TEAD3, implicated by association to grilseing in Atlantic salmon

Citation for published version:
Christensen, KA, Gutierrez Silva, A, Lubieniecki, KP & Davidson, WS 2017, 'TEAD3, implicated by association to grilseing in Atlantic salmon' Aquaculture, vol. 479, pp. 571-578. DOI:
10.1016/j.aquaculture.2017.06.026

Digital Object Identifier (DOI):
10.1016/j.aquaculture.2017.06.026

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published in:
Aquaculture

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 providing details, and we will remove access to the work immediately and investigate your claim.
1. Introduction

New Brunswick Atlantic salmon farming generated an estimated $250 million CAD in 2002 (McClure et al., 2007). Of this, between $11–$24 million CAD was estimated to be lost due to early maturation (known as grilsing in Atlantic salmon) (McClure et al., 2007). Typically, the loss in revenue from early maturation has to do with the flesh quality degradation that accompanies maturation (Aksnes et al., 1986). Understanding why some fish mature earlier than others or how to control the timing of maturity is likely a balance between male mating strategies and genetic diversity (Saunders and Schom, 1985). The diversity maintained by multiple life-histories may also make Atlantic salmon more adaptable to different environments (e.g. low-flow and high-flow streams, reviewed in (Fleming, 1996)) and potentially explains part of the reason why they have been successful at colonizing so many locations.

The mechanism maintaining variation in age at maturity is likely complex, but may reflect a balance between male mating strategies (Fleming, 1996; Gross, 1985). Larger males tend to have an aggressive mating style, fighting for mating opportunities; while the smaller males often use a strategy of wait and sneak-in (Gross, 1985). In environments where there are many large males, a sneak fertilization may be a highly effective strategy of reproduction, but when there are many small fish trying to sneak-in, fighting may be more effective (Fleming, 1996; Gross, 1985).

Genetic variation associated with age at maturity has been seen in quantitative trait loci (QTL) and genome wide association (GWA) studies (Gutierrez et al., 2015, 2014; Johnston et al., 2014). More recently, two research groups have posited that the vestigial-like family member 3 (Vgll3) gene is associated with early maturation, where there are many large males, a sneak fertilization may be a highly effective strategy of reproduction, but when there are many small fish trying to sneak-in, fighting may be more effective (Fleming, 1996; Gross, 1985).

The homologs of the Vgll3 gene have been studied extensively in other organisms. In Drosophila, vestigial was found to be involved in wing formation (Kim et al., 1996). Later, it was found that another protein, scalloped (a homolog of transcription enhancer factor-1) in vertebrates, was needed in wing formation and that the proteins of these two genes interacted to do so (P. Paumard-Rigal, 1998; Simmonds et al., 1998). Simmonds et al. (1998), suggested that vestigial was a tissue-specific co-factor of the transcription factor scalloped.

In vertebrates, three vestigial homologs were characterized that all...
contain a transcription enhancer factor-1 (TEF-1 commonly referred to as TEAD genes) interacting domain (Maeda et al., 2002). The vestigial homologs appear to form heterodimers with TEF-1 (Maeda et al., 2002). This interaction appeared to play a role in skeletal muscle differentiation in a mouse model (Maeda et al., 2002). Again, suggesting that the vestigial homologs act as transcriptional co-factors. It has also been shown that vertebrate TEF-1 shares enough similarity with Drosophila scalloped, that TEF-1 can be substituted for scalloped and still have wingblade development (Deshpande et al., 1997). Similarly, human vestigial homologs can also be substituted for vestigial in Drosophila and wing development will still continue (Vaudin et al., 1999).

