Diet and life stage associated remodeling of lipid metabolism regulation in the duplicated Atlantic salmon genome

Gareth Gillard*, Thomas N. Harvey[1]*, 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, Ås NO-1432, Norway*

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

[3] *Department of Biology, Norwegian University of Science and Technology, N7491 Trondheim, Norway.*

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

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

[6] *AquaGen AS, Ås NO-1432, Norway*

Abstract

Atlantic salmon migrates from rivers to sea to feed, grow and develop gonads before returning to spawn in freshwater. These habitat shifts require great phenotypic plasticity. To address the unresolved question of how the shift in diet between fresh and saltwater affects the regulation of metabolic function, we fed salmon contrasting diets in each of the two life stages. Combining transcriptomics with comparative genomics, we found that lipid metabolism undergoes a concerted shift between fresh- and saltwater stages. Lipogenesis and lipid transport become less active in liver after transition to saltwater, while genes involved in uptake of dietary lipids in gut greatly increases in lipid-rich seawater environments. We assess how the whole-genome duplication ancestral to salmonids has impacted the evolution of lipid metabolism, and find signatures of pathway-specific selection pressure on gene duplicates and a limited number of cases of increased gene dosage.

Introduction

Atlantic salmon lives a 'double life'. It starts its life in rivers, before it transforms its physiology and behavior and migrates 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 remodeling of gene expression. The resulting adaptations to a marine environment include shifts in salt-tolerance, coloration, behavior, growth rates, and metabolism.

Salmon transforms its lipid metabolism function during smoltification (Sheridan, 1989). This is likely related to differences in diet between environments. Salmon in rivers mostly eat invertebrates that are low in essential long-chain polyunsaturated fatty acids (LC-PUFA). Possibly as an adaptation to this (Leaver, Bautista, et al., 2008), salmon has evolved a greater capacity for endogenous production of LC-PUFAs than many other fish species, and several studies have demonstrated that salmon has the ability to increase or decrease endogenous omega-3 synthesis as a response to the dietary availability of LC-PUFAs (Kennedy et al., 2006; Leaver, Villeneuve, et al., 2008; Morais et al., 2011; Ruyter, Røsjø, Måsøval, Einen, & Thomassen, 2000; D.R. Tocher, Bell, MacGlaughlin, McGhee, & Dick, 2001; Douglas R. Tocher et al., 2002; Zheng et al., 2005). In contrast to rivers, marine habitat food chains are high in available LC-PUFAs, and it has therefore been hypothesized that smoltification-associated lipid metabolism changes are linked to preparation to this new dietary situation. However, little is known about the extent and nature of the lipid metabolism remodeling associated with the life stage shift from freshwater to sea.

The evolution of novel traits in salmonids 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). Gene duplication can give rise to new adaptive phenotypes in several ways: through evolution of novel functions or gene regulation, subdivision of function and specialization, or because of an adaptive increase in gene dosage. The Atlantic salmon genome contains ~10,000 pairs of Ss4R gene duplicates (Lien et al., 2016), 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, Zheng, Tocher, & Leaver, 2016). This is believed to have facilitated evolution of novel traits, including flexible phenotypes necessary for an anadromous life history (Stefansson, Björnsson,

68 Ebbesson, & McCormick, 2008). However, no systematic genome wide study has been conducted to
69 assess the importance of the Ss4R in evolution of salmon lipid metabolism.
70
71 In this study, we integrate comparative genomics with transcriptomic data from feeding trials carried out
72 across the fresh to saltwater transition to build a functional annotation of lipid metabolism pathway genes
73 in salmon. We use this annotation to elucidate (i) the nature of the lipid metabolism transformation from
74 freshwater to saltwater life stages of salmon and (ii) the impact of whole genome duplication on evolution
75 of the lipid gene repertoire and metabolic function. Our results indicate a life-stage programmed shift in
76 lipid metabolism after transition to seawater, and identify differences between lipid pathways with respect
77 to selection pressure on gene duplicates from the salmonid whole genome duplication.
78

79 Results and discussion

80 Annotation of lipid metabolism genes

81 To identify genes involved in lipid metabolism in Atlantic salmon, we initially assembled groups of
82 orthologous genes (orthogroups) from a selection of four salmonid species, Northern pike (their closest
83 unduplicated relative), as well as teleost and mammalian outgroups (Figure 1a). Next, we aligned
84 orthogroup proteins and constructed maximum likelihood gene trees. The majority (82-98%) of proteins
85 from each species were represented in 23,782 ortholog gene trees. The salmonid species had significantly
86 higher number of proteins included in ortholog gene trees compared to non-salmonid fish (Figure S1),
87 reflecting the salmonid specific whole genome duplication. We then used the evolutionary distances in
88 gene trees to infer the most likely salmon sequence orthologs of zebrafish genes selected from 19 KEGG
89 pathways involved in lipid metabolism (Table S1). This resulted in the annotation of 1421 salmon lipid
90 metabolism genes, of which 326 (23%) showed a 2:1 ortholog ratio between salmon and zebrafish (Figure
91 1b). Only 87 (6%) of the zebrafish genes could not be assigned a salmon ortholog.

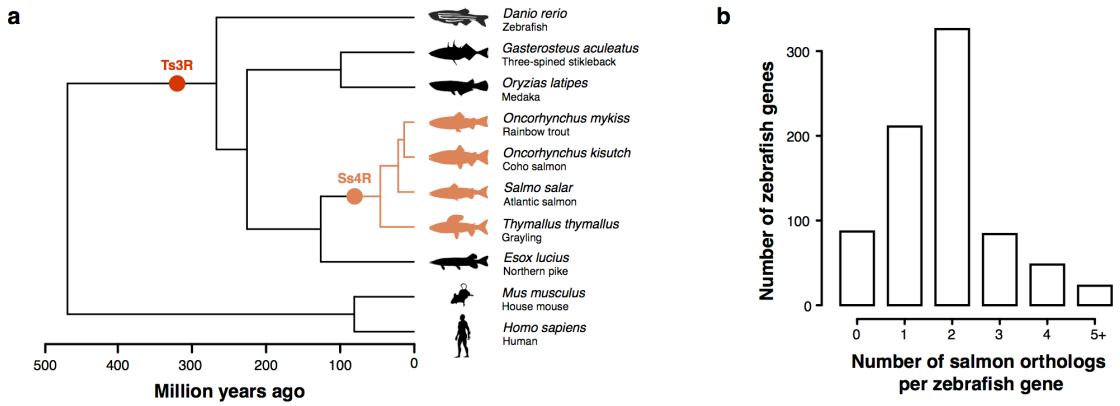


Figure 1: Ortholog annotation (a) Species used to construct ortholog groups and their evolutionary distance. Points in the phylogenetic tree show the time of the teleost specific (Ts3R) and salmonid specific (Ss4R) whole genome duplications. (b) The number of salmon orthologs found (1421 genes in total)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

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 in wild-type Atlantic salmon, and found agreement with known properties of vertebrate lipid metabolism. 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, and heart) (Glatz, Luiken, & Bonen, 2010; Rimoldi, Benedito-Palos, Terova, & Pérez-Sánchez, 2016; Douglas R 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: (url).

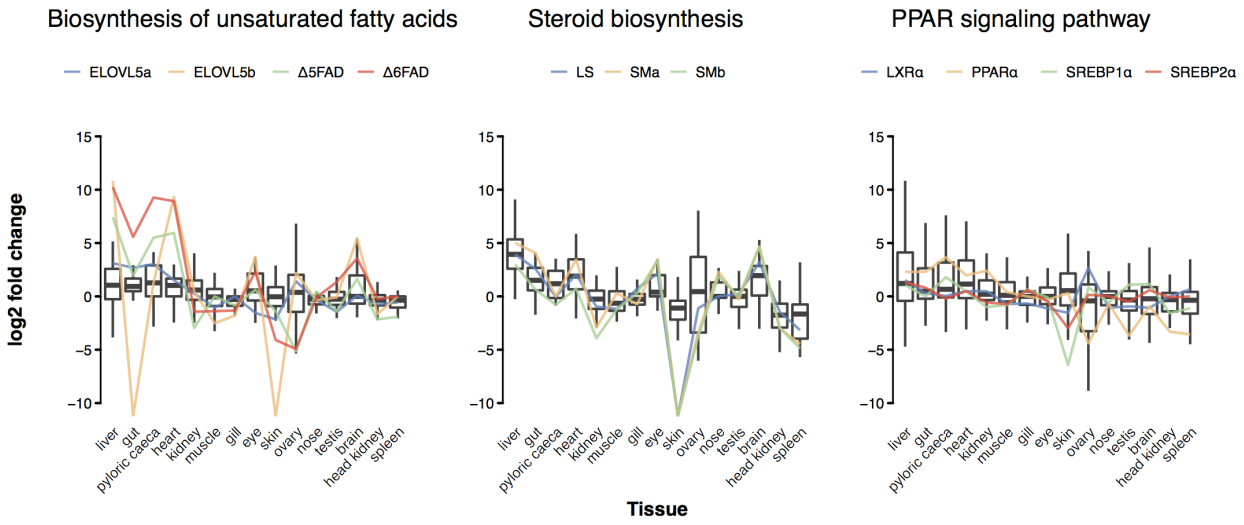


Figure 2: Tissue expression profiles of salmon genes in lipid metabolism pathways

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 for gene details).

