GENOME WIDE ASSOCIATION STUDIES OF PRODUCTION TRAITS IN RAINBOW TROUT USING WHOLE GENOME SEQUENCING

Valentin Thomas, Florian Enez, Pierrick Haffray, Boudjema Imarazene, Mathilde Dupont-Nivet, Florence Phocas

GENOME WIDE ASSOCIATION STUDIES OF PRODUCTION TRAITS IN RAINBOW TROUT USING WHOLE GENOME SEQUENCING

V. Thomas[1], F. Enez[2], P. Haffray[2], B. Imarazene[3], M. Dupont-Nivet[1], P. Phocas[1]

[1] Université Paris-Saclay, INRAE, AgroParisTech, GABI, 78 350 Jouy-en-Josas, France
[2] SYSAAF, Station LPGP-INRAE, 35 042 Rennes, France
[3] Les Sources de l'Avance, Pissos 40 410, France
E-mail : valentin.thomas@inrae.fr

Introduction

Rainbow trout is one of the most commonly farmed salmonid species across the world and the most farmed fish species in France. Trout are reared in France for the production of large fillets which are consumed fresh or smoked. Thus, the main objectives of genetic selection are to improve growth and fillet's yield. To assess growth and fillet's quality various traits are measured, such as body weight, carcass and fillet yields or fat content. Though some QTLs have been found in previous experiments (Blay et al., 2021), we still lack precise information about genes and biologicals mechanisms linked to those traits. Our aim was to estimate genetic parameters of production traits in a commercial line of rainbow trout and to detect and accurately localize associated QTLs.

Materials and methods

The stock was established from 3 generations of a commercial selected line of Les Sources de l'Avance breeding company (Aqualande Group, France) whose breeding program started 12 generations ago. The stock was reared under commercial conditions until harvesting. From the 9th, the 10th and the 12th generations, 2198, 1410 and 824 fish were sampled between 503 and 539 days post fecundation. Fish were measured for fork length (FL, cm), body weight (BW, g), carcass yield (CY), head gutted carcass weight (HGCW, g), head gutted carcass yield (HGCY), viscera weight (VW, g), gut yield (GY), fat content in the muscle (FAT, %) recorded using a Fish Torry Fat-meter®.
Fin samples of 3390 phenotyped fish were genotyped for 57,501 SNPs using the Axiom® Trout Genotyping array by the INRAE genotyping platform Gentyane (Clermont-Ferrand, France). All the parents of phenotyped fish were also genotyped with the 57K array (184, 183 and 80 parents of the 9th generation,10th and 12th generations).
After SNPs quality control including filtering out SNPs with minor allele frequency (MAF) < 1%, and SNP call rate < 97%; 31,968 SNPs were retained for the analysis.
The 99 sires of the phenotyped fish in 10th generation were also sequenced with the NovaSeq6000® paired-end technology (Illumina 2×150bp) at INRAE sequencing platform Get-PlaGe (Toulouse, France). We used the nfcore/sarek 2.7.1 pipeline to call variants. We retained 10,520,443 bi-allelic SNPs with MAF > 1%. The imputation of the 32K genotypes of phenotyped fish into 10,520K genotypes was performed using FIMPUTE3 software using pedigree information. After imputation we kept 1,231,034 SNPs with a MAF above 10%, a mendelian error rate below 3% and linkage disequilibrium r^2 between SNPs < 0.9 in sliding 100kb-windows.
Genetic parameters were estimated using BLUPF90 software (Misztal et al., 2014). GWAS was performed for the 8 traits of interest with a Bayesian Sparse Linear Mixed Model (BSLMM) using GEMMA - 0.98.5 software (Zhou et al., 2013).

Results and discussion

Heritability varied from moderate to high values (TABLE 1). The highest heritability was estimated for FAT (0,63 ± 0,03) and the lowest value for FL (0,16 ± 0,02). Production traits were polygenic, the proportion of genetic variance accounted for by the 150 to 500 SNPs with the largest effects ranged from 75 to 95% depending on the trait.
We detected 17 QTLs (TABLE 2) with very strong evidence based on posterior inclusion probability (PIP) of SNPs in BSLMM and associated Bayes Factor (2lnBF > 14). We did

not re-detect seven of the ten QTLs found in the previous study based on generations 9th and 10th of the same rainbow trout line (Blay et al., 2021).

