Multi-trait genome wide association study in correlated traits: fillet colour and body weight in Atlantic salmon

P. Sae-Lim*, S. Boison, and M. Baranski

Mowi Genetics AS, Sandviksbodene 77A, 5035 Bergen, Norway; *panya.sae-lim@mowi.com

Abstract

The 'redness' of the flesh ('fillet colour') is an economically important trait in Atlantic salmon. Pre-adjustment of fillet colour (measured as the concentration of the pigment astaxanthin) with body weight is usually performed in a genetic analysis to account for the correlation between these two traits, however multi-trait models are thought to offer superior power for genome-wide association studies (GWAS). In this study, allele substitution effects (*b*) and their significance levels from three different GWAS approaches were compared; single-trait (STGWAS), pre-adjusting fillet colour for body weight (STadjGWAS), and a multi-trait (MTGWAS). All GWAS analyses used 3,417 animals with 53,159 markers. All approaches consistently identified the most significant markers on chromosome 2 and 26, however MTGWAS increased *b* of fillet colour by 36.8% and 48.5% compared to STGWAS and STadjGWAS, and is therefore recommended as the optimal method to maximise the power of GWAS for correlated traits of this nature.

Introduction

The red colouration of Atlantic salmon flesh (hereafter referred to as 'fillet colour') is an economically important trait that is caused by the deposition of organic pigments, known as carotenoids. Specifically, it is the uptake of supplementary astaxanthin that is responsible and this process has been shown to be heritable with values ranging from 0.11 to 0.45 (Garber *et al.*, 2019; Kristjánsson *et al*., 2020). Empirical data show that fast growing fish tend to deposit more supplementary astaxanthin. Hence, fillet colour usually correlates positively with body weight, therefore phenotypic variation in fillet colour may be partly explained by variation in body weight. However, pre-adjusting fillet colour by body weight may cause considerable changes in heritability and genetic correlation if fillet colour is not genetically independent from body weight. Kennedy *et al.* (1993) show that, when a derived trait is a linear function of two component traits, it can be optimally selected using an index of the two component traits. Nonetheless, GWAS tries to account for body weight by using it as a covariate when identifying QTL for several focal traits (fillet colour, fillet fat, lice count etc.). We hypothesize that using body weight as a correlated trait in the MTGWAS will be a superior model that will also improve statistical power to detect candidate genes. The aim of this study was to compare allele substitution effects (*b*) of fillet colour and their significance level estimated with three different approaches in GWAS; single-trait (STGWAS), pre-adjusting fillet colour for body weight (STadjGWAS), and a multi-trait (MTGWAS) where fillet colour and body weight were simultaneously analysed.

Materials & Methods

Phenotypes & Genotypes. Fillet colour in this study is defined as the concentration of astaxanthin in the fillet estimated from near infrared spectrometry. The fillet colour and final body weight were obtained from 3,420 Atlantic salmon originating from 341 full-sib families of Mowi Genetics breeding population. The mean (standard deviation) of fillet colour and

body weight was 5.09 (1.38) mg/kg fillet and 4.13 (1.13) kg, respectively. The regression coefficient of fillet colour on body weight indicated that fillet colour increased by 0.49 mg/kg per 1 kg body weight. Variance components of fillet colour and body weight were preliminary estimated with REML using bivariate animal mixed model in ASReml v4.1 prior to the GWAS analysis. The heritability (SE in parenthesis) for fillet colour and body weight was 0.44 (0.04) and 0.35 (0.04), respectively. The genetic correlation (SE in parenthesis) between fillet colour and body weight was 0.35 (0.07) but reduced to 0.08 (0.09) for fillet colour adjusted for body weight. Each individual was genotyped using a ThermoFisher Axiom genotyping array containing 55,725 single nucleotide polymorphism (SNP) markers distributed over the salmon genome. Quality control of SNPs was performed in PLINK v1.9, based on the following criteria: SNPs were removed if (1) locus missing rate or individual missing rate was greater than 10%, (2) SNPs deviated from Hardy–Weinberg equilibrium with a *p*-value cut-off of 1×10^{-100}, and (3) minor allele frequency was lower than 0.01. After the quality control, 53,159 SNPs remained and the phenotypes reduced to 3,417 individuals.

