

This is the peer reviewed version of the following article: Gillard G, Harvey TN, Gjuvsland A, et al. Life-stage-associated remodelling of lipid metabolism regulation in Atlantic salmon. *Molecular Ecology* 2018;27:1200–1213, which has been published in final form at <https://doi.org/10.1111/mec.14533>. This article may be used for non-commercial purposes in accordance With Wiley Terms and Conditions for self-archiving.

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

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

\* shared first authors

† shared corresponding authors

## *Affiliations:*

<sup>1</sup> *Faculty of Chemistry, Biotechnology and Food Sciences, Norwegian University of Life Sciences, NO-1432 Ås, Norway*

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

<sup>3</sup> *Department of Biology, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway.*

<sup>4</sup> *Department of Animal and Aquacultural Sciences, Faculty of Biosciences, Norwegian University of Life Sciences, NO-1432 Ås, Norway*

<sup>5</sup> *Institute of Aquaculture, School of Natural Sciences, University of Stirling, Pathfoot Building, Stirling FK9 4LA, Scotland, UK*

<sup>6</sup> *AquaGen AS, NO-1432 Ås, Norway*

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

## **Corresponding authors:**

Simen Rød Sandve, [simen.sandve@nmbu.no](mailto:simen.sandve@nmbu.no), +47 94870082, Jon Olav Vik, [jonovik@gmail.com](mailto:jonovik@gmail.com), +47 45882998

**Running title:** Adaptive remodeling of lipid metabolism

## 25 Abstract

26 Atlantic salmon migrates from rivers to sea to feed, grow and develop gonads before returning to spawn in  
27 freshwater. The transition to marine habitats is associated with dramatic changes in the environment,  
28 including water salinity, exposure to pathogens and shift in dietary lipid availability. Many changes in  
29 physiology and metabolism occur across this life-stage transition, but little is known about the molecular  
30 nature of these changes. Here, we use a long-term feeding experiment to study transcriptional regulation of  
31 lipid metabolism in Atlantic salmon gut and liver in both fresh- and saltwater. We find that lipid metabolism  
32 becomes significantly less plastic to differences in dietary lipid composition when salmon transitions to  
33 saltwater and experiences increased dietary lipid availability. Expression of genes in liver relating to  
34 lipogenesis and lipid transport decreases overall and becomes less responsive to diet, while genes for lipid  
35 uptake in gut become more highly expressed. Finally, analyses of evolutionary consequences of the  
36 salmonid-specific whole-genome duplication on lipid metabolism reveal several pathways with  
37 significantly different ( $p < .05$ ) duplicate retention or duplicate regulatory conservation. We also find a  
38 limited number of cases where the whole-genome duplication has resulted in an increased gene dosage. In  
39 conclusion, we find variable and pathway-specific effects of the salmonid genome duplication on lipid  
40 metabolism genes. A clear life-stage-associated shift in lipid metabolism regulation is evident, and we  
41 hypothesize this to be, at least partly, driven by nondietary factors such as the preparatory remodelling of  
42 gene regulation and physiology prior to sea migration.

43

## 44 Introduction

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

51  
52 A key difference between freshwater and sea-habitats is the dietary availability of essential long-chain  
53 polyunsaturated fatty acids. Salmon in rivers mostly eat invertebrates that are low in physiologically critical  
54 n-3 and n-6, 20 and 22 carbon long-chain polyunsaturated fatty acids (n-3LC-PUFA and n-6LC-PUFA),  
55 arachidonic acid (20:4n-6), eicosapentaenoic acid (20:5n-3) and docosahexaenoic (22:6n-3), while marine  
56 habitat food chains are high in available LC-PUFAs. Possibly as an adaptation to this (Leaver et al., 2008),  
57 salmon have evolved a high capacity for endogenous production of LC-PUFAs by elongation and  
58 desaturation of essential dietary 18 carbon precursor linoleic and linolenic acids (18:2n-6 and 18:3n-3;  
59 Figure 4) and the ability to increase or decrease endogenous production of n-3LC-PUFA from dietary  
60 18:3n-3 (linolenic acid) as a response to the dietary availability (Kennedy et al., 2006; Leaver et al., 2008;  
61 Morais et al., 2011; Ruyter et al., 2000; Tocher et al., 2001; Tocher et al., 2002; Zheng et al., 2005). During  
62 smoltification and after sea-migration, Atlantic salmon have been shown to undergo transformation of lipid  
63 metabolism function, by decreasing lipid syntheses and increasing lipid breakdown (Sheridan, 1989).  
64 However, very little is known about the molecular nature of this life-stage associated transformation  
65 physiological function.

66

67 The evolution of novel traits in salmonids, such as increased plasticity and the ability to migrate to sea, may  
68 have been facilitated by their ancestral whole genome duplication (called Ss4R) some 80 million years ago  
69 (Allendorf & Thorgaard, 1984; Lorgen et al., 2015; Macqueen & Johnston, 2014; Robertson et al., 2017).  
70 Gene duplication can give rise to new adaptive phenotypes in different ways: through evolution of novel  
71 functions or gene regulation, subdivision and/or specialization of function among duplicates, or via an  
72 adaptive increase in gene dosage. The Atlantic salmon genome contains ~10,000 pairs of Ss4R gene  
73 duplicates, of which ~50% have evolved some novel regulation (Lien et al., 2016; Robertson et al., 2017).  
74 Indeed, in the context of lipid metabolism, it has recently been shown that a Ss4R duplicate of *elov15*, a key  
75 enzyme in LC-PUFA syntheses, has gained expression compared to its ancestral regulation with likely  
76 implications for the ability to synthesize LC-PUFAs (Carmona-Antoñanzas et al., 2016). This is believed  
77 to have facilitated evolution of novel traits, including flexible phenotypes necessary for an anadromous life  
78 history (Stefansson et al., 2008). However, no systematic genome wide study has yet been conducted to  
79 assess the importance of the Ss4R in evolution of salmon lipid metabolism.

80

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

## 88 Materials and methods

### 89 Orthogroup prediction

90 Protein sequences were obtained from seven teleost fish species; *Danio rerio* (zebrafish), *Gasterosteus*  
91 *aculeatus* (three-spined stickleback), *Oryzias latipes* (medaka), *Oncorhynchus mykiss* (Rainbow trout),  
92 *Oncorhynchus kisutch* (coho salmon), *Salmo salar* (Atlantic salmon), *Thymallus thymallus* (grayling), *Esox*  
93 *lucius* (northern pike), and two mammalian outgroup species; *Homo sapiens* (human), *Mus musculus* (house  
94 mouse). Human, mouse, zebrafish, medaka and stickleback protein fasta data were obtained from  
95 ENSEMBL (release 83). Atlantic salmon (RefSeq assembly GCF\_000233375.1, Annotation Release 100)  
96 and northern pike (RefSeq assembly GCF\_000721915.2, Annotation Release 101) proteins were obtained  
97 from NCBI RefSeq. Rainbow trout proteins were obtained from an assembly and annotation of the genome  
98 (Berthelot et al., 2014). Grayling proteins were obtained from an assembly and annotation of the genome  
99 (Varadharajan et al., 2017). The coho salmon transcriptome (Kim, Leong, Koop, & Devlin, 2016) was  
100 obtained from NCBI (GDQG00000000.1). Where transcriptome data was used, protein sequences were  
101 translated using TransDecoder (v2.0.1, <http://transdecoder.github.io/>). Protein fasta files were filtered to  
102 retrieve only the longest protein isoform per gene. Orthofinder (v0.2.8) (Emms et al., 2015) assigned groups  
103 of orthologs based on protein sequence similarity. Proteins within an orthogroups were further aligned using  
104 MAFFT (v7.130) (Katoh et al., 2002) and maximum likelihood trees were estimated using FastTree (v2.1.8)  
105 (Price et al., 2010).

### 106 Annotation of salmon lipid metabolism genes

107 A list of zebrafish proteins obtained from 19 manually selected zebrafish KEGG pathways related to lipid  
108 metabolism (Table S1) were used to search for Atlantic salmon orthologs. Orthogroups that contained a  
109 selected zebrafish protein were identified. Salmon proteins within those orthogroups were assigned as

110 orthologs of the closest zebrafish protein based on the orthogroup tree distance. A lipid metabolism gene  
111 list was created including salmon orthologs to the selected zebrafish genes. Additional salmon genes related  
112 to lipid metabolism not included in KEGG pathways (e.g. regulators or transporters, SREBP, LXR, FABP,  
113 etc.) were manually searched for through NCBI and added to the list.

## 114 Tissue expression

115 Atlantic salmon RNA-seq samples from 15 different tissues (liver, gut, pyloric caeca, heart, kidney, muscle,  
116 gill, eye, skin, ovary, nose, testis, brain, head kidney, spleen) were obtained from NCBI SRA  
117 (PRJNA72713) (Lien et al., 2016). Fastq files were adapter trimmed before alignment to the Atlantic  
118 salmon genome (RefSeq assembly GCF\_000233375.1) (Lien et al., 2016) using STAR (v2.5.2a) (Dobin et  
119 al., 2013). HTSeq-count (v0.6.1p1) (Anders et al., 2015) counted the sum of uniquely aligned reads in exon  
120 regions of each gene in the annotation (RefSeq Annotation Release 100). Gene FPKM values were  
121 calculated based on the gene count over the samples effective library size (see TMM method from edgeR  
122 (Robinson et al., 2010) user manual) and the mean gene transcript isoform length.

