



# THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### Verification of SNPs Associated with Growth Traits in Two Populations of Farmed Atlantic Salmon

**Citation for published version:**

Tsai, HY, Hamilton, A, Guy, DR, Tinch, AE, Bishop, S & Houston, RD 2016, 'Verification of SNPs Associated with Growth Traits in Two Populations of Farmed Atlantic Salmon' International Journal of Molecular Sciences, vol. 17, no. 1, 5. DOI: 10.3390/ijms17010005

**Digital Object Identifier (DOI):**

[10.3390/ijms17010005](https://doi.org/10.3390/ijms17010005)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher's PDF, also known as Version of record

**Published In:**

International Journal of Molecular Sciences

**Publisher Rights Statement:**

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.





Article

# Verification of SNPs Associated with Growth Traits in Two Populations of Farmed Atlantic Salmon

Hsin Y. Tsai <sup>1,\*</sup>, Alastair Hamilton <sup>2</sup>, Derrick R. Guy <sup>2</sup>, Alan E. Tinch <sup>2</sup>, Steve C. Bishop <sup>1</sup> and Ross D. Houston <sup>1</sup>

Received: 17 November 2015; Accepted: 11 December 2015; Published: 22 December 2015

Academic Editor: Jun Li

<sup>1</sup> The Roslin Institute and Royal (Dick) School of Veterinary Studies, the University of Edinburgh, Midlothian EH25 9RG, UK; Stephen.Bishop@roslin.ed.ac.uk (S.C.B.); Ross.Houston@roslin.ed.ac.uk (R.D.H.)

<sup>2</sup> Landcatch Natural Selection Ltd., 15 Beta Centre, Stirling University Innovation Park, Stirling FK9 4NF, UK; Alastair.Hamilton@hendrix-genetics.com (A.H.); Derrick.Guy@hendrix-genetics.com (D.R.G.); Alan.Tinch@hendrix-genetics.com (A.E.T.)

\* Correspondence: hsin-yuan.tsai@roslin.ed.ac.uk; Tel.: +44-0-131-6519-224; Fax: +44-0-131-6519-105

**Abstract:** Understanding the relationship between genetic variants and traits of economic importance in aquaculture species is pertinent to selective breeding programmes. High-throughput sequencing technologies have enabled the discovery of large numbers of SNPs in Atlantic salmon, and high density SNP arrays now exist. A previous genome-wide association study (GWAS) using a high density SNP array (132K SNPs) has revealed the polygenic nature of early growth traits in salmon, but has also identified candidate SNPs showing suggestive associations with these traits. The aim of this study was to test the association of the candidate growth-associated SNPs in a separate population of farmed Atlantic salmon to verify their effects. Identifying SNP-trait associations in two populations provides evidence that the associations are true and robust. Using a large cohort ( $N = 1152$ ), we successfully genotyped eight candidate SNPs from the previous GWAS, two of which were significantly associated with several growth and fillet traits measured at harvest. The genes proximal to these SNPs were identified by alignment to the salmon reference genome and are discussed in the context of their potential role in underpinning genetic variation in salmon growth.

**Keywords:** Atlantic salmon; growth; GWAS; marker-assisted selection; MAS; SNP; PCNT; *MEP1A*

## 1. Introduction

The Food and Agriculture Organization (FAO) reported that the worldwide production of farmed finfish was approximately 66.6 million tonnes from 2011 to 2012, an increase of 26% compared with 2008 to 2009 [1]. The demands for high quality animal proteins are continuously expanding due to global economic development and human population increase. Aquaculture has a major role in fulfilling the increased requirement of protein consumption, and the continuous improvement of farming scale, sustainability and efficiency is required. Selective breeding for key production traits (such as feed efficiency and disease resistance) in finfish and shellfish species is an essential component of this improvement. However, aquaculture breeding schemes are generally fewer and less developed than terrestrial livestock and plants [2,3]. Gjedrem *et al.* [4] indicated that less than 10% of aquaculture production was based on genetically-improved stock. Notably, the annual genetic gain in selective breeding programmes of aquaculture species is typically higher than that of farmed terrestrial species [4], highlighting that genetic improvement of the key economic traits can be readily achieved.

The development of high throughput sequencing technologies has expedited the discovery of millions of genome-wide SNPs, particularly for salmonid species, which have high economic values; e.g., Atlantic salmon [5,6], rainbow trout [7,8] and sockeye salmon [9]. To date, for Atlantic salmon, traits, such as fillet colour, sexual maturation and fat percentage, have been initially studied using genome-wide association (GWA) analyses using an SNP array with approximately 6 K markers [10,11]. Additionally, GWAS for host resistance to sea lice [12], host resistance to *Piscirickettsia salmonis* [13] and early growth traits [14] have been performed using higher density SNP chips (50 or 132 K SNPs). Around 70 to 100 million years ago, the ancestor of modern salmonids underwent a whole genome duplication (WGD) event [15,16], which was followed by extensive modifications of both the genome and transcriptome and is still under the process of returning to diploidy [17,18]. The relics of the duplicated genomes generated by WGD complicate the discovery and interpretation of genomic variation, partly due to the difficulty in distinguishing true segregating polymorphism from paralogous variation [17]. Nonetheless, the vast majority of SNPs discovered to date in salmonid species segregate in a diploid manner [19].