Life-stage dependent remodelling 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. 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 Supplementary table 3 for details on feed composition). VO based diets are also known to contain lower cholesterol levels (Ciftci, Przybylski, & Rudzińska, 2012; Verleyen et al., 2002). In total, 32 and 23 fish were sampled for RNA-seq of liver and gut, respectively, including eight biological replicates from each diet and life-stage (freshwater and saltwater). Fish in the different dietary groups were given FO and VO feed from first feeding (<0.2g) until sampling.

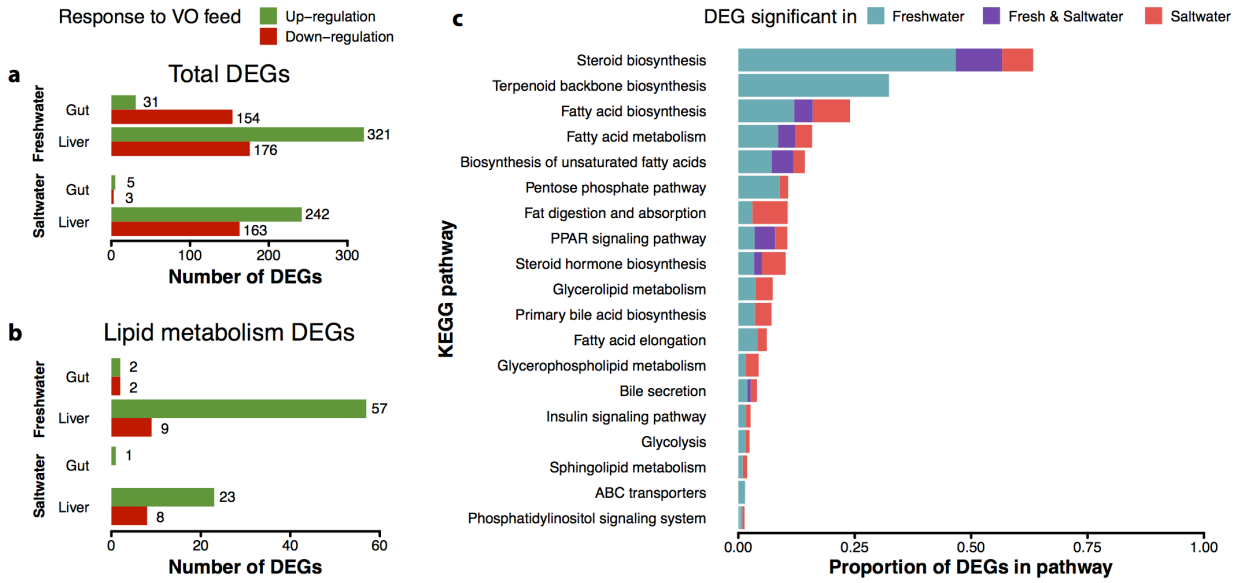
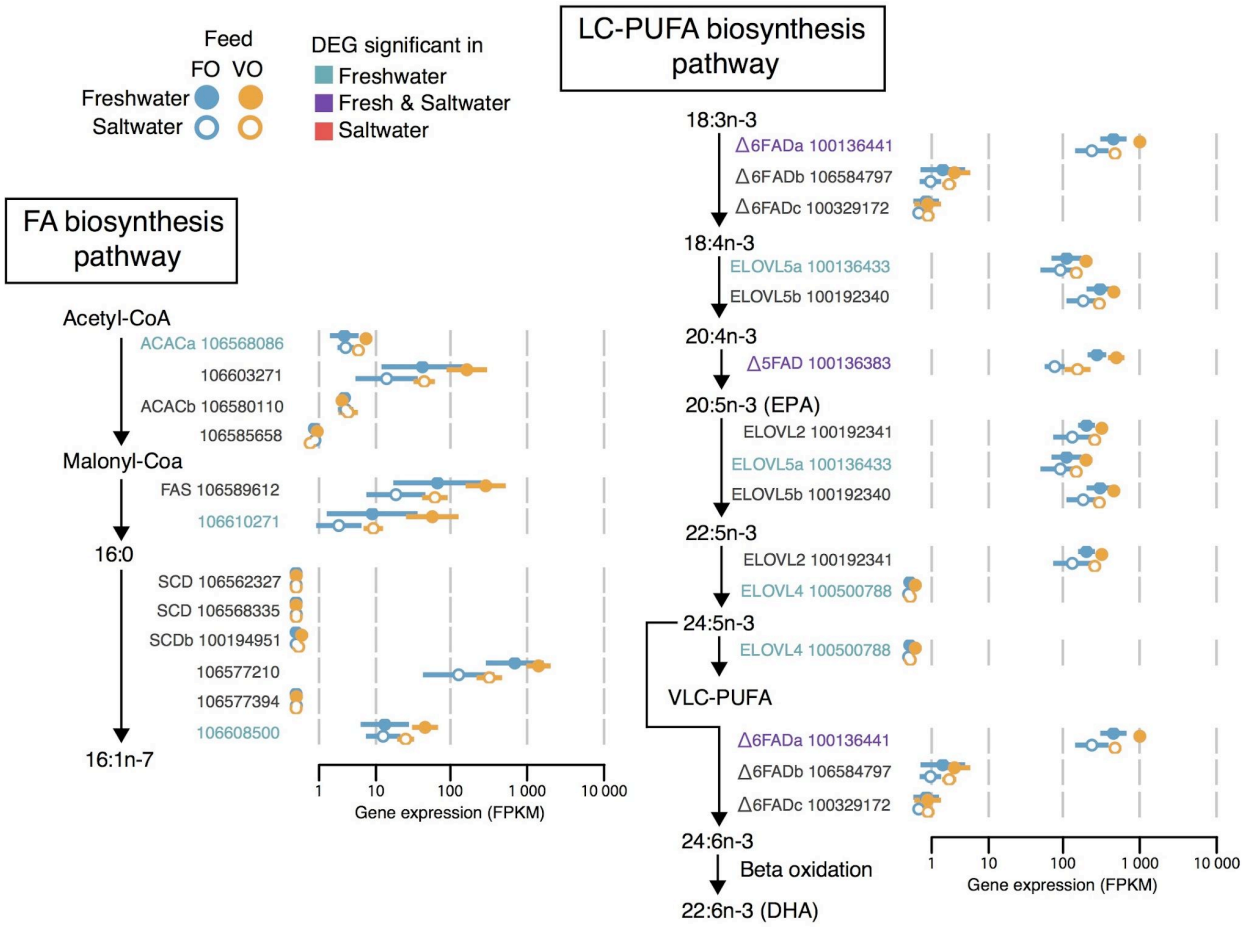


Figure 3: Gene regulation in response to feed type. (a) Total number of significant (FDR < 0.05) differentially expressed genes (DEGs) between fish oil (FO) and vegetable oil (VO) fed salmon in the liver and gut tissues of freshwater and saltwater stage Atlantic salmon. (b) As above, but for lipid-associated 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.

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 increased lipid-metabolism related gene expression compared to FO diets, with 57 genes (86% of the differentially expressed genes, DEGs) upregulated in freshwater and 23 genes (74% of DEGs) upregulated in saltwater (Figure 3b). The increased activity of liver lipid metabolism under VO diets confirm the well-known ability of salmon to regulate endogenous synthesis of LC-PUFA and cholesterol in fresh water (Kortner, Björkhem, Krasnov, Timmerhaus, & Krogdahl, 2014; Leaver, Villeneuve, et al., 2008; Zheng et al., 2005).

