Life-stage associated remodeling of lipid metabolism regulation in Atlantic salmon

Gareth Gillard[1]*, Thomas N. Harvey[2]*, Arne Gjuvsland[1], Yang Jin[3], Magny Thomassen[4], Sigbjørn Lien[2], Michael Leaver[5], Jacob S. Torgersen[6], Torgeir R. Hvidsten[1], Jon Olav Vik[2]†, Simen R. Sandve[2]†

* shared first authors

† shared corresponding authors

Affiliations:

[1] *Faculty of Chemistry, Biotechnology and Food Sciences, Norwegian University of Life Sciences, NO-1432 Ås, Norway*

[2] *Centre for Integrative Genetics (CIGENE), Department of Animal and Aquacultural Sciences, Faculty of Biosciences, Norwegian University of Life Sciences, NO-1432 Ås, Norway*

[3] *Department of Biology, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway.*

[4] *Department of Animal and Aquacultural Sciences, Faculty of Biosciences, Norwegian University of Life Sciences, NO-1432 Ås, Norway*

[5] *Institute of Aquaculture, School of Natural Sciences, University of Stirling, Pathfoot Building, Stirling FK9 4LA, Scotland, UK*

[6] *AquaGen AS, NO-1432 Ås, Norway*

Keywords: *transcriptomics, adaptation, fish, life-stage, metabolism*

Corresponding authors:

Simen Rød Sandve, simen.sandve@nmbu.no, +47 94870082, Jon Olav Vik, jonovik@gmail.com, +47 45882998

Running title: Adaptive remodeling of lipid metabolism

Abstract

Atlantic salmon migrates from rivers to sea to feed, grow and develop gonads before returning to spawn in freshwater. The transition to marine habitats is associated with dramatic changes in the environment, including water salinity, exposure to pathogens, and shift in dietary lipid availability. Many changes in physiology and metabolism occur across this life-stage transition, but little is known about the molecular nature of these changes.

Here we use a long term feeding experiment to study transcriptional regulation of lipid metabolism in Atlantic salmon gut and liver in both fresh- and saltwater. We find that lipid metabolism becomes significantly less plastic to differences in dietary lipid composition when salmon transitions to saltwater and experiences increased dietary lipid availability. Expression of genes in liver relating to lipogenesis and lipid transport decrease overall and become less responsive to diet, while genes for lipid uptake in gut become more highly expressed. Finally, analyses of evolutionary consequences of the salmonid specific whole-genome duplication on lipid metabolism reveals several pathways with significantly different ($p<0.05$) duplicate retention or duplicate regulatory conservation. We also find a limited number of cases where the whole genome duplication has resulted in an increased gene dosage.

In conclusion, we find variable and pathway-specific effects of the salmonid genome duplication on lipid metabolism genes. A clear life-stage associated shift in lipid metabolism regulation is evident, and we hypothesize this to be, at least partly, driven by non-dietary factors such as the preparatory remodeling of gene regulation and physiology prior to sea migration.

Introduction

Atlantic salmon lives a 'double life'. It starts its life in rivers, before transforming its physiology and behavior and migrating to sea to grow and accumulate resources for reproduction. This shift in environment requires preparatory remodeling of physiology prior to sea migration (referred to as smoltification), which encompasses a suite of coordinately regulated processes involving hormonal changes and large scale alteration of gene expression. The resulting adaptations to a marine environment include transformation of salt-tolerance, coloration, behavior, growth rate, and metabolism (reviewed in Stefansson et al., 2008).

A key difference between freshwater and sea-habitats is the dietary availability of essential long-chain polyunsaturated fatty acids. Salmon in rivers mostly eat invertebrates that are low in physiologically critical n-3 and n-6, 20 and 22 carbon long-chain polyunsaturated fatty acids (n-3LC-PUFA and n-6LC-PUFA), arachidonic acid (20:4n-6), eicosapentaenoic acid (20:5n-3) and docosahexaenoic (22:6n-3), while marine habitat food chains are high in available LC-PUFAs. Possibly as an adaptation to this (Leaver et al., 2008), salmon have evolved a high capacity for endogenous production of LC-PUFAs by elongation and desaturation of essential dietary 18 carbon precursor linoleic and linolenic acids (18:2n-6 and 18:3n-3; Figure 4) and the ability to increase or decrease this endogenous production as a response to the dietary availability (Kennedy et al., 2006; Leaver et al., 2008; Morais et al., 2011; Ruyter et al., 2000; Tocher et al., 2001; Tocher et al., 2002; Zheng et al., 2005). During smoltification and after sea-migration, Atlantic salmon have been shown to undergo transformation of lipid metabolism function, by decreasing lipid syntheses and increasing lipid breakdown (Sheridan, 1989). However, very little is known about the molecular nature of this life-stage associated transformation physiological function.

The evolution of novel traits in salmonids, such as increased plasticity and the ability to migrate to sea, may have been facilitated by their ancestral whole genome duplication (called Ss4R) some 80 million years ago (Allendorf & Thorgaard, 1984; Lorgen et al., 2015; Macqueen & Johnston, 2014; Robertson et al., 2017). Gene duplication can give rise to new adaptive phenotypes in different ways: through evolution of novel functions or gene regulation, subdivision and/or specialization of function among duplicates, or via an adaptive increase in gene dosage. The Atlantic salmon genome contains ~10,000 pairs of Ss4R gene duplicates, of which ~50% have evolved some novel regulation (Lien et al., 2016; Robertson et al., 2017). Indeed, in the context of lipid metabolism, it has recently been shown that a Ss4R duplicate of elovl5, a key enzyme in LC-PUFA syntheses, has gained expression compared to its ancestral regulation with likely implications for the ability to synthesize LC-PUFAs (Carmona-Antoñanzas et al., 2016). This is believed to have facilitated evolution of novel traits, including flexible phenotypes necessary for an anadromous life history (Stefansson et al., 2008). However, no systematic genome wide study has yet been conducted to assess the importance of the Ss4R in evolution of salmon lipid metabolism.

In this study, we integrate comparative genomics with transcriptomic data from a feeding trial carried out across the fresh to saltwater transition to build a functional annotation of lipid metabolism pathway genes in salmon. We use this annotation to elucidate (i) the nature of the transformation of lipid metabolism from freshwater to saltwater life-stages and (ii) the impact of whole genome duplication on evolution of the lipid gene repertoire and metabolic function. Our results indicate a striking shift in lipid metabolism after transition to seawater, and show that lipid

pathways differ with respect to selection pressure on gene duplicates from the salmonid whole genome duplication.

Results and discussion

Annotation of lipid metabolism genes

To identify genes involved in lipid metabolism in Atlantic salmon, we initially assembled groups of orthologous genes (orthogroups) using protein sequence similarity. We included proteins from four salmonid species sharing the Ss4R genome duplication, in addition to four non-salmonid fish genomes and two model mammalian outgroup species (Figure 1a) to aid in distinguishing Ss4R copies from other gene duplicates. Next, we aligned orthogroup proteins and constructed maximum likelihood gene trees. The majority (82-98%) of proteins from each species were represented in 23,782 ortholog gene trees. The salmonid species had significantly higher number of proteins included in ortholog gene trees compared to non-salmonid fish (Figure S1), reflecting the salmonid specific whole genome duplication. We then used the evolutionary distances in gene trees to infer the most likely salmon sequence orthologs of zebrafish genes selected from 19 KEGG pathways involved in lipid metabolism (File S1). This resulted in the annotation of 1421 (File S2) salmon lipid metabolism genes, of which 326 (23%) showed a 2:1 ortholog ratio between salmon and zebrafish (Figure 1b). Only 87 (6%) of the zebrafish genes could not be assigned a salmon ortholog.

To validate our ortholog annotation pipeline used to identify lipid metabolism genes, we analyzed the tissue specificity of these genes using gene expression data from 15 tissues (File S3) of Atlantic salmon (Lien et al., 2016). Genes in certain fatty acid metabolism related pathways (*'fatty acid metabolism'*, *'PPAR signaling pathway'*, *'fat digestion and absorption'*) had higher overall expression in tissues known to have high lipid metabolism activity (i.e. pyloric caeca, liver, heart, and brain) (Glatz et al., 2010; Benedito-Palos & Pérez-Sánchez, 2016; Tocher, 2003) (Figure 2). Examples include: 1) Liver was the site of highest expression for all genes in the LC-PUFA biosynthesis pathway (the desaturases Δ6FAD and Δ5FAD, and the elongases elovl5, elovl2 and elovl4). 2) Bile acids are essential for fat digestion in the gut, but are synthesized in liver. As expected, the rate limiting step for bile syntheses, cytochrome P450 7A1 (CYP7A1), has the highest expression in the liver. 3) Cholesterol, an essential component of cell membranes and precursor to bile acids, is known to be synthesized in all tissues, but primarily in liver, intestine, and brain (Brown & Sharpe, 2016). This is reflected in our annotation by high expression of the key cholesterol biosynthesis genes 3-hydroxy-3-methyl-glutaryl-CoA reductase (HMGCR), isopentenyl-diphosphase Δisomerase (IDI1), squalene epoxidase (SM), and lanosterol synthase (LS) in these tissues. 4) Several known regulators of lipid metabolism show high expression in liver, heart, brain and pyloric caeca, as expected, including liver X receptor (LXR), peroxisome proliferator-activated receptor alpha (PPARα), sterol regulatory element binding protein 1 (SREBP1), and sterol regulatory element binding protein 2 (SREBP2). Taken together, the tissue distribution of lipid metabolism gene expression is in line with knowledge about vertebrate physiology in general, and support the validity of our annotation of lipid metabolism genes in salmon. To make all data underlying our annotation easily available, and to facilitate further

refinement through manual community curation, we have created an interactive web-server available online (goo.gl/8Ap89a).

Life-stage associated remodeling of lipid metabolism

We conducted a feeding trial to study how salmon adjusts its lipid metabolism to different levels of LC-PUFA in freshwater and saltwater (see figure 8 for experimental details). Groups of salmon were fed contrasting diets from hatching until after transition to seawater. One feed was vegetable oil based (VO) and hence low in LC-PUFA, similar to river ecosystem diets, whereas the other was based on fish oil (FO) and high in LC-PUFA as expected in a marine-type diet (see Table S2 and S3 for details on feed composition). VO based diets are also low in cholesterol (Ciftci, et al., 2012; Verleyen et al., 2002). The proportion of fat in feed also increased between FW and SW (Table S2), as is standard practice in the aquaculture industry to maintain optimal growth conditions (Storebakken, 2002). Moreover, total lipid availability is also expected to increase between natural riverine and marine ecosystem diets. The contrasting levels of EPA/DHA between FO and VO diets remained constant across life-stages. In total, 32 and 23 fish were sampled for RNA-Seq of liver and gut, respectively, including up to eight biological replicates from each diet and life-stage (freshwater and saltwater, see figure 8c for details). Fish in the different dietary groups were given FO and VO feed from first feeding (<0.2 g body weight) until sampling.

