

Accepted refereed manuscript of:

De Santis C, Bartie K, Olsen RE, Taggart J & Tocher DR (2015) Nutrigenomic profiling of transcriptional processes affected in liver and distal intestine in response to a soybean meal-induced nutritional stress in Atlantic salmon (*Salmo salar*), *Comparative Biochemistry and Physiology - Part D: Genomics and Proteomics*, 15, pp. 1-11.

DOI: [10.1016/j.cbd.2015.04.001](https://doi.org/10.1016/j.cbd.2015.04.001)

© 2015, Elsevier. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International  
<http://creativecommons.org/licenses/by-nc-nd/4.0/>

1 **Nutrigenomic profiling of transcriptional processes affected in liver and distal intestine**  
2 **in response to a soybean meal-induced nutritional stress in Atlantic salmon (*Salmo***  
3 ***salar*).**

4

5 C. De Santis<sup>1\*</sup>, K.L. Bartie<sup>1</sup>, R.E. Olsen<sup>2,3</sup>, J.B. Taggart<sup>1</sup> and D.R. Tocher<sup>1</sup>

6

7 <sup>1</sup> Institute of Aquaculture, School of Natural Sciences, University of Stirling, Stirling FK9  
8 4LA, Scotland

9 <sup>2</sup> Institute of Marine Research, Matre, 5984 Matredal, Norway

10 <sup>3</sup> Norwegian University of Science and Technology, Department of Biology, 7491  
11 Trondheim, Norway

12

13 **Keywords:** dietary protein substitution, transcriptome, liver, intestine, Atlantic salmon, *Salmo*  
14 *salar*, soybean meal, nutrigenomic, nutritional stress

15

16

17 **Abstract**

18 The aim of the present study was to generate an experimental model to characterize the  
19 nutrigenomic profile of a plant-derived nutritional stress. Atlantic salmon (*Salmo salar*) was  
20 used as the model species. The nutritional stress was induced by inclusion of dietary defatted  
21 soybean meal (SBM), as this ingredient had been previously demonstrated to induce  
22 enteropathy in the distal intestine and reduce growth in salmon. Triplicate groups of Atlantic  
23 salmon were fed increasing concentrations (0, 100, 200 and 300 g kg<sup>-1</sup>) of SBM for 12 weeks  
24 and reduced growth performance was used as the indicator of nutritional stress. The  
25 transcriptome was analysed in two tissues, liver and distal intestine, with the hypothesis being  
26 that the liver transcriptome would be characterized by gene expression responses related to  
27 overall growth and health performance, whereas intestinal gene expression would be  
28 dominated by specific responses to SBM. A set of 133 genes was differentially expressed in  
29 liver including 44 genes in common with the intestinal response. The liver specific response  
30 included up-regulation of genes involved in protein digestion, energy metabolism and  
31 immune functions, whereas genes regulated in other metabolic pathways were generally  
32 anabolic and down-regulated. These responses may be more related to general nutritional  
33 stress than to SBM *per se*. The transcriptomic profile in the distal intestine was consistent  
34 with the enteritis response as described previously. This study provides a comprehensive  
35 report on the profiles of liver and distal intestine transcriptomes, highlighting the role of the  
36 former tissue in fish undergoing SBM-induced nutritional stress.

37

## 38 **1. Introduction**

39 Aquaculture is a major player of the global food security program producing approximately  
40 fifty percent of the world seafood supply (The World Bank, 2013). In an effort to provide the  
41 growing population with a consistent supply of high-quality and sustainable seafood,  
42 alternative sources of protein to the traditionally used fishmeal (FM) are now being sought  
43 after, particularly for industries involved in the farming of carnivorous fish species such as  
44 Atlantic salmon (*Salmo salar*). Over the last two decades, significant research investment has  
45 been made to identify alternative sources of protein that could fulfil the requirements of these  
46 fish, provide a competitive growth rate and good flesh properties, and avoid any negative  
47 impact on fish welfare. Thus far, plant-based ingredients have proved to be a potentially  
48 attractive and sustainable alternative mainly due to a competitive market price and  
49 availability. The use of plant protein products such as soybean, pea and other legumes, wheat  
50 and corn gluten is either already established in commercial feed manufacturing protocols or is  
51 proven to be a viable alternative to FM (Gatlin et al., 2007).