Freshwater and saltwater fish shared a relatively small number of DEGs for each pathway (Table S4). We found that some pathways had more DEGs in freshwater (‘*fatty acid biosynthesis*’, ‘*steroid biosynthesis*’, and its precursor ‘*terpenoid backbone biosynthesis*’), whereas others had more DEGs in saltwater (‘*steroid biosynthesis*’, ‘*fatty acid biosynthesis*’, and ‘*fat digestion and absorption*’) (Figure 3c). Out of 71 lipid metabolism DEGs in the dietary contrast, 78% (51 genes) were freshwater specific, 11% (8 genes) saltwater specific, and 11% (8 genes) shared dietary response (Table S4). For example, only two genes in

161 the FA and LC-PUFA biosynthesis pathways (Δ6FADa and Δ5FAD) shared response to dietary
162 availability in fresh- and saltwater (Figure 4). A similar trend was seen for diet-induced expression
163 changes in the pathways responsible for cholesterol biosynthesis (Figure 5). The few genes that showed
164 diet-effects specific to saltwater included bile salt activated lipase, responsible for the hydrolysis of free
165 fatty acids from TAG obtained from the diet (Douglas R Tocher, 2003). Two of these genes, carboxyl
166 ester lipase, tandem duplicate 2a (CEL2a) and b (CEL2b), are highly upregulated in saltwater in response
167 to VO diet. Taken together, our results present clear evidence of life-stage dependent remodeling of
168 dietary-inducible flexibility in endogenous liver lipid metabolism, and is in line with the idea of a post-
169 smoltification phenotype adapted to an environment with surplus of essential lipids in the food-chain.



170

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

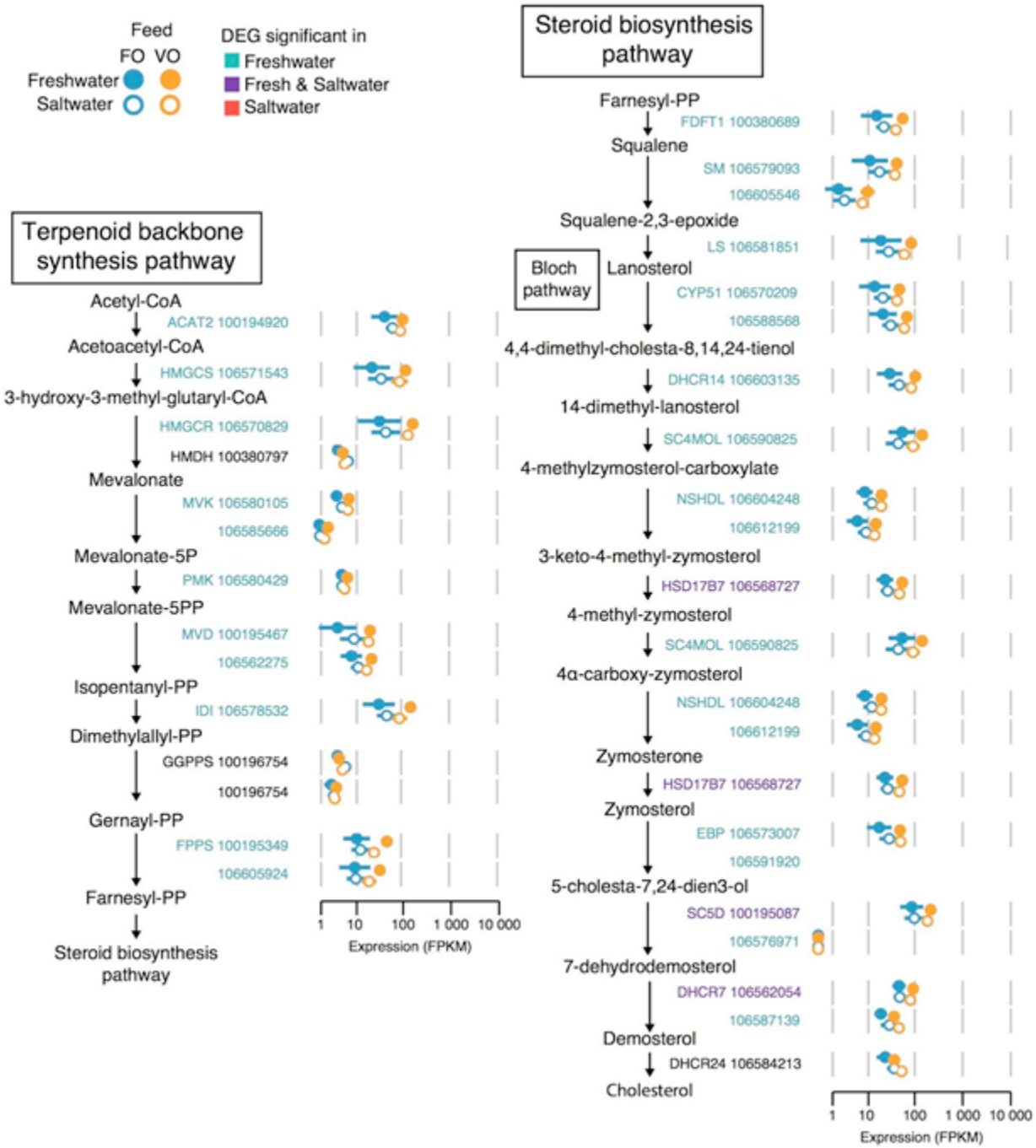
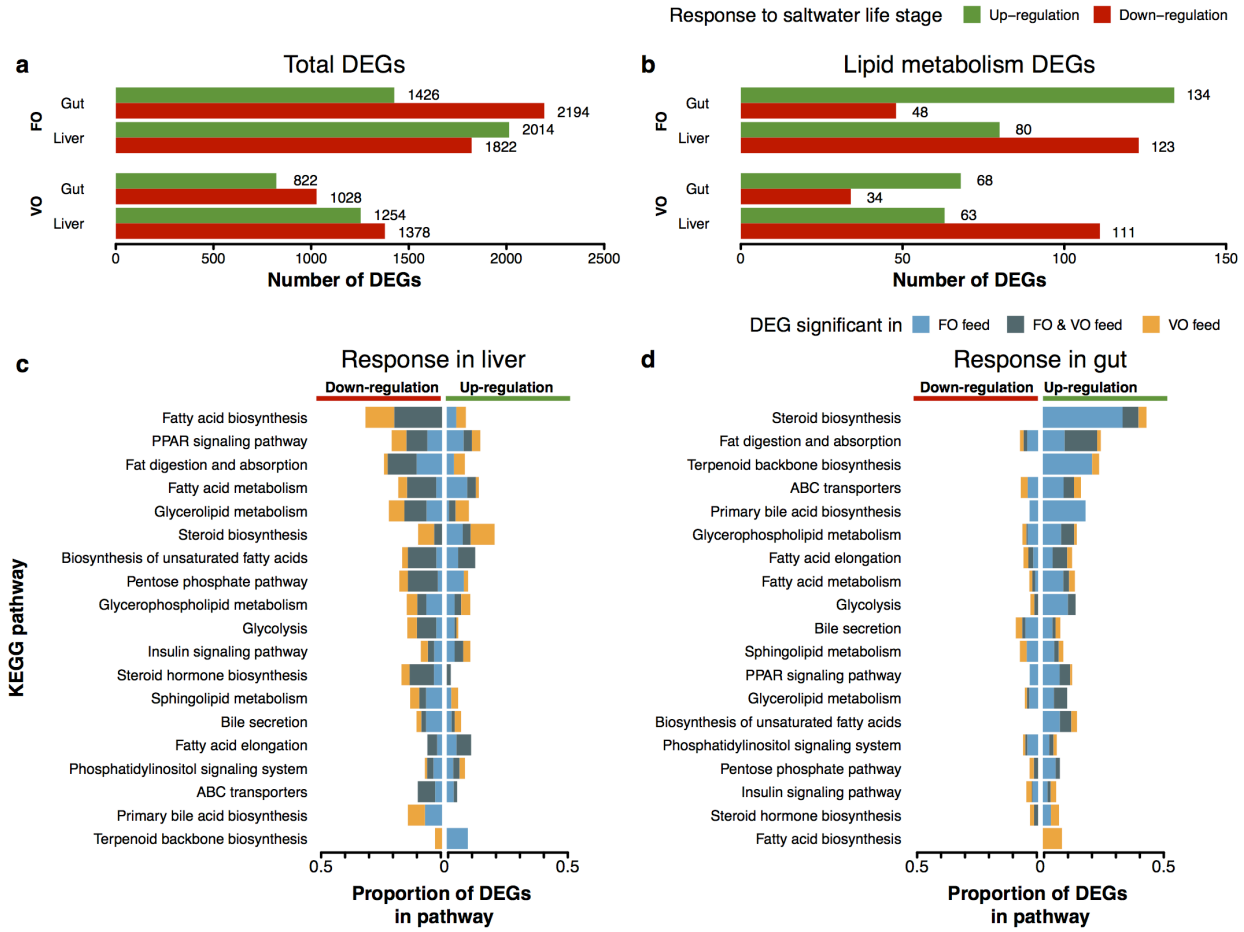


Figure 5: Diet and life stage effects on cholesterol biosynthesis in salmon liver. Terpenoid backbone synthesis and steroid biosynthesis pathways with Atlantic salmon genes annotated to each catalytic step (enzyme names followed by NCBI gene numbers). Gene expression levels are shown as mean (point) 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.