Table 1 Heritability estimates (± standard error) for production traits in rainbow trout

	FL	BW	CY	HGCW	HGCY	VW	GY	FAT
Pedigree based h^2	0.22 (0.03)	0.29 (0.04)	0.59 (0.05)	0.54 (0.05)	0.39 (0.04)	0.61 (0.05)	0.42 (0.05)	0.55 (0.06)
Genomic h^2	0.16 (0.02)	0.21 (0.03)	0.61 (0.03)	0.48 (0.03)	0.36 (0.03)	0.59 (0.03)	0.03 (0.03)	0.63 (0.03)

For FL, 3 QTLs were identified, among them the QTL at 7.1 Mb on chr 22 was also detected for BW and entirely located in the *slc39a10* gene. For BW, 4 other QTLs were detected on chr 4, 5 and 8. The QTL on chr 4 (with peak SNP at 20.2 Mb) was located within *rcan2*. The 2 QTLs on chr 8 and 22 were previously identified in the same trout line (Blay et al., 2021). The QTL on chr 8 spanned the region of *map3k7* and *bach2b* genes. For HGCW and HGCY a common QTL was identified on chr 6 in a region where *top3b* and *ppm1f* genes are annotated. All the other QTLs were in intergenic regions.

Table 2 Characteristics of QTLs for production traits in rainbow trout

Trait	Chr	Peak SNP	Position	PIP	2lnBF	QTL Start	QTL End
FL	3	Affx-1237726497	24,399,565	0.21	15.4	24,399,565	24,399,565
FL	22	AQ227110489	7,110,489	0.30	16.3	7,110,489	7,117,869
FL	22	AQ227564448	7,564,448	0.12	14.0	7,547,490	7,567,760
BW	4	AQ0419964635	19,964,635	0.14	14.0	19,964,635	19,964,635
BW	4	AQ0420198141	20,198,141	0.19	14.8	20,197,462	20,198,141
BW	5	AQ0527505234	27,505,234	0.21	15.1	27,505,234	27,505,234
BW	8	AQ0816947480	16,947,480	0.25	15.5	16,947,480	17,177,997
BW	22	AQ227110489	7,110,489	0.55	18.1	6,995,960	7,110,489
HGCW	6	AQ066609263	6,609,263	0.39	14.9	6,547,380	6,638,958
HGCW	20	AQ20439513	439,513	0.70	17.5	439,513	442,488
HGCW	30	AQ306279631	6,279,631	0.34	14.4	6,227,968	6,296,368
HGCY	6	AQ066609263	6,609,263	0.34	14.2	6,608,668	6,609,263
VW	17	AQ1731988589	31,988,589	0.30	14.0	31,988,589	31,992,982
GY	30	AQ308096466	8,096,466	0.40	15.3	8,096,466	8,096,478
CY	18	AQ1827622382	27,622,382	0.32	14.2	27,622,382	27,644,041
CY	8	Affx-88929033	16,330,345	0.25	13.5	16,318,476	16,330,345
FAT	8	Affx-88929033	16,330,345	0.55	16.8	16,330,261	16,330,345

Conclusion

Thanks to imputation on over 1 million SNPs, we significantly refined the location of 2 QTLs previously identified, as well as we detected 15 new QTLs. These results will be validated in an undergoing study by genotyping the 13th generation of the Aqualande's line to confirm the associations between genotypes and production traits.

Acknowledgements

The European Maritime and Fisheries Fund supported this study (Omega-Truite PFEA470017FA1000008, Phenomix PFEA470019FA1000013, RedOut PFEA470019FA1000007).

References

Blay et al (2021). *Frontiers in Genetics, 12*.
Misztal et al (2014). *Manual for BLUPF90 family of programs*.
Zhou et al (2013). *PLoS Genetics, 9*(2), e1003264.