In general, global gene expression levels were more affected by dietary composition in liver than in gut (which was largely unresponsive), and the effect was more pronounced in freshwater than in saltwater (Figure 3a). VO diets, compared to FO diets, increased lipid-metabolism related gene

156 expression in liver. In freshwater, 66 genes were differentially expressed with 57 (86%) of these
157 upregulated, while in saltwater 31 genes were differentially expressed with 23 (74%) of these
158 upregulated (Figure 3b). The increased activity of liver lipid metabolism under VO diets confirm
159 the well-known ability of salmon to regulate endogenous synthesis of LC-PUFA and cholesterol
160 in response to VO diets (Kortner et al., 2014; Leaver et al., 2008; Zheng et al., 2005).
161
162 Fish sampled in freshwater and saltwater shared a relatively small number of differentially
163 expressed genes (DEGs) for each pathway (Table S4). We found that most pathways had more
164 DEGs in freshwater (‘*fatty acid biosynthesis’*, ‘*steroid biosynthesis*’, and its precursor ‘*terpenoid*
165 *backbone biosynthesis’*), whereas few had more DEGs in saltwater (‘*fat digestion and absorption’*
166 and *‘steroid hormone biosynthesis’*) (Figure 3c). Out of 87 lipid metabolism DEGs in the dietary
167 contrast, 56 (64%) were freshwater specific, 21 (24%) saltwater specific, and 10 (11%) shared
168 dietary response. For example, only two genes in the FA and LC-PUFA biosynthesis pathways
169 (Δ6FADa and Δ5FAD) shared response to diet in fresh- and saltwater (Figure 4). Similarly, in the
170 pathways responsible for cholesterol biosynthesis there were more DEGs between diets in FW (21
171 DEGs in FW, 4 shared, and no SW specific) (Figure 5). The few genes that showed diet-effects
172 specific to saltwater included bile salt activated lipase, responsible for the hydrolysis of free fatty
173 acids from TAG obtained from the diet (Tocher, 2003). Two of these genes, carboxyl ester lipase,
174 tandem duplicate 2a (CEL2a) and b (CEL2b), are highly upregulated in saltwater in response to
175 VO diet. Taken together, our results show higher metabolic plasticity in parr-stage salmon,
176 suggesting a life-stage associated remodeling of lipid metabolism in liver. This corroborates the
177 idea of a post-smoltification phenotype adapted to an environment with a surplus of n-3LC-PUFA.
178

179 To further investigate the life-stage associated changes in lipid metabolism we tested for
180 differential expression between salmon in freshwater and saltwater fed diets with identical n-3LC-
181 PUFA profiles (Figure 6). Liver and gut showed contrasting effects of saltwater on lipid gene
182 expression with extensive downregulation in liver and upregulation in gut (Figure 6b). The number
183 of DEGs in each tissue were similar for the environment comparison (Figure 6a), unlike for the
184 diet comparison (Figure 3).
185
186 Further examination of key lipid metabolism genes revealed that after life-stage transition the
187 system-wide lipid metabolism remodeling represented a concerted shift in the metabolic role of
188 liver and gut. After the salmon entered the marine stage, lipogenic gene expression in the liver was
189 significantly decreased, as evident by the markedly lower expression (2.2-3.3 fold) of the master
190 regulator of lipid metabolism SREBP1, a 5-fold decrease in expression of fatty acid synthase, and
191 a 2-3 fold decrease in rate-limiting enzymes in LC-PUFA synthesis (i.e. Δ5FAD, Δ6FADa) (Figure
192 4). Liver and gut gene expression also indicated increased catabolic activity in saltwater, with
193 upregulation of the carnitine palmitoyltransferase 1 and 2 genes, responsible for uptake of fatty
194 acids into mitochondria for β-oxidation (Lehner & Quiroga, 2016). Finally, expression of lipid
195 transport genes shifted from liver to gut with the transition to seawater (apolipoproteins, pathway
196 "Fat digestion and absorption" in Figure 6). Four apolipoproteins (out of 11 annotated) were
197 differentially regulated in liver between different life-stages, with a 2.4-5 fold decrease in saltwater
198 compared to freshwater. In stark contrast, nine of the diet-regulated apolipoproteins in gut
199 increased their expression in saltwater between 1.8-9.7 fold. The results point to an adaptive shift
200 in lipid metabolism, with increased ability to take up lipids in the gut after Atlantic salmon migrates
201 to sea where lipid availability is higher. Remodeling of lipid metabolism across life-stages is likely

202 the result of a combination of factors, including the direct regulatory effect of dietary fat itself,
203 effect of salinity, and smoltification-induced physiological changes influencing gene regulation.
204 Although the relative importance of these factors is undetermined in our study, the fact that DEGs
205 in the VO versus FO feed contrast were mostly life-stage specific (Figure 3), supports that factors
206 other than the diet itself contribute significantly to the fresh and seawater metabolic phenotypes.
207
208 Interestingly, diet had a strong influence on the number and direction of gene expression changes between
209 freshwater and saltwater (Figure 6). In gut, about twice as many DEGs (with respect to the fresh- to
210 saltwater transition) were observed in salmon when fed FO diet than VO diet (Figure 6a). In liver, the diet
211 effect was less pronounced, with the FO group containing 46% more DEGs than the VO group (Figure 6a).
212 This diet effect pattern was reflected in the lipid metabolism genes with 89% and 16% more DEGs in the
213 FO group for gut and liver, respectively (Figure 6b). As this diet and life-stage interaction is a genome wide
214 trend, and more pronounced in gut tissue than in liver, this pattern could be related to differences in
215 osmoregulation and adaptation to saltwater. Two studies have suggested that Atlantic salmon raised on VO
216 based feeds more closely resembling riverine diets adapt to saltwater sooner and better than salmon raised
217 on FO based diets (Bell et al., 1997; Tocher et al., 2000). Conversely, there has been evidence that VO
218 based diets can reduce markers for stress response upon saltwater challenge, resulting in reduced
219 osmoregulatory capacity (Oxley et al., 2010). Regardless of the effect, it is clear that diet can modulate the
220 smoltification process and could explain the discrepancy between diets in number of life-stage related
221 DEGs. Another possibility is that the different levels of fatty acids in the diets, for example DHA, affect
222 DNA-methylation and thus trigger genome wide divergence in gene regulation (Kulkarni et al., 2011).
223
224 Our results clearly demonstrate very different baseline lipid metabolic functions in pre- and post-
225 smolt salmon, as well as life-stage associated changes in the plasticity of lipid metabolism, e.g. the
226 ability to regulate endogenous LC-PUFA synthesis as a response to changes in diet (i.e. fatty acid

227 composition). As opportunistic carnivores, salmon tend to eat whatever the local environment
228 provides. Thus, in freshwater, insects and amphipods provide variable, mostly low amounts of
229 essential LC-PUFA and total fat (Jonsson & Jonsson, 2011; Sushchik et al., 2003), favoring a
230 metabolic function that can efficiently regulate endogenous lipid synthesis based on dietary
231 availability (Carmona-Antonanzas et al., 2014). Conversely, in marine environments, amphipods
232 and smaller fish provide a higher, more stable source of n-3LC-PUFA and total fat (Jonsson &
233 Jonsson, 2011; Baeza-Rojano, et. al., 2014), promoting a metabolic function that allocates less
234 energy to endogenous synthesis of essential lipids.
235

236 Selection on gene duplicates after whole genome duplication

237 Carmona-Antonanzas et al. (2014, 2016) proposed that the salmonid whole-genome duplication
238 may have adaptively increased the potential for endogenous lipid synthesis. We pursued this
239 hypothesis by searching for distinct signatures of selection pressure on lipid metabolism genes in
240 salmon. Specifically, we compared pathways in terms of their tendency to retain both duplicates
241 of gene pairs, in terms of whether duplicates showed similar regulation (expression patterns across
242 diets and environments), and in terms of total gene dosage (for the one or two genes retained of a
243 pair) in salmon compared to pike, its closest unduplicated sister lineage.
244

245 To assess the level of Ss4R duplicate retention, we first defined 10,752 Ss4R duplicate pairs
246 (21,504 genes) in the NCBI RefSeq annotation using the same approach as Lien et al. (2016). Of
247 the 1,421 annotated lipid metabolism genes, 867 (61%) were retained as duplicated genes after
248 Ss4R (Figure 7a) (in contrast to 47% of the 45,127 salmon genes assigned to ortholog groups).

Moreover, our results showed large variation in the proportion of retained duplicates in each lipid metabolism pathway (Figure 7), with the most extreme case being '*fat digestion and absorption*' with 80% retained duplicates and '*steroid hormone biosynthesis*' with only 27% retained Ss4R duplicates.

The regulatory conservation of the duplicates was then estimated by calculating co-expression correlation between Ss4R duplicates from RNA-Seq data representing a time course of dynamic changes in gene expression and lipid metabolism function in liver. Fish in the same feeding trial were switched from VO to FO feed and vice versa, in both fresh and saltwater conditions (see Figure 8 for details). In total, 38 sampling time points (20 in freshwater and 18 in saltwater) from the feed switch experiment were used. Pathway-level analyses showed that regulatory conservation was not associated with duplicate retention (Figure 7). For example, the *'biosynthesis of unsaturated fatty acids'* pathway had significantly fewer duplicates retained than expected by chance (P-value < 0.0234), but a significant overrepresentation of duplicate pairs that display highly similar regulation (P-value < 0.0142 and < 0.0361 in freshwater and saltwater, respectively). Interestingly, the *'insulin signaling pathway'* also showed higher than expected duplicate co-regulation. This pathway has been shown to be important in regulating uptake and transport of FAs in adipose tissue, liver and muscle of Atlantic salmon (Sánchez-Gurmaches et al., 2011). Other pathways showing signatures of increased duplicate co-regulation were *'terpenoid backbone biosynthesis'*, *'steroid biosynthesis'*, *'fat digestion and absorption'*, and *'fatty acid metabolism'* (Figure 7b-c). Overall, the distinct differences in duplicate retention and conservation of regulatory mechanisms across the lipid metabolism pathways suggest differences in selective pressures shaping duplicate evolution following Ss4R. Moreover, the pathways with highly conserved

duplicate co-regulation were also those that were most responsive to dietary differences in fatty acid composition (Figure 3).

Finally, to link duplicate retention and co-regulation to signals of increased gene dosage following Ss4R, we used RNA-Seq data from the Northern pike (*Esox lucius*), a species that belongs to the unduplicated sister lineage (see methods for details). For each duplicate pair, we computed the ratio between the sum of Ss4R duplicate expression and its non-duplicated ortholog in pike and compared these ratios to those observed for salmon genes that had not retained two Ss4R duplicates. In total 69 duplicate pairs from 18 different lipid-metabolism related pathways displayed a combined dosage increase relative to single copy genes, of which 26 had highly conserved regulation (i.e. correlated expression) (File S8). We saw no systematic effect of gene dosage when comparing the total gene expression of duplicate pairs with that of single-copy genes; nor did co-regulation of duplicates associate with increased gene dosage (Figure 7d). This pattern was also true for most individual lipid pathways (Figure S4-S5), except for *'biosynthesis of unsaturated fatty acids'*, *'fatty acid metabolism'* and *'fatty acid elongation'*. These three pathways showed a link between co-regulation of duplicated genes and higher total gene dosage (Figure S4-S5, Figure 7d). Underlying this link were three genes with co-regulated dosage effects shared between all three pathways; trifunctional enzyme alpha subunit b (hadhab), elovl6, and the previously identified elovl5 (Carmona-Antonanzas et al., 2014; Carmona-Antoñanzas et al., 2016). Only elovl5 is known to be directly involved in core PUFA biosynthesis. Hadhab is involved in mitochondrial β-oxidation/elongation and elovl6 is involved in elongation of saturated and monounsaturated fatty acids (Bond et al., 2016). Although we do not see a general trend of increased gene dosage effects on lipid metabolism genes after whole genome duplication, it is

likely that an increased dosage of elovl5 and the 68 other duplicate pairs has affected the function of lipid metabolism in salmon.