The heritability of growth traits, such as body weight and length, in Atlantic salmon is moderate to high (e.g., [10,20,21]); but these complex traits are usually considered highly polygenic, and the underlying physiological basis for growth is likely to involve networks of many interacting genes. Typically, functional networks regulating growth-related traits involve hundreds of candidate genes [22,23]. Detecting and investigating the function of each individual gene within such complex networks is practically unfeasible. However, clues to the possible roles of particular candidate genes can be determined by associating genomic variation within or close to the gene with phenotypic variation in the trait of interest on a population scale. Herein lies the potential of GWAS to inform the underlying biology of the trait in question, in addition to providing potential markers for selective breeding programmes. Several previous studies of the association between candidate gene polymorphisms and phenotypic variation in salmon populations have focused on well-known candidates with previously-demonstrated physiological roles in the trait of interest (e.g., [19,24,25]). With the advent of high density and high throughput genotyping assays, GWAS and subsequent alignment to a reference genome [26] can identify positional candidate genes in a more systematic manner. However, with all association studies, it is important to assess the robustness of any putative significant result by testing the association between the SNP and the trait in a separate population/study. Therefore, the aims of this study were (i) to test the association of a subset of the most significant SNPs associated with weight and length of juvenile salmon [14] in another large cohort of fish and (ii) to identify and discuss putative candidate genes proximal to the SNPs that may directly contribute to variation in the growth phenotypes.

## 2. Results

### 2.1. Heritability Estimation

The population used in the analysis was a random subset of a larger population (Population 2) measured for overall and component weight traits, colour and fat content. Heritabilities of fillet-related traits were moderate to high (0.52 to 0.53), whereas the waste weights (e.g., head weight) were approximately 0.3. The heritability of fat percentage and fillet colour was slightly lower (0.14 to 0.18). The phenotypic and genetic correlations were high for all of the weight-related traits ( $r \sim 0.96$  to 0.99), but with little correlation between weight traits and fillet colour ( $r = -0.08$ ). A summary of the heritability estimation and general statistics are given in Table 1, and they were consistent with estimates made on the larger population analysed previously [19].

**Table 1.** The summary statistics and heritability estimates for the harvest traits.

Traits	Mean (SD)	Heritability (SE)
Harvest weight (kg)	2.65 (0.72)	0.52 (0.05)
Head weight (kg)	0.30 (0.12)	0.21 (0.03)
Body waste weight (kg)	0.34 (0.15)	0.15 (0.02)
Total waste weight (kg)	0.67 (0.21)	0.32 (0.04)
Gutted weight (kg)	2.42 (0.65)	0.53 (0.05)
Deheaded weight (kg)	2.11 (0.57)	0.52 (0.05)
Fillet weight (kg)	1.76 (0.48)	0.53 (0.05)
Fat percentage (%)	13.2 (5.98)	0.18 (0.03)
Fillet colour (20–34)	28.9 (0.74)	0.14 (0.03)
Gut weight (kg)	0.22 (0.08)	0.30 (0.04)

### 2.2. Association between SNPs and Traits of Interest

Based on the results of the 2007 year group (Population 1) GWA analysis, 16 nominally significant SNPs were selected for genotyping in the 1999 year group (Population 2). These SNPs were chosen from QTL regions on chromosomes 16, 21 and 28 for weight and chromosomes 5, 16, 17 and 20 for length (Figures 1 and 2). Assays failed for six SNPs, and two more were monomorphic (details of selected markers were tabulated in Table S1). Of the remaining eight successfully genotyped SNPs, two were significantly associated with several growth traits (Table 2).

**Table 2.** Results of the association analysis including the predicted mean value (and standard error) and proportion of additive genetic variance due to SNP (PVE) for each trait and genotype class.

Traits	AX88141678 (Gene: MEPIA)				AX88270804 (Gene: PCNT)			
	A/A	A/G	G/G	PVE (%)	A/A	A/G	G/G	PVE (%)
# of fish	<i>n</i> = 651	<i>n</i> = 436	<i>n</i> = 52		<i>n</i> = 281	<i>n</i> = 581	<i>n</i> = 265	
Harvest weight	2.59 (0.04)	2.63 (0.05)	2.33 (0.1) **	0.3	2.66 (0.05)	2.60 (0.04)	2.50 (0.06) *	1
Head weight	0.30 (0.01)	0.30 (0.01)	0.26 (0.02) **	1.3	0.31 (0.01)	0.3 (0.01)	0.28 (0.01) **	1
Body waste weight	0.33 (0.01)	0.35 (0.01)	0.32 (0.02)	0.1	0.34 (0.01)	0.35 (0.01)	0.31 (0.01) **	3
Total waste weight	0.65 (0.01)	0.67 (0.01)	0.61 (0.03)	0	0.66 (0.02)	0.67 (0.01)	0.61 (0.02) **	2
Gutted weight	2.37 (0.04)	2.39 (0.04)	2.19 (0.09) *	0.2	2.41 (0.05)	2.38 (0.04)	2.28 (0.05) **	1
Deheaded weight	2.07 (0.03)	2.10 (0.03)	1.94 (0.08)	0.05	2.10 (0.04)	2.09 (0.03)	1.99 (0.04) *	1
Fillet weight	1.71 (0.03)	1.76 (0.03)	1.59 (0.07) **	0	1.76 (0.04)	1.72 (0.03)	1.67 (0.04)	1
Fat percentage	13.19 (0.27)	13.12 (0.31)	12.41 (0.84)	0.4	13.65 (0.39)	13.17 (0.28)	12.45 (0.4) *	4
Fillet colour	28.96 (0.04)	28.90 (0.05)	29.03 (0.12)	0.1	28.98 (0.06)	28.90 (0.04)	28.97 (0.06)	0.02
Gut weight	0.21 (0)	0.22 (0)	0.20 (0.01)	0.01	0.22 (0.01)	0.22 (0)	0.20 (0.01) **	3