Genome-wide association study (GWAS). We identified SNPs associated with fillet colour using three different approaches. For STGWAS and STadjGWAS, the analysis was performed using GCTA (Yang *et al*., 2011), where the fixed covariates in the model were cages, sex, and principle components (PCA, number of PCA used until the inflation factor (λ) approached 1). For the MTGWAS, the analysis was performed using multivariate linear mixed model in GEMMA (Zhou and Stephens, 2014). The number of PCA dimensions in MTGWAS was set to be the same as STGWAS. A subset of top ranked markers associated with QTL for fillet colour from different approaches was taken from all the markers based on ten highest $-\log_{10}(0.05/\text{number of markers})$. The magnitude of marker effect estimates and their ranking due to *p*-value across different approaches were compared. The difference in marker estimates was quantified by calculating root mean square error (RMSE).

Results

The STGWAS identified the most significant SNPs associated with fillet colour on chromosome 2 and 26 (Figure 1). On chromosome 2, marker AX-88047706 was the most significant with an allele substitution effect estimate (*b*) of 0.26 mg/kg. However, a different marker (AX-96373699) in the same region was the most significant from the STadjGWAS and MTGWAS models. The *b* (0.35 mg/kg) of AX-96373699 from MTGWAS was greater than *b* (0.30 mg/kg) from STadjGWAS.

On chromosome 26, marker AX-87369251 was the most significant from all the three models however, *b* from MTGWAS increased by 36.8% compared to STGWAS. Significant level based on *p*-value increased from 8.73×10^{-25} (STGWAS) to 2.61×10^{-29} (MTGWAS). The λ was equal to 1.1 in both STGWAS and STadjGWAS while λ reduced to 0.93 for MTGWAS.

The proportion of *b* to the mean of fillet colour of the top ten SNPs indicated the reduction in *b* for STadjGWAS compared to STGWAS. The MTGWAS revealed greater *b* in fillet colour for all top SNPs than STGWAS and STadjGWAS (Figure 2). The RMSE of *b* between STGWAS and MTGWAS of the ten most significant markers was 0.092 mg/kg and it slightly increased to 0.114 mg/kg for RMSE between STadjGWAS and MTGWAS.