## 123 Feed trial

124 Atlantic salmon fry were obtained from AquaGen Breeding Centre, Kyrksæterøra, Norway and reared in  
125 the Norwegian Institute for Water Research (NIVA), Solbergstranda, Norway in four partitioned 1000  
126 liter tanks on vegetable oil (VO) or fish oil (FO) based diets continuously from first feeding (fry weight  
127 <0.2 g). Daily feed amount was calculated based on total biomass in each tank and decreased as the fish  
128 grew, from 3% at first feeding to 1.2% by the end of the trial. Fish were euthanized periodically  
129 throughout the experiment to maintain appropriate levels of dissolved oxygen. VO based feeds contained  
130 a combination of linseed oil and palm oil at a ratio of 1.8:1 and FO based feeds contained only North  
131 Atlantic fish oil. All feeds were formulated and produced by EWOS innovation (Supplementary File 3).  
132 Local groundwater was UV sterilized for use in the freshwater life-stage and water from the Oslofjord

133 taken from 60 meters below sea surface (~3-3.5% salinity) was UV sterilized for use in the saltwater life-  
134 stage. Fish were raised under constant light and water temperature (~12°C) for 26 weeks. Then, 40 pre-  
135 smolt salmon (~50g) from each control tank (~240 fish per control tank) were switched to the contrasting  
136 diet (VO to FO and vice versa) by physically moving them to the empty partition of the tank receiving the  
137 appropriate feed (Figure 8a). Five fish from each of the control tanks (2 VO tanks and 2 FO tanks) were  
138 sampled before switching feeds (D0), then fish from both control and feed switch conditions were  
139 similarly sampled 1, 2, 5, 9, 16, and 20 days after switching feeds (5 fish x 2 replicate tanks x 4  
140 conditions = 40 fish per time point, figure 8b). Two weeks after freshwater sampling (31 weeks after first  
141 feeding), smoltification was triggered by 5 weeks of winter-like conditions with decreased light (12 hours  
142 per day) and water temperature (~8°C), immediately followed by 5 weeks of spring-like conditions,  
143 returning to normal light (24 hours per day) and water temperature (~12°C). All salmon from the control  
144 groups (VO or FO) were then switched to saltwater and allowed to acclimate for 3 weeks. The feed  
145 switch was repeated in saltwater by transferring half (~40 fish) of the post-smolt salmon (~200 g) from  
146 each control tank to the contrasting feed condition. Again, pre-switch control samples were taken (D0)  
147 followed by sampling 1, 2, 6, 9, 16, and 20 days post-diet switch (Figure 8b). For both freshwater and  
148 saltwater samplings, feeding was stopped in the mornings of each of the sampling days. All fish were  
149 euthanized by a blow to the head and samples of liver and midgut (gut section between pyloric caeca and  
150 hindgut) were flash frozen in liquid nitrogen and stored under -80 °C. A subset of the samples taken were  
151 used for further RNA-seq analysis (see figure 8c for details).

## 152 RNA-sequencing

153 Total RNA was extracted from selected feed trial samples (see figure 8c for details) using the RNeasy Plus  
154 Universal kit (QIAGEN). Quality was determined on a 2100 Bioanalyzer using the RNA 6000 nano kit  
155 (Agilent). Concentration was determined using a Nanodrop 8000 spectrophotometer (Thermo Scientific).  
156 cDNA libraries were prepared using the TruSeq Stranded mRNA HT Sample Prep Kit (Illumina). Library



157 mean length was determined by running on a 2100 Bioanalyzer using the DNA 1000 kit (Agilent) and  
158 library concentration was determined with the Qbit BR kit (Thermo Scientific). Single end sequencing of  
159 sample libraries was completed on an Illumina HiSeq 2500 with 100 bp reads.

## 160 Differential expression analysis between feed conditions and life-stages

161 To analyze gene expression differences between feed conditions and life-stages, samples from the feed trial  
162 were selected for RNA-seq. Liver and gut tissue RNA were sequenced from fish fed each of the feeds (FO,  
163 VO) at day 0 of the diet switch, both before (freshwater) and after (saltwater) smoltification (See figure 8c  
164 for the number of RNA-seq replicates and sampling details). Fastq files were processed to produce gene  
165 count and FPKM data using the same protocol described under the tissue expression method section. For  
166 the feed comparison, changes in gene expression were tested between FO and VO feed conditions for both  
167 freshwater and saltwater samples, and liver and gut tissues. For the life-stage comparison, changes in gene  
168 expression were tested between freshwater and saltwater stages for both FO and VO feed conditions, and  
169 liver and gut tissues. Using RNA-seq gene count data, lowly expressed genes were filtered prior to testing,  
170 retaining genes with a minimum of one read count per million (CPM) in two or more samples. Differential  
171 expression analysis was carried out using a standard edgeR (Robinson et al., 2010) protocol. Effective  
172 library sizes were calculated using the edgeR TMM-normalisation procedure allowing effective comparison  
173 of expression data between different sample types (see edgeR user manual). An exact test between  
174 expression levels of a pair of conditions gave the log<sub>2</sub> fold change, P-value and false discovery rate (FDR)  
175 for each gene. Genes with FDR < 0.05 were considered differentially expressed genes (DEGs).

## 176 Identification of Ss4R duplicates

177 To identify putative gene duplicates stemming from the Ss4R, we used the same approach as in Lien et al.  
178 (2016). All-vs-all protein blast was run with e-value cutoff of 1e-10 and pident (percentage of identical  
179 matches) ≥80 and blast hit coverage of ≥50% of protein length. Only the best protein hits between the 98

180 defined syntenic blocks (see Lien et al., 2016) were considered as putative Ss4R duplicates. Blast result  
181 ranking was done using the product of pident times bitscore to avoid spurious ‘best blast matches’ with low  
182 pident (<85) but high bitscore.

## 183 Duplicate analysis

184 Genes from the lipid metabolism gene list were paired together with their putative Ss4R duplicates  
185 identified above. The retention of gene duplicates (i.e. whether both genes in a pair were retained, or just  
186 one) was compared between all identified duplicates in the salmon genome annotation and the lipid  
187 metabolism gene list. Pathway-level retention was explored by comparing the number of genes in each of  
188 the 19 selected KEGG pathways (Table S1) in a duplicate pairing to that of the total list of lipid genes, to  
189 find pathways with significantly less or more duplicate retention (Fisher’s exact test, P-value < 0.05).  
190 Regulatory conservation of lipid gene duplicates was explored by correlation of gene expression changes  
191 between duplicates over the course of the feed trial described above. RNA-seq data was generated from  
192 liver samples of salmon from 38 sampling time points (19 in freshwater and 19 in saltwater). Fastq files  
193 were processed to produce gene count and FPKM data using the same protocol described under the tissue  
194 expression method section. For each duplicate pair, mean FPKM values were retrieved for each time point  
195 and used to calculate a freshwater and saltwater correlation value. Duplicates with Pearson correlation  $\geq$   
196 0.6 were considered correlated (P-value < 0.003 from 19 sample points). The number of duplicates with  
197 correlated expression profiles was counted for each pathway and compared to all lipid genes to find  
198 pathways with significantly less or more correlated duplicates (Fisher’s exact test, P-value < 0.05). The  
199 effect of gene duplication on gene dosage was estimated by calculating a dosage ratio between the FPKM  
200 value of a salmon ortholog (sum of gene expression in duplicate pairs) over the FPKM value of the non-  
201 duplicated ortholog from northern pike. For salmon, the RNA-seq data from the freshwater and saltwater  
202 FO feed trial was used (samples used in differential expression analysis section). For pike, RNA-seq from  
203 livers of four individuals were aligned (see tissue expression section for protocol) to their respective

204 genomes (see genomes in ortholog prediction section). RSEM (v1.2.31) (Li & Dewey, 2011) was used to  
205 generate FPKM values for genes so that non-uniquely mapped reads between salmon duplicate genes were  
206 not ignored but instead assigned proportionately to each gene to match the proportions of uniquely mapped  
207 reads between the genes. Gene dosage levels for duplicate pairs with correlated expression (see above),  
208 non-correlated expression and single genes were compared for all lipid metabolism genes and for each  
209 pathway.

210

## 211 Results and discussion

### 212 Annotation of lipid metabolism genes

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

226

227 To validate our ortholog annotation pipeline used to identify lipid metabolism genes, we analyzed the tissue  
228 specificity of these genes using gene expression data from 15 tissues (File S3) of Atlantic salmon (Lien et  
229 al., 2016). Genes in certain fatty acid metabolism related pathways (*'fatty acid metabolism'*, *'PPAR*  
230 *signaling pathway'*, *'fat digestion and absorption'*) had higher overall expression in tissues known to have  
231 high lipid metabolism activity (i.e. pyloric caeca, liver, heart, and brain) (Glatz et al., 2010; Benedito-Palos  
232 & Pérez-Sánchez, 2016; Tocher, 2003) (Figure 2). Examples include: 1) Liver was the site of highest  
233 expression for all genes in the LC-PUFA biosynthesis pathway (the desaturases  $\Delta 6$ FAD and  $\Delta 5$ FAD, and  
234 the elongases elovl5, elovl2 and elovl4). 2) Bile acids are essential for fat digestion in the gut, but are  
235 synthesized in liver. As expected, the rate limiting step for bile syntheses, cytochrome P450 7A1  
236 (CYP7A1), has the highest expression in the liver. 3) Cholesterol, an essential component of cell  
237 membranes and precursor to bile acids, is known to be synthesized in all tissues, but primarily in liver,  
238 intestine, and brain (Brown & Sharpe, 2016). This is reflected in our annotation by high expression of the  
239 key cholesterol biosynthesis genes 3-hydroxy-3-methyl-glutaryl-CoA reductase (HMGCR), isopentenyl-  
240 diphosphate  $\Delta$ isomerase (IDI1), squalene epoxidase (SM), and lanosterol synthase (LS) in these tissues. 4)  
241 Several known regulators of lipid metabolism show high expression in liver, heart, brain and pyloric caeca,  
242 as expected, including liver X receptor (LXR), peroxisome proliferator-activated receptor alpha (PPAR $\alpha$ ),  
243 sterol regulatory element binding protein 1 (SREBP1), and sterol regulatory element binding protein 2  
244 (SREBP2) . Taken together, the tissue distribution of lipid metabolism gene expression is in line with  
245 knowledge about vertebrate physiology in general, and support the validity of our annotation of lipid  
246 metabolism genes in salmon. To make all data underlying our annotation easily available, and to facilitate  
247 further refinement through manual community curation, we have created an interactive web-server  
248 available online ([goo.gl/8Ap89a](http://goo.gl/8Ap89a)).