184 To further investigate the life-stage associated changes in lipid metabolism we tested for differential
185 expression between freshwater and saltwater in salmon fed the same diets (Figure 6). Liver and gut
186 showed contrasting effects of saltwater on lipid gene expression: Mostly downregulation in liver,
187 upregulation in gut (Figure 6b). The number of DEGs in each tissue were similar for the environment
188 comparison (Figure 6a), unlike for the diet comparison (Figure 3).
189
190 Further examination of key lipid metabolism genes revealed that the system-wide lipid metabolism
191 remodeling represented a concerted shift in the metabolic role of liver and gut. As the salmon entered the
192 marine stage, lipogenic activity in the liver was significantly ramped down, as evident by the markedly
193 lower expression (2.2-3.3 fold) of the master regulator of lipid metabolism SREBP1, a 5-fold decrease in
194 expression of fatty acid synthase, and a 2-3 fold decrease in rate-limiting enzymes in LC-PUFA syntheses
195 (i.e. Δ5FAD, Δ6FADa) (Figure 4). Liver and gut also increased their catabolic activity in saltwater,
196 upregulating the uptake of fatty acids into mitochondria for β-oxidation (uptake carried out by carnitine
197 palmitoyltransferase 1 and 2) (Lehner & Quiroga, 2016). Finally, expression of lipid transport genes
198 shifted from liver to gut with the transition to seawater (apolipoproteins, pathway "Fat digestion and
199 absorption" in Figure 1). Four apolipoproteins (out of 11 annotated) were differentially regulated in liver
200 between different diets; these all decreased their expression 2.4-5 fold in saltwater compared to
201 freshwater. In stark contrast, nine of the diet-regulated apolipoproteins in gut increased their expression in
202 saltwater between 1.8-9.7 fold. These results suggest that salmon's decreased ability to synthesize LC-
203 PUFAs after fresh-to-saltwater transition is compensated by an increased ability to take up lipids in the
204 gut.
205
206 Interestingly, diet had a strong influence on the number and direction of gene expression changes between
207 freshwater and saltwater (Figure 6). In gut, the transcriptional changes between fresh- and seawater was
208 twice as high when fed FO diet than VO diet (Figure 6a). In liver, the diet effect was less pronounced,
209 with the FO group containing 46% more DEGs than the VO group (Figure 6a). Identical patterns were
210 seen for lipid metabolism genes with 89% and 16% more DEGs in FO group for gut and liver,
211 respectively (Figure 6b). As this diet and life-stage interaction is a genome wide trend, and more
212 pronounced in external barrier gut tissue than in liver, this pattern could be related to differences in
213 osmoregulation and adaptation to saltwater. The higher levels of essential omega-6 FA linoleic acid (18:2
214 n-6, LA) in VO diets could result in increased levels of arachidonic acid (20:4 n-6, ARA). ARA is a
215 precursor to eicosanoids with a multitude of biological functions including osmoregulation (Mustafa &
216 Srivastava, 1989; Oxley et al., 2010). Another possibility is that the different levels of fatty acids in the

217 diets, for example DHA, affect DNA-methylation and thus trigger genome wide divergence in gene
218 regulation (Kulkarni et al., 2011).



219

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

225

226 Our results clearly demonstrate very different baseline lipid metabolic functions in pre- and post-smolt
227 salmon, as well as life-stage associated changes in the plasticity of lipid metabolism, e.g. the ability to
228 regulate endogenous LC-PUFA synthesis as a response to changes in diet (i.e. lipid content). As
229 opportunistic carnivores, salmon tend to eat whatever the local environment provides. Thus, in
230 freshwater, insects and amphipods provide variable, mostly low amounts of essential LC-PUFA (Jonsson
231 & Jonsson, 2011; Sushchik, Gladyshev, Moskvichova, Makhutova, & Kalachova, 2003), favoring a
232 metabolic function that can efficiently regulate endogenous lipid synthesis based on dietary availability

233 (Carmona-Antonanzas, Tocher, Martinez-Rubio, & Leaver, 2014). Conversely, marine environments
234 provide a stable source of essential dietary lipids, promoting a metabolic function that allocate less energy
235 to endogenous synthesis of essential lipids.

236 Selection on gene duplicates after whole genome duplication

237 Carmona-Antonanzas et al (2014, 2016) proposed that the salmonid whole-genome duplication may have
238 adaptively increased the potential for endogenous lipid synthesis. We pursued this hypothesis by
239 searching for distinct signatures of selection pressure on lipid metabolism genes in salmon. Specifically,
240 we compared pathways in terms of their tendency to retain both duplicates of gene pairs, in terms of
241 whether duplicates showed similar regulation (expression patterns across tissues, diets and environments),
242 and in terms of total gene dosage (for the one or two genes retained of a pair) in salmon compared to pike,
243 its closest unduplicated relative.
244
245 To assess the level of Ss4R duplicate retention, we first defined 10,752 Ss4R duplicate pairs (21,504
246 genes) in the NCBI refseq annotation using the same approach as Lien et al. (2016). Of the 1,421
247 annotated lipid metabolism genes, 867 (61%) were retained as duplicated genes after Ss4R (Figure 7a) (in
248 contrast to 47% of the 45,127 salmon genes assigned to ortholog groups). Moreover, our results showed
249 large variation in the proportion of retained duplicates in each lipid metabolism pathway (Figure 7), with
250 the most extreme case being '*fat digestion and absorption*' with 80% retained duplicates and '*steroid*
251 *hormone biosynthesis*' with only 27% retained Ss4R duplicates.
252
253 The regulatory conservation of the duplicates was then estimated from RNA-seq data representing a time
254 course of dynamic changes in gene expression and lipid metabolism function in liver. Fish in the same
255 feeding trial as described above were switched from VO to FO feed and vice versa, in both fresh and
256 saltwater conditions (see methods and supplementary data for details). In total, 38 sampling time points
257 (19 in freshwater and 19 in saltwater) from the switch experiment were used to calculate co-expression
258 correlation for Ss4R duplicate pairs. Pathway-level analyses showed that conservation at the regulatory
259 level was not associated with duplicate retention (Figure 7). For example, the *'biosynthesis of unsaturated*
260 *fatty acids'* pathway had significantly fewer duplicates retained than expected by chance (P-value <
261 0.0234), but a significant overrepresentation of duplicate pairs that display highly similar regulation (P-
262 value < 0.0142 and < 0.0361 in freshwater and saltwater, respectively). Other pathways showing
263 signatures of increased duplicate co-regulation were *'insulin signalling pathway'*, *'terpenoid backbone*
264 *biosynthesis'*, *'steroid biosynthesis'*, *'fat digestion and absorption'*, and *'fatty acid metabolism'* (Figure
265 7b-c). Over all, 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 was also those that were most responsive to dietary differences in fatty acids? (Figure 3).

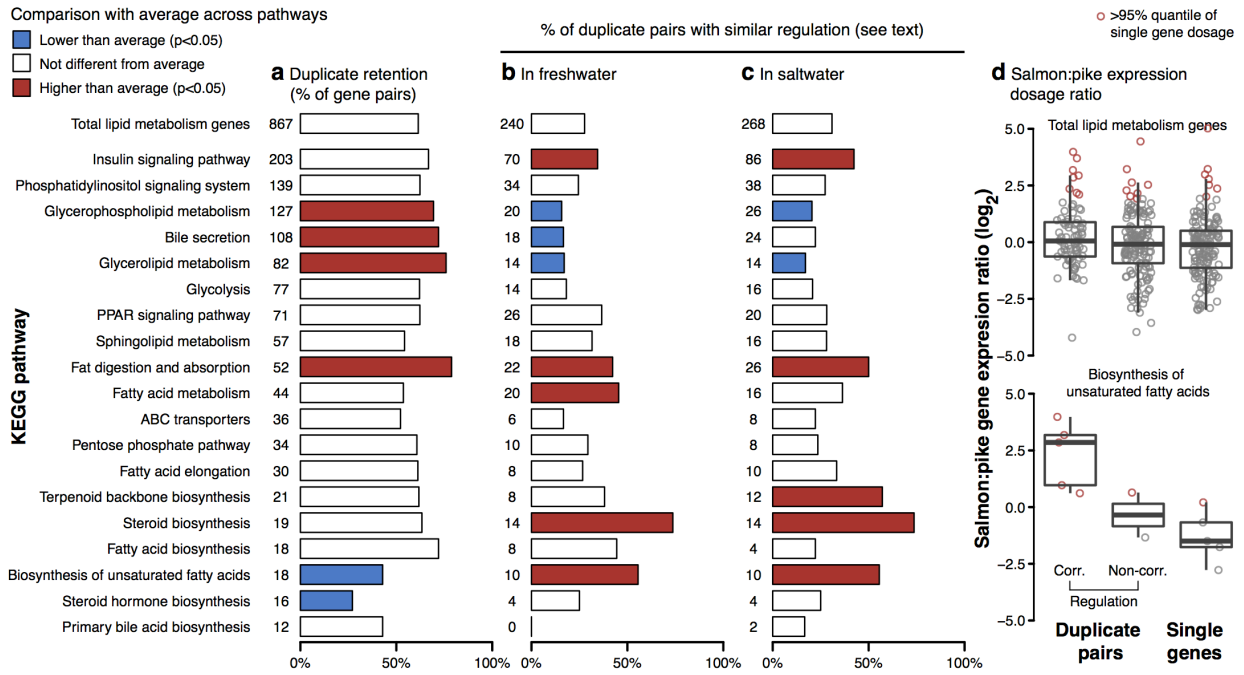


Figure 7 - Gene duplication in lipid metabolism pathways. For the total list of lipid metabolism genes in Atlantic salmon, and sets of genes belonging to different KEGG pathways: (a) Number and percentage of genes with a duplicate homolog from the Ss4R duplication. (b) Number and percentage of duplicate 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.

