ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

SHORT COMMUNICATION

doi: 10.1111/j.1365-2052.2012.02387.x

Assessing footprints of selection in commercial Atlantic salmon populations using microsatellite data

V. Martinez*, P. Dettleff*, P. Lopez*, G. Fernandez*, A. Jedlicki*, J. M. Yañez* and W. S. Davidson†

*FAVET-INBIOGEN, Faculty of Veterinary Sciences, University of Chile, Avda. Santa Rosa 11735, Santiago, Chile. †Department of Molecular Biology and Biochemistry, Simon Fraser University, 8888 University Drive, Burnaby, BC, V5A1S6, Canada

Summary

Relatively large rates of response to traits of economic importance have been observed in different selection experiments in salmon. Several QTL have been mapped in the salmon genome, explaining unprecedented levels of phenotypic variation. Owing to the relatively large selection intensity, individual loci may be indirectly selected, leaving molecular footprints of selection, together with increased inbreeding, as its likely relatives will share the selected loci. We used population differentiation and levels of linkage disequilibrium in chromosomes known to be harbouring QTL for body weight, infectious pancreatic necrosis resistance and infectious salmon anaemia resistance to assess the recent selection history at the genomic level in Atlantic salmon. The results clearly suggest that the marker *SSA0343BSFU* on chromosome 3 (body weight QTL) showed strong evidence of directional selection. It is intriguing that this marker is physically mapped to a region near the coding sequence of *DVL2*, making it an ideal candidate gene to explain the rapid evolutionary response of this chromosome to selection for growth in *Salmo salar*. Weak evidence of diversifying selection was observed in the QTL associated with infectious pancreatic necrosis and infectious salmon anaemia resistance. Overall, this study showed that artificial selection has produced important changes in the Atlantic salmon genome, validating QTL in commercial salmon populations used for production purposes according to the recent selection history.

Keywords body weight, footprints of selection, genetic selection, QTL, salmon.

Selection can have important effects on genome variability through either population differentiation (as calculated from F_{ST}) or increasing levels of linkage disequilibrium (LD). Even though several demographic parameters can explain these population parameters, attributed to 'hitch-hiking', loci near genes explaining quantitative genetic variation (QTL) will also be affected, generating significant LD and allele frequency changes (Martinez 2010). Because of the reproductive biology of fish species, the effects of artificial selection are likely to be of great importance in shaping the genomic variability around molecular markers harbouring QTL, as a result of indirect increased selection intensity. This is likely to be the case in many breeding programs in which large selection responses are observed, especially for body weight (15% per generation; Martinez *et al.* 2006). We assessed the effects of selection by considering chromosomes harbouring QTL for economically important traits in Atlantic salmon using populations that differ in their recent population history. We used individuals from four populations: the first two were mainly selected for growth, based on grading at different stages [POP_1 ($n = 64$) and POP_2 ($n = 50$)]; a third [POP_3 ($n = 48$)] was maintained with closed matings but apparently without a formal selection scheme; and the fourth population was a naturalised population ($n = 30$) obtained from a river with natural spawning occurring for more than a decade (the Quiman river in southern Chile) (Martinez *et al.* 2012). This last population was included as a negative control of recent artificial selection for production traits.

Genomic DNA was extracted from muscle tissue from all individuals using a DNeasy Blood & Tissue Kit (Qiagen). Atlantic salmon chromosomes (SSA) were selected based on the previous studies that showed segregation of QTL related

Address for correspondence

V. Martinez, FAVET-INBIOGEN, Faculty of Veterinary Sciences, University of Chile, Avda. Santa Rosa 11735, Santiago, Chile. Website: www.genetica-animal.uchile.cl.
E-mail: vmartine@u.uchile.cl

Accepted for publication 23 April 2012

to infectious salmon anaemia resistance (SSA15, Moen *et al.* 2007), infectious pancreatic necrosis resistance (SSA26, Moen *et al.* 2009) and growth rate (SSA03, Reid *et al.* 2005). PCR annealing temperatures, primers, repeat motifs and genetic locations of the microsatellite markers are presented in Table S1, and the PCR protocols are shown in Appendix S1. Fragment sizes and alleles were determined using GENEMAPPER after standardising allele size.