249

## 250 Life-stage dependent remodeling of lipid metabolism

251 We conducted a feeding trial to study how salmon adjusts its lipid metabolism to different levels of LC-  
252 PUFA in freshwater and saltwater (See figure 8 for experimental details). Groups of salmon were fed  
253 contrasting diets from hatching until after transition to seawater. One feed was vegetable oil based (VO)  
254 and hence low in LC-PUFA, similar to river ecosystem diets, whereas the other was based on fish oil (FO)  
255 and high in LC-PUFA as expected in a marine-type diet (see Table S2 for details on feed composition). VO  
256 based diets are also low in cholesterol (Ciftci, et al., 2012; Verleyen et al., 2002). In total, 32 and 23 fish  
257 were sampled for RNA-seq of liver and gut, respectively, including up to eight biological replicates from  
258 each diet and life-stage (freshwater and saltwater, see figure 8c for details). Fish in the different dietary  
259 groups were given FO and VO feed from first feeding (<0.2 g body weight) until sampling.

260

261 In general, global gene expression levels were more affected by dietary composition in liver than in gut  
262 (which was largely unresponsive), and the effect was more pronounced in freshwater than in saltwater  
263 (Figure 3a). VO diets, compared to FO diets, increased lipid-metabolism related gene expression in liver.  
264 In freshwater, 66 genes were differentially expressed with 57 (86%) of these upregulated, while in saltwater  
265 31 genes were differentially expressed with 23 (74%) of these upregulated (Figure 3b). The increased  
266 activity of liver lipid metabolism under VO diets confirm the well-known ability of salmon to regulate  
267 endogenous synthesis of LC-PUFA and cholesterol in response to VO diets (Kortner et al., 2014; Leaver et  
268 al., 2008; Zheng et al., 2005).

269

270 Fish sampled in freshwater and saltwater shared a relatively small number of differentially expressed genes  
271 (DEGs) for each pathway (Table S3). We found that most pathways had more DEGs in freshwater (*fatty*  
272 *acid biosynthesis*, *steroid biosynthesis*, and its precursor *terpenoid backbone biosynthesis*), whereas  
273 few had more DEGs in saltwater (*fat digestion and absorption* and *steroid hormone biosynthesis*)  
274 (Figure 3c). Of 87 lipid metabolism DEGs in the dietary contrast, 56 (64%) were freshwater specific, 21

275 (24%) saltwater specific, and 10 (11%) shared dietary response. For example, only two genes in the FA and  
276 LC-PUFA biosynthesis pathways (D6FADa and D5FAD) shared response to diet in freshwater and  
277 saltwater (Figure 4). Similarly, in the pathways responsible for cholesterol biosynthesis, there were more  
278 DEGs between diets in FW (21 DEGs in FW, 4 shared and no SW-specific; Figure 5). The few genes that  
279 showed diet effects specific to saltwater included bile salt-activated lipase, responsible for the hydrolysis  
280 of free fatty acids from TAG obtained from the diet (Tocher, 2003). Two of these genes, carboxyl ester  
281 lipase, tandem duplicate 2a (CEL2a) and b (CEL2b), are highly upregulated in saltwater in response to VO  
282 diet. Taken together, our results show higher metabolic plasticity in parr stage salmon, suggesting a life-  
283 stage-associated remodelling of lipid metabolism in liver. This corroborates the idea of a postsmoltification  
284 phenotype adapted to an environment with a surplus of n-3LCPUFA.

285

286 To further investigate the life-stage-associated changes in lipid metabolism, we tested for differential  
287 expression between salmon in freshwater and saltwater fed diets with identical n-3LC-PUFA profiles  
288 (Figure 6). Liver and gut showed contrasting effects of saltwater on lipid gene expression with extensive  
289 downregulation in liver and upregulation in gut (Figure 6b). The number of DEGs in each tissue was similar  
290 for the environment comparison (Figure 6a), unlike for the diet comparison (Figure 3).

291

292 Further examination of key lipid metabolism genes revealed that after life-stage transition, the system wide  
293 lipid metabolism remodeling represented a concerted shift in the metabolic role of liver and gut. As the  
294 salmon entered the marine stage, lipogenic gene expression in the liver was significantly decreased, as  
295 evident by the markedly lower expression (2.2-3.3 fold) of the master regulator of lipid metabolism  
296 SREBP1, a 5-fold decrease in expression of fatty acid synthase, and a 2-3 fold decrease in rate-limiting  
297 enzymes in LC-PUFA synthesis (i.e.  $\Delta 5$ FAD,  $\Delta 6$ FADa) (Figure 4). Liver and gut gene expression also  
298 indicated increased catabolic activity in saltwater, with upregulation of the carnitine palmitoyltransferase 1  
299 and 2 genes, responsible for uptake of fatty acids into mitochondria for  $\beta$ -oxidation (Lehner & Quiroga,  
300 2016). Finally, expression of lipid transport genes shifted from liver to gut with the transition to seawater

301 (apolipoproteins, pathway "Fat digestion and absorption" in Figure 6). Four apolipoproteins (out of 11  
302 annotated) were differentially regulated in liver between different life-stages, with a 2.4-5 fold decrease in  
303 saltwater compared to freshwater. In stark contrast, nine of the diet-regulated apolipoproteins in gut  
304 increased their expression in saltwater between 1.8-9.7 fold. The results point to an adaptive shift in lipid  
305 metabolism, with increased ability to take up lipids in the gut after Atlantic salmon migrates to sea where  
306 lipid availability is higher. Remodelling of lipid metabolism across life stages is likely the result of a  
307 combination of factors, including the direct regulatory effect of dietary fat itself, effect of salinity and  
308 smoltification-induced physiological changes influencing gene regulation. Although the relative  
309 importance of these factors is undetermined in our study, the fact that DEGs in the VO versus FO feed  
310 contrast were mostly lifestage- specific (Figure 3) supports that factors other than the diet itself contribute  
311 significantly to the freshwater and seawater metabolic phenotypes.

312

313 Interestingly, diet had a strong influence on the number and direction of gene expression changes between  
314 freshwater and saltwater (Figure 6). In gut, about twice as many DEGs (with respect to the fresh- to  
315 saltwater transition) were observed in salmon when fed FO diet than VO diet (Figure 6a). In liver, the diet  
316 effect was less pronounced, with the FO group containing 46% more DEGs than the VO group (Figure 6a).  
317 This diet effect pattern was reflected in the lipid metabolism genes with 89% and 16% more DEGs in the  
318 FO group for gut and liver, respectively (Figure 6b). As this diet and life-stage interaction is a genome wide  
319 trend, and more pronounced in gut tissue than in liver, this pattern could also be related to differences in  
320 osmoregulation and adaptation to saltwater. Two studies have suggested that Atlantic salmon raised on VO  
321 based feeds more closely resembling riverine diets adapt to saltwater sooner and better than salmon raised  
322 on FO based diets (Bell et al., 1997; Tocher et al., 2000). Conversely, there has been evidence that VO  
323 based diets can reduce markers for stress response upon saltwater challenge, resulting in reduced  
324 osmoregulatory capacity (Oxley et al., 2010). Regardless of the effect, it is clear that diet can modulate the  
325 smoltification process and could explain the discrepancy between diets in number of life-stage related

326 DEGs. Another possibility is that the different levels of fatty acids in the diets, for example DHA, affect  
327 DNA-methylation and thus trigger genome wide divergence in gene regulation (Kulkarni et al., 2011).

328

329 Our results clearly demonstrate very different baseline lipid metabolic functions in pre- and post-smolt  
330 salmon, as well as life-stage associated changes in the plasticity of lipid metabolism, e.g. the ability to  
331 regulate endogenous LC-PUFA synthesis as a response to changes in diet (i.e. fatty acid composition). As  
332 opportunistic carnivores, salmon tend to eat whatever the local environment provides. Thus, in freshwater,  
333 insects and amphipods provide variable, mostly low amounts of essential LC-PUFA and total fat (Jonsson  
334 & Jonsson, 2011; Sushchik et al., 2003), favoring a metabolic function that can efficiently regulate  
335 endogenous lipid synthesis based on dietary availability (Carmona-Antonanzas et al., 2014). Conversely,  
336 in marine environments, amphipods and smaller fish provide a higher, more stable source  
337 of n-3LC-PUFA and total fat (Baeza-Rojano, Hachero-Cruzado, & Guerra-Garcia, 2014; Jonsson &  
338 Jonsson, 2011), promoting a metabolic function that allocates less energy to endogenous synthesis  
339 of essential lipids.