Early estimates of the number of TEF-1 genes in mammalian genomes was three to five (Kaneko and DePamphilis, 1998). It was shown that each gene contains a TEA DNA (TEAD) binding domain, with nearly identical amino acid sequence. This allows each of them to bind conserved non-coding DNA elements (Kaneko and DePamphilis, 1998). Expression of these genes varies from tissues, but nearly every tissue expresses at least one TEAD gene. These observations led researchers to conclude that in order for TEAD genes to differentially regulate genes in separate tissues, there must a mechanism for ubiquitous TEAD proteins to only bind a subset of the possible conserved non-coding DNA elements (Kaneko and DePamphilis, 1998). All these pieces of evidence support the idea that TEAD genes and vestigial homologs (Vgll1–4) are involved in regulating transcription in both vertebrates and invertebrates, and that vestigial homologs are tissue-specific co-factors for the more general TEF-1.

More recently, four TEAD genes and four Vgll genes have been characterized in vertebrates (reviewed in Pobbat and Hong, 2013). Along with the TEA DNA binding domain, each of the TEAD genes also contains a transactivation domain, allowing them to interact with several coactivators, including: yes-associated protein (YAP/TAZ), Myb-binding protein 1A (p160), and Vgll (Pobbat and Hong, 2013). TEAD1 and TEAD2 appear to be especially important for differentiation in cardiac muscle, TEAD4 is vital for placental development, and major functions for TEAD3 have yet to be elucidated (Pobbat and Hong, 2013). In mammals, Vgll1 and Vgll4 are expressed in multiple tissues, Vgll2 is mainly expressed in skeletal muscle, and Vgll3 expression is found almost entirely in the placenta (Pobbat and Hong, 2013).

An interesting feature of the interplay between TEAD and Vgll genes is the ability of TEAD proteins to be able to bind multiple coactivators. In the Hippo signaling pathway, involved in cell proliferation and differentiation, TEAD proteins interact with YAP/TAZ to stimulate cell growth and proliferation (reviewed in Ashraf and Pervaiz, 2015). YAP/TAZ may be activated by external stimuli such as reactive oxygen species (Ashraf and Pervaiz, 2015). Competition for the TEAD transactivation domain, may create a dynamic interplay between YAP/TAZ and Vgll. Variation in any of these genes may alter the dynamic between YAP/TAZ and Vgll; this is highlighted in various cancers (Ashraf and Pervaiz, 2015).

As previously stated, genetic variation in Vgll3 in Atlantic salmon appears to partially control the timing of when a salmon will return to fresh water (age at maturity). Based on previous research, a likely hypothesis is that Vgll3 does so by interacting with a TEAD protein. The genetic variation in Vgll3, associated with age at maturation, (Ayillon et al., 2015; Barson et al., 2015) is not found in the Vg_Tdu super family domain of Atlantic salmon's Vgll3-based on a comparison of the Vgll3 sequence to the conserved domain database (Marchler-Bauer et al., 2015). This suggests that the variation does not disrupt binding of the TEAD and Vgll3 genes.

We suspected that if variation in the Vgll3 gene could influence age at maturity in Atlantic salmon, then variation in a Vgll3 interacting gene may also alter timing of maturation. We used samples and datasets from Gutierrez et al. (2014) and Gutierrez et al. (2015), first to determine if Vgll3 was responsible for any of the QTL or GWA's in these studies. Next, we used these datasets and new Vgll3 genotype information to identify possible interacting QTL. The Atlantic salmon genome was used to identify if any of the markers near QTL were physically linked to binding partners of Vgll3.

2. Materials and methods

2.1. Population and dataset

Five Atlantic salmon families utilized by Gutierrez et al. (2014) were used for QTL analysis, GWA analysis, and association testing. These families are from the Mowi strain and were bred for Mainstream Canada (now Cermaq) as part of their broodstock program (Gutierrez et al., 2014). Fertilization occurred in late 2005, the fry from these families were pooled in February 2006, and phenotypes were scored as late as 2009 (Gutierrez et al., 2014). An additional 160 individuals (80 Grilse and 80 Non-Grilse) and 192 parents from the 2005 broodstock year were utilized in a GWA analysis. These individuals were originally a part of a GWA study conducted by Gutierrez et al. (2015). Genotype information was also collected by Gutierrez et al. (2014) and Gutierrez et al. (2015).