The basic allele frequency analysis using GENEPOP (Rousset 2008) showed that, to a large extent, markers exhibited an excess of homozygous genotypes. The average inbreeding did not differ significantly between populations, but there were significant differences between chromosomes (SSA03 showing the greatest average inbreeding, $F_{IS} = 0.23$; $P < 0.05$). The average F_{ST} analysis revealed that overall there is little differentiation between populations (between $F_{ST} = 0.02$ and $F_{ST} = 0.03$) and there were no significant differences between the chromosomes analysed (Fig. 1).

Outlier loci detection was carried out using BAYESCAN, which modelled the F_{ST} considering a population and a specific locus component under the assumption of different effective population sizes (Ne) and migration (Foll & Gaggiotti 2008). Concomitantly, we used LOSITAN, which simulates the joint distribution between F_{ST} and heterozygosity (He) under neutrality to obtain the empirical distribution under the null hypothesis of no selection considering an island model (Antao *et al.* 2008). Several markers show strong evidence of either balancing or directional selection (Fig. 2). However, both methods confirmed that marker *SSA0343BSFU* showed outlier behaviour for directional selection. It is interesting to note that the highest F_{ST} between pairs of populations were observed between one of the selected populations and the rest (POP_2), with F_{ST} values ranging between 0.16 and 0.20. Other marker loci provide strong evidence of balancing selection, but because there might be admixture in the samples, this result should be considered with caution.

The value of D' was calculated between pairs of syntenic marker loci to estimate the impact of selection on LD levels (Martinez 2010). This parameter was calculated using ldmax for the analysis of non-family genotypic data (Abecasis & Cookson 2000). The results show that there is extensive LD spanning relatively large distances in the chromosomes (Fig. S1). It is important to note that relatively high LD was observed between markers located at distal positions (4.3 and 87 cM) of SSA03 (see Fig. S1).

