`

Proceedings, 10th World Congress of Genetics Applied to Livestock Production

Genome-wide association study (GWAS) for growth rate and sexual maturation in Atlantic salmon (*Salmo salar*)

Alejandro P. Gutierrez[1], Jose M. Yañez[2], Steve Fukui[3], Bruce Swift[4] and William S. Davidson[1]
[1]Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia, Canada. [2]Faculty of Veterinary and Animal Sciences, University of Chile, Av. Santa Rosa 11735, La Pintana, Santiago, Chile. [3]Cermaq Canada, 203-919 Island Highway, Campbell River, British Columbia, Canada. [4]TRI-GEN Fish Improvement Ltd., 2244 Wilson Road, Agassiz, British Columbia, Canada.

Abstract: Early sexual maturation is a serious drawback for Atlantic salmon aquaculture as it retards growth, affecting flesh quality and increases production times. Both growth and sexual maturation are complex traits; however, selection has been accomplished since the beginning of Atlantic salmon selective breeding programs. 480 individuals from the Cermaq Canada broodstock program were genotyped using a 6.5K single-nucleotide polymorphism (SNP) array. The GenABEL library in R statistical environment allowed us to identify markers showing a significant association with growth, late maturation and grilsing. The most significant associations were found for late maturation on markers located in Ssa28, Ssa01 and Ssa21. A lower level of association was detected with growth on Ssa13 and with grilsing on Ssa02 and Ssa14. Candidate genes were found to be associated with these traits and some of them show a direct relationship with developmental processes, especially for those in association with sexual maturation.

Introduction

Growth and age at sexual maturation are some of the most important traits in Atlantic salmon aquaculture, with fish having been continuously selected for these traits since the beginning of selective breeding programs in Norway in the 1970s (Gjedrem, 2012). Early sexual maturation is considered a serious drawback for aquaculture as it retards growth for several months while affecting flesh quality and overall production times (Nævdal, 1983; Thorpe, 1994). These two traits, growth and sexual maturation, are complex physiological processes controlled by genetic and environmental factors. In spite of the high number of environmental factors that may directly or indirectly influence growth rates, the heritability of both growth and age at maturity have been reported to show moderate levels in most cases (Garcia de Leaniz, et al., 2007; Gjedrem, 2012; Gjerde, et al., 1994). This scenario makes artificial selection for these traits plausible, which allows their improvement by means of selective breeding.

The ability to determine the chromosomal regions that affect economically important traits has led to the implementation of selective breeding based on genetic selection practices by identifying animals with favorable genotypes. In Atlantic salmon aquaculture, marker assisted selection (MAS) could be a valuable addition to current selective breeding programmes by improving the accuracy of selection, and therefore the genetic gain.

With the emergence of new sequencing technologies it is possible to obtain thousands of single nucleotide polymorphism (SNP) markers for population genotyping, which has allowed the construction of high density genetic maps (Goddard, Hayes, 2009). The current SNP array technologies provide better tools for the identification of QTL and specific markers associated with traits of interest than was possible using microsatellite markers. For Atlantic salmon in particular, a 6.5k SNP was developed (Gidskehaug, et al., 2011; Kent, et al., 2009) resulting in a SNP linkage map with ~ 5,500 markers (Lien, et al., 2011). We previously conducted QTL analyses by making use of this SNP array (Gutierrez, et al., 2013; Gutierrez, et al., 2012); however, the study of the genomic regions controlling production traits in Atlantic salmon has not yet been carried out using a genome-wide association study (GWAS) approach. GWAS evaluates the association with a trait relying on the levels of linkage disequilibrium (LD) between the markers and the genetic variation affecting the trait, testing for association of each SNP and therefore, making possible the identification of specific alleles affecting the trait. The aim of this study was to analyze growth and sexual maturation in this species using a GWAS approach in order to refine our previous mapping carried out using standard QTL analysis.

Materials and methods

Samples and phenotype data

Families were part of a commercial breeding program developed by Cermaq (formerly Mainstream) Canada and based on the Mowi strain of Atlantic salmon. Five full-sib families were selected for analysis, comprising 279 individuals (including parents) as described in Gutierrez et al. (2012). In order to increase the power to detect association, in the present study we also added 192 Atlantic salmon parents from the same broodstock year (BY) 2005.