340

## 341 Selection on gene duplicates after whole genome duplication

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

349



350 To assess the level of Ss4R duplicate retention, we first defined 10,752 Ss4R duplicate pairs (21,504 genes)  
351 in the NCBI refseq annotation using the same approach as Lien et al. (2016). Of the 1,421 annotated lipid  
352 metabolism genes, 867 (61%) were retained as duplicated genes after Ss4R (Figure 7a) (in contrast to 47%  
353 of the 45,127 salmon genes assigned to ortholog groups). Moreover, our results showed large variation in  
354 the proportion of retained duplicates in each lipid metabolism pathway (Figure 7), with the most extreme  
355 case being '*fat digestion and absorption*' with 80% retained duplicates and '*steroid hormone biosynthesis*'  
356 with only 27% retained Ss4R duplicates.

357

358 The regulatory conservation of the duplicates was then estimated by calculating co-expression correlation  
359 between Ss4R duplicates from RNA-seq data representing a time course of dynamic changes in gene  
360 expression and lipid metabolism function in liver. Fish in the same feeding trial were switched from VO to  
361 FO feed and vice versa, in both fresh and saltwater conditions (see Figure 8 for details). In total, 38 sampling  
362 time points (20 in freshwater and 18 in saltwater) from the feed switch experiment were used. Pathway-  
363 level analyses showed that regulatory conservation was not associated with duplicate retention (Figure 7).  
364 For example, the '*biosynthesis of unsaturated fatty acids*' pathway had significantly fewer duplicates  
365 retained than expected by chance (P-value < 0.0234), but a significant overrepresentation of duplicate pairs  
366 that display highly similar regulation (P-value < 0.0142 and < 0.0361 in freshwater and saltwater,  
367 respectively). Interestingly, the '*insulin signaling pathway*' also showed higher than expected duplicate co-  
368 regulation. This pathway has been shown to be important in regulating uptake and transport of FAs in  
369 adipose tissue, liver and muscle of Atlantic salmon (Sánchez-Gurmaches et al., 2011). Other pathways  
370 showing signatures of increased duplicate co-regulation were '*terpenoid backbone biosynthesis*', '*steroid*  
371 '*biosynthesis*', '*fat digestion and absorption*', and '*fatty acid metabolism*' (Figure 7b-c). Overall, the distinct  
372 differences in duplicate retention and conservation of regulatory mechanisms across the lipid metabolism  
373 pathways suggest differences in selective pressures shaping duplicate evolution following Ss4R. Moreover,  
374 the pathways with highly conserved duplicate co-regulation were also those that were most responsive to  
375 dietary differences in fatty acid composition (Figure 3).

376

377 Finally, to link duplicate retention and co-regulation to signals of increased gene dosage following Ss4R,  
378 we used RNA-seq data from the Northern pike (*Esox lucius*), a species that belongs to the unduplicated  
379 sister lineage (see methods for details). For each duplicate pair, we computed the ratio between the sum of  
380 Ss4R duplicate expression and its non-duplicated ortholog in pike and compared these ratios to those  
381 observed for salmon genes that had not retained two Ss4R duplicates. In total 69 duplicate pairs from 18  
382 different lipid-metabolism related pathways displayed a combined dosage increase relative to single copy  
383 genes, of which 26 had highly conserved regulation (i.e. correlated expression) (File S8). We saw no  
384 systematic effect of gene dosage when comparing the total gene expression of duplicate pairs with that of  
385 single-copy genes; nor did co-regulation of duplicates associate with increased gene dosage (Figure 7d).  
386 This pattern was also true for most individual lipid pathways (Figure S4-S5), except for '*biosynthesis of*  
387 '*unsaturated fatty acids*', '*fatty acid metabolism*' and '*fatty acid elongation*'. These three pathways showed  
388 a link between co-regulation of duplicated genes and higher total gene dosage (Figure S4-S5, Figure 7d).  
389 Underlying this link were three genes with co-regulated dosage effects shared between all three pathways;  
390 trifunctional enzyme alpha subunit b (hadhab), elovl6, and the previously identified elovl5 (Carmona-  
391 Antonanzas et al., 2014; Carmona-Antoñanzas et al., 2016). Only elovl5 is known to be directly involved  
392 in core PUFA biosynthesis. Hadhab is involved in mitochondrial  $\beta$ -oxidation/elongation and elovl6 is  
393 involved in elongation of saturated and monounsaturated fatty acids (Bond et al., 2016). Although we do  
394 not see a general trend of increased gene dosage effects on lipid metabolism genes after whole genome  
395 duplication, it is likely that an increased dosage of elovl5 and the 68 other duplicate pairs has affected the  
396 function of lipid metabolism in salmon.

## 397 Conclusion

398 Atlantic salmon needs great plasticity of physiology and behaviour to adapt for migration between  
399 freshwater and sea. By analysing transcriptomic changes through the transition from fresh- to saltwater and

400 the associated increase in dietary lipids, we identified an overall remodelling of lipid metabolism, with liver  
401 reflecting higher lipid metabolic plasticity and higher capacity of endogenous synthesis of LC-PUFAs in  
402 freshwater, while gut lipid uptake genes become more active in saltwater. These results indicate adaptive  
403 optimization of the Atlantic salmon lipid metabolism to account for life-stage-specific dietary availability.  
404 Moreover, we found signatures of pathway-specific selection pressure on gene duplicates, including a gene  
405 dosage increase in three genes involved in fatty acid metabolism. This illustrates possible adaptive  
406 consequences of the salmonid whole-genome duplication for the evolution of lipid metabolism. Future  
407 studies should attempt to decipher how the life-stage-related metabolic reprogramming is controlled (e.g.,  
408 through epigenetic mechanisms). Understanding this will have important implications for understanding  
409 evolution of genome regulatory processes in anadromous salmonids and potentially have economically  
410 important implications for Atlantic salmon aquaculture

411

## 412 Acknowledgements

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

416

## 417 Author Contributions

418 S.R.S., J.O.V., A.G., and J.S.T. conceived of the study and designed the feeding trial experiment. S.R.S.,  
419 J.O.V., and A.G. carried out orthogroup prediction and lipid metabolism annotation. T.N.H., Y.J., J.S.T.,  
420 and J.O.V. carried out the feeding trial. T.N.H. and Y.J. prepared the samples for RNA sequencing. M.T.,  
421 S.L., and M.L. provided input on the experimental design and helped interpreting the results from the  
422 transcriptomic and gene duplicate analysis. G.G. and T.R.H. performed the bioinformatic analyses. T.N.H.,  
423 G.G., S.R.S., and T.R.H. wrote the manuscript. All authors reviewed the final manuscript draft.

## 424 References

- 425 Allendorf, F. W., & Thorgaard, G. H. (1984). Tetraploidy and the Evolution of Salmonid Fishes. In  
426 *Evolutionary Genetics of Fishes* (pp. 1–53). Boston, MA: Springer US. [http://doi.org/10.1007/978-](http://doi.org/10.1007/978-1-4684-4652-4_1)  
427 [1-4684-4652-4\\_1](http://doi.org/10.1007/978-1-4684-4652-4_1)
- 428 Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq--a Python framework to work with high-throughput  
429 sequencing data. *Bioinformatics*, *31*(2), 166–169. <http://doi.org/10.1093/bioinformatics/btu638>
- 430 Bell, J. G., Tocher, D. R., Farndale, B. M., Cox, D. I., McKinney, R. W., & Sargent, J. R. (1997). The  
431 effect of dietary lipid on polyunsaturated fatty acid metabolism in Atlantic salmon (*Salmo salar*)  
432 undergoing parr-smolt transformation. *Lipids*, *32*(5), 515–525. [http://doi.org/10.1007/s11745-997-](http://doi.org/10.1007/s11745-997-0066-4)  
433 [0066-4](http://doi.org/10.1007/s11745-997-0066-4)
- 434 Berthelot, C., Brunet, F., Chalopin, D., Juanchich, A., Bernard, M., Noël, B., ... Guiguen, Y. (2014). The  
435 rainbow trout genome provides novel insights into evolution after whole-genome duplication in  
436 vertebrates. *Nature Communications*, *5*, 3657. <http://doi.org/10.1038/ncomms4657>
- 437 Bond, L. M., Miyazaki, M., O'Neill, L. M., Ding, F., & Ntambi, J. M. (2016). Fatty Acid Desaturation  
438 and Elongation in Mammals. In *Biochemistry of Lipids, Lipoproteins and Membranes* (pp. 185–  
439 208). Elsevier. <http://doi.org/10.1016/B978-0-444-63438-2.00006-7>
- 440 Brown, A. J., & Sharpe, L. J. (2016). Cholesterol Synthesis. In *Biochemistry of Lipids, Lipoproteins and*  
441 *Membranes* (pp. 327–358). Elsevier. <http://doi.org/10.1016/B978-0-444-63438-2.00011-0>
- 442 Carmona-Antonanzas, G., Tocher, D. R., Martinez-Rubio, L., & Leaver, M. J. (2014). Conservation of  
443 lipid metabolic gene transcriptional regulatory networks in fish and mammals. *Gene*, *534*(1), 1–9.  
444 <http://doi.org/10.1016/j.gene.2013.10.040>
- 445 Carmona-Antoñanzas, G., Zheng, X., Tocher, D. R., & Leaver, M. J. (2016). Regulatory divergence of  
446 homeologous Atlantic salmon elov15 genes following the salmonid-specific whole-genome  
447 duplication. *Gene*, *591*(1), 34–42. <http://doi.org/10.1016/j.gene.2016.06.056>
- 448 Ciftci, O. N., Przybylski, R., & Rudzińska, M. (2012). Lipid components of flax, perilla, and chia seeds.