2.2. Genotyping Vgll3

Two genotyping assays were used to interrogate the Vgll3 genotypes of all of the individuals (Fig. 1). The preliminary assay utilized restriction fragment length polymorphisms (RFLP) to determine if variation existed at the same loci found in previous research (Ayillon et al., 2015; Barson et al., 2015). Once both variants were identified in the population, a TaqMan assay was used for all of the samples.

For the RFLP assay, the following reagents were used to amplify two segments of the Vgll3 gene on chromosome 25 of the Atlantic salmon genome in a 10 μL reaction:

<table>
<thead>
<tr>
<th>Assay Type</th>
<th>Primer/Probe Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>200 nM primer</td>
<td>Amino Acid 54 Assay</td>
</tr>
<tr>
<td>Forward Primer: TGA TCG TTG TGG TGC TGA CTA</td>
<td></td>
</tr>
<tr>
<td>Reverse Primer: AGC CCT GGA GAA ATG CTC ATC</td>
<td></td>
</tr>
<tr>
<td>Amino Acid 323 Assay</td>
<td></td>
</tr>
<tr>
<td>Forward Primer: TCC TGC CTG GTG TGA AGA CT</td>
<td></td>
</tr>
<tr>
<td>Reverse Primer: AGA AAC GTA GAC ATG GGA GAG AG</td>
<td></td>
</tr>
<tr>
<td>1 × PCR Buffer with Mg2+</td>
<td></td>
</tr>
<tr>
<td>200 μM dNTP Mix</td>
<td></td>
</tr>
<tr>
<td>1 U Taq DNA Polymerase</td>
<td></td>
</tr>
</tbody>
</table>

An initial denaturation of 95 °C for 5 min was followed by 35 rounds of 95 °C for 30 s, 62.5 °C for 30 s, and 72 °C for 30 s. This was followed by a final extension at 72 °C for 10 min. The 10 μL PCR product was then digested either using NalII (amino acid 54) or AluI (amino acid 323), for 4 h using 5 units per reaction at 37 °C. This was followed by a heat-inactivation of 80 °C for 30 min. The reaction was then loaded onto a 1.5% agarose gel and ran for 40 min at 110 V and scored for markers.

The same amino acid positions of the Vgll3 gene were verified using a TaqMan assay. For amino acid position 54 of the Vgll3 gene on chromosome 25 of the Atlantic Salmon, the following primers were used to amplify the segment around the amino acid and the following probes were used to verify the allele at that position:

<table>
<thead>
<tr>
<th>Assay Type</th>
<th>Primer/Probe Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward Primer: CTC CCT CTC TCT CGG TCT ATT T</td>
<td></td>
</tr>
<tr>
<td>Reverse Primer: CGC TGC TGT TGC TGT CT</td>
<td></td>
</tr>
<tr>
<td>Probe (C): /5HEX/CAGC + A + C + GGAG + CA/3IAkFQ/</td>
<td></td>
</tr>
<tr>
<td>Probe (T): /56-FAM/AC + AGC + A + T + GGA + GC/3IAkFQ/</td>
<td></td>
</tr>
</tbody>
</table>

For amino acid position 323 of the Vgll3 gene, the following primers and probes were used:

<table>
<thead>
<tr>
<th>Assay Type</th>
<th>Primer/ Probe Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward Primer: AGA GCC CAG GGA TAC ACA</td>
<td></td>
</tr>
</tbody>
</table>
Reverse Primer: GAT ATC CAT GGG TGT GTG TAG AG
Probe (G): /5HEX/TG + TGA + A + G + ACA + GAGG/3IABkFQ/
Probe (C): /56-FAM/TGTGA + A + C + ACA + G + AGG/3IABkFQ/

PrimeTime Gene Expression Master Mix (1×) was used with the primers (500 nM), and probes (150 nM) to amplify DNA samples (~25 ng) using an initial denaturation period of 3 min at 95 °C and 40 cycles of denaturation at 95 °C for 5 s and annealing/extensions at 60 °C for 30 s. The genotypes for amino acid position 54 and position 323 were nearly identical (data not shown) and so only the results from positions 54 were reported.