Body weight measurements were taken as described in our previous analysis (Gutierrez, et al., 2012). Based on the weight measurements taken at times during the production cycle, the number of days required to reach

`

a market weight of 5 kg was calculated for all fish. Maturation times were classified as: precocious (≤ 12 months of age), grilse (36 months of age, at 1st sea winter (SW), normally maturing (48 to 60 months of age at second SW or third SW), and late-maturing fish (>60 months). For

SNP array and linkage mapping

All 471 individuals were selected for SNP genotyping at CIGENE, Norwegian University of Life Sciences, Ås, Norway using an Atlantic salmon 6.5K Illumina iSelect SNP-array (Gidskehaug, et al., 2011; Kent, et al., 2009). Analyses were based on an Atlantic salmon linkage map which contains ~5,650 SNP (Lien, et al., 2011).

Genome-wide association analysis

GWA analysis was carried out using the GenABEL library implemented in R. Considering the presence of closely related fish in our sample, we used the GRAMMAS approach (Genome-wide association using Mixed Model and Score test) (Amin, et al., 2007; Aulchenko, et al., 2007). Quality control (QC) was performed and those markers with a call rate lower than 95% and a minimum allele frequency lower than 0.05 were filtered and excluded from the analysis.

Linkage disequilibrium

The levels of linkage disequilibrium as r^2 were calculated using the GenABEL package and calculated for all adjacent marker pairs. The values of mean and median were obtained by the use of SPSS (IBM).

Results and discussion

A total of 466 samples and 3,908 markers passed the QC and consequently GWA analysis was carried out.

Only one marker (GCR_cBin15343_Ctg1_36) located on chromosome 13 (Ssa13) was found to be significantly associated with growth (see Figure 1), but only at the chromosome-wide level of significance ($p < 2.55e-4$ for Ssa13). Within 10 kb upstream of the location of the SNP we found MAGI-1 (membrane-associated guanylate kinase, WW and PDZ domain-containing protein 1), a protein related to cell-cell contact such as neuronal synaptic and epithelial tight junctions (Hata, et al., 1998) The detection of a single marker in significant association was certainly unexpected and is in disagreement with previous analyses in terms of the number of regions controlling this complex trait. However, it agrees in part with results from our previous analysis where we found a genome-wide significant QTL on Ssa13 associated with growth in Atlantic salmon (Gutierrez, et al., 2012)

Analysis for grilsing identified seven markers with a chromosome-wide significant association to the trait. Most of these markers (6 of 7) are located on Ssa02 (ESTNV_16766_113, GCR_cBin2733_Ctg1_528, GCR_cBin28152_Ctg1_60, GCR_cBin24972_Ctg1_42, ESTNV_36998_1021 & GCR_cBin28152_Ctg1_185). Three of these markers (GCR_cBin28152_Ctg1_60, GCR_cBin24972_Ctg1_42 and GCR_cBin28152_Ctg1_185) were assigned to AGKD01052475 comprising a small region of 2 cM, but they were not associated with a particular gene. The most significantly associated marker ESTNV_16766_113 was positioned on Ssa02 ($p < 3.91e-04$ for Ssa02), and is located within the coding region of RING finger protein 146, a RING-domain E3 ubiquitin ligase, that acts as a positive regulator of Wnt signaling (Zhang, et al., 2011), which plays essential roles in embryonic development and adult tissue homeostasis. GCR_cBin2733_Ctg1_528 also on Ssa02, was located in the intronic region of a myoglobin gene (MYG). The other five markers that showed a chromosome-wide significant association with grilsing were located in uncharacterized genes. To our knowledge these chromosomes have not been described as containing sexual maturation QTL in salmonids. However, numerous QTL related to growth have been identified on Ssa02 (Baranski, et al., 2010; Boulding, et al., 2008; Gutierrez, et al., 2012; Houston, et al., 2009; Reid, et al., 2005). Regions of this chromosome share homology with linkage groups from other salmonid species such as rainbow trout (RT-27), where body-weight and early maturation QTL have been described (Haidle, et al., 2008; Wringe, et al., 2010)