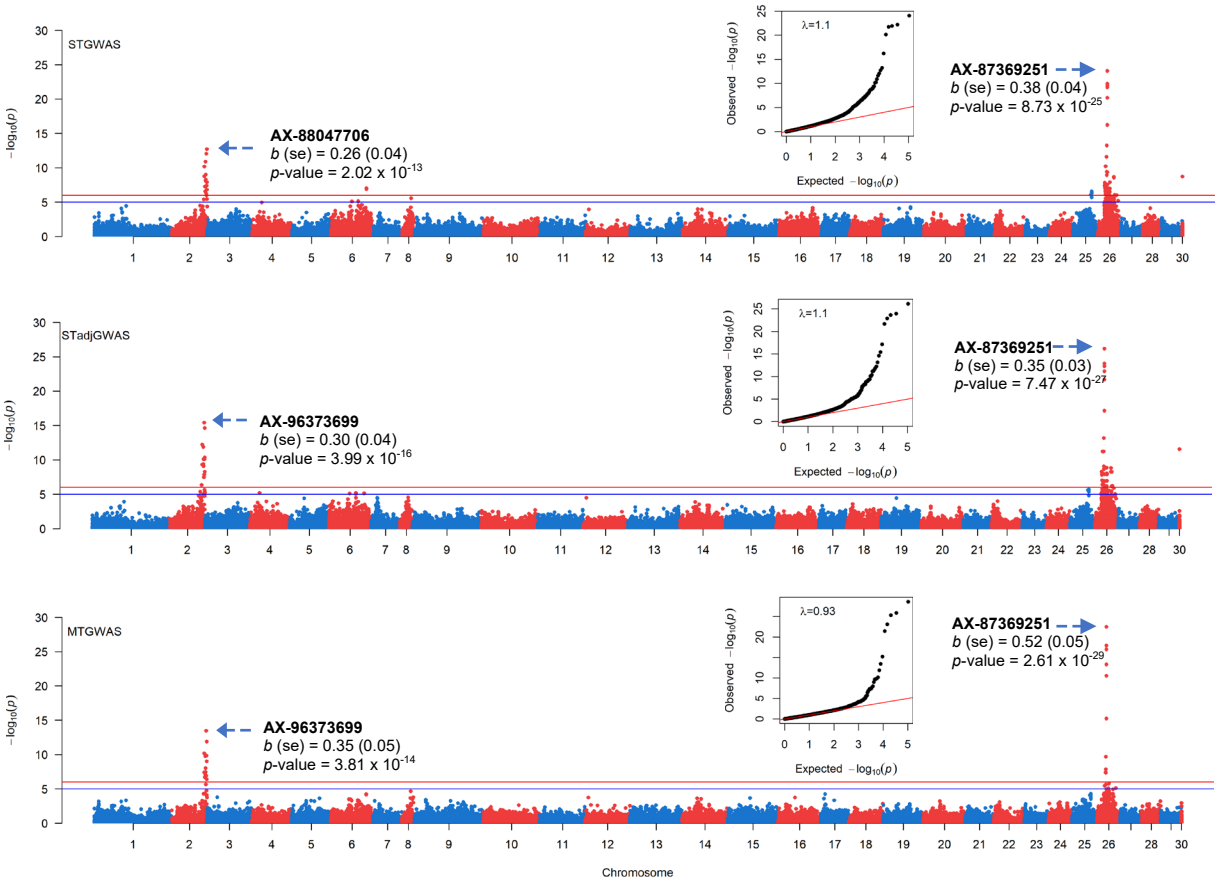

Figure 1. Single trait (STGWAS or STadjGWAS) and multi-trait genome-wide association (MTGWAS) on fillet colour and body weight in Norwegian Atlantic salmon.

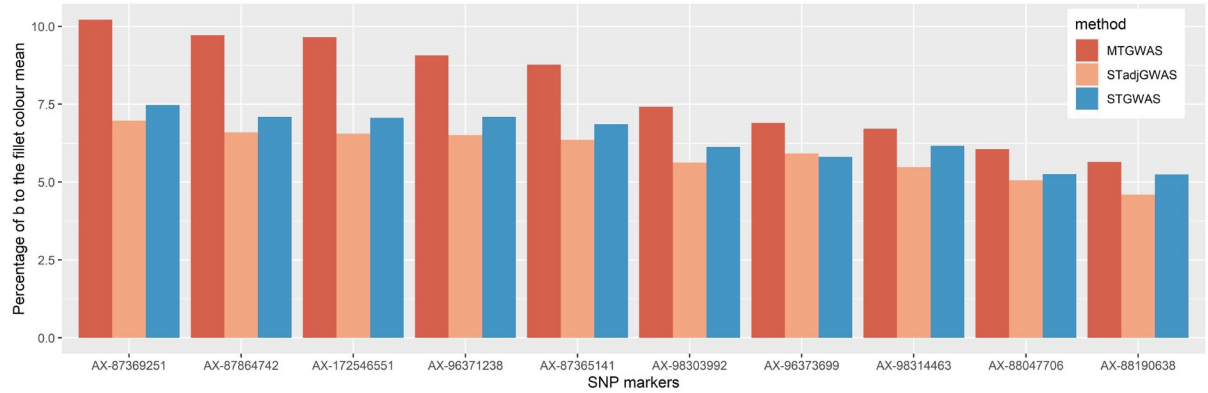

Figure 2. Percentage of allele substitution effect (*b*) to fillet colour mean of the top ten (ranked from highest to lowest) significant SNPs based on different methods of genome-wide association study.