52 One limitation of plant-based products, however, is the presence of chemical compounds  
53 known as antinutritional factors (ANFs), endogenously produced by the plant with a variety  
54 of functions including structural (e.g. fibres), storage (e.g. phytic acid) or as defence  
55 mechanisms (e.g. saponins, lectins and protease inhibitors; reviewed by Francis et al., 2001,  
56 Gatlin et al., 2007, Krogdahl et al., 2010). Processes to concentrate protein in plant products  
57 can lower the concentration of ANFs, but ANFs can also co-purify with proteins and be  
58 present in protein concentrates utilized as an alternative to FM. Fish are generally not able to  
59 metabolize ANFs and when these are present in their diets they can cause a number of effects  
60 primarily associated with digestive physiology, health and metabolism, impacting on fish  
61 welfare and resulting in reduced productivity (Francis et al., 2001). For example, substances  
62 such as fibre and phytic acid can affect the uptake and utilization of micronutrients, while

63 protease inhibitors, saponins and phytosterols can interfere with the digestion and/or  
64 absorption of proteins, lipids and cholesterol, respectively (Gatlin et al., 2007). In addition,  
65 some ANFs such as saponins and lectins can have important health consequences primarily  
66 affecting the intestinal mucosa, causing inflammation and increasing the permeability to  
67 pathogens and other unwanted substances (Krogdhal et al., 2010). It follows that the  
68 nutritional profile of plant-based protein raw materials does not always reflect the true value  
69 available to the animal for absorption and utilization, but varies depending on the presence  
70 and abundance of ANFs.

71 Understanding how fish respond to and utilize dietary plant ingredients has therefore been a  
72 key topic of research in recent years. The study of individual ANFs has been a useful  
73 approach to unveil mechanisms behind the biological response of fish to these chemicals and  
74 research in this direction will continue to provide crucial information (Buttle et al., 2001,  
75 Francis et al., 2002, Krogdahl et al., 2010). However, from a commercial and applied  
76 perspective it is critical to maintain a holistic approach particularly as it is increasingly  
77 evident that certain effects are only visible when two or more ANFs are present in  
78 combination (Kortner et al., 2012, Krogdahl et al., 2010). The high-demand for alternative  
79 protein sources continues to be a strong driver for research testing novel raw materials  
80 individually or in combination (Azaza et al., 2009, Booth et al. 2001, Glencross et al., 2005,  
81 Hartviksen et al., 2014, Kaushik et al., 2004, Kortner et al., 2013, Øverland et al., 2009,  
82 Panserat et al., 2009). In addition, novel processing technologies are contributing to the  
83 introduction of new products from the same raw materials but with improved nutritional  
84 characteristics and reduced levels of ANFs. When testing ingredients with reduced levels of  
85 ANFs, a macroscopic physiological response might not be measurable in short-term  
86 experimental trials, whereas cumulative adverse effects, gut pathologies or other detrimental  
87 physiological conditions might accrue under long-term or more challenging commercial  
88 conditions (Krogdahl et al., 2010).

90 It is therefore increasingly important at experimental levels to have sensitive tools at our  
91 disposal that are able to detect subtle physiological changes and help predict long-term  
92 detrimental effects of dietary modification. This approach is the foundation of nutrigenomics,  
93 a science still in its infancy that studies the influence of food or food constituents on the  
94 transcriptome. In nutrigenomics, nutrients are considered signals through which cells interpret  
95 information about the environment (diet) and respond, according to necessity, by modifying  
96 metabolic pathways through regulation of gene and protein expression towards homeostasis  
97 (Muller and Kersten, 2003). Nutrigenomics can provide a means to interpret how dietary  
98 ingredients are perceived by fish at the tissue level and also provide a molecular snapshot of  
99 the physiological response of specific tissues. Nutrigenomic studies have utilized both  
100 candidate gene and high-throughput approaches to investigate the physiological responses of  
101 tissues and mechanisms of adaptation to plant-based dietary ingredients in fish or to  
102 understand the molecular mechanisms underlying pathologies such as enteritis caused by  
103 these ingredients (Gu et al., 2014, Kortner et al., 2012, Kortner et al., 2013, Panserat et al.,  
104 2009, Tacchi et al., 2012, Wacyk et al., 2012). However, in order to develop a practical  
105 diagnostic use of nutrigenomic approaches, it is essential to determine “dietary signatures”  
106 that characterize the physiological response to nutritional stressors.