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) (Supplementary file 10). Compared to single copy genes, the analyses of expression sum of duplicate pairs did not reveal a systematic increase in gene dosage, 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 links 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, Miyazaki, O'Neill, Ding, & Ntambi, 2016). Although we do not see a general trend of increased gene dosage of 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

Seagoing salmon needs great plasticity of physiology and behaviour to adapt for migration between freshwater and sea. By analyzing transcriptomic changes through the transition from fresh- to saltwater, we identified an overall remodeling of lipid metabolism, with liver being more active in freshwater to gut being more active in saltwater. This baseline remodelling was modulated by diet: Our experiment also included life-long dietary contrasts showing that lipid metabolism was significantly more responsive to dietary differences in lipid composition in fresh water than 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 this life-stage related metabolic reprogramming is controlled, e.g. to what extent it reflects a "hard coded" physiological transformation before sea migration (for example through epigenetic remodeling).

Materials and methods

Orthogroup prediction

Protein sequences were obtained from seven teleost fish species; *Danio rerio* (zebrafish), *Gasterosteus aculeatu* (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.*, awaiting publication of data). 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, Misawa, Kuma, & Miyata, 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 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, such as genes 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.

346 # Tissue expression

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

356 # Feed trial

357 Atlantic salmon fry were obtained from AquaGen Breeding Centre, Kyrksæterøra, Norway and reared in
358 the Norwegian Institute for Water Research (NIVA), Solbergstranda, Norway on vegetable oil (VO) or
359 fish oil (FO) based diets from first feeding (fry weight <0.2g). VO based feeds contained a combination
360 of linseed oil and palm oil at a ratio of 1.8:1 and FO based feeds contained only North Atlantic fish oil.
361 All feeds were formulated and produced by EWOS innovation. Smoltification was triggered by 5 weeks
362 of winter-like conditions with 12 hours of light per day followed by spring-like conditions with 24 hours
363 of light per day. Salmon were then immediately switched to saltwater and allowed to acclimate for 3
364 weeks before first sampling. Pre-smolt salmon in freshwater (~50g) and post-smolt salmon in saltwater
365 (~200g) were either switched to the contrasting feed condition (VO to FO and vice versa) or maintained
366 on the same feed as a control. Fish samples were taken on days 0, 1, 2, 6, 9, 16, and 20 post each diet
367 switching, with five fish from each of two replicate tanks per treatment (5*2). All fish were euthanized by
368 a blow to the head and samples of liver and midgut (gut section between pyloric caeca and hindgut) were
369 flash frozen in liquid nitrogen and stored under -80℃ for further analysis.

370 # RNA sequencing

371 Total RNA was extracted from selected feed trial samples using the RNeasy Plus Universal kit
372 (QIAGEN). Quality was determined on a 2100 Bioanalyzer using the RNA 6000 nano kit (Agilent).
373 Concentration was determined using a Nanodrop 8000 spectrophotometer (Thermo Scientific). cDNA
374 libraries were prepared using the TruSeq Stranded mRNA HT Sample Prep Kit (Illumina). Library mean
375 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 100bp reads.

Differential expression analysis of feed and life stages

To analyse gene expression differences between feed types and life stages, samples from the feed trial were selected for RNA-seq. The liver and gut tissue of a number of replicate fish were sequenced for each of the feeds (FO, VO) at day 0 of the diet switch, both before (freshwater) and after (saltwater) smoltification. Fastq files were processed to produce gene count and FPKM data using the same protocol described under the tissue expression method section. For feed comparison, changes in gene expression were tested between FO and VO feeds for both freshwater and saltwater samples, and liver and gut tissues. For life stage (water condition) comparison, changes in gene expression were tested between freshwater and saltwater stages for both FO and VO feed samples, 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 an FDR < 0.05 were considered a differentially expression gene (DEG).

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 $\geq 50\%$ 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

Gene from the lipid metabolism gene list were paired together with their putative Ss4R duplicate identified above. The retention of gene duplicates (number of genes in duplicate pairings) 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

406 KEGG pathways (Table S1) in a duplicate pairing to that of the total list of lipid genes to find pathways
407 with significantly less or more duplicate retention (Fisher's exact test, P-value < 0.05). Regulatory
408 conservation of lipid gene duplicates was explored by correlation of gene expression changes between
409 duplicates over the course of the feed trial described above. RNA-seq data was generated from liver
410 samples of salmon from 38 sampling time points (19 in freshwater and 19 in saltwater). Fastq files were
411 processed to produce gene count and FPKM data using the same protocol described under the tissue
412 expression method section. For each duplicate pair, mean FPKM values were retrieved for each time
413 point and used to calculate a freshwater and saltwater correlation value. Duplicates with Pearson
414 correlation ≥ 0.6 were considered correlated (P-value < 0.003 from 19 sample points). The number of
415 duplicates with correlated expression profiles was counted for each pathway and compared to all lipid
416 genes to find pathways with significantly less or more correlated duplicates (Fisher's exact test, P-value <
417 0.05). The effect of gene duplication on gene dosage was estimated by calculating a dosage ratio between
418 the FPKM value of a salmon duplicate (sum of gene expression in duplicate pairs) over the FPKM value
419 of the non-duplicated ortholog from northern pike. For salmon, the RNA-seq data from the freshwater and
420 saltwater FO fed trial was used (samples used in differential expression analysis section). For pike, RNA-
421 seq from livers of four individuals were aligned (see tissue expression section for protocol) to their
422 respective genomes (see genomes in ortholog prediction section). RSEM (v1.2.31) (Li & Dewey, 2011)
423 was used to generate FPKM values for genes so that non-uniquely mapped reads between duplicate genes
424 were not ignored but instead assigned to each gene to match the proportions of uniquely mapped reads
425 between the genes. Gene dosage levels for duplicate pairs with correlated expression (see above), non-
426 correlated expression and single genes was compared for all lipid metabolism genes and for each
427 pathway.

428 References

429 Allendorf, F. W., & Thorgaard, G. H. (1984). Tetraploidy and the Evolution of Salmonid Fishes. In *Evolutionary Genetics of*
430 *Fishes* (pp. 1–53). Boston, MA: Springer US. http://doi.org/10.1007/978-1-4684-4652-4_1
431 Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data.
432 *Bioinformatics*, *31*(2), 166–169. http://doi.org/10.1093/bioinformatics/btu638
433 Berthelot, C., Brunet, F., Chalopin, D., Juanchich, A., Bernard, M., Noël, B., … Guiguen, Y. (2014). The rainbow trout genome
434 provides novel insights into evolution after whole-genome duplication in vertebrates. *Nature Communications*, *5*, 3657.
435 http://doi.org/10.1038/ncomms4657
436 Bond, L. M., Miyazaki, M., O'Neill, L. M., Ding, F., & Ntambi, J. M. (2016). Fatty Acid Desaturation and Elongation in
437 Mammals. In *Biochemistry of Lipids, Lipoproteins and Membranes* (pp. 185–208). Elsevier. http://doi.org/10.1016/B978-
438 0-444-63438-2.00006-7
439 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*, *41*(5), 423–436. http://doi.org/10.1007/s11745-006-5116-4

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

Kortner, T. M., Björkhem, I., Krasnov, A., Timmerhaus, G., & Krogdahl, Å. (2014). Dietary cholesterol supplementation to a plant-based diet suppresses the complete pathway of cholesterol synthesis and induces bile acid production in Atlantic salmon (Salmo salar L.). *British Journal of Nutrition*, *111*(12), 2089–2103. http://doi.org/10.1017/S0007114514000373

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

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

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

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

Li, B., & Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference

genome. *BMC Bioinformatics*, *12*(1), 323. http://doi.org/10.1186/1471-2105-12-323

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

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

Macqueen, D. J., & Johnston, I. A. (2014). A well-constrained estimate for the timing of the salmonid whole genome duplication reveals major decoupling from species diversification. *Proceedings of the Royal Society of London B: Biological Sciences*, *281*(1778).

