Genome-wide Association Study for body weight in rainbow trout
(*Oncorhynchus mykiss*)

R.V. Reis Neto1, G. Yoshida2, Jean P. Lhorente3, J.M. Yáñez3*

1Departamento de Engenharia de Pesca, Campus Experimental de Registro, Universidade Estadual Paulista “Julio de Mesquita Filho”, Av. Nelson Brihi Badur 430, #11900-000 Registro SP, Brasil
2Departamento de Zootecnia, FCAV, Universidade Estadual Paulista “Julio de Mesquita Filho”, Jaboticabal, SP, Brasil
3Facultad de Ciencias Veterinarias y Pecuarias, Universidad de Chile, Avenida Santa Rosa 11735, 8820808, La Pintana, Santiago, Chile
4Aquainnovo, Cardonal s/n, Puerto Montt, Chile

jmayanez@uchile.cl (Corresponding Author)

Summary

Linking genomic and phenotypic variation of growth-related traits may help on better understanding these important traits. The use of this information can be used to make breeding programs more efficient. Genome-Wide Association Studies (GWAS) using a dense SNP panel was performed in a commercial rainbow trout population with the objective of verifies the association between regions of the genome and growth traits. A total of 4596 rainbow trout from 105 full-sib families belonging to the 2011 year class of the Aguas Claras breeding program were used. An average of 47 animals from each family were tagged and distributed in six different tanks at an average density of 834 fish per tank where they were reared for approximately 18 months. Body weight at tagging (BWT = 7.09 ± 1.44g) and at 18 months (BW18M = 169.39 ± 36.19g) were recorded for each fish. Genotyping was performed using an Affymetrix® Axiom® 57K SNP panel. GWAS for BWT and BW18M were performed using the PREGSF90, AIRMFLF90, BLUPF90 and POSTGSF90 programs from the BLUPF90 family. The analyses were carried out using sliding windows of 5 SNPs. The five most important windows for BWT were found on chromosomes 16, 2, 23 and 1, and together explained 0.93% of genetic variance, whereas the top five windows for BW18M are on chromosomes 15, 21, 7, 25 and 4, and explain 1.13% of the genetic variance. We identified important regions of the genome related to growth in rainbow trout. The next step will be to identify possible candidate genes in the genome regions harboring top windows.

Keywords: SNP, growth, GWAS

Introduction

In almost all regions of the world, where climatic conditions are favorable, rainbow trout farming takes place. In 2014, world production exceeded 812 thousand tons generated about 3.9 billion dollars (FAO, 2016). Chile is one of the major producers of rainbow trout with 152 thousand tons produced in 2015 (SERNAPESCA, 2016).
Salmonids breeding programs are one of the major responsible for successful global spread of salmon and trout aquaculture; growth related-trait are usually the most economically important and, therefore, are used as selection criterion in programs (Leeds et al., 2016). Therefore, to associate genomic and phenotypic variation may help on a better understanding of these important traits, and this information can be used to plan strategies to make breeding programs more efficient (Yañez et al., 2015).

Recent advances in molecular biotechnologies have made possible to obtain panels with thousands of SNP markers for population genotyping (Goddard & Hayes, 2009). These panels are composed of a large number of markers, very close to each other, and distributed throughout the entire genome, so they are very efficient tools for identifying QTLs or specific markers associated with phenotypes of interest (Gutierrez et al., 2015).

In genome wide association studies (GWAS), a group of individuals, with a given phenotype measured, is genotyped to statistically verify the association between the markers and the trait of interest (Bush & Moore, 2012). In addition to verify the association, GWAS allows us to quantify the genetic variation that affects the traits of interest and to identify polymorphisms useful in marker-assisted selection schemes (Sodeland et al., 2013).

GWAS analyses using a high density SNP panel were performed in a commercial population of rainbow trout aiming to identify the association between genomic regions and growth-related-trait.

**Material and methods**

**Rainbow trout population and traits**

For analysis, 4596 rainbow trout from 105 full-sib families were used. The animals belonged to the year class 2011 from Aguas Claras (Puerto Montt, Chile) trout breeding program. The eggs of each full-sib family were incubated and reared separately until tagging. An average of 47 animals of each family (minimum 32 and maximum 59) were tagged (PIT - Passive Integrated Transponder) and distributed in six different tanks with an average of 834 fish per tank (ranging from 793 to 880), where fish were reared for approximately 18 months. Body weight at tagging (BWT = 7.09 ± 1.44g) and at 18 months (BW18M = 169.39 ± 36.19g) were recorded for each fish.

**Genotyping and Statistical Analysis**

DNA was extracted from caudal fin samples of 4596 using commercial extraction kit, and genotyping was performed using an Affymetrix® Axiom® 57K SNP panel (Palti et al., 2015). The quality of the genotypes was tested considering the Hardy- Weinberg Equilibrium (p-value < 1 × 10⁻³), Minor Allele Frequency (MAF >0.05) and genotyping call rate for SNP and samples> 0.90. Quality control was performed by PREGS90 software from the BLUPF90 family. GWAS for BWT and BW18M was performed by single-step GBLUP approach (ssGWAS) using the BLUPF90, PREGS90, AIRMLF90, BLUPF90 and POSTGSF90, also from BLUPF90 family. The association analysis followed the general linear mixed model \( y = Xb + Za + e \), where \( y \) is the vector of phenotypic values (BWT or BW18M); \( b \) is the vector for fixed effect (tanks); \( a \) is the vector additive direct genetic effects considering a covariance structure among individuals established by the genomic relationship matrix; and \( e \) is the vector for random error; \( X, Z \) are the incidence matrices for fixed and individual random effects respectively.