\* Overall SNP  $p < 0.1$ ; \*\* overall SNP  $p < 0.05$ .

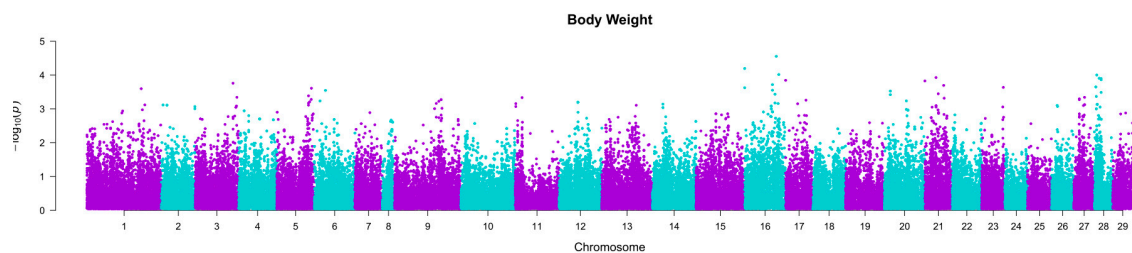
The SNP AX88270804 was significantly associated ( $p < 0.05$ ) with most of the fillet and waste traits, including a suggestive association with fat content ( $p < 0.1$ ). The adenine allele corresponds to higher trait means for the carcass weight and fatness traits. The SNP AX88141678 was associated with overall harvest weight, head weight and gutted weight ( $p < 0.05$ ). At this SNP, the adenine allele was also associated with higher trait means for the carcass and overall weight traits. The estimation of the additive genetic variation explained by the SNPs indicated that AX88270804 explained a small percentage of the overall variation in fillet traits (~1%), waste traits (2% to 3%) and fat percentage (4%). The SNP AX88141678 explained approximately 1% of the additive genetic variation in the weight-related traits (Table 2). To account for variation in the overall size of the fish when analysing component traits, Model (1) was preformed, including harvest weight as a covariate. In this analysis, most of the SNP-trait associations were no longer significant, but SNP AX88270804 showed an association with body waste weight and total waste weight.

### 2.3. QTL Region Characterization and Putative Gene Identification

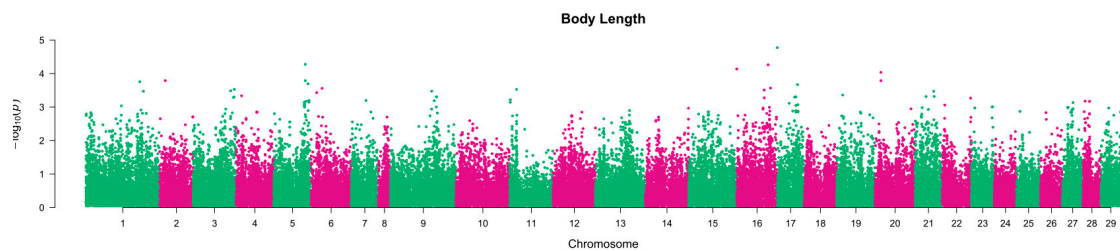
The corresponding flanking sequences for the two significant SNPs were aligned with the reference genome (assembly GCA\_000233375.4), and the putative genes proximal to the SNPs were

identified, indicating that the loci AX88141678 (chr. 5) and AX88270804 (chr. 16) were located within *MEP1A* (meprin A subunit beta-like) and *PCNT* (pericentrin), respectively. AX88270804 was located in an exon (non-synonymous), whereas AX88141678 was in a non-coding region (Table 2). The details of all SNPs tested in the current study are given in Table S1.

The main results of the GWA analysis in Population 1 are given in Tsai *et al.* [14]. However, due to the recent availability of a chromosome-anchored reference genome sequence assembly for Atlantic salmon (GCA\_000233375.4), we used BlastN to align the flanking sequence of the SNPs on the array with the assembly to identify their putative chromosome and position. This information was used to draw Manhattan plots to view the QTL regions from which the candidate SNPs were chosen (Figures 1 and 2).



**Figure 1.** The Manhattan plot of body weight in the GWAS of Population 1 [14]. The Bonferroni genome-wide significance threshold is  $p \sim 4.50 \times 10^{-7}$ .



**Figure 2.** The Manhattan plot of body length in the GWAS of Population 1 [14]. The Bonferroni genome-wide significance threshold is  $p \sim 4.50 \times 10^{-7}$ .