107 The overarching aim of the present study was to generate an experimental model to  
108 characterize the nutrigenomic profile of a plant-derived nutritional stress. A nutritional stress  
109 was induced in Atlantic salmon using dietary defatted soybean meal (SBM), which has been  
110 demonstrated to induce enteropathy in the distal intestine (Baeverfjord and Kroghdahl, 1996,  
111 Urán et al., 2008, Urán et al., 2009). Impaired growth was taken as an indicator of pronounced  
112 nutritional stress. Transcriptomic analyses were performed in two tissues, liver and distal  
113 intestine. Liver was chosen for being arguably the key metabolically active tissue responding

114 to diet, while the distal intestine has been studied for the development of specific pathologies  
115 associated with plant ingredients such as SBM (Baeverfjord and Krogdahl, 1996, Kortner et  
116 al., 2012). The results provided: *a*) an insight into the molecular signatures resulting from a  
117 nutritional stress response to which future nutrigenomic studies can refer to; and *b*) a platform  
118 for the identification of candidate genes for the molecular phenotyping of several  
119 physiological parameters in liver and distal intestine. These data were specifically analysed to  
120 test the hypothesis that the liver transcriptome would include gene expression responses that  
121 could be more general and related to overall growth/health performance, whereas intestinal  
122 gene expression would be dominated by specific responses to SBM.

## 123 **2. Materials and Methods**

### 124 *2.1. Nutritional trial, diets and sampling*

125 The nutritional trial was conducted at the Institute of Marine Research, Matre, Norway using  
126 Atlantic salmon (*Salmo salar*) of the commercial Aquagen strain (Aquagen Ltd,  
127 Kyrksæterøra, Norway). The fish were normal smolts (1+) produced in 5 m indoor tanks  
128 under natural photoperiod (June 2012). A total of 540 fish were anaesthetized in Finquel,  
129 tagged with PIT tags, and equally distributed into 12 white fiberglass tanks each holding  
130 400 L (95 cm x 95 cm x 60 cm, L x W x H). The tanks were closed with lids, and supplied  
131 with two fluorescent light tubes (18 W each) and an automatic feeder (ARVO-TEC T Drum  
132 2000, Arvotec, Huutokoski, Finland). After two days, the salinity was gradually increased  
133 from freshwater to full salinity (35 ppt) and the temperature adjusted to 12 °C (aerated and  
134 temperature-controlled seawater). Water flow was fixed at 20 L min<sup>-1</sup>. Fish were acclimatized  
135 for 4-weeks during which time they were all fed the same feed (S0, Table 1). Two days prior  
136 to the start of the trial, the individual fish were weighed (group mean 175 g) and returned to  
137 the tanks. Four dietary treatments were investigated with progressively increasing inclusions  
138 of solvent-extracted soybean meal (SBM; 0 g kg<sup>-1</sup>, 100 g kg<sup>-1</sup>, 200 g kg<sup>-1</sup>, 300 g kg<sup>-1</sup>)

139 substituting other protein sources (FM, corn gluten, sunflower cake and horsebeans) referred  
140 to as diets S0, S10, S20 and S30, respectively (Table 1). Each experimental feed was fed to  
141 triplicate tanks. All feeds were formulated to meet the nutritional requirement of salmon  
142 (National Research Council, NRC, 2011). The feeding trial lasted for 87 days and at the end  
143 of this period individual weights and lengths of all fish were measured and specific growth  
144 rate (SGR) calculated using the following equation [ $SGR = \ln WT_{fin} - \ln WT_{in} / \text{days}$ ]. In  
145 addition, the same tip of the liver and a portion of the distal (posterior) intestine were  
146 dissected from six individuals per dietary treatment (two per tank replicate). Tissue samples  
147 were immediately placed in RNALater (Life Technologies, Paisley, UK) and processed as per  
148 manufacturer's instructions before being stored at -20 °C prior to molecular analyses.