Late maturation analysis detected association with five markers, located on Ssa28, Ssa01 and Ssa21 (ESTNV_22894_922, ESTNV_35192_247, GCR_cBin47084_Ctg1_67, ESTNV_31055_861 and ESTNV_27268_490), which showed a genome-wide significant association with the trait according to the Bonferroni thresholds (only ESTNV_22894_922 and ESTNV_35192_247 reached genome-wide significance ($p<0.05$) using the permutation method).The two most significant markers are identified as ESTNV_22894_922 and ESTNV_35192_247, and are located on chromosomes Ssa28 and Ssa01, respectively. ESTNV_22894_922 is a SNP that by annotation was positioned in the coding region of FRA10AC1, a gene found within the rare FRA10A folate-sensitive fragile. ESTNV_35192_247 is located in the coding region of a CPEB-associated factor Maskin putative protein, which has been mainly associated to the control of oocyte maturation in *Xenopus* by repression and derepression of mRNAs (Stebbins-Boaz, et al., 1999). Previous studies analysing sexual maturation related traits in other salmonid species have identified QTL in similar regions. In rainbow trout for example, genome-wide significant QTL associated with early maturation was found on RT-17, which shares homology with Ssa28, (Haidle, et al., 2008), but also chromosome-wide significant QTL linked

`

to developmental rates (Easton, et al., 2011). In addition to QTL found on Ssa28, previous studies have also described QTL on Ssa01 for other salmonid species. For instance, QTL for age at sexual maturation was described in AC-9 that shares homology with Ssa01 (Küttner, et al., 2011) and also QTL for condition factor (Moghadam, et al., 2007). In the case of Ssa21, we recently described QTL for grilsing located in the same chromosome; however, in this analysis the QTL was detected for late maturation instead, which suggests that these regions controlling sexual development contain genes that work in both processes

Given the differences found between regular QTL analysis and GWA analysis, we believe that such discrepancies could be explained by the different approaches for these analyses. Standard QTL analyses make use of the amount of recombination or linkage between individuals to detect association between markers and trait, which makes it a powerful method when using a low number of markers. On the other hand, the statistical power of GWA is a function of sample size, effect size and marker allele frequency (Stranger, et al., 2011), depending on the level of linkage disequilibrium between the genetic markers to detect association. Being that said, the low power of detection specially observed in our analysis of growth (days to 5 kg) and grilsing could be attributed to the low number of markers and samples analysed (Sodeland, et al., 2013). Accordingly, Sodeland, et al. (2013) using the same SNP chip to analyse quality traits in Atlantic salmon, recently showed that the levels of LD were not sufficiently high to explain a considerable proportion of the genetic variation for carcass quality traits in Atlantic salmon, even with a larger sample size. A similar result regarding low values of LD between adjacent markers was found in our analyzed data, suggesting that a higher density SNP panel will be needed to fully exploit LD in association mapping studies.

The levels of LD detected using the markers available was on average = 0.22 with a median = 0.11, which is higher than the critical value of $r^2 = 0.2$ needed to capture the effect of genomic regions affecting quantitative traits using LD between two markers; however the decay of linkage disequilibrium is correlated with the distance between markers. These results could indicate that the kinship between our sampled individuals is high, but also shows that the low power of detection is due to small effects and the experimental design. Nevertheless, the density of the SNPs used here could be improved in order to increase the power to detect association between traits and markers.

The estimated levels of heritability calculated using a pedigree-based relationship matrix, for grilsing and late maturation in this population are low: grilse $h^2= 0.09$, late maturation $h^2= 0.11$ but in agreement the current literature that has described levels of heritability for age at sexual maturity traits in Atlantic salmon ranging from 0.04 to 0.17 as reviewed by Garcia de Leaniz, et al. (2007). Heritability of growth (days to 5 Kg) on the other hand, was found to be significantly higher ($h^2= 0.2$) also in agreement with previous findings which estimated heritabilities for growth rate and body weight range from 0.04 – 0.26 and 0.05–0.44 respectively (Garcia de Leaniz, et al., 2007). Thus, selection for these traits is possible and it has been done since the early 70's (Gjedrem, 2012). Selective breeding programs have been effective in increasing body size while also controlling undesired early sexual maturity in farmed fish (Gjedrem 2000), suggesting that the main genomic regions controlling these traits are different. The use of molecular marker information could be a valuable tool to improve conventional broodstock selection programs by the identification of affected alleles that could then be screened into different populations

References

Amin, N., van Duijn, C.M., Aulchenko, Y.S., 2007. PLoS ONE. 2, e1274.

Aulchenko, Y.S., Ripke, S., Isaacs, A., van Duijn, C.M., 2007. Bioinformatics. 23, 1294-1296.