449 *European Journal of Lipid Science and Technology*, 114(7), 794–800.  
450 <http://doi.org/10.1002/ejlt.201100207>

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

454 Emms, D. M., Kelly, S., Alexeyenko, A., Tamas, I., Liu, G., Sonnhammer, E., ... Kellis, M. (2015).  
455 OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves  
456 orthogroup inference accuracy. *Genome Biology*, 16(1), 157. [http://doi.org/10.1186/s13059-015-](http://doi.org/10.1186/s13059-015-0721-2)  
457 [0721-2](http://doi.org/10.1186/s13059-015-0721-2)

458 Glatz, J. a N. F. C., Luiken, J. J. F. P., & Bonen, A. (2010). Membrane Fatty Acid Transporters as  
459 Regulators of Lipid Metabolism : Implications for Metabolic Disease. *Physiological Reviews*, 90,  
460 367–417. <http://doi.org/10.1152/physrev.00003.2009>.

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

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

466 Kennedy, S. R., Leaver, M. J., Campbell, P. J., Zheng, X., Dick, J. R., & Tocher, D. R. (2006). Influence  
467 of dietary oil content and conjugated linoleic acid (CLA) on lipid metabolism enzyme activities and  
468 gene expression in tissues of Atlantic salmon (*Salmo salar* L.). *Lipids*, 41(5), 423–436.  
469 <http://doi.org/10.1007/s11745-006-5116-4>

470 Kim, J.-H., Leong, J. S., Koop, B. F., & Devlin, R. H. (2016). *Multi-tissue transcriptome profiles for coho*  
471 *salmon (Oncorhynchus kisutch), a species undergoing rediploidization following whole-genome*  
472 *duplication. Marine Genomics (Vol. 25)*. <http://doi.org/10.1016/j.margen.2015.11.008>

473 Kortner, T. M., Björkhem, I., Krasnov, A., Timmerhaus, G., & Krogdahl, Å. (2014). Dietary cholesterol  
474 supplementation to a plant-based diet suppresses the complete pathway of cholesterol synthesis and

475 induces bile acid production in Atlantic salmon (*Salmo salar* L.). *British Journal of Nutrition*,  
476 111(12), 2089–2103. <http://doi.org/10.1017/S0007114514000373>

477 Kulkarni, A., Dangat, K., Kale, A., Sable, P., Chavan-Gautam, P., & Joshi, S. (2011). Effects of Altered  
478 Maternal Folic Acid, Vitamin B12 and Docosahexaenoic Acid on Placental Global DNA  
479 Methylation Patterns in Wistar Rats. *PLoS ONE*, 6(3), e17706.  
480 <http://doi.org/10.1371/journal.pone.0017706>

481 Leaver, M. J., Bautista, J. M., Björnsson, B. T., Jönsson, E., Krey, G., Tocher, D. R., & Torstensen, B. E.  
482 (2008). Towards Fish Lipid Nutrigenomics: Current State and Prospects for Fin-Fish Aquaculture.  
483 *Reviews in Fisheries Science*, 16(April), 73–94. <http://doi.org/10.1080/10641260802325278>

484 Leaver, M. J., Villeneuve, L. A., Obach, A., Jensen, L., Bron, J. E., Tocher, D. R., & Taggart, J. B.  
485 (2008). Functional genomics reveals increases in cholesterol biosynthetic genes and highly  
486 unsaturated fatty acid biosynthesis after dietary substitution of fish oil with vegetable oils in Atlantic  
487 salmon (*Salmo salar*). *BMC Genomics*, 9, 299. <http://doi.org/10.1186/1471-2164-9-299>

488 Lehner, R., & Quiroga, A. D. (2016). Fatty Acid Handling in Mammalian Cells. In *Biochemistry of*  
489 *Lipids, Lipoproteins and Membranes* (pp. 149–184). Elsevier. [http://doi.org/10.1016/B978-0-444-](http://doi.org/10.1016/B978-0-444-63438-2.00005-5)  
490 [63438-2.00005-5](http://doi.org/10.1016/B978-0-444-63438-2.00005-5)

491 Li, B., & Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or  
492 without a reference genome. *BMC Bioinformatics*, 12(1), 323. [http://doi.org/10.1186/1471-2105-12-](http://doi.org/10.1186/1471-2105-12-323)  
493 [323](http://doi.org/10.1186/1471-2105-12-323)

494 Lien, S., Koop, B. F., Sandve, S. R., Miller, J. R., Kent, M. P., Nome, T., ... Davidson, W. S. (2016). The  
495 Atlantic salmon genome provides insights into rediploidization. *Nature*, 533(7602), 200–205.  
496 <http://doi.org/10.1038/nature17164>

497 Lorgen, M., Casadei, E., Krol, E., Douglas, A., Birnie, M. J., Ebbesson, L. O. E., ... Martin A.m, S.  
498 (2015). Functional divergence of type 2 deiodinase paralogs in the Atlantic salmon. *Current*  
499 *Biology*, 25(7), 936–941. <http://doi.org/10.1016/j.cub.2015.01.074>

500 Macqueen, D. J., & Johnston, I. A. (2014). A well-constrained estimate for the timing of the salmonid

501 whole genome duplication reveals major decoupling from species diversification. *Proceedings of the*  
502 *Royal Society of London B: Biological Sciences*, 281(1778).

503 Morais, S., Pratoomyot, J., Taggart, J. B., Bron, J. E., Guy, D. R., Bell, J. G., & Tocher, D. R. (2011).  
504 Genotype-specific responses in Atlantic salmon (*Salmo salar*) subject to dietary fish oil replacement  
505 by vegetable oil: a liver transcriptomic analysis. *BMC Genomics*, 12(1), 255.  
506 <http://doi.org/10.1186/1471-2164-12-255>

507 Oxley, A., Jolly, C., Eide, T., Jordal, A.-E. O., Svoldal, A., & Olsen, R.-E. (2010). The combined impact  
508 of plant-derived dietary ingredients and acute stress on the intestinal arachidonic acid cascade in  
509 Atlantic salmon (*Salmo salar*). *The British Journal of Nutrition*, 103(6), 851–61.  
510 <http://doi.org/10.1017/S0007114509992467>

511 Price, M. N., Dehal, P. S., Arkin, A. P., Nawrocki, E., Kolbe, D., Eddy, S., ... Meyer, F. (2010). FastTree  
512 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE*, 5(3), e9490.  
513 <http://doi.org/10.1371/journal.pone.0009490>

514 Rimoldi, S., Benedito-Palos, L., Terova, G., & Pérez-Sánchez, J. (2016). Wide-targeted gene expression  
515 infers tissue-specific molecular signatures of lipid metabolism in fed and fasted fish. *Reviews in Fish*  
516 *Biology and Fisheries*, 26(1), 93–108. <http://doi.org/10.1007/s11160-015-9408-8>

517 Robertson, F. M., Gundappa, M. K., Grammes, F., Hvidsten, T. R., Redmond, A. K., Martin, S. A. M., ...  
518 Macqueen, D. J. (2017). Lineage-specific rediploidization is a mechanism to explain time-lags  
519 between genome duplication and evolutionary diversification. *Genome Biology*.  
520 <http://doi.org/doi:10.1101/098582>

521 Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: a Bioconductor package for  
522 differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)*,  
523 26(1), 139–40. <http://doi.org/10.1093/bioinformatics/btp616>

524 Ruyter, B., Røsjø, C., Måsøval, K., Einen, O., & Thomassen, M. S. (2000). Influence of dietary n-3 fatty  
525 acids on the desaturation and elongation of [ 1- 14 C ] 18 : 2 n-6 and [ 1- 14 C ] 18 : 3 n-3 in Atlantic  
526 salmon hepatocytes, 151–158.