### 3. Discussion

Abundant SNPs discovered by modern sequencing technologies and bioinformatics tools have allowed us to better understand the association between genomic variation and production traits in aquatic species [27]. In a recent study, we applied a high density SNP array (~132 K) [6] to identify candidate markers associated with weight and length traits in a farmed salmon population measured at one year of age [14]. To test a subset of promising SNPs from the previous study in a different population, we successfully genotyped eight SNPs in a population of 1152 salmon with growth and harvest-related traits measured at three years of age. Two SNPs were found to be significantly associated with several growth and harvest traits in the second population, implying that these SNPs are linked to QTL with effects on growth at multiple stages of the salmon production cycle. For the remaining six SNPs where no significant association was detected, this may reflect false positives in the initial study or false negatives in the current study. Alternatively, SNPs may have specific lifecycle stage-specific effects on growth that were not observed in both studies due to the difference in age at which the salmon were measured (one and three years respectively). While only weight and length were measured in the GWAS [14], there were eight fillet- and carcass-related traits measured in the current study. Therefore, for the two SNPs that were validated in the current study, the use of these additional measurements helps to determine a more specific growth phenotype associated with the SNP effects. For example, the SNP AX88270804 was associated with fat percentage in the

current study, which indicates that the faster growth associated with the favourable allele also leads to increased fat content of the fish.

Alignment of the SNP flanking sequences with the Atlantic salmon reference genome predicted that AX88270804 was a synonymous exonic SNP within the *PCNT* gene and showed a significant association with several muscle and skeletal growth traits ( $p < 0.05$ ) in Population 2 (current study) and growth traits ( $p \sim 10^{-4}$ ) in Population 1 [14]. The SNP explained between 1% and 4% of the genetic variation in various harvest traits. In humans, the *PCNT* gene encodes the centrosome protein pericentrin, which contributes to the organisation of the mitotic spindle for the segregation of the chromosomes during cell division, thus influencing cell cycle progression. Mitotic centrosome dysfunction caused by pericentrin mutations can be expected to cause disturbances in cell division and is known to result in seriously stunted growth of the body and brain [28,29]. Interestingly, the SNP within the *PCNT* gene in salmon also has a suggestive association with fat percentage, explaining approximately 4% of the genetic variation. As expected, the allele associated with faster growth is also associated with increased fatness (Table 2). Major mutations in the *PCNT* gene in humans also affect adipocyte differentiation and can result in dyslipidemia as part of a wider insulin resistance syndrome. The fact that *PCNT* function is necessary for normal growth and lipid regulation in humans raises the possibility that further minor genetic variation within and around the gene may contribute to phenotypic variation in these traits. However, the role of the pericentrin in salmonid species has not yet been established.

The SNP AX88141678 was found in the intronic region of the *MEP1A* gene, which encodes meprin A subunit alpha. Meprins are zinc metalloendopeptidases that are predominantly found in kidney and intestinal brush border membranes in mammals and are known to play a role in protein metabolism [30]. Like *PCNT*, little is known about the function of *MEP1A* in Atlantic salmon, but interestingly, diet manipulation in another salmonid species (rainbow trout (*O. mykiss*)) has been shown to result in marked expression changes of *MEP1A* in the intestine [31]. In addition, *MEP1A* expression was shown to differ between domesticated and wild brook char (*Salvelinus fontinalis*) and its putative effect on growth factors was postulated to be the underlying mechanism for the higher expression in selected fish [32]. Therefore, while the association with growth traits may be due to variation in nearby candidate genes, the association of an SNP within the *MEP1A* gene and growth traits and its postulated functional connection to the growth traits raise the possibility that the causative effect underlying this association may be mediated via the *MEP1A* gene itself. It is worth noting that the genotype means for the SNP suggest an overdominance effect, which may explain why the additive variation explained is very small (Table 2).

Loci AX88141678 and AX88270804 were mapped to chr. 5 and chr. 16 using sire-based linkage mapping, respectively [6], and alignment with the reference genome assembly. A recent quantitative trait loci (QTL) mapping study by our group [20] in the same population as the current study showed that chr. 16 harbours loci affecting several growth traits with chromosome-wide significance in a sire-based analysis, although no QTL were detected on chr. 5. To date, there is a lack of consistency between the locations of the QTL affecting growth traits in different studies and commercial salmon populations [20,33–35]; therefore, the growth traits are considered to be regulated by population-specific and polygenic factors. Further, while the association between the *PCNT* and *MEP1A* candidate gene polymorphisms and growth-related phenotypes measured in two different populations of salmon is encouraging, the direction of the allelic effects between the two studies was generally not consistent (see Table S2). For both SNPs in the current study, fish carrying two copies of the adenine allele had better growth performance than other genotypes, whereas in Tsai *et al.* [14], this genotype was associated with lower weight and length values. This may be due to opposing effects in different lifecycles and environments (freshwater *versus* seawater). A genotype by environment interaction has been shown to be evident for the direction of association of individual SNPs (e.g., [36]). Alternatively, these SNPs may be marking QTL some distance away, and the relationship between marker and QTL may vary from population to population. The QTL regions

identified in the GWAS cover a relatively large region of the chromosomes (Figures 1 and 2). As such, while identifying chromosomal regions and putative genes harbouring variation contributing to growth phenotypes in salmon is of biological interest, it is unlikely that specific marker-assisted selection for these individual loci will be of high value, in particular for growth traits, which are directly measurable on the selection candidates themselves. This is particularly the case because genomic prediction using relatively few genome-wide markers can lead to very accurate prediction of breeding values for complex traits, such as growth (e.g., accuracy ~0.7 for juvenile weight and length in [14]). Therefore, genomic selection-based breeding schemes are likely to be increasingly utilised for the improvement of polygenic traits as genotyping technology becomes more affordable [14,37], especially for those traits with high economic value and that are difficult to be visualized (e.g., milk yield in dairy and fillet weight in fish).