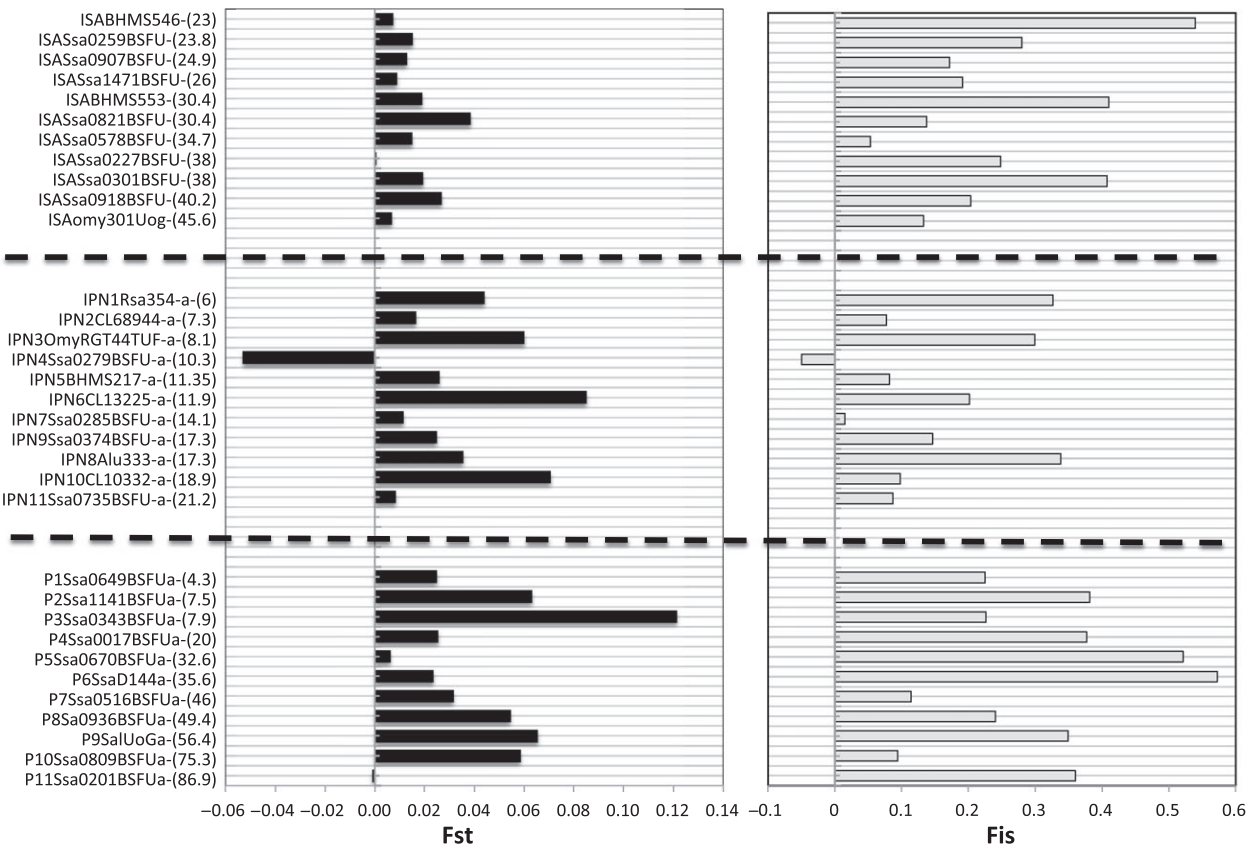

Figure 1 Estimates of F_{ST} (black bars) and F_{IS} (grey bars) over all populations as obtained from GENEPOP for each marker locus [microsatellite marker (in parenthesis, male female average cM)] for each chromosome analysed (separated by black dashed lines).

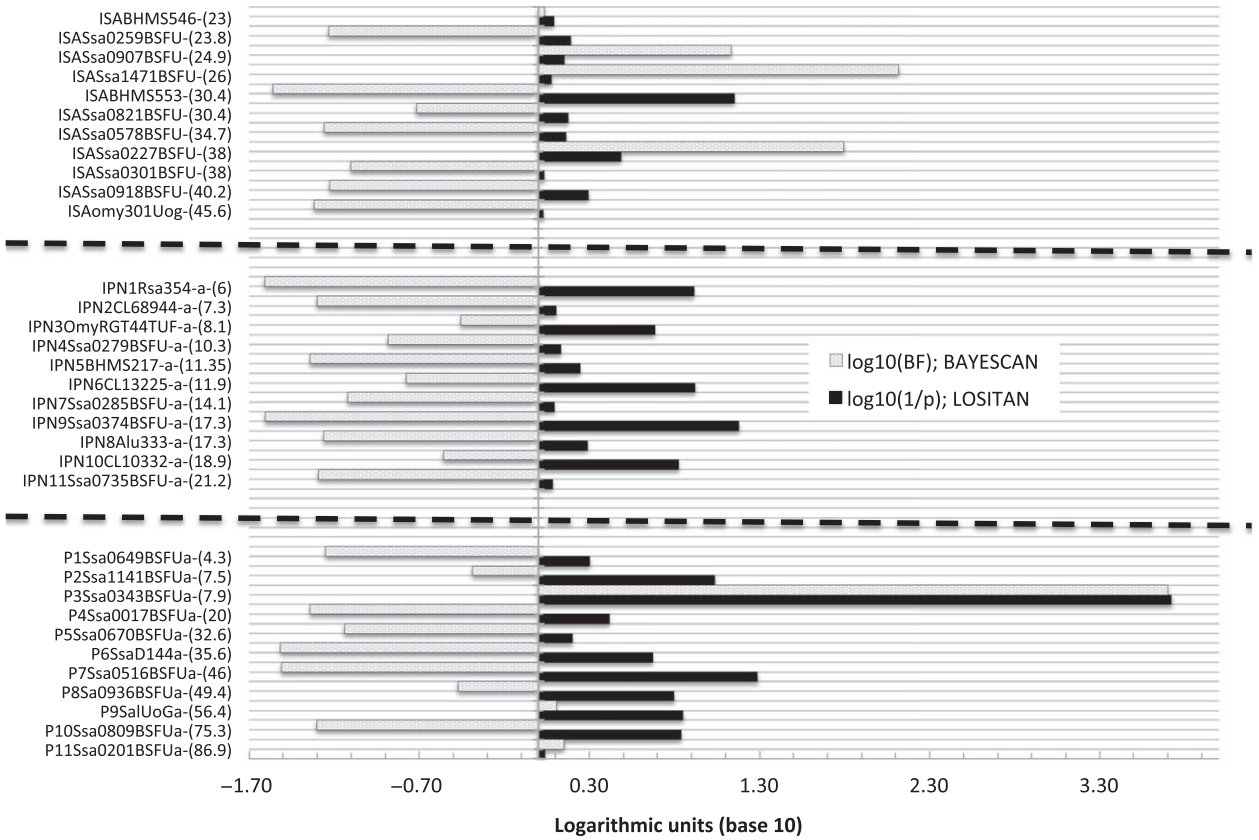

Figure 2 Outlier loci detection for selection. Black bars represent the log10 of the Bayes factor of a model considering or not considering a locus's [microsatellite marker (in parenthesis, male female average cM)] specific component with negative (positive) values indicating balancing (directional selection) and grey bars representing log10 of (1/p) as obtained from the empirical distributions under neutrality as obtained from the simulations in LOSITAN. Different chromosomes analysed are separated by black dashed lines.