527 Sánchez-Gurmaches, J., Østbye, T.-K., Navarro, I., Torgersen, J., Hevrøy, E. M., Ruyter, B., &  
528 Torstensen, B. E. (2011). In vivo and in vitro insulin and fasting control of the transmembrane fatty  
529 acid transport proteins in Atlantic salmon (*Salmo salar*). *American Journal of Physiology.*  
530 *Regulatory, Integrative and Comparative Physiology*, 301(4), R947-57.  
531 <http://doi.org/10.1152/ajpregu.00289.2011>

532 Sheridan, M. A. (1989). Alterations in lipid metabolism accompanying smoltification and seawater  
533 adaptation of salmonid fish. *Aquaculture*, 82(1–4), 191–203. [http://doi.org/10.1016/0044-](http://doi.org/10.1016/0044-8486(89)90408-0)  
534 [8486\(89\)90408-0](http://doi.org/10.1016/0044-8486(89)90408-0)

535 Stefansson, S. O., Björnsson, B. T., Ebbesson, L. O., & McCormick, S. D. (2008). Smoltification. *Fish*  
536 *Larval Physiology*, 639–681. [http://doi.org/DOI: 10.1111/j.1095-8649.2009.02440\\_2.x](http://doi.org/DOI:10.1111/j.1095-8649.2009.02440_2.x)

537 Sushchik, N. N., Gladyshev, M. I., Moskvichova, A. V., Makhutova, O. N., & Kalachova, G. S. (2003).  
538 Comparison of fatty acid composition in major lipid classes of the dominant benthic invertebrates of  
539 the Yenisei river. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular*  
540 *Biology*, 134(1), 111–122. [http://doi.org/10.1016/S1096-4959\(02\)00191-4](http://doi.org/10.1016/S1096-4959(02)00191-4)

541 Tocher, D. R. (2003). Metabolism and Functions of Lipids and Fatty Acids in Teleost Fish. *Reviews in*  
542 *Fisheries Science*, 11(2), 107–184. <http://doi.org/10.1080/713610925>

543 Tocher, D. R., Bell, J. G., Dick, J. R., Henderson, R. J., McGhee, F., Michell, D., & Morris, P. C. (2000).  
544 Polyunsaturated fatty acid metabolism in Atlantic salmon (*Salmo salar*) undergoing parr-smolt  
545 transformation and the effects of dietary linseed and rapeseed oils. *Fish Physiology and*  
546 *Biochemistry*, 23(1), 59–73. <http://doi.org/10.1023/A:1007807201093>

547 Tocher, D. R., Bell, J. G., MacGlaughlin, P., McGhee, F., & Dick, J. R. (2001). Hepatocyte fatty acid  
548 desaturation and polyunsaturated fatty acid composition of liver in salmonids: effects of dietary  
549 vegetable oil. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular*  
550 *Biology*, 130(2), 257–270. [http://doi.org/10.1016/S1096-4959\(01\)00429-8](http://doi.org/10.1016/S1096-4959(01)00429-8)

551 Tocher, D. R., Fonseca-Madrigal, J., Bell, J. G., Dick, J. R., Henderson, R. J., & Sargent, J. R. (2002).  
552 Effects of diets containing linseed oil on fatty acid desaturation and oxidation in hepatocytes and



553 intestinal enterocytes in Atlantic salmon (*Salmo salar*). *Fish Physiology and Biochemistry*, 26(2),  
554 157–170. <http://doi.org/10.1023/A:1025416731014>

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

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

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

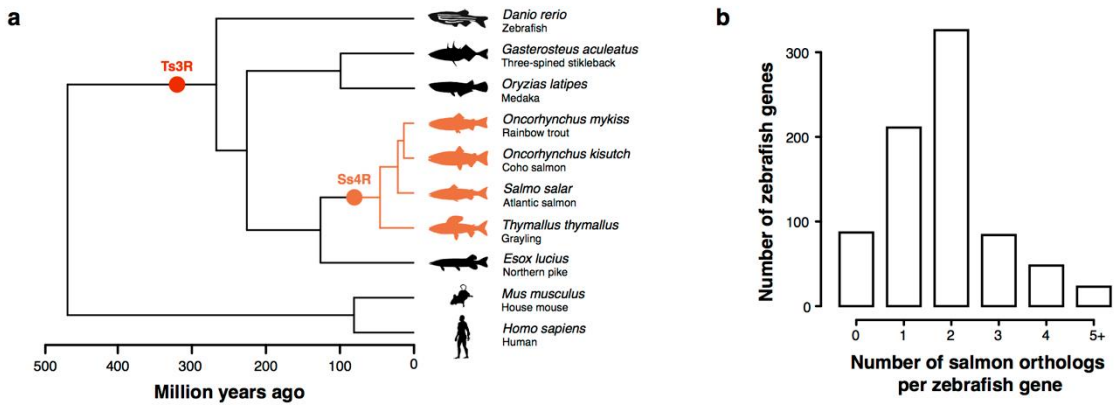
566

## 567 Data Accessibility:

- 568 - Supplementary files have been deposited to datadryad.org under the accession:  
569 [doi:10.5061/dryad.j4h65](https://doi.org/10.5061/dryad.j4h65)
- 570 - All gene expression results can be accessed through the interactive shiny web server:  
571 <https://goo.gl/8Ap89a>
- 572 - Lipid metabolism gene annotation can be accessed from <https://goo.gl/VVUVWr>
- 573 - Raw RNA-Seq data have been deposited into European Nucleotide Archive (ENA) under the  
574 project Accession no. PRJEB24480

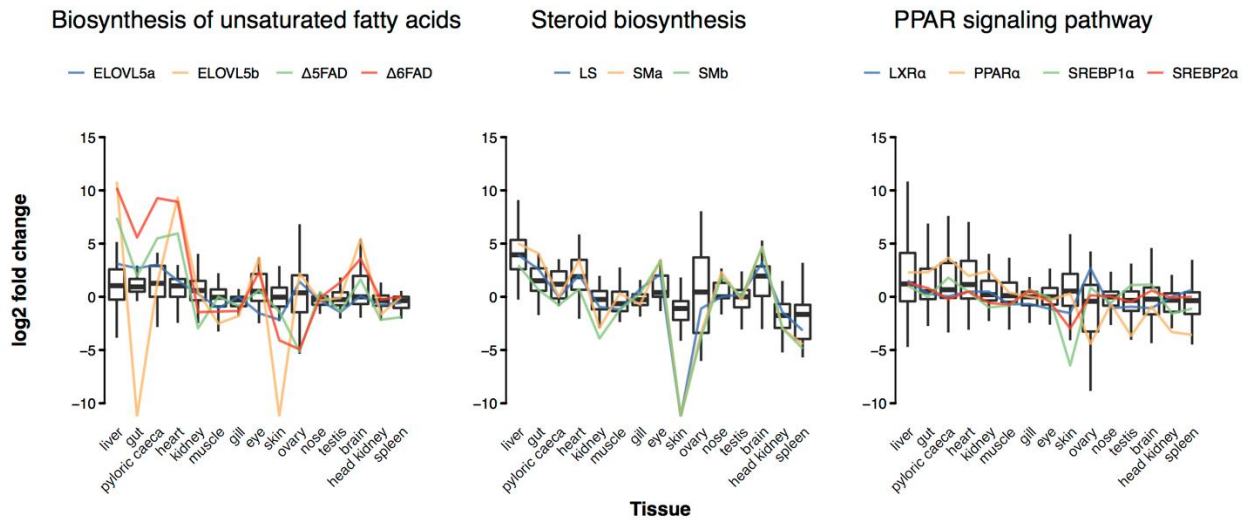
575

### Figures



576

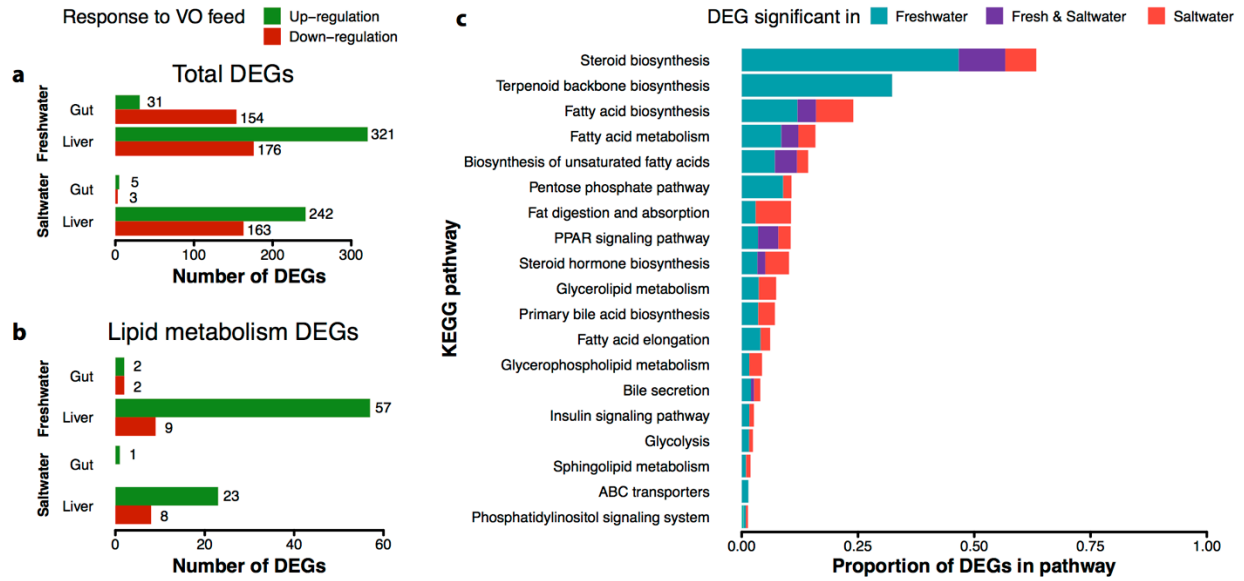
577 **Figure 1: Ortholog annotation** (a) Species used to construct ortholog groups and their evolutionary distance. Points  
 578 in the phylogenetic tree show the time of the teleost specific (Ts3R) and salmonid specific (Ss4R) whole genome  
 579 duplications. (b) The number of salmon orthologs found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