## 4. Experimental Section

### 4.1. Animals

The GWAS used to identify the SNPs with putative association with growth in commercial salmon populations was based on the 2007 year group population of the Landcatch Natural Selection (LNS; Ormsary, UK) broodstock that were measured for weight and length at the end of the freshwater period (~1 year old; "Population 1") [14]. To test the candidate SNPs in a new population, 1152 individuals were randomly selected from a larger population ( $n \sim 5000$ ) comprising the 1999 year group of LNS broodstock that were measured for weight and other fillet traits at harvest ("Population 2"). The 1152 genotyped fish were across 191 full sibling families from 131 sires and 185 dams. The phenotypes were measured by LNS at harvest (approximately 3 years old), including overall harvest weight (kg), gutted weight (kg), deheaded weight (kg), fillet weight (kg), head weight (kg), gut weight (kg), body waste weight (kg) and total waste weight (kg), fat percentage (% as estimated using a Torry Fatmeter (Distell Ltd., Aberdeen, Scotland)) and fillet colour (assessed visually using the Roche SalmoFan scale (Hoffmann-La Roche, West Sussex, UK), ranging from 20 (Yellow) to 34 (Red)). The body waste weight was calculated as deheaded weight minus fillet weight (weight of vertebrae and caudal fin), and total waste weight was by head weight plus body waste weight. Details of the population and phenotype measurement are given in Tsai *et al.* and Peñaloza *et al.* [20,25]. An adipose fin tissue sample of each individual was clipped and retained for DNA extraction using DNeasy-96 tissue DNA extraction kits (Qiagen, Crawley, UK).

All animals were reared in accordance with all relevant national and EU legislation concerning health and welfare. Landcatch is an accredited participant in the RSPCA (Royal Society for the Prevention of Cruelty to Animals) Freedom Foods standard, the Scottish Salmon Producers Organization Code of Good Practice and the EU Code-EFABAR (<http://www.responsiblebreeding.eu/>) Code of Good Practice for Farm Animal Breeding and Reproduction Organizations.

### 4.2. SNP Selection and Genotyping

The candidate SNPs were selected based on two relevant studies [14,20]. Firstly, a GWA analysis was performed in Population 1 to select the candidate markers for genotyping [14], and a proportion of the SNPs surpassing a nominal significance ( $p \sim 10^{-3}$ ) were selected. Secondly, chromosome 20 was identified as containing loci affecting growth and fillet-related traits in Population 2 [20]. Therefore, two SNPs with nominally significant association with weight and length ( $p \sim 10^{-2}$ ) [14] from this QTL region were also included in the shortlist for further investigation. The details of candidate SNPs are given in Table S1. In total, sixteen candidate SNPs were selected for assay design and genotyping in Population 2, of which eight were successfully genotyped and showed segregation. Candidate SNP markers and their flanking sequences were provided to LGC Genomics (Herts, UK) for the design of "kompetitive allele-specific PCR (KASP)" assays (see KASP technique details at [38]) for genotyping with 1152 offspring in Population 2.

### 4.3. Statistical Analysis

#### 4.3.1. Heritability Estimation and SNP Associations

The heritability of the traits was calculated as described previously [20]. The simple animal model (Model (1)) was used to estimate the additive genetic effect of each SNP genotype ( $G$ ):

$$Y = \mu + G + A + e \quad (1)$$

where  $Y$  represents the observed phenotype,  $\mu$  is the overall mean of the trait,  $G$  is the fixed effect of the SNP genotype,  $A$  is the additive genetic effect and  $e$  is the residual error. For estimating heritability, the equivalent model was used, but without the SNP effect ( $G$ ) using the model:

$$h^2_a = \sigma_a^2 / \sigma_p^2 \quad (2)$$

where  $\sigma_a^2$  is the additive genetic variance and  $\sigma_p^2$  is the total phenotypic variance. The analysis was performed by ASReml 3.0 [39].

#### 4.3.2. Allelic Substitution Assessment

The allelic substitution effects of informative SNPs were estimated using Model (1) performed by ASReml 3.0 [39]. The SNP genotype was fitted as the fixed effect in the analysis. The additive effect of the candidate marker was calculated as the difference of the predicted phenotypic means of two homozygotes divided by two, which was given as  $(AA - BB)/2$ , and the dominance effect was  $AB - ((AA + BB)/2)$ , where the  $AB$  represents the predicted phenotypic means from heterozygote and  $AA$  or  $BB$  are from homozygote in the statistical analysis. The proportion of genetic variance due to SNP (PVE) was also estimated, by the following equation [40]:

$$PVE = [2pq(\alpha + \delta(q - p))^2] / V_A \quad (3)$$

where  $\alpha$  and  $\delta$  are the additive and dominance effect, respectively,  $p$  is the frequency of the most frequent allele,  $q$  is the frequency of the minor allele and  $V_A$  is the total additive genetic variance of the trait obtained when no SNP effects are included in the model.