Conclusion

Atlantic salmon needs great plasticity of physiology and behavior to adapt for migration between freshwater and sea. By analyzing transcriptomic changes through the transition from fresh- to saltwater and the associated increase in dietary lipids, we identified an overall remodeling of lipid metabolism, with liver reflecting higher lipid metabolic plasticity and higher capacity of endogenous synthesis of LC-PUFAs in freshwater, while gut lipid uptake genes become more active in saltwater. These results indicate adaptive optimization of the Atlantic salmon lipid metabolism to account for life-stage specific dietary availability. Moreover, we found signatures of pathway-specific selection pressure on gene duplicates, including a gene dosage increase in three genes involved in fatty acid metabolism. This illustrates possible adaptive consequences of the salmonid whole-genome duplication for the evolution of lipid metabolism. Future studies should attempt to decipher how the life-stage related metabolic reprogramming is controlled (for example through epigenetic mechanisms). Understanding this will have important implications for understanding evolution of genome regulatory processes in anadromous salmonids and potentially have economically important implications for Atlantic salmon aquaculture.

Materials and methods

Orthogroup prediction

Protein sequences were obtained from seven teleost fish species; *Danio rerio* (zebrafish), *Gasterosteus aculeatus* (three-spined stickleback), *Oryzias latipes* (medaka), *Oncorhynchus mykiss* (Rainbow trout), *Oncorhynchus kisutch* (coho salmon), *Salmo salar* (Atlantic salmon), *Thymallus thymallus* (grayling), *Esox lucius* (northern pike), and two mammalian outgroup species; *Homo sapiens* (human), *Mus musculus* (house mouse). Human, mouse, zebrafish, medaka and stickleback protein fasta data were obtained from ENSEMBL (release 83). Atlantic salmon (RefSeq assembly GCF_000233375.1, Annotation Release 100) and northern pike (RefSeq assembly GCF_000721915.2, Annotation Release 101) proteins were obtained from NCBI RefSeq. Rainbow trout proteins were obtained from an assembly and annotation of the genome (Berthelot et al., 2014). Grayling proteins were obtained from an assembly and annotation of the genome (Varadharajan et al., 2017). The coho salmon transcriptome (Kim, Leong, Koop, & Devlin, 2016) was obtained from NCBI (GDQG00000000.1). Where transcriptome data was used, protein sequences were translated using TransDecoder (v2.0.1, http://transdecoder.github.io/). Protein fasta files were filtered to retrieve only the longest protein isoform per gene. Orthofinder (v0.2.8) (Emms et al., 2015) assigned groups of orthologs based on protein sequence similarity. Proteins within an orthogroups were further aligned using MAFFT (v7.130) (Katoh et al., 2002) and maximum likelihood trees were estimated using FastTree (v2.1.8) (Price et al., 2010).

Annotation of salmon lipid metabolism genes

A list of zebrafish proteins obtained from 19 manually selected zebrafish KEGG pathways related to lipid metabolism (Table S1) were used to search for Atlantic salmon orthologs. Orthogroups that contained a selected zebrafish protein were identified. Salmon proteins within those orthogroups were assigned as orthologs of the closest zebrafish protein based on the orthogroup tree distance. A lipid metabolism gene list was created including salmon orthologs to the selected zebrafish genes. Additional salmon genes related to lipid metabolism not included in KEGG pathways (e.g. regulators or transporters, SREBP, LXR, FABP, etc.) were manually searched for through NCBI and added to the list.

Tissue expression

Atlantic salmon RNA-Seq samples from 15 different tissues (liver, gut, pyloric caeca, heart, kidney, muscle, gill, eye, skin, ovary, nose, testis, brain, head kidney, spleen) were obtained from NCBI SRA (PRJNA72713) (Lien et al., 2016). Fastq files were adapter trimmed before alignment to the Atlantic salmon genome (RefSeq assembly GCF_000233375.1) (Lien et al., 2016) using STAR (v2.5.2a) (Dobin et al., 2013). HTSeq-count (v0.6.1p1) (Anders et al., 2015) counted the sum of uniquely aligned reads in exon regions of each gene in the annotation (RefSeq Annotation Release 100). Gene FPKM values were calculated based on the gene count over the samples effective library size (see TMM method from edgeR (Robinson et al., 2010) user manual) and the mean gene transcript isoform length.

Feed trial

Atlantic salmon fry were obtained from AquaGen Breeding Centre, Kyrksæterøra, Norway and reared in the Norwegian Institute for Water Research (NIVA), Solbergstranda, Norway in four partitioned 1000 liter tanks on vegetable oil (VO) or fish oil (FO) based diets continuously from first feeding (fry weight <0.2 g). Daily feed amount was calculated based on total biomass in each tank and decreased as the fish grew, from 3% at first feeding to 1.2% by the end of the trial. Fish were euthanized periodically throughout the experiment to maintain appropriate levels of dissolved oxygen. VO based feeds contained a combination of linseed oil and palm oil at a ratio of 1.8:1 and FO based feeds contained only North Atlantic fish oil. Percent protein in feed decreased with fish size from 56% at first feeding to 41% at the end of the trial. This corresponded with an increase in percent lipid from 16% at first feeding to 31% at the end of the trial. At the time of sampling, the proportion of lipid in the feed was 22% in freshwater and 31% in saltwater (Table S2). Increasing lipid proportion in feed with fish size is standard practice in the aquaculture industry as this maintains optimal growing conditions by decreasing the digestible protein to digestible energy ratio (Storebakken, 2002). All feeds were formulated and produced by EWOS innovation (Supplementary File 3). Local groundwater was UV sterilized for use in the freshwater life-stage and water from the Oslofjord taken from 60 meters below sea surface (~3-3.5% salinity) was UV sterilized for use in the saltwater life-stage. Fish were raised under constant light and water temperature (~12°C) for 26 weeks. Then, 40 pre-smolt salmon (~50g) from each control tank (~240 fish per control tank) were switched to the contrasting diet (VO to FO and vice versa) by physically moving them to the empty partition of the tank receiving the appropriate feed (Figure 8a). Five fish from each of the control tanks (2 VO tanks and 2 FO tanks) were sampled before switching feeds (D0), then fish from both control and feed switch conditions were similarly

sampled 1, 2, 5, 9, 16, and 20 days after switching feeds (5 fish x 2 replicate tanks x 4 conditions = 40 fish per time point, figure 8b). Two weeks after freshwater sampling (31 weeks after first feeding), smoltification was triggered by 5 weeks of winter-like conditions with decreased light (12 hours per day) and water temperature (~8°C), immediately followed by 5 weeks of spring-like conditions, returning to normal light (24 hours per day) and water temperature (~12°C). All salmon from the control groups (VO or FO) were then switched to saltwater and allowed to acclimate for 3 weeks. The feed switch was repeated in saltwater by transferring half (~40 fish) of the post-smolt salmon (~200 g) from each control tank to the contrasting feed condition. Again, pre-switch control samples were taken (D0) followed by sampling 1, 2, 6, 9, 16, and 20 days post-diet switch (Figure 8b). For both freshwater and saltwater samplings, feeding was stopped in the mornings of each of the sampling days. All fish were euthanized by a blow to the head and samples of liver and midgut (gut section between pyloric caeca and hindgut) were flash frozen in liquid nitrogen and stored under -80 °C. A subset of the samples taken were used for further RNA-Seq analysis (see figure 8c for details).

RNA-sequencing

Total RNA was extracted from selected feed trial samples (see figure 8c for details) using the RNeasy Plus Universal kit (QIAGEN). Quality was determined on a 2100 Bioanalyzer using the RNA 6000 nano kit (Agilent). Concentration was determined using a Nanodrop 8000 spectrophotometer (Thermo Scientific). cDNA libraries were prepared using the TruSeq Stranded mRNA HT Sample Prep Kit (Illumina). Library mean length was determined by running on a 2100 Bioanalyzer using the DNA 1000 kit (Agilent) and library concentration was determined with the

Qbit BR kit (Thermo Scientific). Single end sequencing of sample libraries was completed on an Illumina HiSeq 2500 with 100 bp reads.

Differential expression analysis between feed conditions and life-stages

To analyze gene expression differences between feed conditions and life-stages, samples from the feed trial were selected for RNA-Seq. Liver and gut tissue RNA were sequenced from fish fed each of the feeds (FO, VO) at day 0 of the diet switch, both before (freshwater) and after (saltwater) smoltification (See figure 8c for the number of RNA-Seq replicates and sampling details). Fastq files were processed to produce gene count and FPKM data using the same protocol described under the tissue expression method section. For the feed comparison, changes in gene expression were tested between FO and VO feed conditions for both freshwater and saltwater samples, and liver and gut tissues. For the life-stage comparison, changes in gene expression were tested between freshwater and saltwater stages for both FO and VO feed conditions, and liver and gut tissues. Using RNA-Seq gene count data, lowly expressed genes were filtered prior to testing, retaining genes with a minimum of one read count per million (CPM) in two or more samples. Differential expression analysis was carried out using a standard edgeR (Robinson et al., 2010) protocol. Effective library sizes were calculated using the edgeR TMM-normalisation procedure allowing effective comparison of expression data between different sample types (see edgeR user manual). An exact test between expression levels of a pair of conditions gave the log2 fold change, P-value and false discovery rate (FDR) for each gene. Genes with FDR < 0.05 were considered differentially expressed genes (DEGs).

Identification of Ss4R duplicates

To identify putative gene duplicates stemming from the Ss4R, we used the same approach as in Lien et al. (2016). All-vs-all protein blast was run with e-value cutoff of 1e-10 and pident (percentage of identical matches) ≥80 and blast hit coverage of ≥50% of protein length. Only the best protein hits between the 98 defined synteny blocks (see Lien et al., 2016) were considered as putative Ss4R duplicates. Blast result ranking was done using the product of pident times bitscore to avoid spurious 'best blast matches' with low pident (<85) but high bitscore.

