8), 1007–1022.

Vela-Avitúa, S., Meuwissen, T.H.E., Luan, T., Ødegård, J., 2015. Accuracy of genomic selection for a sib-evaluated trait using identity-by-state and identity-by-descent relationships. Genet. Sel. Evol. 47. https://doi.org/10.1186/s12711-014-0084-2.

Wray, G.A., Hahn, M.W., Abouheif, E., Balhoff, J.P., Pizer, M., Rockman, M.V., Romano,

L.A., 2003. The evolution of transcriptional regulation in eukaryotes. Mol. Biol. Evol. 20 (9), 1377–1419. https://doi.org/10.1093/molbev/msg140.

Yan, H., Yuan, W., Velculescu, V.E., Vogelstein, B., Kinzler, K.W., 2002. Allelic variation in human gene expression. Science 297 (5584), 1143. https://doi.org/10.1126/science.1072545.

Yang, J., Lee, S.H., Goddard, M.E., Visscher, P.M., 2011. GCTA: a tool for genome-wide complex trait analysis. Am. J. Hum. Genet. 88 (1), 76–82. https://doi.org/10.1016/j.ajhg.2010.11.011.

Yoshida, G.M., Carvalheiro, R., Lhorente, J.P., Correa, K., Figueroa, R., Houston, R.D., Yáñez, J.M., 2018. Accuracy of genotype imputation and genomic predictions in a two-generation farmed Atlantic salmon population using high-density and low-density SNP panels. Aquaculture 491, 147–154. https://doi.org/10.1016/j.aquaculture.2018.03.004.

Ytrestoyl, T., Aas, T.S., Asgard, T., 2015. Utilisation of feed resources in production of Atlantic salmon (Salmo salar) in Norway. Aquaculture 448, 365–374. https://doi.org/10.1016/j.aquaculture.2015.06.023.