581

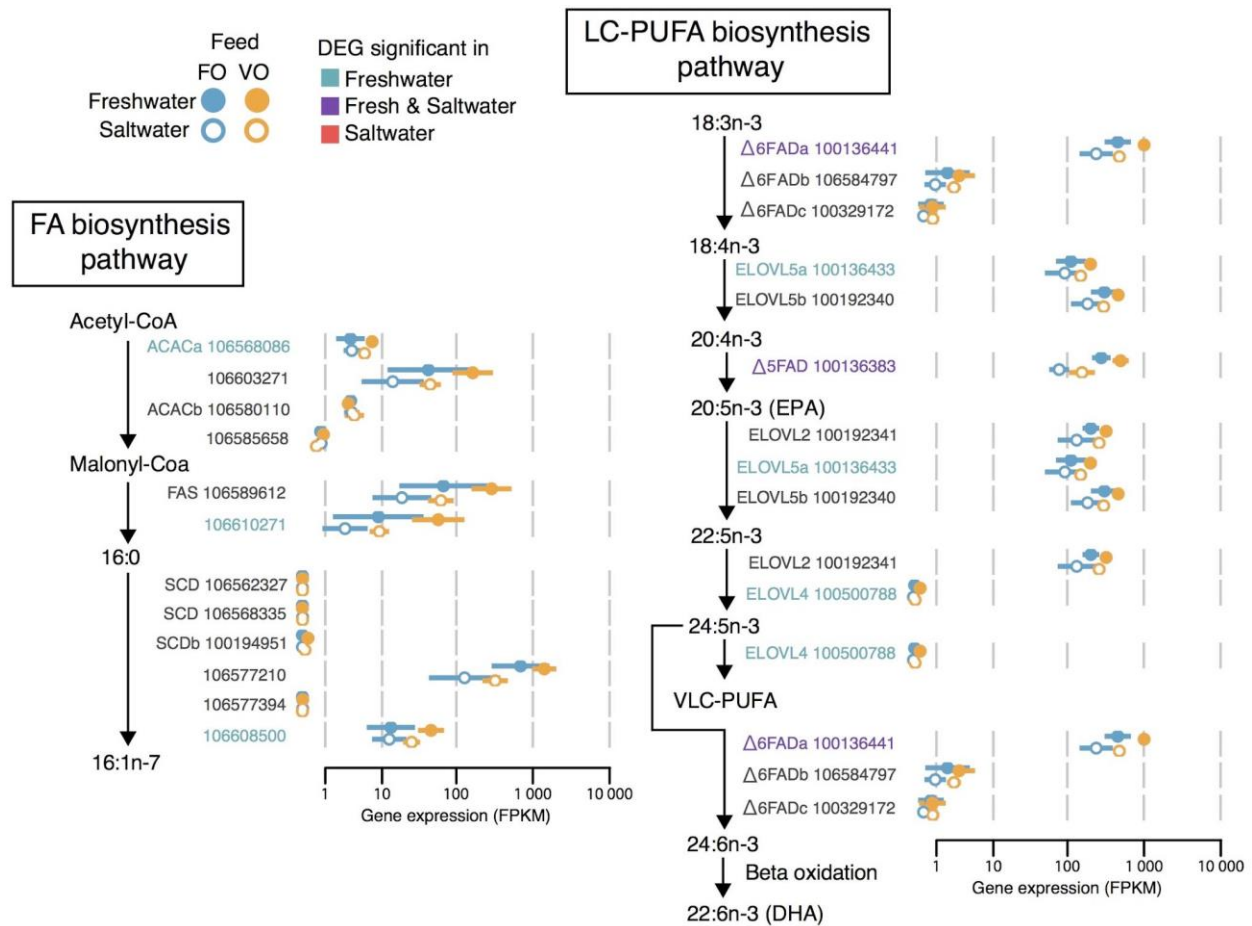
582 **Figure 2: Tissue expression profiles of salmon genes in lipid metabolism pathways**

583 Tissue expression profiles of our annotated lipid metabolism genes were consistent with expectations. Gene expression  
 584 levels are shown as the log<sub>2</sub> fold change difference between the FPKM value of each tissue and the median FPKM  
 585 across all tissues. Expression profiles for selected genes in each pathway are shown (see Figure S2 and S3 for all  
 586 pathways and gene details).