Duplicate analysis

Genes from the lipid metabolism gene list were paired together with their putative Ss4R duplicates identified above. The retention of gene duplicates (i.e. whether both genes in a pair were retained, or just one) was compared between all identified duplicates in the salmon genome annotation and the lipid metabolism gene list. Pathway-level retention was explored by comparing the number of genes in each of the 19 selected KEGG pathways (Table S1) in a duplicate pairing to that of the total list of lipid genes, to find pathways with significantly less or more duplicate retention (Fisher's exact test, P-value < 0.05). Regulatory conservation of lipid gene duplicates was explored by correlation of gene expression changes between duplicates over the course of the feed trial described above. RNA-Seq data was generated from liver samples of salmon from 38 sampling time points (19 in freshwater and 19 in saltwater). Fastq files were processed to produce gene count and FPKM data using the same protocol described under the tissue expression method section. For each duplicate pair, mean FPKM values were retrieved for each time point and used to calculate a freshwater and saltwater correlation value. Duplicates with Pearson correlation ≥ 0.6 were

considered correlated (P-value < 0.003 from 19 sample points). The number of duplicates with correlated expression profiles was counted for each pathway and compared to all lipid genes to find pathways with significantly less or more correlated duplicates (Fisher's exact test, P-value < 0.05). The effect of gene duplication on gene dosage was estimated by calculating a dosage ratio between the FPKM value of a salmon ortholog (sum of gene expression in duplicate pairs) over the FPKM value of the non-duplicated ortholog from northern pike. For salmon, the RNA-Seq data from the freshwater and saltwater FO feed trial was used (samples used in differential expression analysis section). For pike, RNA-Seq from livers of four individuals were aligned (see tissue expression section for protocol) to their respective genomes (see genomes in ortholog prediction section). RSEM (v1.2.31) (Li & Dewey, 2011) was used to generate FPKM values for genes so that non-uniquely mapped reads between salmon duplicate genes were not ignored but instead assigned proportionately to each gene to match the proportions of uniquely mapped reads between the genes. Gene dosage levels for duplicate pairs with correlated expression (see above), non-correlated expression and single genes were compared for all lipid metabolism genes and for each pathway.

References

Allendorf, F. W., & Thorgaard, G. H. (1984). Tetraploidy and the Evolution of Salmonid Fishes. In *Evolutionary Genetics of Fishes* (pp. 1–53). Boston, MA: Springer US. http://doi.org/10.1007/978-1-4684-4652-4_1

Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics*, *31*(2), 166–169. http://doi.org/10.1093/bioinformatics/btu638

Baeza-Rojano, E., Hachero-Cruzado, I., & Guerra-García, J. M. (2014). Nutritional analysis of freshwater and marine amphipods from the Strait of Gibraltar and potential aquaculture applications. Journal of Sea Research, 85, 29–36. http://doi.org/10.1016/J.SEARES.2013.09.007

Bell, J. G., Tocher, D. R., Farndale, B. M., Cox, D. I., McKinney, R. W., & Sargent, J. R. (1997). The effect of dietary lipid on polyunsaturated fatty acid metabolism in Atlantic salmon (Salmo salar) undergoing parr-smolt transformation. *Lipids, 32*(5), 515–525. http://doi.org/10.1007/s11745-997-0066-4

Berthelot, C., Brunet, F., Chalopin, D., Juanchich, A., Bernard, M., Noël, B., … Guiguen, Y. (2014). The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. *Nature Communications*, *5*, 3657. http://doi.org/10.1038/ncomms4657

Bond, L. M., Miyazaki, M., O'Neill, L. M., Ding, F., & Ntambi, J. M. (2016). Fatty Acid Desaturation and Elongation in Mammals. In *Biochemistry of Lipids, Lipoproteins and Membranes* (pp. 185–208). Elsevier. http://doi.org/10.1016/B978-0-444-63438-2.00006-7

Brown, A. J., & Sharpe, L. J. (2016). Cholesterol Synthesis. In *Biochemistry of Lipids, Lipoproteins and Membranes* (pp. 327–358). Elsevier. http://doi.org/10.1016/B978-0-444-63438-2.00011-0

Carmona-Antonanzas, G., Tocher, D. R., Martinez-Rubio, L., & Leaver, M. J. (2014). Conservation of lipid metabolic gene transcriptional regulatory networks in fish and mammals. *Gene*, *534*(1), 1–9. http://doi.org/10.1016/j.gene.2013.10.040

Carmona-Antoñanzas, G., Zheng, X., Tocher, D. R., & Leaver, M. J. (2016). Regulatory divergence of homeologous Atlantic salmon elovl5 genes following the salmonid-specific

whole-genome duplication. *Gene*, *591*(1), 34–42. http://doi.org/10.1016/j.gene.2016.06.056

Ciftci, O. N., Przybylski, R., & Rudzińska, M. (2012). Lipid components of flax, perilla, and chia seeds. *European Journal of Lipid Science and Technology*, *114*(7), 794–800. http://doi.org/10.1002/ejlt.201100207

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., … Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England)*, *29*(1), 15–21. http://doi.org/10.1093/bioinformatics/bts635

Emms, D. M., Kelly, S., Alexeyenko, A., Tamas, I., Liu, G., Sonnhammer, E., … Kellis, M. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology*, *16*(1), 157. http://doi.org/10.1186/s13059-015-0721-2

Glatz, J. a N. F. C., Luiken, J. J. F. P., & Bonen, A. (2010). Membrane Fatty Acid Transporters as Regulators of Lipid Metabolism : Implications for Metabolic Disease. *Physiological Reviews*, *90*, 367–417. http://doi.org/10.1152/physrev.00003.2009.

Jonsson, B., & Jonsson, N. (2011). *Ecology of Atlantic Salmon and Brown Trout - Habitat as a template for life histories*. *Fish and Fisheries Series* (Vol. 33). http://doi.org/10.1007/978-94-007-1189-1

Katoh, K., Misawa, K., Kuma, K., & Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, *30*(14), 3059–66. http://doi.org/10.1093/NAR/GKF436

Kennedy, S. R., Leaver, M. J., Campbell, P. J., Zheng, X., Dick, J. R., & Tocher, D. R. (2006). Influence of dietary oil content and conjugated linoleic acid (CLA) on lipid metabolism enzyme activities and gene expression in tissues of Atlantic salmon (Salmo salar L.). *Lipids*,

503 *41*(5), 423–436. http://doi.org/10.1007/s11745-006-5116-4
504 Kim, J.-H., Leong, J. S., Koop, B. F., & Devlin, R. H. (2016). *Multi-tissue transcriptome profiles*
505 *for coho salmon (Oncorhynchus kisutch), a species undergoing rediploidization following*
506 *whole-genome duplication*. *Marine Genomics* (Vol. 25).
507 http://doi.org/10.1016/j.margen.2015.11.008
508 Kortner, T. M., Björkhem, I., Krasnov, A., Timmerhaus, G., & Krogdahl, Å. (2014). Dietary
509 cholesterol supplementation to a plant-based diet suppresses the complete pathway of
510 cholesterol synthesis and induces bile acid production in Atlantic salmon (Salmo salar L.).
511 *British Journal of Nutrition*, *111*(12), 2089–2103.
512 http://doi.org/10.1017/S0007114514000373
513 Kulkarni, A., Dangat, K., Kale, A., Sable, P., Chavan-Gautam, P., & Joshi, S. (2011). Effects of
514 Altered Maternal Folic Acid, Vitamin B12 and Docosahexaenoic Acid on Placental Global
515 DNA Methylation Patterns in Wistar Rats. *PLoS ONE*, *6*(3), e17706.
516 http://doi.org/10.1371/journal.pone.0017706
517 Leaver, M. J., Bautista, J. M., Björnsson, B. T., Jönsson, E., Krey, G., Tocher, D. R., & Torstensen,
518 B. E. (2008). Towards Fish Lipid Nutrigenomics: Current State and Prospects for Fin-Fish
519 Aquaculture. *Reviews in Fisheries Science*, *16*(April), 73–94.
520 http://doi.org/10.1080/10641260802325278
521 Leaver, M. J., Villeneuve, L. A., Obach, A., Jensen, L., Bron, J. E., Tocher, D. R., & Taggart, J.
522 B. (2008). Functional genomics reveals increases in cholesterol biosynthetic genes and highly
523 unsaturated fatty acid biosynthesis after dietary substitution of fish oil with vegetable oils in
524 Atlantic salmon (Salmo salar). *BMC Genomics*, *9*, 299. http://doi.org/10.1186/1471-2164-9-
525 299

526 Lehner, R., & Quiroga, A. D. (2016). Fatty Acid Handling in Mammalian Cells. In *Biochemistry*
527 *of Lipids, Lipoproteins and Membranes* (pp. 149–184). Elsevier.
528 http://doi.org/10.1016/B978-0-444-63438-2.00005-5
529 Li, B., & Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with
530 or without a reference genome. *BMC Bioinformatics*, *12*(1), 323.
531 http://doi.org/10.1186/1471-2105-12-323
532 Lien, S., Koop, B. F., Sandve, S. R., Miller, J. R., Kent, M. P., Nome, T., … Davidson, W. S.
533 (2016). The Atlantic salmon genome provides insights into rediploidization. *Nature*,
534 *533*(7602), 200–205. http://doi.org/10.1038/nature17164
535 Lorgen, M., Casadei, E., Krol, E., Douglas, A., Birnie, M. J., Ebbesson, L. O. E., … Martin A.m,
536 S. (2015). Functional divergence of type 2 deiodinase paralogs in the Atlantic salmon.
537 *Current Biology*, *25*(7), 936–941. http://doi.org/10.1016/j.cub.2015.01.074
538 Macqueen, D. J., & Johnston, I. A. (2014). A well-constrained estimate for the timing of the
539 salmonid whole genome duplication reveals major decoupling from species diversification.
540 *Proceedings of the Royal Society of London B: Biological Sciences*, *281*(1778).
541 Morais, S., Pratoomyot, J., Taggart, J. B., Bron, J. E., Guy, D. R., Bell, J. G., & Tocher, D. R.
542 (2011). Genotype-specific responses in Atlantic salmon (Salmo salar) subject to dietary fish
543 oil replacement by vegetable oil: a liver transcriptomic analysis. *BMC Genomics*, *12*(1), 255.
544 http://doi.org/10.1186/1471-2164-12-255
545 Oxley, A., Jolly, C., Eide, T., Jordal, A.-E. O., Svardal, A., & Olsen, R.-E. (2010). The combined
546 impact of plant-derived dietary ingredients and acute stress on the intestinal arachidonic acid
547 cascade in Atlantic salmon (Salmo salar). *The British Journal of Nutrition*, *103*(6), 851–61.
548 http://doi.org/10.1017/S0007114509992467

549 Price, M. N., Dehal, P. S., Arkin, A. P., Nawrocki, E., Kolbe, D., Eddy, S., … Meyer, F. (2010).
550 FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE*,
551 *5*(3), e9490. http://doi.org/10.1371/journal.pone.0009490
552 Rimoldi, S., Benedito-Palos, L., Terova, G., & Pérez-Sánchez, J. (2016). Wide-targeted gene
553 expression infers tissue-specific molecular signatures of lipid metabolism in fed and fasted
554 fish. *Reviews in Fish Biology and Fisheries*, *26*(1), 93–108. http://doi.org/10.1007/s11160-
555 015-9408-8
556 Robertson, F. M., Gundappa, M. K., Grammes, F., Hvidsten, T. R., Redmond, A. K., Martin, S.
557 A. M., … Macqueen, D. J. (2017). Lineage-specific rediploidization is a mechanism to
558 explain time-lags between genome duplication and evolutionary diversification. *Genome*
559 *Biology*. http://doi.org/doi:10.1101/098582
560 Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: a Bioconductor package for
561 differential expression analysis of digital gene expression data. *Bioinformatics (Oxford,*
562 *England)*, *26*(1), 139–40. http://doi.org/10.1093/bioinformatics/btp616
563 Ruyter, B., Røsjø, C., Måsøval, K., Einen, O., & Thomassen, M. S. (2000). Influence of dietary n-
564 3 fatty acids on the desaturation and elongation of [1- 14 C] 18 : 2 n-6 and [1- 14 C] 18 : 3
565 n-3 in Atlantic salmon hepatocytes, 151–158.
566 Sánchez-Gurmaches, J., Østbye, T.-K., Navarro, I., Torgersen, J., Hevrøy, E. M., Ruyter, B., &
567 Torstensen, B. E. (2011). In vivo and in vitro insulin and fasting control of the transmembrane
568 fatty acid transport proteins in Atlantic salmon (Salmo salar). *American Journal of*
569 *Physiology. Regulatory, Integrative and Comparative Physiology, 301*(4), R947-57.
570 http://doi.org/10.1152/ajpregu.00289.2011
571 Sheridan, M. A. (1989). Alterations in lipid metabolism accompanying smoltification and seawater