## 149 *2.2. Transcriptomic analysis*

150 Transcriptomic analysis was conducted using an Atlantic salmon custom-designed 4 x 44K  
151 oligo microarray (Agilent Technologies, Wokingham, UK; ArrayExpress accession no. A-  
152 MEXP-2065) as described in detail previously (Tacchi et al., 2011). The salmon custom array  
153 and laboratory procedures utilized have been widely used and extensively validated in  
154 previous studies (e.g. Bicskei et al., 2014, Martinez-Rubio et al., 2012, Morais et al., 2012a,  
155 Morais et al., 2012b, Tacchi et al., 2011, Tacchi et al., 2012). Briefly, total RNA was  
156 extracted from individual samples using TRI Reagent according to the manufacturer's  
157 instructions (Sigma-Aldrich, Dorset, UK), including a high salt precipitation as recommended  
158 for polysaccharide-rich tissues such as liver (Chomczynski and Mackey, 1995). RNA  
159 quantity, integrity and purity were assessed by agarose gel electrophoresis and  
160 spectrophotometry (NanoDrop ND-1000, Thermo Scientific, Wilmington, USA). Equal  
161 amounts of RNA from six individual fish liver and intestine tissues were analyzed. The  
162 resulting RNA samples were amplified using a TargetAmp™ 1-Round Aminoallyl-aRNA  
163 Amplification Kit, (Epicentre Technologies Corporation, Madison, Wisconsin, USA)



164 following recommended procedures and purified through a RNeasy Mini spin column  
165 (Qiagen, Manchester, UK). Aminoallyl-amplified RNA (aRNA) samples were individually  
166 labelled with Cy3 dye (GE HealthCare Life Sciences, Buckinghamshire, UK), while a pool of  
167 all aRNA samples was labelled with Cy5 dye (GE HealthCare Life Sciences) and used as a  
168 common reference. Unincorporated dye was removed by purifying the aRNA samples with an  
169 Illustra AutoSeq G-50 dye terminator column (GE HealthCare Life Sciences). Successful dye  
170 incorporation and sample integrity was assessed for 0.4  $\mu$ L aliquots of labelled samples by  
171 agarose gel electrophoresis followed by fluorescent detection of aRNA products (Typhoon  
172 scanner, GE Healthcare Life Sciences). Cy dye concentration and aRNA quantification were  
173 measured by Nanodrop mediated spectrophotometry (Thermo Scientific, Wilmington, USA).

174 Labelled aRNA samples were hybridized to the custom-made array. A dual-label common  
175 reference design was adopted, where equal amounts of each individual aRNA sample and the  
176 common reference pool were competitively hybridized on an array. The common reference  
177 design allowed standardization of inter- and intra-array variability. A total of 24 samples (two  
178 tissues x two treatments x six replicates) were processed with the Gene Expression  
179 Hybridization Kit (Agilent Technologies), applied to the arrays and immediately incubated  
180 using SureHyb hybridization chambers in a DNA Microarray Hybridization Oven (Agilent  
181 Technologies) at 65 °C for 17 h. Throughout the experiment samples were randomized,  
182 preventing samples from the same treatment being overrepresented in a particular batch in  
183 order to avoid unintentional biases. Scanning was performed using a GenePix 4200 AL  
184 Scanner (Molecular Devices (UK) Ltd., Wokingham, UK) and the resulting images analyzed  
185 with Agilent Feature Extraction Software v.9.5 (Agilent Technologies) to extract the intensity  
186 values and identify the features. The foreground intensity was computed as the mean value of  
187 pixels, considered a better estimator as being less susceptible to distortion from outlier values  
188 (Russell et al., 2009), while background intensities were computed as the median value of

189 pixels. Details of the microarray experiment have been submitted to ArrayExpress and  
190 assigned the accession number E-MTAB-3291.