Discussion

Two genomic regions on chromosome 2 and 26 were associated with fillet colour. The QTL peak on chromosome 26 has also been reported by Baranski *et al*. (2010) and Helgeland *et al*. (2019). Helgeland *et al*. (2019) reported *beta-carotene oxygenase 1 like* (*bco1l*) at chromosome 26 as the functional gene for fillet colour. In our study, because AX-87369251 is located not more than 500-100 bp from SNPs rs863442990 and rs863785818 reported by Helgeland *et al*. (2019), it is very likely that AX-87369251 is also associated with *bco1l*. The

most significant marker on chromosome 2 was located at 70,851,651 bp and was close to the gene *ppa1b* (*Salmo salar* inorganic pyrophosphatase-like, LOC106608436). This enzyme hydrolyses inorganic pyrophosphate which could be a key process in the biosynthesis of astaxanthin (Muthuraman *et al*., 2021).

Using a multi-trait GWAS to account for the relationship (genetic covariance) between body weight and fillet colour resulted in significant increase in allele substitution effect (*b*) and statistical significance (*p*-value). The increase in *b* could be explained by additional information from the genetically correlated trait (body weight). This phenomenon is similar to a change in breeding values when genetic evaluation is based on multiple correlated traits instead of a single trait. Furthermore, an increase in statistical power from MTGWAS models have been reported by Yoshida *et al*. (2021) for body traits in aquaculture; Broadaway *et al*. (2016) for morphological and physiological traits in cotton and Hackinger and Zeggini, (2017) in humans. Lastly, MTGWAS models may help uncover potential pleiotropic effects (Cai *et al*., 2020; Chhetri *et al*., 2019; Hackinger and Zeggini, 2017).

In conclusion, two QTL regions on chromosome 2 and 26 were identified for fillet colour. The MTGWAS approach to account for relationship between fillet colour and body weight increased the magnitude of allele substitution effects of the most significant SNPs and also increased statistical power to detect the QTLs. Therefore, the MTGWAS model is recommended to identifying QTLs when a derived trait is a linear function of two component traits.

References

Baranski M., Moen T., and Våge, D. (2010) Genet. Sel. Evol. 42:17. https://doi.org/10.1186/1297-9686-42-17

Cai Z., Dusza M., Guldbrandtsen B., Sandø Lund M., and Sahana, G. (2020) Genet. Sel. Evol. 52:19. https://doi.org/10.1186/s12711-020-00538-6

Chhetri H.B., Macaya-Sanz D., Kainer D., Biswal A.K., Evans L.M., *et al.* (2019) New. Phytol. 223: 293–309. https://doi.org/10.1111/NPH.15777

Garber A.F., Amini F., Gezan S.A., Swift B.D., Hodkinson S.E., *et al.* (2019) Aquaculture 503:242–253. https://doi.org/https://doi.org/10.1016/j.aquaculture.2019.01.001

Helgeland H., Sodeland M., Zoric N., Seilø T.J., Grammes F., *et al.* (2019) Sci. Rep. 9:20061. https://doi.org/10.1038/s41598-019-56438-3

Kennedy B.W., van der Werf J.H., and Meuwissen T.H. (1993) J. Anim. Sci. 71:3239–3250. https://doi.org/10.2527/1993.71123239x

Kristjánsson Ó.H., Gjerde B., Ødegård J., and Lillehammer M. (2020) Front. Genet. 11:573265. https://doi.org/10.3389/FGENE.2020.573265

Muthuraman A., Shaikh S.A., Ramesh M., and Sikarwar M.S. (2021) Stud. Nat. Prod. Chem. 70: 151-194. https://doi.org/10.1016/B978-0-12-819489-8.00013-2

Yang J., Lee S.H., Goddard M.E., and Visscher P.M. (2011) Am. J. Hum. Genet. 88:76–82. https://doi.org/10.1016/j.ajhg.2010.11.011

Yoshida G.M., Lhorente J.P., Carvalheiro R., and Yáñez J.M. (2017) Anim. Genet. 48:698–703. https://doi.org/10.1111/age.12621

Zhou X., and Stephens M. (2014) Nat. Methods. 11:407–409. https://doi.org/10.1038/nmeth.2848