572 adaptation of salmonid fish. *Aquaculture*, *82*(1–4), 191–203. http://doi.org/10.1016/0044-
573 8486(89)90408-0
574 Stefansson, S. O., Björnsson, B. T., Ebbesson, L. O., & McCormick, S. D. (2008). Smoltification.
575 *Fish Larval Physiology*, 639–681. http://doi.org/DOI: 10.1111/j.1095-8649.2009.02440_2.x
576 Storebakken, T. (2002). Atlantic salmon, Salmo salar. In Nutrient requirements and feeding of
577 finfish for aquaculture (pp. 79–102). Wallingford: CABI.
578 http://doi.org/10.1079/9780851995199.0079
579 Sushchik, N. N., Gladyshev, M. I., Moskvichova, A. V., Makhutova, O. N., & Kalachova, G. S.
580 (2003). Comparison of fatty acid composition in major lipid classes of the dominant benthic
581 invertebrates of the Yenisei river. *Comparative Biochemistry and Physiology Part B:*
582 *Biochemistry and Molecular Biology*, *134*(1), 111–122. http://doi.org/10.1016/S1096-
583 4959(02)00191-4
584 Tocher, D. R. (2003). Metabolism and Functions of Lipids and Fatty Acids in Teleost Fish.
585 *Reviews in Fisheries Science*, *11*(2), 107–184. http://doi.org/10.1080/713610925
586 Tocher, D. R., Bell, J. G., Dick, J. R., Henderson, R. J., McGhee, F., Michell, D., & Morris, P. C.
587 (2000). Polyunsaturated fatty acid metabolism in Atlantic salmon (Salmo salar) undergoing
588 parr-smolt transformation and the effects of dietary linseed and rapeseed oils. *Fish Physiology*
589 *and Biochemistry, 23*(1), 59–73. http://doi.org/10.1023/A:1007807201093
590 Tocher, D. R., Bell, J. G., MacGlaughlin, P., McGhee, F., & Dick, J. R. (2001). Hepatocyte fatty
591 acid desaturation and polyunsaturated fatty acid composition of liver in salmonids: effects of
592 dietary vegetable oil. *Comparative Biochemistry and Physiology Part B: Biochemistry and*
593 *Molecular Biology*, *130*(2), 257–270. http://doi.org/10.1016/S1096-4959(01)00429-8
594 Tocher, D. R., Fonseca-Madrigal, J., Bell, J. G., Dick, J. R., Henderson, R. J., & Sargent, J. R.

(2002). Effects of diets containing linseed oil on fatty acid desaturation and oxidation in hepatocytes and intestinal enterocytes in Atlantic salmon (Salmo salar). *Fish Physiology and Biochemistry*, *26*(2), 157–170. http://doi.org/10.1023/A:1025416731014

Varadharajan, S., Sandve, S. R., Tørresen, O. K., Lien, S., Vollestad, L. A., Jentoft, S., Nederbragt, A.J., Jakobsen, K. S. (2017). The grayling genome reveals selection on gene expression regulation after whole genome duplication. *bioRxiv*. https://doi.org/10.1101/153270

Verleyen, T., Forcades, M., Verhe, R., Dewettinck, K., Huyghebaert, A., & De Greyt, W. (2002). Analysis of free and esterified sterols in vegetable oils. *Journal of the American Oil Chemists' Society*, *79*(2), 117–122. http://doi.org/10.1007/s11746-002-0444-3

Zheng, X., Torstensen, B. E., Tocher, D. R., Dick, J. R., Henderson, R. J., & Bell, J. G. (2005). Environmental and dietary influences on highly unsaturated fatty acid biosynthesis and expression of fatty acyl desaturase and elongase genes in liver of Atlantic salmon (Salmo salar). Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids, 1734(1), 13–24. http://doi.org/10.1016/j.bbalip.2005.01.006

Data Accessibility:

- Supplementary files have been deposited to datadryad.org under the accession: doi:10.5061/dryad.j4h65
- All gene expression results can be accessed through the interactive shiny web server: https://goo.gl/8Ap89a
- Lipid metabolism gene annotation can be accessed from https://goo.gl/VVUVWr
- Raw RNA-Seq data has been deposited into European Nucleotide Archive (ENA) under the project accession number PRJEB24480

Acknowledgements:

This work was financed by the projects DigiSal NFR 248792 and GenoSysFat NFR 244164. We also thank anonymous reviewers for comments on the earlier version of the manuscript as well as Dominic Nanton and Cargill Innovation Center Dirdal for producing the fish trial feeds.

Legends Figures

Figure 1: Ortholog annotation (a) Species used to construct ortholog groups and their evolutionary distance. Points in the phylogenetic tree show the time of the teleost specific (Ts3R) and salmonid specific (Ss4R) whole genome duplications. (b) The number of salmon orthologs found (1421 genes in total) per zebrafish gene in 19 selected KEGG pathways involved in lipid metabolism.

Figure 2: Tissue expression profiles of salmon genes in lipid metabolism pathways

Tissue expression profiles of our annotated lipid metabolism genes were consistent with expectations. Gene expression levels are shown as the log2 fold change difference between the FPKM value of each tissue and the median FPKM across all tissues. Expression profiles for selected genes in each pathway are shown (see Figure S2 and S3 for all pathways and gene details).

Figure 3: Gene regulation in response to feed type. (a) Total number of significant (FDR < 0.05) differentially expressed genes (DEGs) between fish oil (FO) and vegetable oil (VO) fed salmon in the liver and gut tissues of freshwater and saltwater stage Atlantic salmon (see Files S4 (liver) and S5 (gut) for underlying data). (b) As above, but for lipid-associated genes only. (c) Proportions of genes in each KEGG pathway that had significantly different liver expression between the two feed types only in freshwater, only in saltwater, or in both stages.

Figure 4: Diet and life-stage effects on FA and LC-PUFA biosynthesis in salmon liver. Core fatty acid (FA) biosynthesis and biosynthesis of unsaturated fatty acids pathways with Atlantic salmon genes annotated to each catalytic step (enzyme names followed by NCBI gene numbers). Gene expression levels are shown as mean (point) and standard deviation (line) of expression in eight samples (measured in log(FPKM + 1)) from each diet (FO, VO feeds) and life-stage (freshwater, saltwater) combination. Genes significantly (FDR<0.05) differentially expressed (DEG) between diets in a life-stage are highlighted.

Figure 5: Diet and life-stage effects on cholesterol biosynthesis in salmon liver. Terpenoid backbone synthesis and steroid biosynthesis pathways with Atlantic salmon genes annotated to each catalytic step (enzyme names followed by NCBI gene

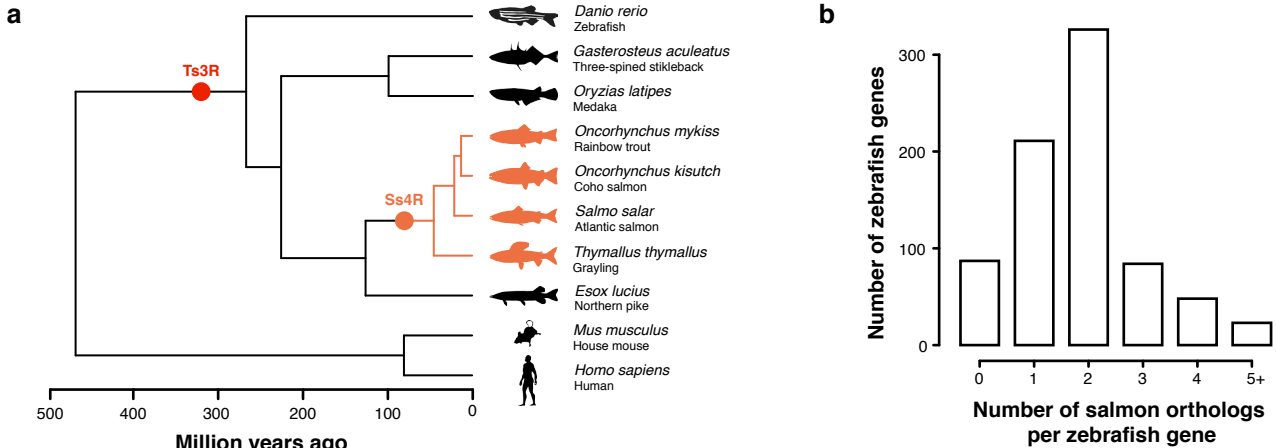
numbers). Gene expression levels are shown as mean (point) and standard deviation (line) of expression in eight samples (measured in log(FPKM + 1)) from each diet (FO, VO feeds) and life-stage (freshwater, saltwater) combination. Genes significantly (FDR<0.05) differentially expressed (DEG) between diets in a life-stage are highlighted.

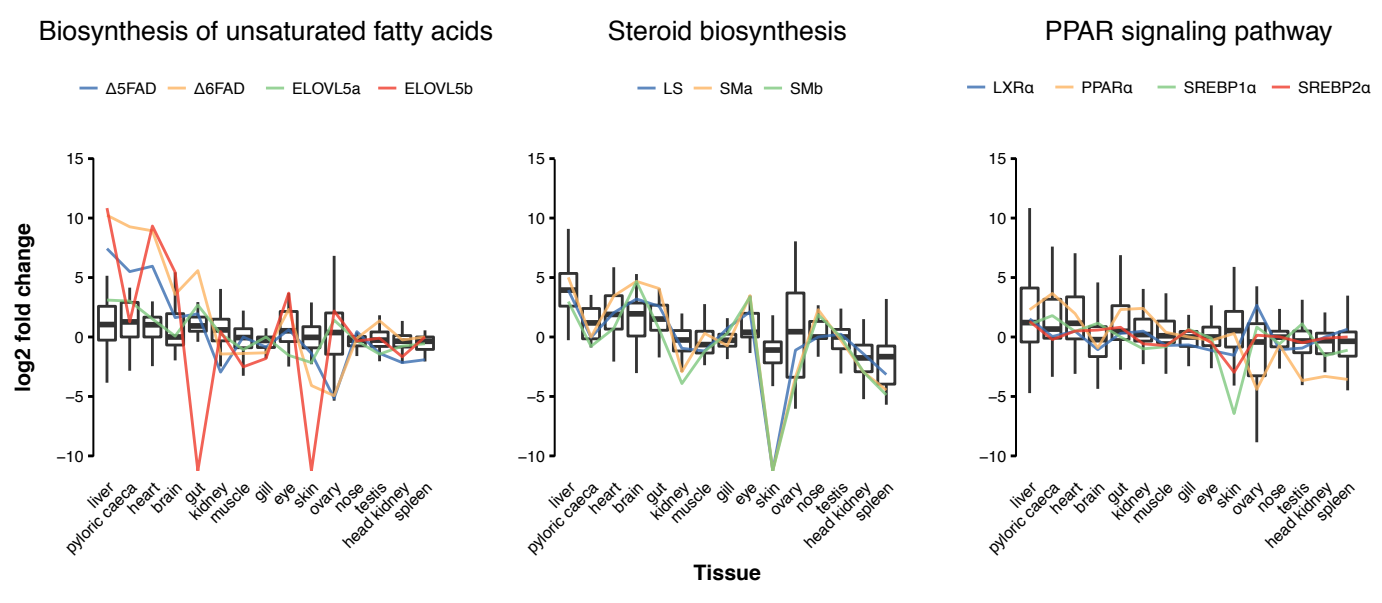
Figure 6: Gene regulation in response to life-stage. (a) Total number of significant (FDR < 0.05) differentially expressed genes (DEGs) between freshwater and saltwater life-stages in the liver and gut tissues of Atlantic salmon fed fish oil (FO) or vegetable oil (VO) diets (see Files S6 and S7 for underlying data). (b) As above, but for lipid metabolism DEGs. (c) Proportion of genes in each KEGG pathway that are DEGs in liver and (d) gut, colored by DEG significance in only FO, only VO, or both diets, and separated into up- or down-regulation in saltwater samples.