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

Mustafa, T., & Srivastava, K. C. (1989). Prostaglandins (Eicosanoids) and Their Role in Ectothermic Organisms (pp. 157–207). Springer Berlin Heidelberg. http://doi.org/10.1007/978-3-642-74510-2_6

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

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

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

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

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

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

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

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

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

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

Tocher, D. R., Bell, J. G., MacGlaughlin, P., McGhee, F., & Dick, J. R. (2001). Hepatocyte fatty acid desaturation and polyunsaturated fatty acid composition of liver in salmonids: effects of dietary vegetable oil. *Comparative Biochemistry*

524 *and Physiology Part B: Biochemistry and Molecular Biology*, *130*(2), 257–270. http://doi.org/10.1016/S1096-
525 4959(01)00429-8
526 Tocher, D. R., Fonseca-Madrigal, J., Bell, J. G., Dick, J. R., Henderson, R. J., & Sargent, J. R. (2002). Effects of diets containing
527 linseed oil on fatty acid desaturation and oxidation in hepatocytes and intestinal enterocytes in Atlantic salmon (Salmo
528 salar). *Fish Physiology and Biochemistry*, *26*(2), 157–170. http://doi.org/10.1023/A:1025416731014
529 Verleyen, T., Forcades, M., Verhe, R., Dewettinck, K., Huyghebaert, A., & De Greyt, W. (2002). Analysis of free and esterified
530 sterols in vegetable oils. *Journal of the American Oil Chemists' Society*, *79*(2), 117–122. http://doi.org/10.1007/s11746-
531 002-0444-3
532 Zheng, X., Torstensen, B. E., Tocher, D. R., Dick, J. R., Henderson, R. J., & Bell, J. G. (2005). Environmental and dietary
533 influences on highly unsaturated fatty acid biosynthesis and expression of fatty acyl desaturase and elongase genes in liver
534 of Atlantic salmon (Salmo salar). *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, *1734*(1),
535 13–24. http://doi.org/10.1016/j.bbalip.2005.01.006
536

537

538

539

540

541

542

Supplemental material

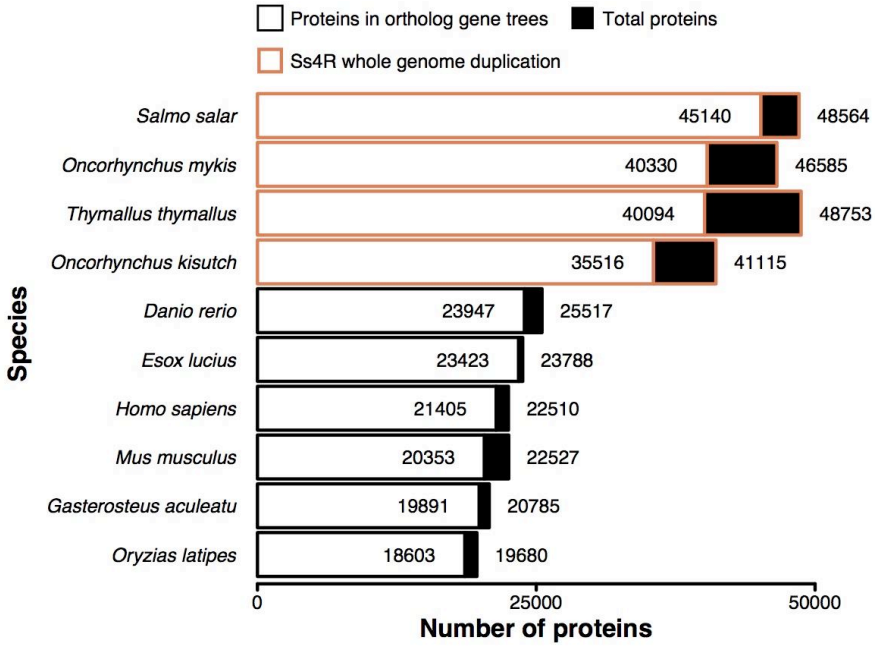


Figure S1: Number of proteins in ortholog gene trees for each species

553
554

KEGG ID	Pathway name
dre00010	Glycolysis
dre00561	Glycerolipid metabolism
dre04070	Phosphatidylinositol signaling system
dre04910	Insulin signaling pathway
dre00600	Sphingolipid metabolism
dre01212	Fatty acid metabolism
dre03320	PPAR signaling pathway
dre00564	Glycerophospholipid metabolism
ko04976	Bile secretion
dre02010	ABC transporters
dre00900	Terpenoid backbone biosynthesis
ko04975	Fat digestion and absorption
dre00100	Steroid biosynthesis
dre00061	Fatty acid biosynthesis
dre00140	Steroid hormone biosynthesis
dre00062	Fatty acid elongation
dre00120	Primary bile acid biosynthesis
dre01040	Biosynthesis of unsaturated fatty acids
dre00030	Pentose phosphate pathway

555
556 Table S1: KEGG pathways related to lipid metabolism selected for gene annotation
557

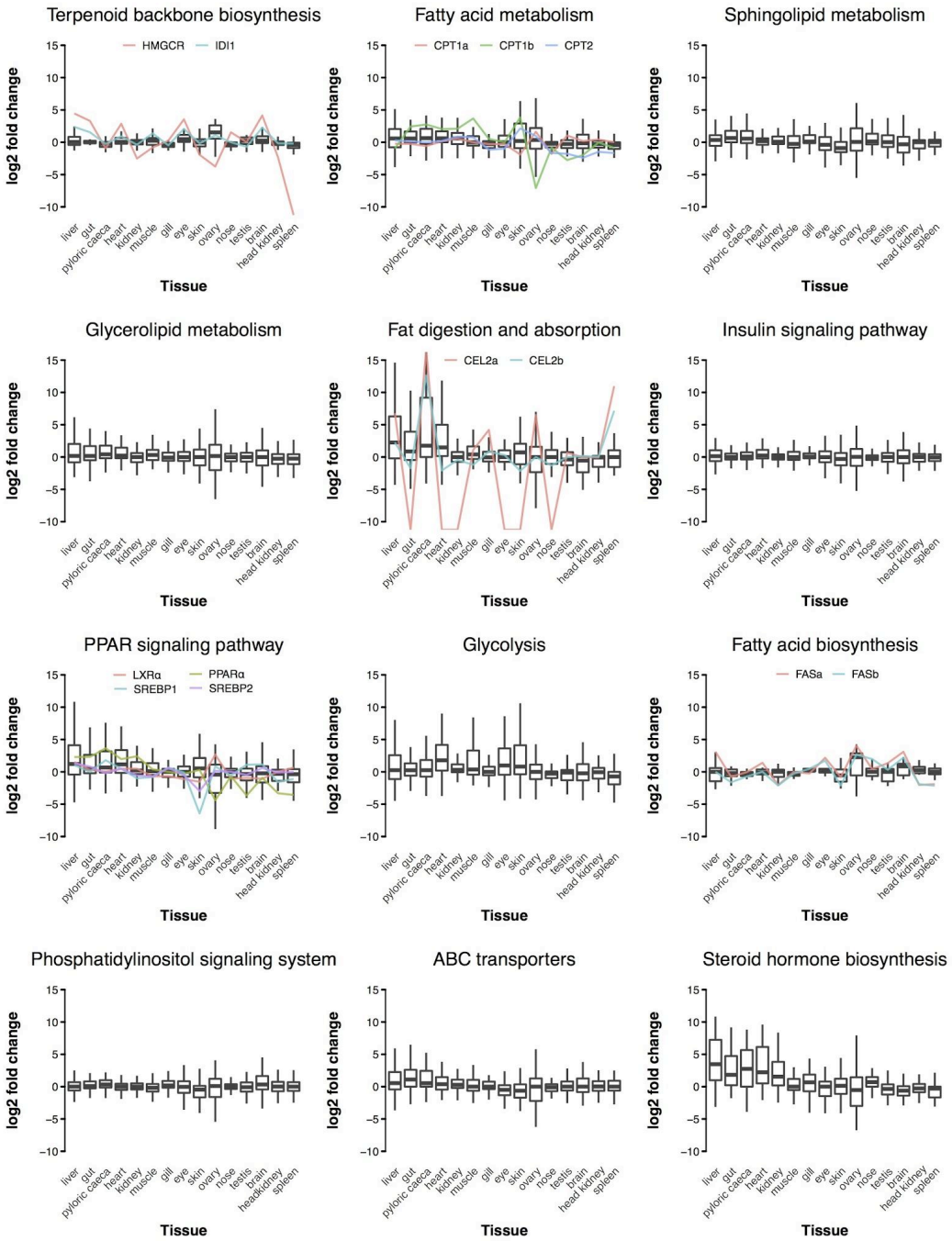


Figure S2: Tissue expression profiles of salmon genes in lipid metabolism pathways.