Taken together, the results of the F_{ST} and LD analysis reveal that a low Ne together with strong selection has facilitated a rapid evolution of the salmon genome in a region harbouring a QTL for body weight. To determine which genes are adjacent to *SSA0343BSFU*, we used the BAC-end sequence (S0023P18_SP6, http://www.asalbase.org) surrounding the microsatellite marker. The highest BLASTX hit (e-value = $1.624e^{-09}$) corresponded to teleost *DVL2* and *DVL3*, dishevelled homologues, which are involved in Wnt signalling and cell proliferation pathways. However, the *DVL2* alignment was deemed to be more reliable. An analysis of the DVL2 protein for conserved domain positions identified three domains: DIX, PDZ and DEP (Habas & Dawid 2005). The *in silico* translated S0023P18_SP6 sequence aligns to between positions 302 and 354 of *DVL2* (NP_997813.1) and corresponds to the end of the PDZ domain. Moreover, the microsatellite $(TG)_n$ sequence is found 16 bp upstream of the coding sequence of protein DVL2 in zebrafish. The proximity of the microsatellite to the coding sequence of *DVL2* makes it a strong candidate gene to explain the rapid selection for the body weight QTL on chromosome 3 in Atlantic salmon.

Acknowledgements

This work was financially supported by the Chilean Government through CONICYT (FONDECYT 1090632).

References

Abecasis G.R. & Cookson W.O. (2000) GOLD–GRAPHICAL OVERVIEW OF LINKAGE DISEQUILIBRIUM. *Bioinformatics* **16**, 182–3.

Antao T., Lopes A., Lopes R.J., Beja-Pereira A. & Luikart G. (2008) LOSITAN: a workbench to detect molecular adaptation based on a F_{ST} outlier method. *BMC Bioinformatics* **9**, 323.

Foll M. & Gaggiotti O.E. (2008) A genome scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* **180**, 977–93.

Habas R. & Dawid I.B. (2005) Dishevelled and Wnt signalling: is the nucleus the final frontier? *Journal of Biology* **4**, 2.

Martinez V. (2010) Modelling linkage disequilibrium in commercial salmon populations using dense SNP marker information: towards detecting selection signatures. *Proceedings WCGALP 2010*, Leipzig, Germany.

Martinez V., Kause A., Mäntysaari E. & Mäki-Tanila A. (2006) The use of alternative breeding schemes to enhance genetic improvement in rainbow trout. II. Two-stage selection. *Aquaculture* **254**, 195–202.

Martinez V., Moen T., Jedlicki A., Kent M. & Lien S. (2012) Genomic assessment for molecular breeding of disease resistance in *Salmo salar*: World-wide genome-wide genetic variability and linkage disequilibrium. First International Conference on Integrative Salmonid Biology. June 14–20, Oslo Norway.

Moen T., Sonesson A.K., Hayes B., Lien S., Munck H. & Meuwissen T.H.E. (2007) Mapping of a quantitative trait locus for resistance against infectious salmon anaemia in Atlantic salmon (*Salmo salar*): comparing survival analysis with analysis on affected/resistant data. *BMC Genetics* **8**, 53.

Moen T., Baranski M., Sonesson A.K. & Kjøglum S. (2009) Confirmation and fine-mapping of a major QTL for resistance to infectious pancreatic necrosis in Atlantic salmon (*Salmo salar*): population-level associations between markers and trait. *BMC Genomics* **10**, 368.

Reid D.P., Szanto A., Glebe B., Danzmann R.G. & Ferguson M.M. (2005) QTL for body weight and condition factor in Atlantic salmon (*Salmo salar*): comparative analysis with rainbow trout (*Oncorhynchus mykiss*) and Arctic charr (*Salvelinus alpinus*). *Heredity* **94**, 166–72.

Rousset F. (2008) GENEPOP'007: a complete reimplementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* **8**, 103–6.

Supporting information

Additional supporting information may be found in the online version of this article.

Appendix S1 PCR protocols.

Figure S1 Significant LD (*D'*) between pairs of markers (distance in cM) spanning chromosomes SSA15, SSA03 and SSA26.

Table S1 PCR annealing temperatures, primers, repeat motifs and genetic locations of the microsatellite markers.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.