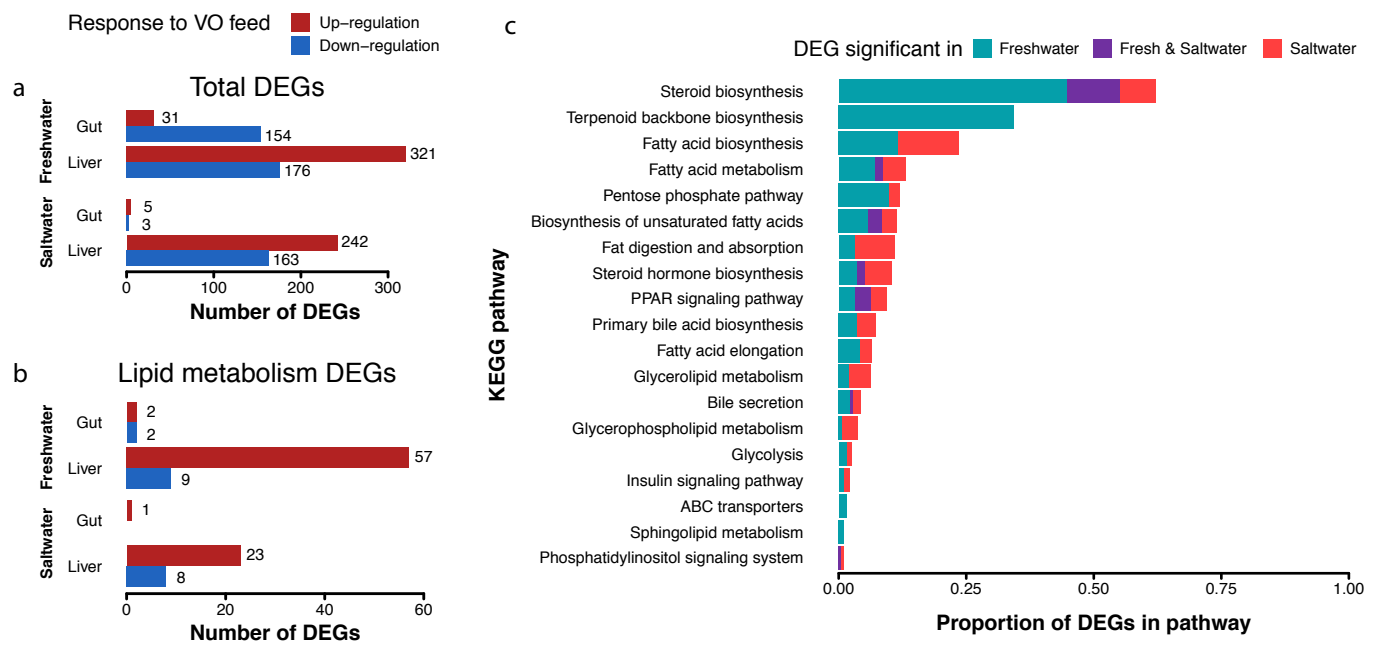
Figure 7: Gene duplication in lipid metabolism pathways. For the total list of lipid metabolism genes in Atlantic salmon, and sets of genes belonging to different KEGG pathways: (a) Number and percentage of genes with a duplicate homolog from the Ss4R duplication. (b) Number and percentage of duplicate genes with correlated liver expression response to feed in freshwater and (c) saltwater (Correlation >= 0.6, P-value < 3.306e-3, using 19 time points from feed trial for each water condition). Fisher's exact test was used to detect pathways with significant enrichment compared to all gene (P-value < 0.05) (d) Log2 gene dosage ratios (salmon:pike) in liver from fish in freshwater, where the ratio is computed between expression in the salmon duplicates (FPKM, sum of the two duplicates) and the expression of the corresponding pike ortholog. Ratios were computed for all lipid metabolism genes and genes in the pathway *'biosynthesis of unsaturated fatty acids'*. For comparison, ratios were also computed for genes without retained duplicates, i.e. with a 1:1 orthology between salmon and pike. Duplicates were grouped into correlated (corr.) or non-correlated (non-corr.) based on saltwater correlation result in (c). Dosage ratios (points) greater than the 95% quantile of single gene dosages are marked in red.

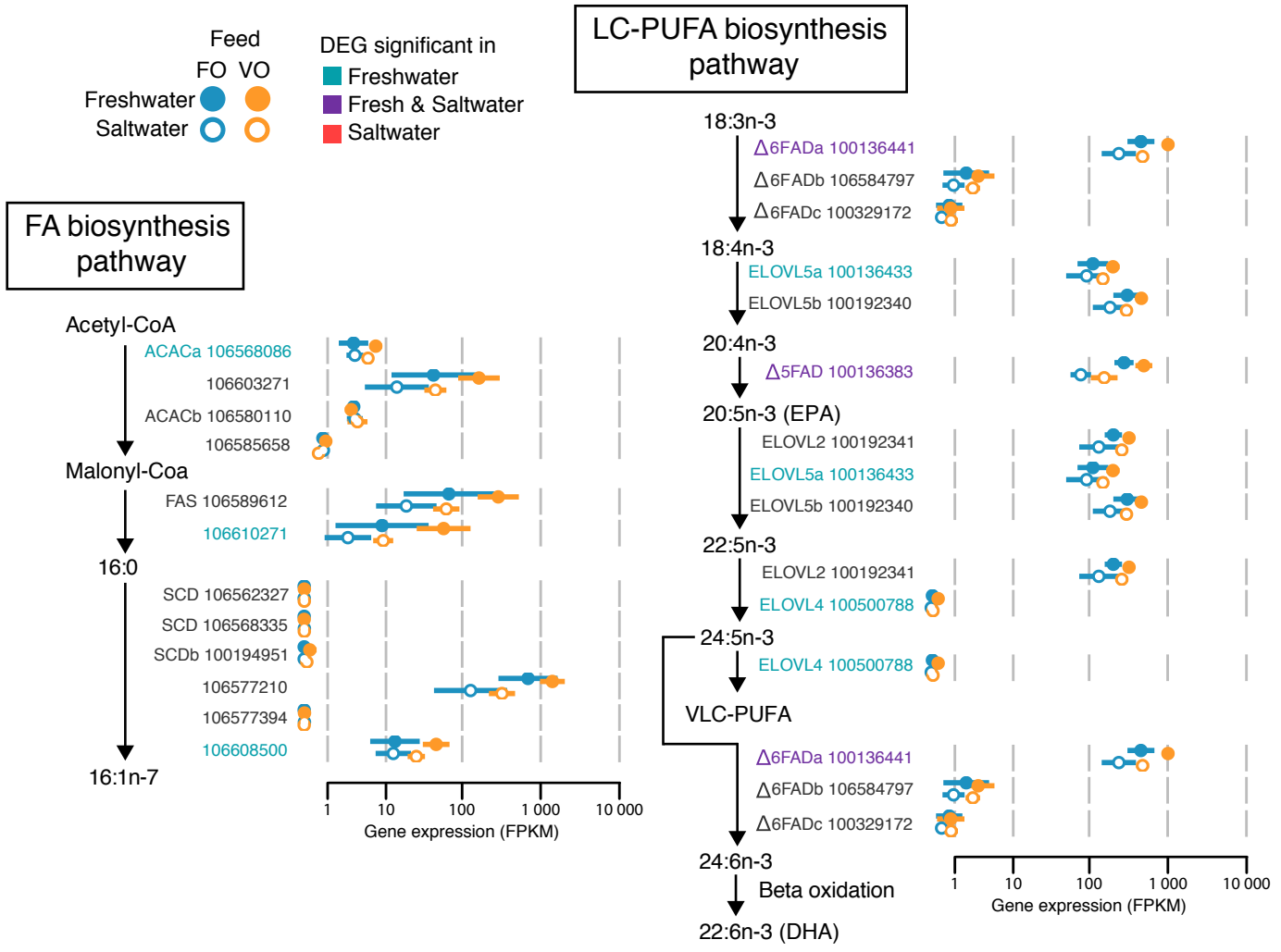
Figure 8: Overview of feed trial experiment.