Baranski, M., Moen, T., Vage, D., 2010. Genet. Sel. Evol. 42, 17.

Boulding, E.G., Culling, M., Glebe, B., Berg, P.R., Lien, S., Moen, T., 2008. Heredity. 101, 381-391.

Easton, A.A., Moghadam, H.K., Danzmann, R.G., Ferguson, M.M., 2011. J. Fish Biol. 78, 602-623.

Garcia de Leaniz, C., Fleming, I.A., Einum, S., Verspoor, E., Jordan, W.C., Consuegra, S., Aubin-Horth, N., Lajus, D., Letcher, B.H., Youngson, A.F., Webb, J.H., Vøllestad, L.A., Villanueva, B., Ferguson, A., Quinn, T.P., 2007. Biological Reviews. 82, 173-211.

Gidskehaug, L., Kent, M., Hayes, B.J., Lien, S., 2011. Bioinformatics. 27, 303-310.

Gjedrem, T., 2012. Aquaculture. 344–349, 12-22.

Gjerde, B., Simianer, H., Refstie, T., 1994. Livestock Production Science. 38, 133-143.

Goddard, M.E., Hayes, B.J., 2009. Nat Rev Genet. 10, 381-391.

Gutierrez, A.P., Lubieniecki, K.P., Fukui, S., Withler, R.E., Swift, B., Davidson, W.S., 2013. Mar. Biotechnol., 1-8.

Gutierrez, A.P., Lubieniecki, K.P., Davidson, E.A., Lien, S., Kent, M.P., Fukui, S., Withler, R.E., Swift, B., Davidson, W.S., 2012. Aquaculture. 358–359, 61-70.

Haidle, L., Janssen, J.E., Gharbi, K., Moghadam, H.K., Ferguson, M.M., Danzmann, R.G., 2008. Mar. Biotechnol. 10, 579-592.

Hata, Y., Nakanishi, H., Takai, Y., 1998. Neuroscience Research. 32, 1-7.

Houston, R.D., Bishop, S.C., Hamilton, A., Guy, D.R., Tinch, A.E., Taggart, J.B., Derayat, A., McAndrew, B.J., Haley, C.S., 2009. Anim. Genet. 40, 753-755.

Kent, M.P., Hayes, B., Xiang, Q., Berg, P.R., Gibbs, R.A., Lien, S., 2009. Development of 16.5K SNP-Chip for Atlantic salmon, Plant & Animal Genomes XVII Conference, San Diego, CA.

Küttner, E., Moghadam, H.K., Skúlason, S., Danzmann, R.G., Ferguson, M.M., 2011. Mol. Genet. Genomics, 1-13.

Lien, S., Gidskehaug, L., Moen, T., Hayes, B., Berg, P., Davidson, W., Omholt, S., Kent, M., 2011. BMC Genomics. 12, 615.

Moghadam, H.K., Poissant, J., Fotherby, H., Haidle, L., Ferguson, M.M., Danzmann, R.G., 2007. Mol. Genet. Genomics. 277, 647-661.

Nævdal, G., 1983. Aquaculture. 33, 97-106.

`

Reid, D.P., Szanto, A., Glebe, B., Danzmann, R.G., Ferguson, M.M., 2005. Heredity. 94, 166-172.
Sodeland, M., Gaarder, M., Moen, T., Thomassen, M., Kjøglum, S., Kent, M., Lien, S., 2013. Aquaculture. 408–409, 169-174.
Stebbins-Boaz, B., Cao, Q., de Moor, C.H., Mendez, R., Richter, J.D., 1999. Mol. Cell. 4, 1017-1027.
Stranger, B.E., Stahl, E.A., Raj, T., 2011. Genetics. 187, 367-383.
Thorpe, J.E., 1994. Aquacult. Res. 25, 77-87.
Wringe, B.F., Devlin, R.H., Ferguson, M.M., Moghadam, H.K., Sakhrani, D., Danzmann, R.G., 2010. BMC Genet. 11, 63.
Zhang, Y., Liu, S., Mickanin, C., Feng, Y., Charlat, O., Michaud, G.A., Schirle, M., Shi, X., Hild, M., Bauer, A., Myer, V.E., Finan, P.M., Porter, J.A., Huang, S.-M.A., Cong, F., 2011. Nat Cell Biol. 13, 623-629.