Tissue expression profiles of salmon genes in KEGG pathways for *Salmo salar* tissues; pyloric caeca, liver, heart, gut, skin, kidney, muscle, eye, ovary, brain, nose, gill, testis, spleen, and head kidney. Gene expression levels are shown as the log2 fold change difference between each tissue FPKM to the medium FPKM across all tissues. Expression of select genes of interest are overlaid. These are 3-hydroxy-3-methyl-glutaryl-CoA reductase (HMGCR), isopentenyl-diphosphase Δisomerase (IDI1), carnitine palmitoyltransferase 1 (CPT1a and CPT1b) and 2 (CPT2), carboxyl ester lipase, tandem duplicate 2 (CEL2a and CEL2b), liver x receptor alpha (LXRαa), peroxisome proliferator-activated receptor alpha (PPARαa), sterol regulatory element binding protein 1 (SREBP1) and 2 (SREBP2), and fatty acid synthase (FASa and FASb).

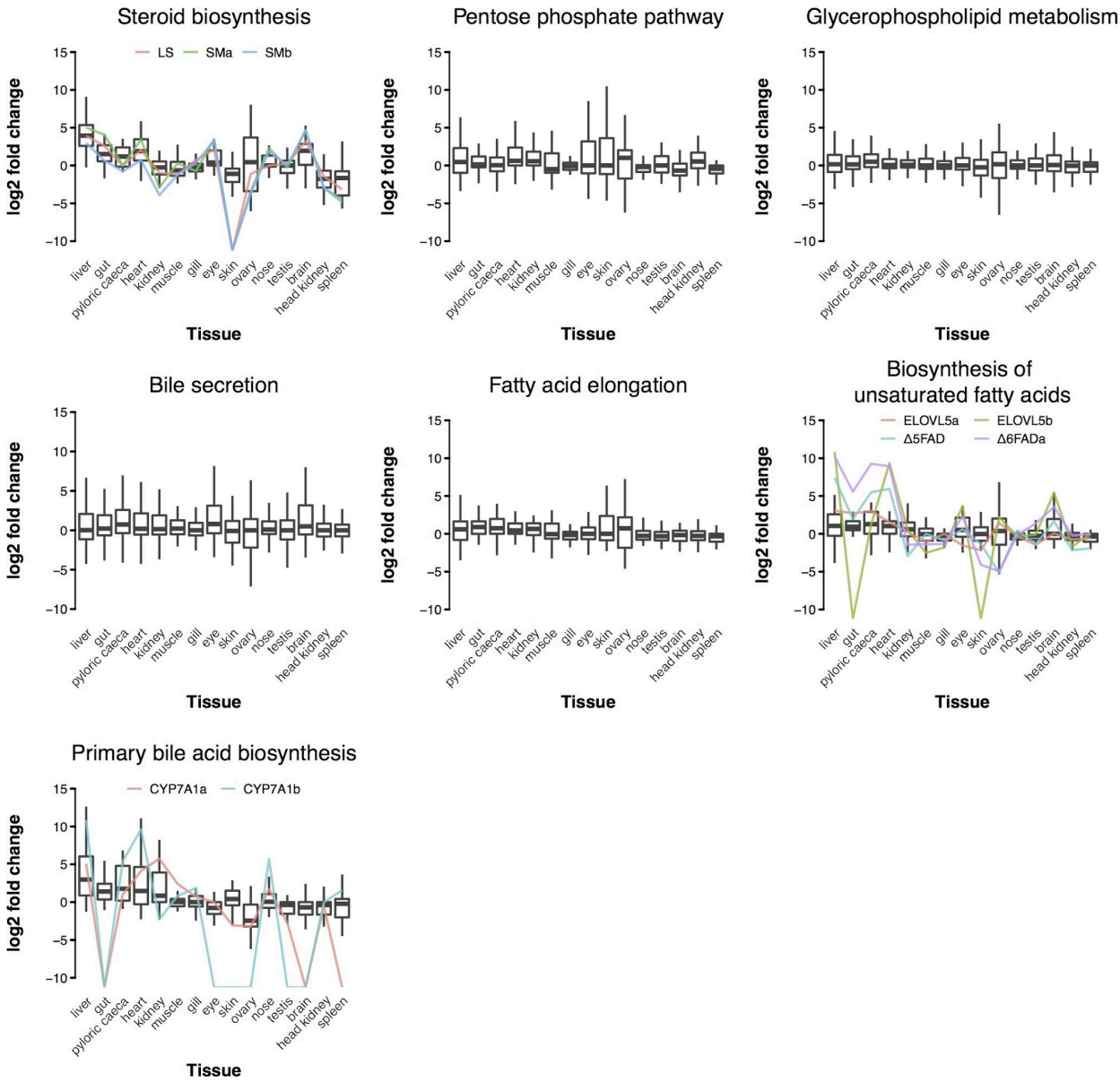


Figure S3: Tissue expression profiles of salmon genes in lipid metabolism pathways. Tissue expression profiles of salmon genes in KEGG pathways for *Salmo salar* tissues; pyloric caeca, liver, heart, gut, skin, kidney, muscle, eye, ovary, brain, nose, gill, testis, spleen, and head kidney. Gene expression levels are shown as the log2 fold change difference between each tissue FPKM to the medium FPKM across all tissues. Expression of select genes of interest are overlaid. These are lanosterol synthase (LS), squalene epoxidase (SMa and SMb), fatty acid elongase 5 (ELOVL5a and ELOVL5b), delta 5 fatty acid desaturase (Δ5FAD), delta 6 fatty acid desaturase (Δ6FAD), and cytochrome P450 7A1 (CYP7A1a and CYP7A1b).

	Freshwater		Saltwater	
Fatty Acid	MA	VO	MA	VO
14:0	6.4	1.5	6.5	1.6
15:0	0.5	0.1	0.6	0.1
16:0	15.4	18.5	15.3	19.5
18:0	2.8	3.8	3.8	4.8
20:0	0.1	0.2	0.1	0.1
Total Saturates	25.2	24.1	26.3	26.1
16:1n7	4.1	0.9	3.8	0.8
18:1n9	16.9	23.5	14.9	23.5
20:1n9	8	1.4	8.2	1.4
22:1n11	12	2.2	12.4	2.1
Total Monounsaturates	41	28	39.3	27.8
18:2n6	4.6	12.7	4.2	12.5
18:3n6	0.1	0.1	0.1	0.1
20:2n6	2.7	0.5	3	0.5
20:4n6	0.3	0.1	0.3	0.1
Total n-6 PUFA	7.7	13.4	7.6	13.2
18:3n3	2.11	28.8	2.5	26.8
20:3n3	0.2	0.1	0.3	0.1
20:5n3	5.8	1.3	6.2	1.4
22:5n3	0.9	0.2	0.8	0.2
22:6n3	10.7	2.3	11.7	2.7
Total n-3 PUFA	19.71	32.7	21.5	31.2
EPA+DHA	16.5	3.6	17.9	4.1

Table S3: FA composition of feeds. All values are expressed as mass percentage of lipid fraction.

pathway_name	total	FW.sig	SW.sig	FW.SW.sig	FW.p	SW.p	FW.SW.p
Phosphatidylinositol signaling system	223	2	2	1	0.0090	0.0090	0.0045
ABC transporters	69	1	NA	NA	0.0145	NA	NA
Sphingolipid metabolism	105	1	1	NA	0.0095	0.0095	NA
Glycolysis	124	2	1	NA	0.0161	0.0081	NA
Insulin signaling pathway	304	5	3	NA	0.0164	0.0099	NA
Bile secretion	150	4	3	1	0.0267	0.0200	0.0067
Glycerophospholipid metabolism	183	3	5	NA	0.0164	0.0273	NA
Fatty acid elongation	49	2	1	NA	0.0408	0.0204	NA
Primary bile acid biosynthesis	28	1	1	NA	0.0357	0.0357	NA
Glycerolipid metabolism	108	4	4	NA	0.0370	0.0370	NA
Steroid hormone biosynthesis	59	3	4	1	0.0508	0.0678	0.0169
PPAR signaling pathway	114	9	8	5	0.0789	0.0702	0.0439
Fat digestion and absorption	66	2	5	NA	0.0303	0.0758	NA
Pentose phosphate pathway	56	5	1	NA	0.0893	0.0179	NA
Biosynthesis of unsaturated fatty acids	42	5	3	2	0.1190	0.0714	0.0476
Fatty acid metabolism	82	10	6	3	0.1220	0.0732	0.0366
Fatty acid biosynthesis	25	4	3	1	0.1600	0.1200	0.0400
Terpenoid backbone biosynthesis	34	11	NA	NA	0.3235	NA	NA
Steroid biosynthesis	30	17	5	3	0.5667	0.1667	0.1000