### 4.4. Candidate Gene Identification

To identify candidate genes near the significant SNPs, the flanking sequence was aligned to the Atlantic salmon reference genome assembly (GCA\_000233375.4), and the corresponding genome contig and position of the SNPs were noted. Approximately 20 kb of sequence surrounding the SNPs were repeat masked (retrieved from [41]), and a BlastX analysis was used to detect putative genes within the vicinity of the SNPs.

## 5. Conclusions

In genome-wide association studies of complex and polygenic traits, the significant SNPs identified are likely to contain a mix of true associations and false positives. Therefore, verification of GWAS findings in a separate population is an important validation step, and SNP associations identified in more than one population are more likely to be reflecting real QTL. We identified two (out of eight successfully genotyped) SNPs that showed an association with growth traits in two different populations, and two different lifecycle stages, in Atlantic salmon. The SNPs are within the pericentrin and meprin alpha genes, which both have potentially relevant functional connections to the growth and harvest traits studied. Further investigation of these candidate genes may be merited to identify putative causative variation.



**Supplementary Materials:** Supplementary materials can be found at <http://www.mdpi.com/1422-0067/17/1/5/s1>.

**Acknowledgments:** The authors acknowledge funding from the Biotechnology and Biological Sciences Research Council (BBSRC) via Grant Number BB/H022007/1 and by Institute Strategic Funding Grants to the Roslin Institute (BB/J004235/1 and BB/J004324/1). Hsin Y. Tsai is supported by the Studying Funding from the Ministry of Education, Taiwan.

**Author Contributions:** The experiments were designed by Ross D. Houston and Steve C. Bishop. Tissue samples and phenotype measurements were provided by Alan E. Tinch, Derrick R. Guy and Alastair Hamilton. Data analyses were performed by Hsin Y. Tsai. The manuscript was written by Hsin Y. Tsai and Ross D. Houston. All authors approved and read the final manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

GWAS: genome-wide association study

MEP1A: meprin A subunit beta-like

PCNT: pericentrin

QTL: quantitative trait loci

SNP: single nucleotide polymorphism

## References

1. FAO. *Yearbook. Fishery and Aquaculture Statistics*; F.A.O.: Rome, Italy, 2010.
2. Yáñez, J.M.; Newman, S.; Houston, R.D. Genomics in aquaculture to better understand species biology and accelerate genetic progress. *Front. Genet.* **2015**, *6*, 128. [[PubMed](#)]
3. Bishop, S.C.; Woolliams, J.A. Genomics and disease resistance studies in livestock. *Livest. Sci.* **2014**, *166*, 190–198. [[CrossRef](#)] [[PubMed](#)]
4. Gjedrem, T.; Robinson, N.; Rye, M. The importance of selective breeding in aquaculture to meet future demands for animal protein: A review. *Aquaculture* **2012**, *350–353*, 117–129. [[CrossRef](#)]
5. Lien, S.; Gidskehaug, L.; Moen, T.; Hayes, B.J.; Berg, P.R.; Davidson, W.S.; Omholt, S.W.; Kent, M.P. A dense SNP-based linkage map for Atlantic salmon (*Salmo salar*) reveals extended chromosome homeologies and striking differences in sex-specific recombination patterns. *BMC Genom.* **2011**, *12*, 615. [[CrossRef](#)] [[PubMed](#)]
6. Houston, R.D.; Taggart, J.B.; Cézard, T.; Bekaert, M.; Lowe, N.R.; Downing, A.; Talbot, R.; Bishop, S.C.; Archibald, A.L.; Bron, J.E.; *et al.* Development and validation of a high density SNP genotyping array for Atlantic salmon (*Salmo salar*). *BMC Genom.* **2014**, *15*, 90.
7. Palti, Y.; Gao, G.; Liu, S.; Kent, M.P.; Lien, S.; Miller, M.R.; Rexroad, C.E.; Moen, T. The Development and Characterization of a 57K SNP Array for Rainbow Trout. *Mol. Ecol. Resour.* **2015**, *15*, 662–672. [[CrossRef](#)] [[PubMed](#)]
8. Sánchez, C.C.; Smith, T.P.L.; Wiedmann, R.T.; Vallejo, R.L.; Salem, M.; Yao, J.; Rexroad, C.E. Single nucleotide polymorphism discovery in rainbow trout by deep sequencing of a reduced representation library. *BMC Genom.* **2009**, *10*, 559. [[CrossRef](#)] [[PubMed](#)]
9. Everett, M.V.; Seeb, J.E. Detection and mapping of QTL for temperature tolerance and body size in Chinook salmon (*Oncorhynchus tshawytscha*) using genotyping by sequencing. *Evol. Appl.* **2014**, *7*, 480–492.
10. Gutierrez, A.P.; Yáñez, J.M.; Fukui, S.; Swift, B.; Davidson, W.S. Genome-Wide Association Study (GWAS) for Growth Rate and Age at Sexual Maturation in Atlantic Salmon (*Salmo salar*). *PLoS ONE* **2015**, *10*, e0119730. [[CrossRef](#)] [[PubMed](#)]
11. Sodeland, M.; Gaarder, M.; Moen, T.; Thomassen, M.; Kjøglum, S.; Kent, M.; Lien, S. Genome-wide association testing reveals quantitative trait loci for fillet texture and fat content in Atlantic salmon. *Aquaculture* **2013**, *408–409*, 169–174. [[CrossRef](#)]
12. Houston, R.D.; Bishop, S.C.; Guy, D.R.; Tinch, A.E.; Taggart, J.B.; Bron, J.E.; Downing, A.; Stear, M.J.; Gharbi, K.; Hamilton, A. Genome Wide Association Analysis for Resistance to Sea Lice in Atlantic Salmon: Application of a Dense SNP Array. In Proceedings of the 10th World Congress on Genetics Applied to Livestock Production, Vancouver, BC, Canada, 17–22 August 2014; pp. 10–12.