2.3. QTL and GWA analyses

Using custom Perl scripts and LepMap2 (Rastas et al., 2015), Vgll3 was added to the genetic map produced by (Lien et al., 2011) and used by Gutierrez et al. (2014). Then genotype and map information was converted to GridQTL (Seaton et al., 2006) format using Perl scripts and a sib-pair analysis was ran for the five families. The sib-pair analysis was ran either using sex as a fixed-effect or not and with Vgll3 as a cofactor or not. A chromosome-wide significance level was chosen for the QTL analysis instead of an experiment-wide significance level. The lower threshold was chosen because we were interested in specific loci and the corrections for multiple testing would likely be too conservative at the experiment-wide significance level.

Similar to the QTL analysis, Perl scripts were used to convert genotype, pedigree, and phenotype information from Gutierrez et al. (2014) and Lien et al. (2011) to a format accepted by GenABEL (Aulchenko et al., 2007) in the R environment (R Core Team, 2015). In GenABEL, problematic markers were removed using the function check.markers. Next, kinship was estimated by the ibs function with the weight variable set to freq. This information was used to estimate a polygenic model of age at maturity. Then a GRAMMAS genome wide analysis was performed using the polygenic relationship model with a permutation test of 1000 to identify the significance level. This model included sex as a factor influencing age at maturity and either included Vgll3 information or this was left out to see if Vgll3 was influencing other candidate associations.

2.4. Genotyping TEAD3

To identify variants in the TEAD3 gene, the sequence of this gene was downloaded from the NCBI (XM_014133832.1), and was aligned to an Atlantic salmon RNA-seq dataset from the NCBI SRA database (SRR1422850); with the max target sequence set to 5000. These alignments were downloaded as a SAM file and visualized with IGV (Robinson et al., 2011; Thorvaldsdóttir et al., 2013). The nucleotide position 2306 of XM_014133832.1 appeared to be variable (C vs. G). An allele-specific PCR assay was designed for this location (Fig. 2).

Reverse Primer: TCA TGC TAT GCC CTT GCT GT
Forward Primers: 2306(C) GTG CCT TTC TCC AAA GAC ATG
2306(G) GTG CCT TTC TCC AAA GAC ATC
1× CoralLoad (Qiagen)
1× PCR Buffer with Mg2+ (Applied Biological Materials Inc.)
200 μM dNTP Mix (Applied Biological Materials Inc.)
1 U Taq DNA Polymerase (Applied Biological Materials Inc.)

An initial denaturation of 95 °C for 5 min was followed by 30 rounds of 95 °C for 30 s, 60.5 °C for 30 s, and 72 °C for 30 s. This was followed by a final extension at 72 °C for 1 min. The reaction was then loaded.
onto a 1.5% agarose gel and electrophoresed for 40 min at 110 V. The genotype information was collected for the five families and the 80 grilse and 80 non-grilse individuals.

The genotypes at the 2306 position of the TEAD3 gene were verified using a TaqMan assay.

Forward Primer: GAT GGC AAC CAA GTG CCT
Reverse Primer: CAG AGT GGC TGG CAA CTG
Probe (G): TCC AAA GAC ATG TAT TGA T
Probe (C): TCC AAA GAC ATC TAT TGA T

TaqMan® Genotyping Master Mix (1×) was used with TaqMan assay solution (1×) to amplify DNA samples (~25 ng) using an initial denaturation period of 10 min at 95 °C and 40 cycles of denaturation at 95 °C for 15 s and annealing/extension at 60 °C for 60 s.