590
591 Table S4: Table of total and DEG gene counts for each lipid metabolism pathway
592
593

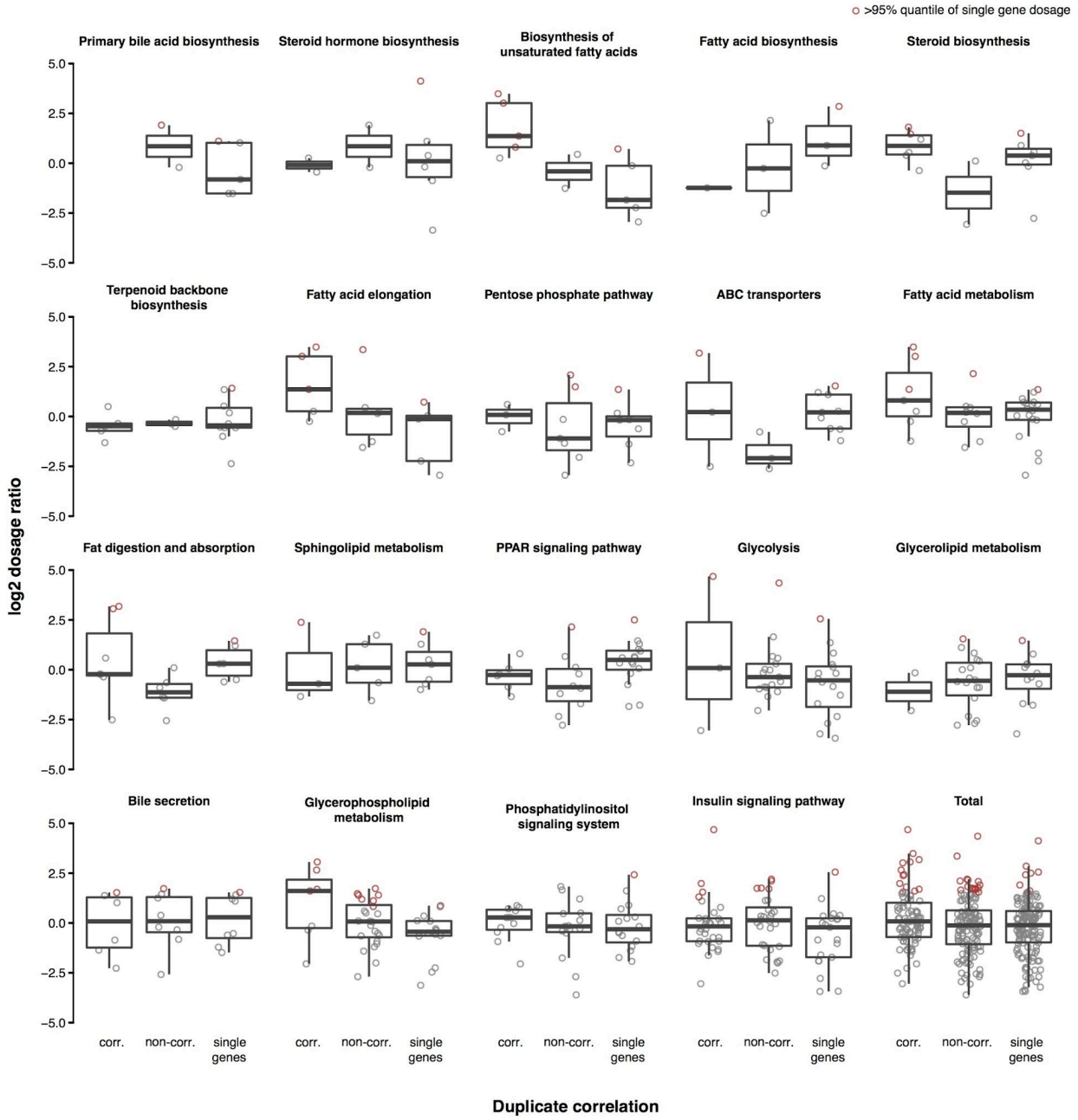


Figure S4: Gene dosage ratios and duplicate correlation for saltwater conditions
Log2 of gene dosage ratios of saltwater salmon ortholog expression in liver (FPKM, summed for duplicates) to expression of a pike ortholog, for total lipid metabolism genes and genes 19 KEGG pathways. Duplicates were grouped into correlated (corr.) or non-correlated (non-corr.) based on saltwater correlation results. Dosage ratios (points) greater than the 95% quantile of single gene dosages in the plot are marked in red.

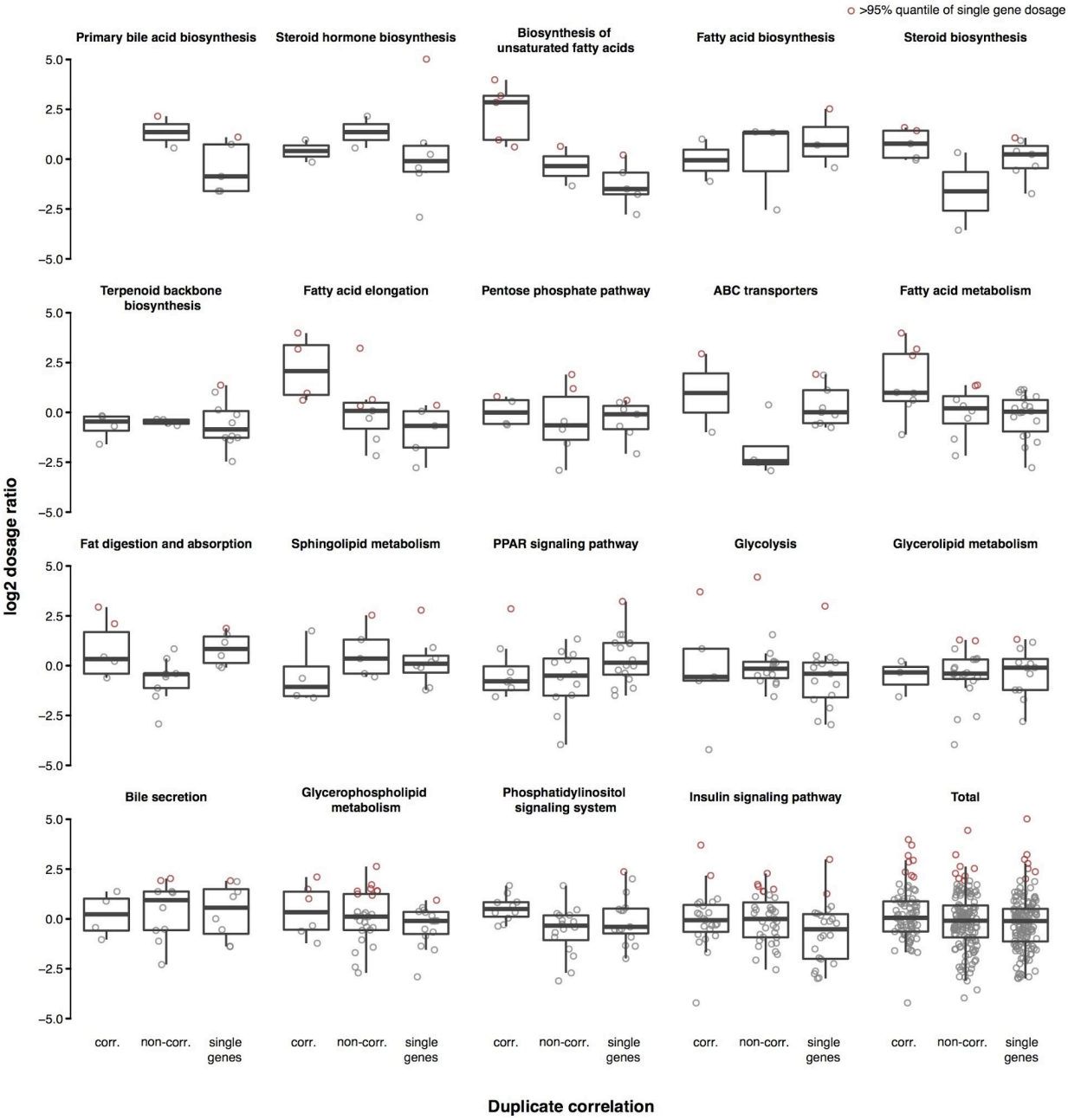


Figure S5: Gene dosage ratios and duplicate correlation for freshwater conditions
Log2 of gene dosage ratios of freshwater salmon ortholog expression in liver (FPKM, summed for duplicates) to expression of a pike ortholog, for total lipid metabolism genes and genes 19 KEGG pathways. Duplicates were grouped into correlated (corr.) or non-correlated (non-corr.) based on freshwater correlation results. Dosage ratios (points) greater than the 95% quantile of single gene dosages in the plot are marked in red.