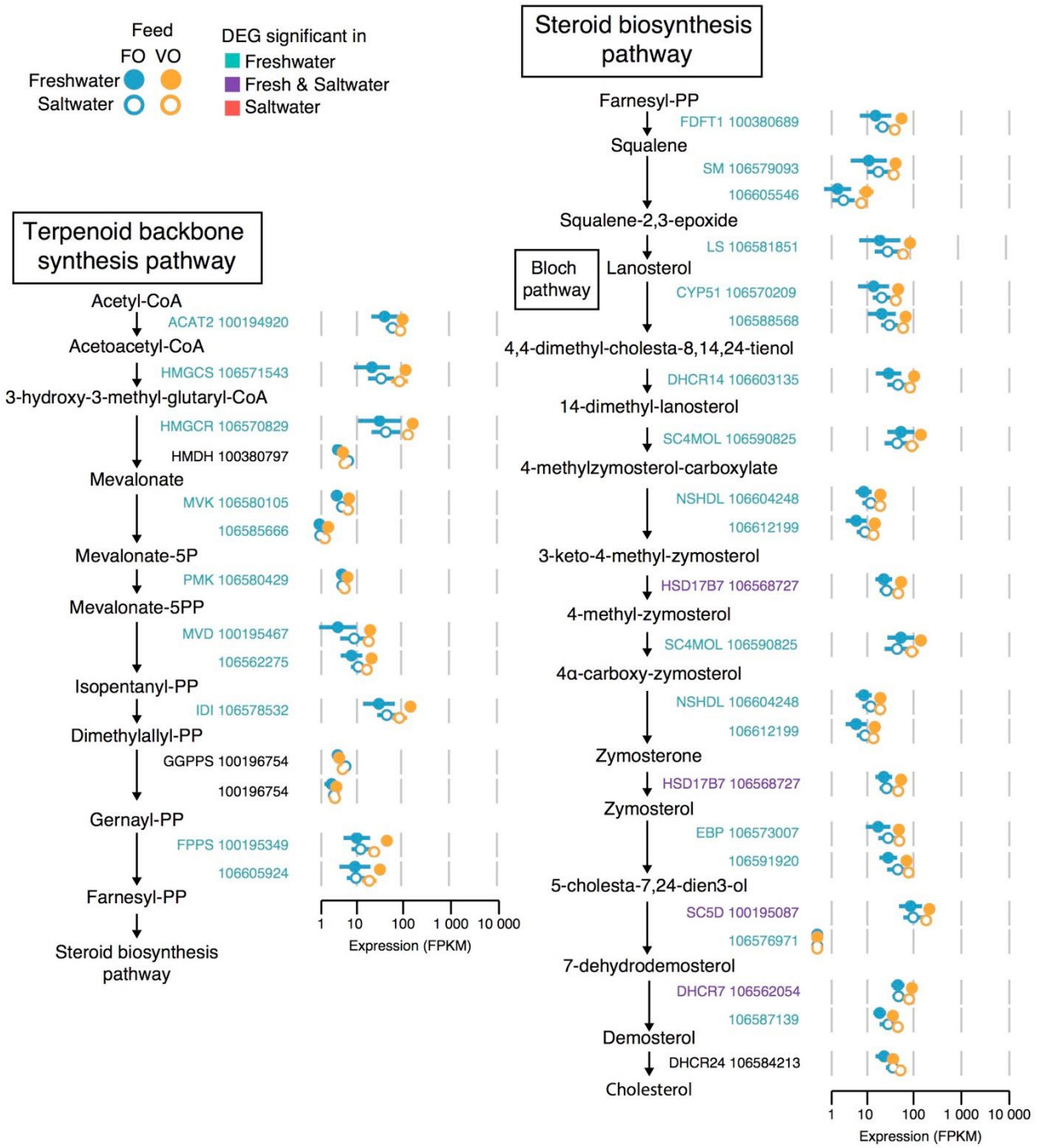
587

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



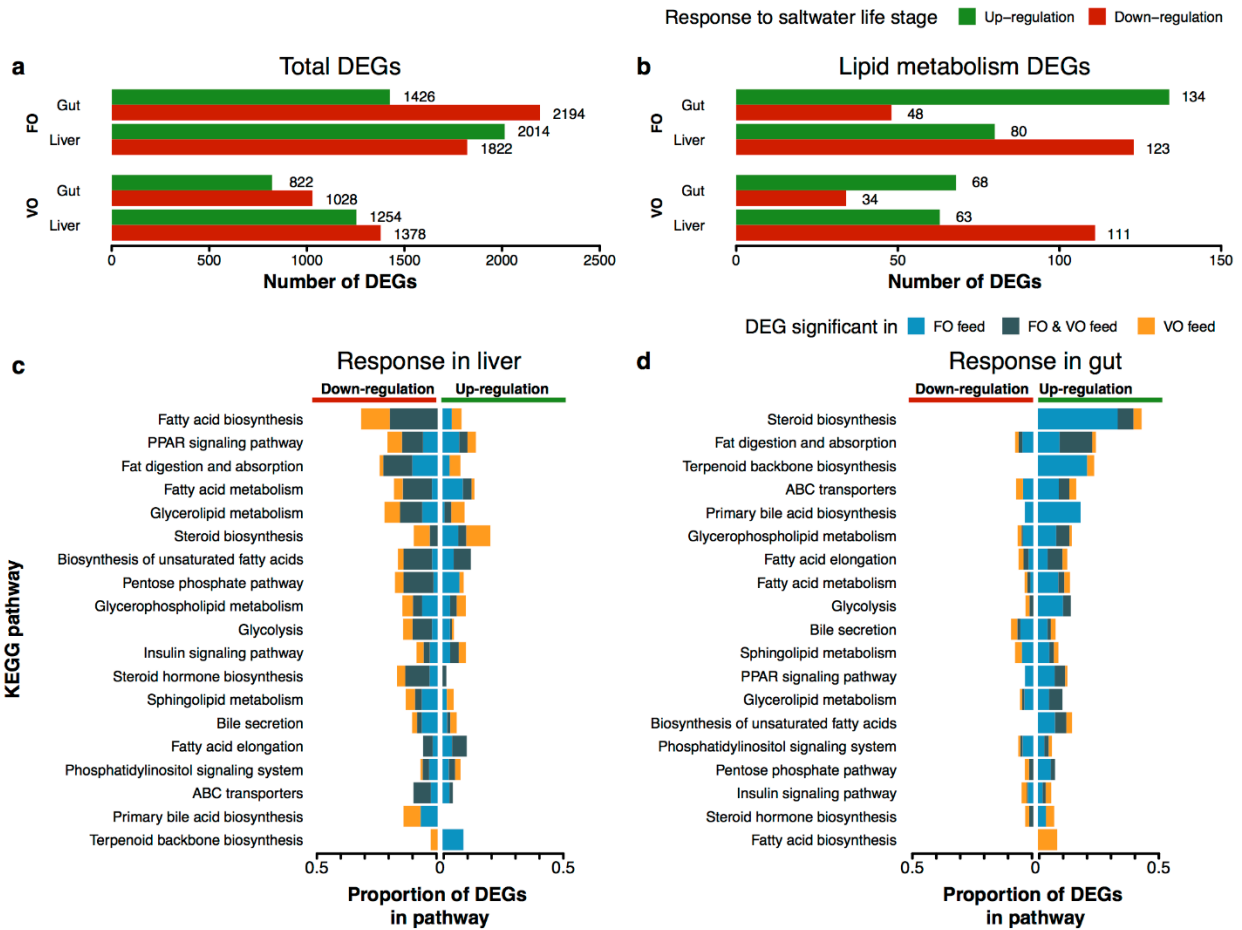
593

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



600

601 **Figure 5: Diet and life-stage effects on cholesterol biosynthesis in salmon liver.** Terpenoid backbone synthesis and steroid  
 602 biosynthesis pathways with Atlantic salmon genes annotated to each catalytic step (enzyme names followed by NCBI gene  
 603 numbers). Gene expression levels are shown as mean (point) and standard deviation (line) of expression in eight samples  
 604 (measured in log(FPKM + 1)) from each diet (FO, VO feeds) and life-stage (freshwater, saltwater) combination. Genes  
 605 significantly (FDR<0.05) differentially expressed (DEG) between diets in a life-stage are highlighted.



606

607

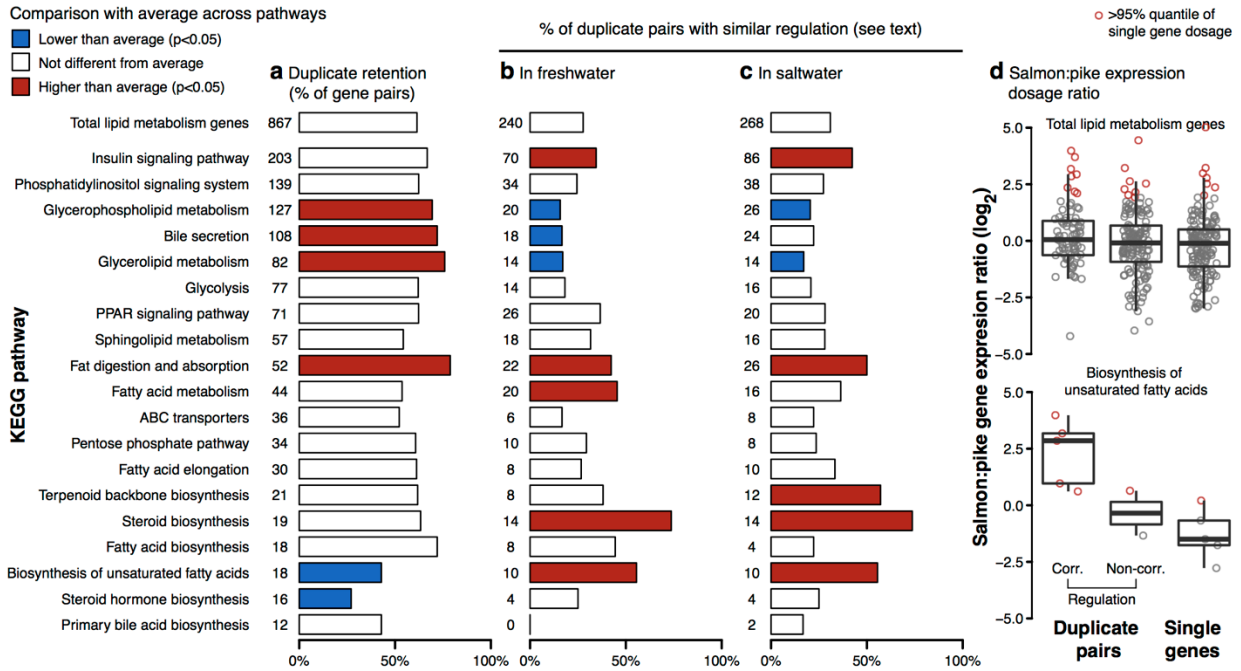
608

609

610

611

**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.



612

613

614

615

616

617

618

619

620

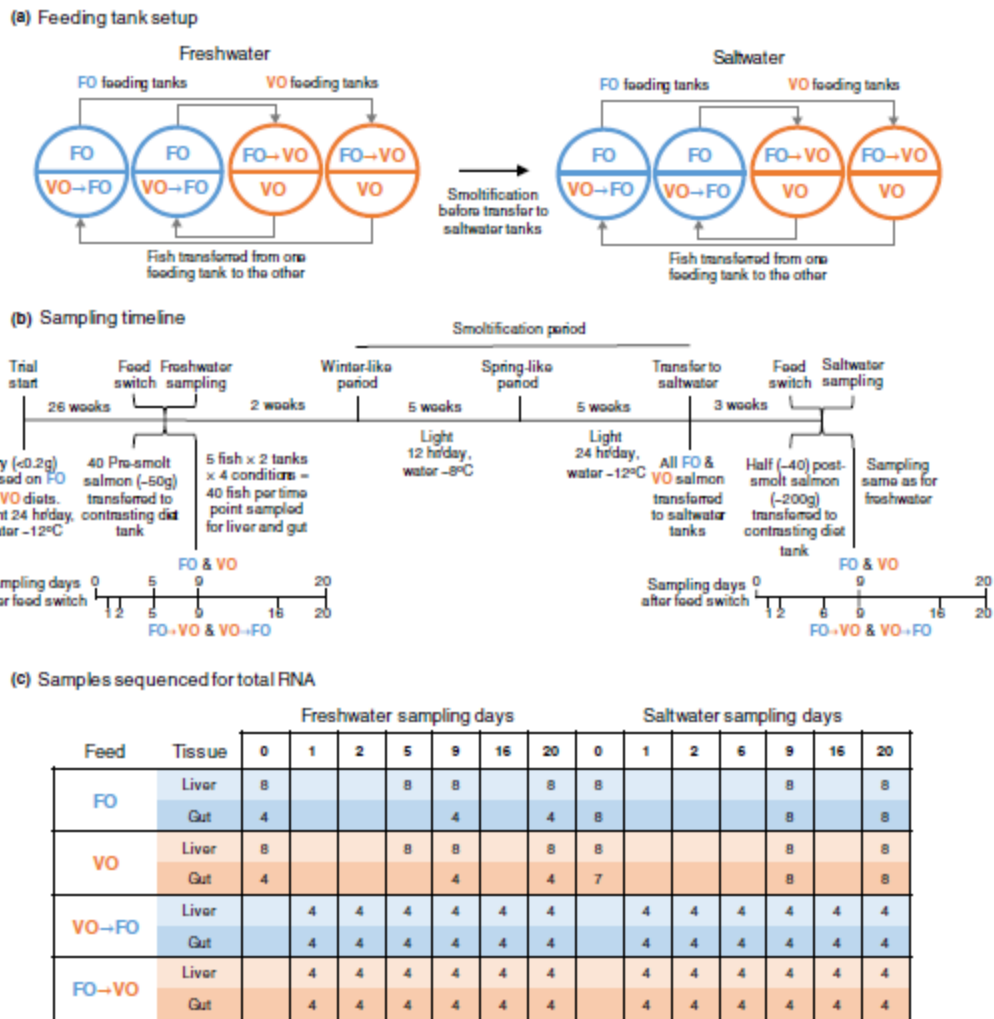
621

622

623

**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  $\geq 0.6$ ,  $P$ -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)  $\log_2$  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.





624

625 **Figure 8: Overview of feed trial experiment.**

626 (a) Atlantic salmon fry were reared in 4 feeding tanks containing freshwater; 2 continuously fed fish oil (FO) and 2 vegetable oil

627 (VO). A feed switch involved the transfer of fish from one tank to an empty partition of another tank fed the opposite diet. After

628 smoltification fish from FO and VO tanks were transferred to 4 new feeding tanks containing saltwater and the feed switch was

629 repeated. (b) Timeline of feed trial showing fish sampling and smoltification periods. Fish were sampled before (D0) and up to 20

630 days after the fresh- or saltwater feed switch. (c) Total RNA was sequenced from select fish tissue samples. The number of RNA-

631 seq replicates are shown for each, tissue, condition and time point.

632

633