13. Correa, K.; Lhorente, J.; Lopez, M.; Bassini, L.; Naswa, S.; Deeb, N.; di Genova, A.; Maas, A.; Davidson, W.; Yáñez, J. Genome-wide association analysis reveals loci associated with resistance against *Piscirickettsia salmonis* in two Atlantic salmon (*Salmo salar* L.) chromosomes. *BMC Genom.* **2015**, *16*, 854. [[CrossRef](#)] [[PubMed](#)]
14. Tsai, H.Y.; Hamilton, A.; Tinch, A.E.; Guy, D.R.; Gharbi, K.; Stear, M.J.; Oswald, M.; Bishop, S.C.; Houston, R.D. Genome wide association and genomic prediction for growth traits in juvenile farmed Atlantic salmon using a high density SNP array. *BMC Genom.* **2015**, *16*, 969. [[CrossRef](#)] [[PubMed](#)]
15. Macqueen, D.J.; Johnston, I.A. A well-constrained estimate for the timing of the salmonid whole genome duplication reveals major decoupling from species diversification. *Proc. R. Soc. B* **2014**, *281*, 1778. [[CrossRef](#)] [[PubMed](#)]
16. Berthelot, C.; Brunet, F.; Chalopin, D.; Juanchich, A.; Bernard, M.; Noël, B.; Bento, P.; Da Silva, C.; Labadie, K.; Alberti, A.; *et al.* The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. *Nat. Commun.* **2014**, *5*, 3657.
17. Gidskehaug, L.; Kent, M.; Hayes, B.J.; Lien, S. Genotype calling and mapping of multisite variants using an Atlantic salmon iSelect SNP array. *Bioinformatics* **2011**, *27*, 303–310. [[CrossRef](#)] [[PubMed](#)]
18. Wolfe, K.H. Yesterday's polyploids and the mystery of diploidization. *Nat. Rev. Genet.* **2001**, *2*, 333–341. [[CrossRef](#)] [[PubMed](#)]
19. Tsai, H.Y.; Hamilton, A.; Guy, D.R.; Houston, R.D. Single nucleotide polymorphisms in the insulin-like growth factor 1 (IGF1) gene are associated with growth-related traits in farmed Atlantic salmon. *Anim. Genet.* **2014**, *45*, 709–715. [[CrossRef](#)] [[PubMed](#)]
20. Tsai, H.Y.; Hamilton, A.; Guy, D.R.; Tinch, A.E.; Bishop, S.C.; Houston, R.D. The genetic architecture of growth and fillet traits in farmed Atlantic salmon (*Salmo salar*). *BMC Genet.* **2015**, *16*, 51. [[CrossRef](#)] [[PubMed](#)]
21. Powell, J.; White, I.; Guy, D.; Brotherstone, S. Genetic parameters of production traits in Atlantic salmon (*Salmo salar*). *Aquaculture* **2008**, *274*, 225–231. [[CrossRef](#)]
22. Mackay, T.F.C.; Stone, E.A.; Ayroles, J.F. The genetics of quantitative traits: challenges and prospects. *Nat. Rev. Genet.* **2009**, *10*, 565–577. [[CrossRef](#)] [[PubMed](#)]
23. Warde-Farley, D.; Donaldson, S.L.; Comes, O.; Zuberi, K.; Badrawi, R.; Chao, P.; Franz, M.; Grouios, C.; Kazi, F.; Lopes, C.T.; *et al.* The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res.* **2010**, *38*, W214–W220.
24. Moghadam, H.K.; Poissant, J.; Fotherby, H.; Haidle, L.; Ferguson, M.M.; Danzmann, R.G. Quantitative trait loci for body weight, condition factor and age at sexual maturation in Arctic charr (*Salvelinus alpinus*): comparative analysis with rainbow trout (*Oncorhynchus mykiss*) and Atlantic salmon (*Salmo salar*). *Mol. Genet. Genom.* **2007**, *277*, 647–661. [[CrossRef](#)] [[PubMed](#)]
25. Peñaloza, C.; Hamilton, A.; Guy, D.R.; Bishop, S.C.; Houston, R.D. A SNP in the 5' flanking region of the myostatin-1b gene is associated with harvest traits in Atlantic salmon (*Salmo salar*). *BMC Genet.* **2013**, *14*, 112. [[CrossRef](#)] [[PubMed](#)]
26. Davidson, W.S.; Koop, B.F.; Jones, S.J.M.; Iturra, P.; Vidal, R.; Maass, A.; Jonassen, I.; Lien, S.; Omholt, S.W. Sequencing the genome of the Atlantic salmon (*Salmo salar*). *Genome Biol.* **2010**, *11*, 403. [[PubMed](#)]
27. Yáñez, J.M.; Houston, R.D.; Newman, S. Genetics and genomics of disease resistance in salmonid species. *Front. Genet.* **2014**, *5*, 415. [[PubMed](#)]
28. Rauch, A.; Thiel, C.T.; Schindler, D.; Wick, U.; Crow, Y.J.; Ekici, A.B.; van Essen, A.J.; Goecke, T.O.; Al-Gazali, L.; Chrzanowska, K.H.; *et al.* Mutations in the Pericentrin (PCNT) Gene Cause Primordial Dwarfism. *Science* **2008**, *319*, 816–819.
29. Müller, T.; Rumpel, E.; Hradetzky, S.; Bollig, F.; Wegner, H.; Blumenthal, A.; Greinacher, A.; Endlich, K.; Endlich, N. Non-muscle myosin IIA is required for the development of the zebrafish glomerulus. *Kidney Int.* **2011**, *80*, 1055–1063. [[CrossRef](#)] [[PubMed](#)]
30. Sterchi, E.E.; Stöcker, W.; Bond, J.S. Meprins, membrane-bound and secreted astacin metalloproteinases. *Mol. Asp. Med.* **2010**, *29*, 309–328. [[CrossRef](#)] [[PubMed](#)]
31. Kirchner, S.; McDaniel, N.K.; Sugiura, S.H.; Soteropoulos, P.; Tian, B.; Fletcher, J.W.; Ferraris, R.P. Salmonid microarrays identify intestinal genes that reliably monitor P deficiency in rainbow trout aquaculture. *Anim. Genet.* **2007**, *38*, 319–331. [[CrossRef](#)] [[PubMed](#)]