2.5. Statistical analysis

A custom Perl script was written to perform a permutation test because many of the expected values calculated for a chi-square analysis were below five (i.e., it violates the assumptions of the chi-square analysis). For this permutation test: the allele frequencies were determined from the observed data, two random distributions were then created from the allele frequency (and the observed total count), and a chi-square value was calculated by comparing these two random distributions. The chi-square value calculation was one permutation, and represents an expected value if the genotypes were randomly distributed for both categories (i.e., there was not an association with the genotypes and categories).

The chi-square value calculation was repeated 20,000 times (2 million for the combined genotype analysis) in order to find the distribution of expected chi-square values if the data was randomly distributed based on the observed allele frequencies and the observed total count. The observed chi-square value was then calculated and compared to the distribution. A p-value was assigned to the comparison based on the percentage of random chi-square values that had a greater chi-square value than the observed chi-square value (one-tailed).

Six comparisons were made in this manner. The male grilse were compared to male non-grilse for the Vgll3, TEAD3, and combined genotypes. There were no female grilse, so a comparison was performed between the 2 sea-winter and 3 sea-winter females for all genotypes. The 80 grilse and 80 non-grilse had an infrequent minor allele frequency (~8%) and were not analyzed for this reason.

In order to test if there was a significant interaction between the TEAD3 and Vgll3 loci, a logistic regression was performed in PLINK (version 1.9) (Purcell et al., 2007) using the –logistic interaction and –condition commands. To use the logistic regression analysis, maturation was treated as a binary trait. A simple logistic regression analysis, without covariates and interaction terms, was also performed to verify that treating the trait as binary, gave similar results as the permutation tests.

3. Results

When evaluating five Atlantic salmon families for QTLs, sex was included as a fixed-effect and Vgll3 genotype information was used as a co-factor. The result of this analysis can be seen in Fig. 3. Four QTL were significant at the chromosome-wide significance level. On chromosome 25, the most significant marker was near 53 cM and not the Vgll3 marker, which is at ~50 cM. When Vgll3 is removed as a co-factor, there is no change in position of the most significant marker. These results are different from Gutierrez et al. (2014), likely due to the inclusion of all markers instead of the subset of markers that were polymorphic in all families used in that study.

When using all of the datasets for a GWA analysis (Fig. 4), again the most significant marker on chromosome 25 was not Vgll3 (position 28,653,293), but a marker near it (position 27,031,877). It should be noted that Vgll3 was removed during quality control, but if quality control was not used, Vgll3 still was not the most significant marker. When the Vgll3 genotype was used in the GRAMMAS model, markers on
chromosome 25 were no longer significantly associated with age at maturity. The remaining significant markers appeared anomalous since surrounding markers did not show any association with age at maturity. Association with surrounding markers might be expected because of linkage disequilibrium.

Of the four QTL, only the one on chromosome 12 was located near a potential binding partner of Vgll3. This QTL was located at 115 cM and the marker located at this position aligned to position 75,694,886 on chromosome 12. TEAD3 meanwhile is located at position 79,047,649, which is closer to markers near 117 cM (78,652,735 to 79,202,674) of the genetic map (Fig. 3). This association was not seen in the GWA analysis. Johnston et al. (2014), found a GWA relatively near this location (chromosome 12, position 61,511,101) for sea-winter age. Pedersen et al. (2013), found a QTL on chromosome 12 for parr maturation, which was the most significant QTL for that trait.

To see if TEAD3 was associated with age at maturity, the five families were genotyped for a SNP in TEAD3. Using a permutation test, the male genotypes were significantly different ($p = 0.046$) between grilse and non-grilse individuals; suggesting that there is an association between this genotype and age at maturity in males. The genotypes for the various phenotype categories are shown in Figs. 5, 6, and 7 for the five families genotyped for Vgll3, TEAD3, and for both genotypes. Allele frequencies for Vgll3 and TEAD3 are shown in Fig. S1.