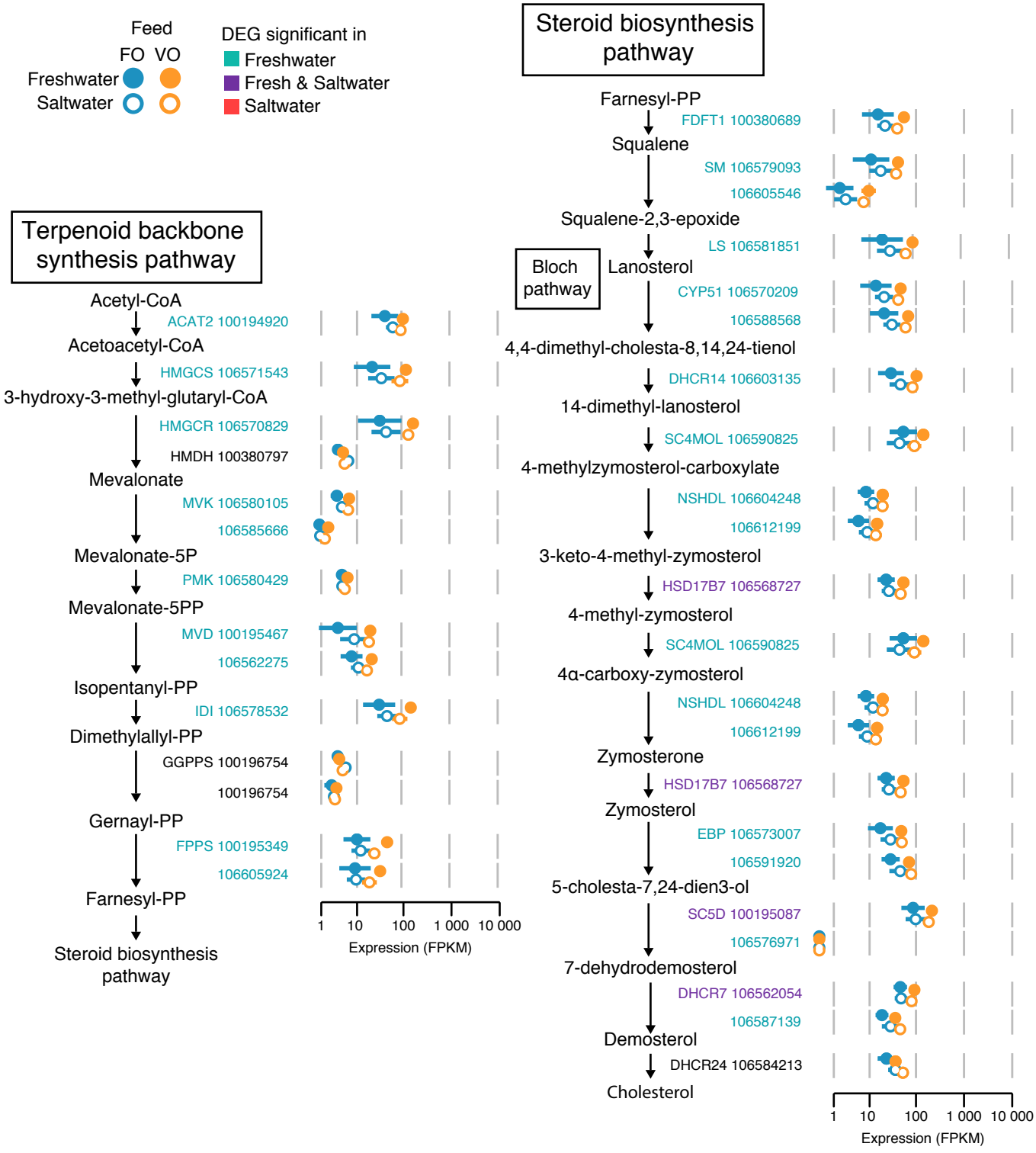
(a) Atlantic salmon fry were reared in 4 feeding tanks containing freshwater; 2 continuously fed fish oil (FO) and 2 vegetable oil (VO). A feed switch involved the transfer of fish from one tank to an empty partition of another tank fed the opposite diet. After smoltification fish from FO and VO tanks were transferred to 4 new feeding tanks containing saltwater and the feed switch was repeated. (b) Timeline of feed trial showing fish sampling and smoltification periods. Fish were sampled before (D0) and up to 20 days after the fresh- or saltwater feed switch. (c) Total RNA was sequenced from select fish tissue samples. The number of RNA-Seq replicates are shown for each, tissue, condition and time point.

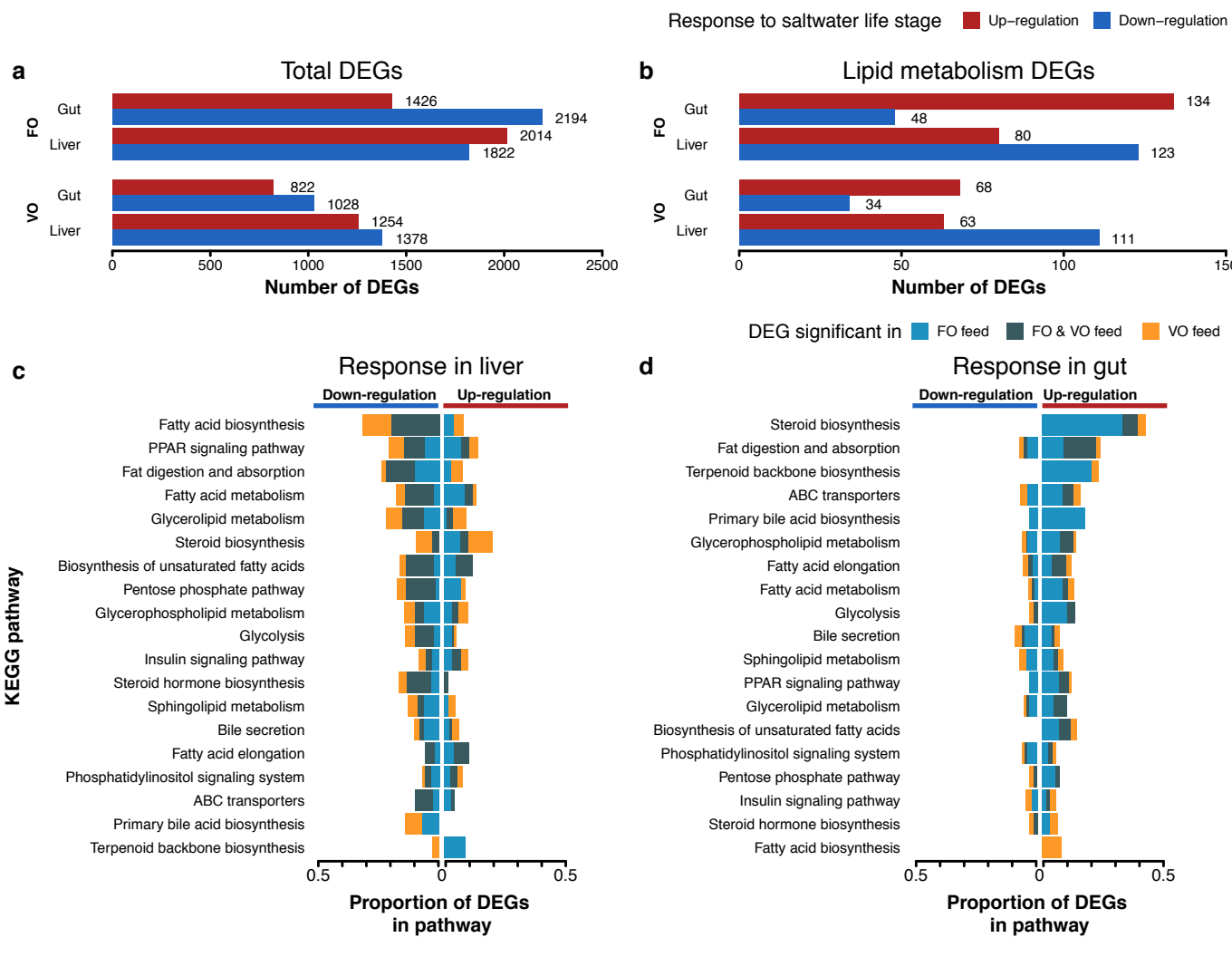

a
Ts3R
Ss4R
Danio rerio
Zebrafish
Gasterosteus aculeatus
Three-spined stickleback
Oryzias latipes
Medaka
Oncorhynchus mykiss
Rainbow trout
Oncorhynchus kisutch
Coho salmon
Salmo salar
Atlantic salmon
Thymallus thymallus
Grayling
Esox lucius
Northern pike
Mus musculus
House mouse
Homo sapiens
Human
500
400
300
200
100
0
Million years ago
b
Number of zebrafish genes
300
200
100
0
0
1
2
3
4
5+
Number of salmon orthologs per zebrafish gene

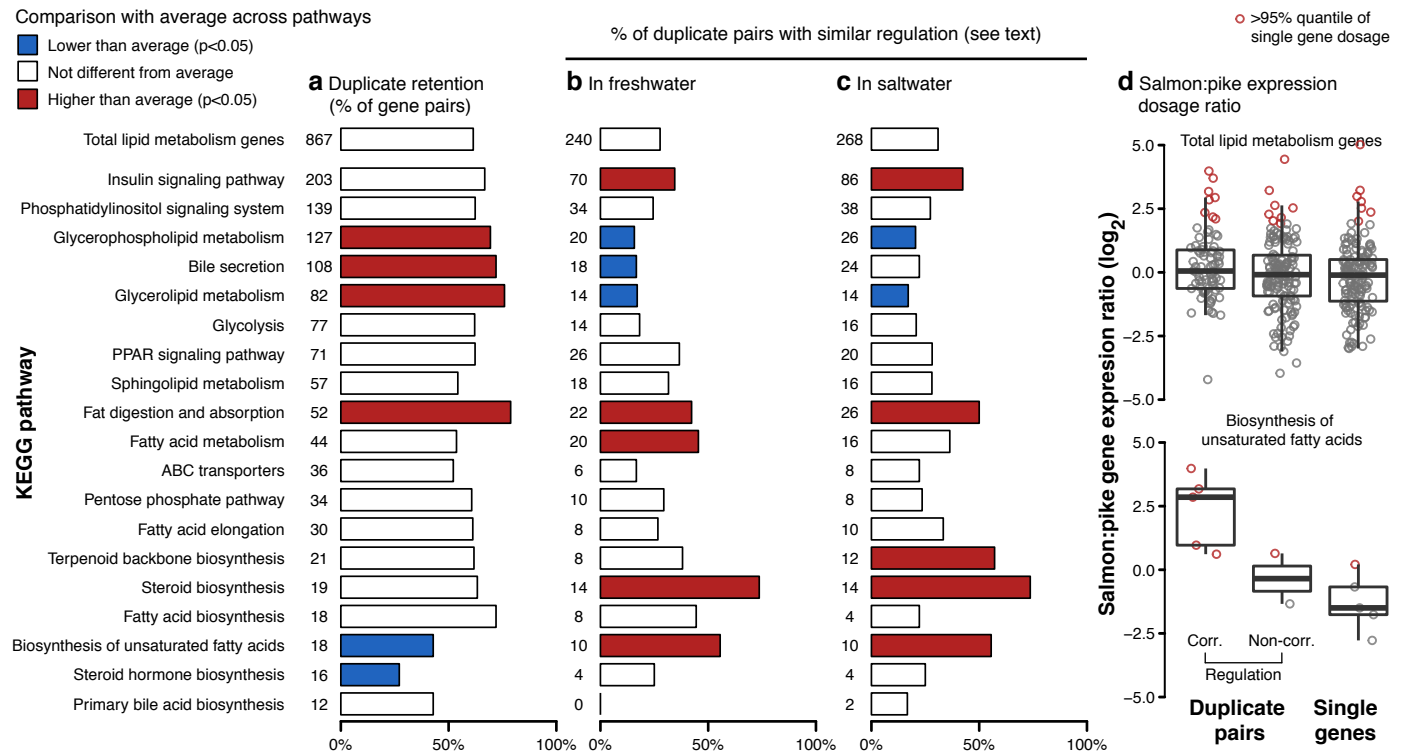

Biosynthesis of unsaturated fatty acids
Δ5FAD
Δ6FAD
ELOVL5a
ELOVL5b
Steroid biosynthesis
LS
SMa
SMb
PPAR signaling pathway
LXRα
PPARα
SREBP1α
SREBP2α
log2 fold change
15
10
5
0
−5
−10
liver
pyloric caeca
heart
brain
gut
kidney
muscle
gill
eye
skin
ovary
nose
testis
head kidney
spleen
Tissue