32. Sauvage, C.; Derôme, N.; Normandeau, E.; St-Cyr, J.; Audet, C.; Bernatchez, L. Fast transcriptional responses to domestication in the brook charr *Salvelinus fontinalis*. *Genetics* **2010**, *185*, 105–112. [[CrossRef](#)] [[PubMed](#)]
33. Gutierrez, A.P.; Lubieniecki, K.P.; Davidson, E.A.; Lien, S.; Kent, M.P.; Fukui, S.; Withler, R.E.; Swift, B.; Davidson, W.S. Genetic mapping of quantitative trait loci (QTL) for body-weight in Atlantic salmon (*Salmo salar*) using a 6.5K SNP array. *Aquaculture* **2012**, *358–359*, 61–70. [[CrossRef](#)]
34. Baranski, M.; Moen, T.; Våge, D.I. Mapping of quantitative trait loci for flesh colour and growth traits in Atlantic salmon (*Salmo salar*). *Genet. Sel. Evol.* **2010**, *42*, 17. [[CrossRef](#)] [[PubMed](#)]
35. 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. Detection of QTL affecting harvest traits in a commercial Atlantic salmon population. *Anim. Genet.* **2009**, *40*, 753–755. [[CrossRef](#)] [[PubMed](#)]
36. Bastiaansen, J.W.M.; Bovenhuis, H.; Lopes, M.S.; Silv, F.F.; Megens, H.J.; Calu, M.P.L. SNP Effects Depend on Genetic and Environmental Context. In Proceedings of the 10th World Congress on Genetics Applied to Livestock Production, Vancouver, BC, Canada, 17–22 August 2014; pp. 356–362.
37. Odegård, J.; Moen, T.; Santi, N.; Korsvoll, S.A.; Kjøglum, S.; Meuwissen, T.H.E. Genomic prediction in an admixed population of Atlantic salmon (*Salmo salar*). *Front. Genet.* **2014**, *5*, 402. [[PubMed](#)]
38. LGC Genomics KASP (kompetitive allele-specific PCR) assays. Available online: <http://www.lgcgroup.com/products/kasp-genotyping-chemistry/kasp-technicalresources/#.VWC3TU9VhXt> (accessed on 1 November 2015).
39. Gilmour, A.R.; Gogel, B.J.; Cullis, B.R.; Thompson, R. *ASReml User Guide*, 4th ed.; VSN International Ltd: Hemel Hempstead, UK, 2014.
40. Falconer, D.S.; Mackay, T.F.C. *Introduction to Quantitative Genetics*, 4th ed.; Longmans Green: Harlow, Essex, UK, 1996.
41. Centre for biomedical research, RepeatMasker, University of Victoria. Available online: [http://lucy.ceh.uvic.ca/repeatmasker/cbr\\_repeatmasker.py](http://lucy.ceh.uvic.ca/repeatmasker/cbr_repeatmasker.py) (accessed on 30 August 2015).



© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).