In Fig. 5, it can be seen that 3% of the grilse were homozygous for the Vgll3 “CC” genotype, while 29% of the 2SW males had the “CC” genotype. In Fig. 7, it can be seen that 55% of the male grilse have a “CT × GG” haplotype, which is 36% more than the “CT × GG” haplotype seen in 2SW males. The difference between the two categories appears greater than that between the categories when considering a single gene (Figs. 5 and 6). The $p$-value for the combined genotypes is also smaller for the combined analysis, and may suggest that the genotypes are more predictive in combination.

In order to determine if there was a significant interaction between the two loci, a logistic regression analysis was performed and included TEAD3 as a cofactor and an interaction coefficient between the TEAD3 and Vgll3 loci. The interaction term was not significant ($p = 0.95$) and suggests that the interaction does not affect the phenotype more than would be expected by the linear addition of each loci independently. For this analysis, the phenotype was treated as a binary trait. To determine if this would influence the association between the phenotype and genotypes, another regression analysis was performed without cofactor or interaction terms. Only Vgll3 was significantly associated in the logistic regression analysis ($p = 0.013$), but TEAD3 had a low $p$-value ($p = 0.064$) similar to the permutation test.

4. Discussion

While researching the mechanisms governing age at maturation has fairly apparent applications in aquaculture (i.e. marker assisted selection), less obvious benefits may be to conservation efforts. As described in 1. Introduction, various ages at maturity may benefit a species by preserving genetic diversity that would otherwise be lost due to environmental constraints (e.g. low-flow and high-flow streams). Conservation efforts which monitor age at maturity may be more
effective than otherwise for this reason. Genotyping for markers related to age at maturity offers a convenient and cost-effective means of this type of monitoring; as well as a means for selection against undesirable maturation ages in aquaculture.

With the discovery that Vgll3 was associated with age at maturity, the question of how it might do so now remains to be discovered. Based on previous research on vestigial and other Vgll genes in other vertebrates, a logical hypothesis is that it does so by interacting with a TEAD gene. Here we tested if variation in TEAD genes might also be involved with determining age at maturity, and we found evidence suggesting that TEAD3 is involved.

Previous evidence that Vgll and TEAD proteins physically interact, implicates transcription initiation from the Vgll3/TEAD3 putative complex playing a role in determining age at maturation. There are some preliminary studies done with pigs that may illuminate what types of genes might be activated by TEAD3 in Atlantic salmon. Evidence from several quantitative trait loci (QTL) analyses, led researchers to suggest that the accumulation of androstenone in pigs was related to the TEAD3 gene (Lee et al., 2005; Quintanilla et al., 2003; Robic et al., 2012, 2011). Because of the conservation that has been seen with TEAD genes (i.e. human TEAD protein substitution in Drosophila still producing wings), one could speculate that TEAD3 may play
a similar role in Atlantic salmon. With the generation of these hypotheses, an avenue of research can now be explored to further refine our understanding of age at maturation.

An interesting note about the position of the QTL and GWA analyses was that the Vgll3 marker was not the most significant marker in the datasets, even though it was close to the most significant marker on chromosome 25. Phenotyping and genotyping errors may explain the position discrepancy. Another note of interest is that the TEAD3 genotype was unlikely to be the causative variant since it was randomly chosen based on sequence alignments that were unrelated to age at maturation rather than from sequences from individuals with variation in the age at maturation phenotype. It is more likely that this particular variant is in linkage disequilibrium with the causative mutation.

A few lines of investigation that may be fruitful to pursue in the future include: 1) confirming Vgll3 and TEAD3 interact, 2) identifying in which tissues they occur, and 3) performing QTL analyses in other populations to understand if TEAD3 influences age at maturation in other populations.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.aquaculture.2017.06.026.

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. We would like to thank the anonymous reviewer for their helpful comments and suggestions.

References


Vaudin, P., Delanoue, R., Davidson, I., Silber, J., Zider, A., 1999. TONDU (TDU), a novel human protein related to the product of vestigial (vg) gene of Drosophila melanogaster interacts with vertebrate TEF factors and substitutes for Vg function in wing forma-

Development 126, 4807–4816.