For each evaluated trait, the percentage of the genetic variance explained by five consecutive SNPs was calculated for each sliding window and plotted by chromosomes.
Results

After quality control, 35,630 SNPs remained for association analysis. Manhattan plots show that the top window for BWT is found on chromosome 16 and explained 0.21% of the genetic variance (Figure 1 and Table 1), while the top window for BW18M explained 0.26% of the genetic variance and was located on chromosome 15 (Figure 2 and Table 1). The top five windows for BWT were found on chromosomes 16, 2, 23 and 1, and together explain 0.93% of the genetic variance (Figure 1 and Table 1), while the top five windows for BW18M were on chromosomes 15, 21, 7, 25 and 4, and explained 1.13% of the genetic variance (Figure 2 and Table 1).

Table 1. Top five SNP windows (sliding windows of 5 SNP) associated with body weight at tagging (BWT) and at 18 months (BW18M) in rainbow trout (Oncorhynchusmykiss)

<table>
<thead>
<tr>
<th>Windows BWT</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Window</td>
<td>60</td>
<td>1545</td>
<td>1544</td>
<td>859</td>
<td>27</td>
</tr>
<tr>
<td>Chromosome</td>
<td>16</td>
<td>2</td>
<td>2</td>
<td>23</td>
<td>1</td>
</tr>
<tr>
<td>Map Position (cM)</td>
<td>317</td>
<td>16632</td>
<td>16632</td>
<td>8621</td>
<td>157</td>
</tr>
<tr>
<td>Var (%)1</td>
<td>0.207</td>
<td>0.182</td>
<td>0.181</td>
<td>0.178</td>
<td>0.176</td>
</tr>
<tr>
<td>Acc.variance (%)2</td>
<td>0.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Windows BW18M</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Window</td>
<td>1091</td>
<td>394</td>
<td>652</td>
<td>481</td>
<td>126</td>
</tr>
<tr>
<td>Chromosome</td>
<td>15</td>
<td>21</td>
<td>7</td>
<td>25</td>
<td>4</td>
</tr>
<tr>
<td>Map Position (cM)</td>
<td>14272</td>
<td>5633</td>
<td>8642</td>
<td>6802</td>
<td>1277</td>
</tr>
<tr>
<td>Var (%)</td>
<td>0.26</td>
<td>0.23</td>
<td>0.22</td>
<td>0.21</td>
<td>0.21</td>
</tr>
<tr>
<td>Acc. variance (%)</td>
<td>1.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 Percentage of genetic variance explained. 2 Percentage of the accumulated genetic variance explained by the top 5 windows.

![Manhattan Plot SNP Variance explained by 5 adjacents SNP window](image)

Figure 1 Manhattan plot of genetic variance explained by 5 SNP windows for body weight at tagging (BWT).
Figure 2 Manhattan plot of genetic variance explained by 5 SNP windows for body weight at 18 months (BW18M).

Discussion

The percentage of genetic variance explained by SNP windows in both traits can be considered low, as expected because growth-related traits are widely known to have a strong polygenic nature. Gonzales-Pena et al. (2016), evaluated 875 rainbow trout using the same 57K SNP chip, and also observed a low percentage of genetic variance explained by the SNP markers on body weight at 10 and 12 months, confirming that growth is controlled by a large number of genes.

Despite the low percentage of variation explained, regions of the top five windows should be checked for candidate genes. Yoshida et al. (2017) found important SNPs on chromosomes 6, 9, 21, 23 and 25 in a GWAS for body weight at 25 months in Atlantic salmon (Salmo salar L.). The authors point to myosin light chain kinase (MYLK); growth factor beta receptor type 3 (TGFBR3) genes and the myosin light chain 1 (MYL1) gene, as the strong functional candidate genes.

In general, with the analysis carried out in this research, it was possible to identify important regions of the genome related to the growth of rainbow trout, mainly on chromosomes 16, 2, 23 and 1 for BWT and 15, 21, 7, 25 and 4 for BW18M. Certainly, the next step for the present study will be to identify possible candidate genes in the genome regions of top windows.

List of References


FAO, 2016. Fisheries and Aquaculture Information and Statistical Branch.


Gutierrez AP, Yáñez JM, Fukui S, Swift B, Davidson WS. 2015. Genome-Wide Association
Study (GWAS) for Growth Rate and Age at Sexual Maturation in Atlantic Salmon (Salmo salar). PLoS ONE 10(3):e0119730.

Acknowledgements

Authors acknowledge Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP processes numbers 2017/00123-6) for fellowship. This study was partially funded by FONDECYT1171720.