Response to VO feed
Up−regulation
Down−regulation
a
Total DEGs
Freshwater
Saltwater
Gut
Liver
31
154
321
176
5
3
242
163
0
100
200
300
Number of DEGs
b
Lipid metabolism DEGs
2
2
57
9
1
23
8
0
20
40
60
Number of DEGs
c
DEG significant in
Freshwater
Fresh & Saltwater
Saltwater
KEGG pathway
Steroid biosynthesis
Terpenoid backbone biosynthesis
Fatty acid biosynthesis
Fatty acid metabolism
Pentose phosphate pathway
Biosynthesis of unsaturated fatty acids
Fat digestion and absorption
Steroid hormone biosynthesis
PPAR signaling pathway
Primary bile acid biosynthesis
Fatty acid elongation
Glycerolipid metabolism
Bile secretion
Glycerophospholipid metabolism
Glycolysis
Insulin signaling pathway
ABC transporters
Sphingolipid metabolism
Phosphatidylinositol signaling system
0.00
0.25
0.50
0.75
1.00
Proportion of DEGs in pathway

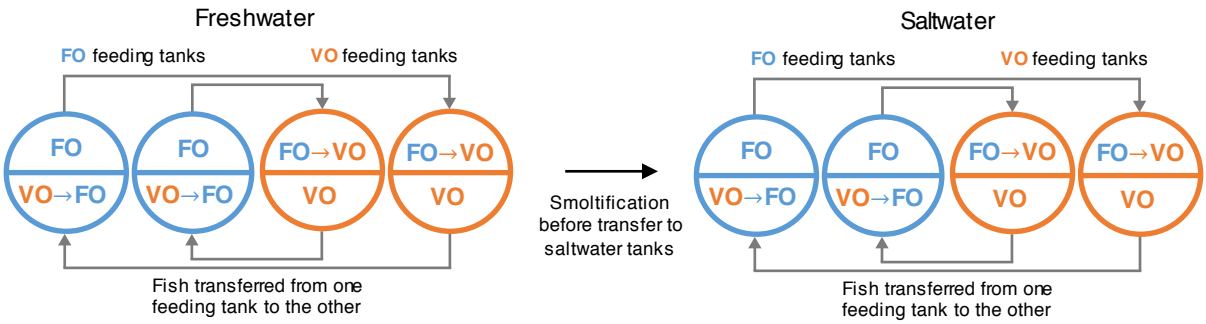

Feed
FO VO
Freshwater
Saltwater
DEG significant in
Freshwater
Fresh & Saltwater
Saltwater
FA biosynthesis pathway
Acetyl-CoA
ACACa 106568086
106603271
ACACb 106580110
106585658
Malonyl-Coa
FAS 106589612
106610271
16:0
SCD 106562327
SCD 106568335
SCDb 100194951
106577210
106577394
106608500
16:1n-7
1
10
100
1 000
10 000
Gene expression (FPKM)
LC-PUFA biosynthesis pathway
18:3n-3
Δ6FADa 100136441
Δ6FADb 106584797
Δ6FADc 100329172
18:4n-3
ELOVL5a 100136433
ELOVL5b 100192340
20:4n-3
Δ5FAD 100136383
20:5n-3 (EPA)
ELOVL2 100192341
ELOVL5a 100136433
ELOVL5b 100192340
22:5n-3
ELOVL2 100192341
ELOVL4 100500788
24:5n-3
ELOVL4 100500788
VLC-PUFA
Δ6FADa 100136441
Δ6FADb 106584797
Δ6FADc 100329172
24:6n-3
Beta oxidation
22:6n-3 (DHA)
1
10
100
1 000
10 000
Gene expression (FPKM)


Feed
FO VO
Freshwater
Saltwater
DEG significant in
Freshwater
Fresh & Saltwater
Saltwater
Terpenoid backbone synthesis pathway
Acetyl-CoA
ACAT2 100194920
Acetoacetyl-CoA
HMGCS 106571543
3-hydroxy-3-methyl-glutaryl-CoA
HMGCR 106570829
HMDH 100380797
Mevalonate
MVK 106580105
106585666
Mevalonate-5P
PMK 106580429
Mevalonate-5PP
MVD 100195467
106562275
Isopentanyl-PP
IDI 106578532
Dimethylallyl-PP
GGPPS 100196754
100196754
Gernayl-PP
FPPS 100195349
106605924
Farnesyl-PP
Steroid biosynthesis pathway
1 10 100 1 000 10 000
Expression (FPKM)
Steroid biosynthesis pathway
Farnesyl-PP
FDFT1 100380689
Squalene
SM 106579093
106605546
Squalene-2,3-epoxide
LS 106581851
Bloch pathway
Lanosterol
CYP51 106570209
106588568
4,4-dimethyl-cholesta-8,14,24-tienol
DHCR14 106603135
14-dimethyl-lanosterol
SC4MOL 106590825
4-methylzymosterol-carboxylate
NSHDL 106604248
106612199
3-keto-4-methyl-zymosterol
HSD17B7 106568727
4-methyl-zymosterol
SC4MOL 106590825
4α-carboxy-zymosterol
NSHDL 106604248
106612199
Zymosterone
HSD17B7 106568727
Zymosterol
EBP 106573007
106591920
5-cholesta-7,24-dien3-ol
SC5D 100195087
106576971
7-dehydrodemosterol
DHCR7 106562054
106587139
Demosterol
DHCR24 106584213
Cholesterol
1 10 100 1 000 10 000
Expression (FPKM)

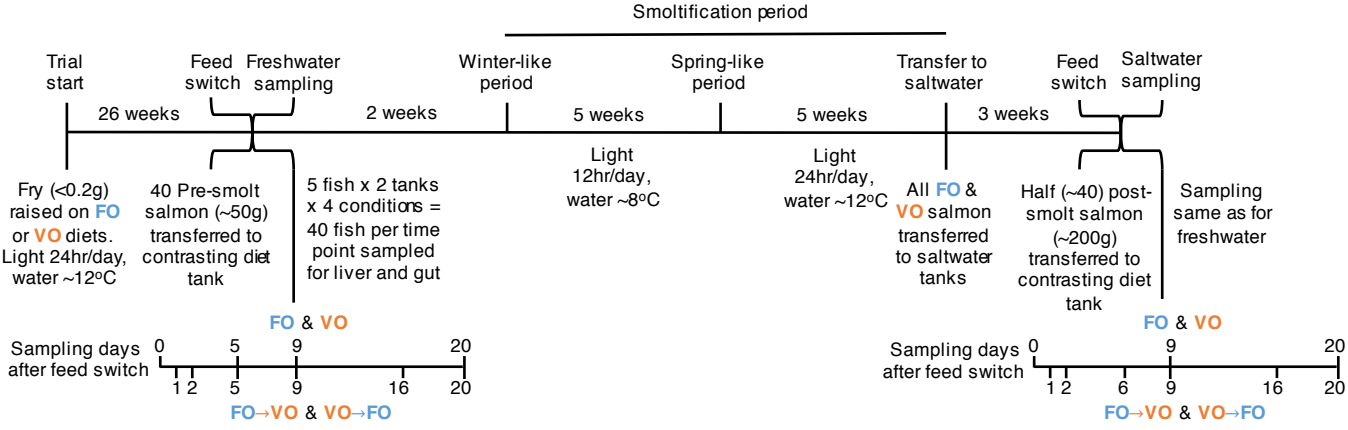

Response to saltwater life stage
Up−regulation
Down−regulation
a
Total DEGs
FO
Gut
Liver
VO
Gut
Liver
1426
2194
2014
1822
822
1028
1254
1378
0
500
1000
1500
2000
2500
Number of DEGs
b
Lipid metabolism DEGs
FO
Gut
Liver
VO
Gut
Liver
134
48
80
123
68
34
63
111
0
50
100
150
Number of DEGs
DEG significant in
FO feed
FO & VO feed
VO feed
c
Response in liver
Down-regulation
Up-regulation
KEGG pathway
Fatty acid biosynthesis
PPAR signaling pathway
Fat digestion and absorption
Fatty acid metabolism
Glycerolipid metabolism
Steroid biosynthesis
Biosynthesis of unsaturated fatty acids
Pentose phosphate pathway
Glycerophospholipid metabolism
Glycolysis
Insulin signaling pathway
Steroid hormone biosynthesis
Sphingolipid metabolism
Bile secretion
Fatty acid elongation
Phosphatidylinositol signaling system
ABC transporters
Primary bile acid biosynthesis
Terpenoid backbone biosynthesis
0.5
0
0.5
Proportion of DEGs in pathway
d
Response in gut
Down-regulation
Up-regulation
Steroid biosynthesis
Fat digestion and absorption
Terpenoid backbone biosynthesis
ABC transporters
Primary bile acid biosynthesis
Glycerophospholipid metabolism
Fatty acid elongation
Fatty acid metabolism
Glycolysis
Bile secretion
Sphingolipid metabolism
PPAR signaling pathway
Glycerolipid metabolism
Biosynthesis of unsaturated fatty acids
Phosphatidylinositol signaling system
Pentose phosphate pathway
Insulin signaling pathway
Steroid hormone biosynthesis
Fatty acid biosynthesis
0.5
0
0.5
Proportion of DEGs in pathway


Comparison with average across pathways
Lower than average (p<0.05)
Not different from average
Higher than average (p<0.05)
% of duplicate pairs with similar regulation (see text)
>95% quantile of single gene dosage
a Duplicate retention (% of gene pairs)
b In freshwater
c In saltwater
d Salmon:pike expression dosage ratio
KEGG pathway
Total lipid metabolism genes 867 240 268
Insulin signaling pathway 203 70 86
Phosphatidylinositol signaling system 139 34 38
Glycerophospholipid metabolism 127 20 26
Bile secretion 108 18 24
Glycerolipid metabolism 82 14 14
Glycolysis 77 14 16
PPAR signaling pathway 71 26 20
Sphingolipid metabolism 57 18 16
Fat digestion and absorption 52 22 26
Fatty acid metabolism 44 20 16
ABC transporters 36 6 8
Pentose phosphate pathway 34 10 8
Fatty acid elongation 30 8 10
Terpenoid backbone biosynthesis 21 8 12
Steroid biosynthesis 19 14 14
Fatty acid biosynthesis 18 8 4
Biosynthesis of unsaturated fatty acids 18 10 10
Steroid hormone biosynthesis 16 4 4
Primary bile acid biosynthesis 12 0 2
0% 50% 100%
Total lipid metabolism genes
Biosynthesis of unsaturated fatty acids
Salmon:pike gene expresion ratio (log2)
5.0
2.5
0.0
-2.5
-5.0
Corr.
Non-corr.
Regulation
Duplicate pairs
Single genes

a Feeding tank setup



b Sampling timeline



c Samples sequenced for total RNA

		Freshwater sampling days							Saltwater sampling days						
Feed	Tissue	0	1	2	5	9	16	20	0	1	2	6	9	16	20
FO	Liver	8			8	8		8	8				8		8
	Gut	4				4		4	8				8		8
VO	Liver	8			8	8		8	8				8		8
	Gut	4				4		4	7				8		8
VO→FO	Liver		4	4	4	4	4	4		4	4	4	4	4	4
	Gut		4	4	4	4	4	4		4	4	4	4	4	4
FO→VO	Liver		4	4	4	4	4	4		4	4	4	4	4	4
	Gut		4	4	4	4	4	4		4	4	4	4	4	4