191

192 2.3. Data pre-processing and differential expression analysis

193 Transcriptomic data analysis was performed using R v.3.0.1 and Bioconductor v.2.13  
194 (Gentleman et al., 2004, R Core Team, 2013). Quality control, data pre-processing and  
195 identification of differentially expressed features/genes were conducted using the software  
196 package limma (Smyth, 2004). Array quality was assessed by visualizing and comparing  
197 boxplots of red (R) and green (G) background and foreground intensities, MA plots [ $M =$   
198  $\log_2(G/R)$ ;  $A = 1/2\log_2(R*G)$ ] and spatial heterogeneity to reveal the presence of any  
199 technical bias. Following this quality assessment, all arrays were retained for further analyses.  
200 Features considered outliers in two or more replicates within at least one treatment were  
201 excluded from further analyses. Foreground intensities were background-corrected using the  
202 normexp approach (maximum likelihood variant “mle”, offset = 50) as previously reported as  
203 the most reliable method for two-color microarrays where background estimates are available  
204 (Ritchie et al., 2007, Silver et al., 2009). Data were log-transformed and normalized using the  
205 function *normalizeWithinArrays* (method = “loess”) and *normalizeBetweenArrays* (method =  
206 “RQuantile”) (Smyth and Speed, 2003). Controls and features expressed just above  
207 background (A values < 6) were also removed.

208 Features of the array were annotated using BLAST 2.2.29+ (blastx) against the entire non-  
209 redundant protein database as well as using the KEGG Automatic Annotation Server to obtain  
210 functional annotations (Altschul et al., 1990, Moriya et al., 2007). A total of 89.6 % of all  
211 probes were returned with a BLAST annotation (annotation date May 2014) with e-value  
212 < 0.001, while 59 % of probes were returned with a functional annotation (KEGG identifier)  
213 using the KAAS server. Differentially expressed features between treatments were estimated  
214 by least squares fitting of linear models on a probe-by-probe basis using the entire pre-  
215 processed dataset. The function *lmFit* was used to compute differential expression and  
216 statistics were extracted using *ebayes* (trend = TRUE), both limma functions (Smyth, 2004).

217 Features representing the same target gene as implied from KEGG annotation were merged  
218 into a unique value obtained by selecting the feature with the highest F-value. A new dataset  
219 was therefore generated for further analyses where each gene was only represented by one  
220 feature. Merging resulted in a dataset of 6729 annotated features targeting unique genes.

## 221 *2.4 Data Mining*

222 *Hierarchical Clustering.* Hierarchical cluster analysis was performed on gene expression  
223 normalized values to evaluate overall similarity between samples and analyzed using the R  
224 package *pvclust* (Suzuki and Shimodaira, 2011). “Correlation” was used as the distance  
225 measure using 1000 bootstrap replication.

226 *Overview of differential expression.* Differentially expressed genes were plotted using the R  
227 package *ggplot2* (Wickham, 2009). For figures involving functional information, the KEGG  
228 database was used as the chosen classification system. Venn diagrams were generated using  
229 the function *VennDiagram* from the *limma* software package and modified using *ggplot2*.

230 *Gene-Set Enrichment Analysis (GSEA).* Unique annotated sequences were analyzed using the  
231 R function *gage* of the software package GAGE (Generally Applicable Gene-set Enrichment,  
232 Luo et al., 2009) to identify mechanistic changes as suggested by coordinated expression  
233 changes in gene-sets. For completeness, two types of test were performed: *1 direction (1d)*,  
234 testing all genes in a gene-set moving towards the same direction; and *2 directions (2d)*,  
235 testing genes in a gene-set that move towards both directions. Gene-sets with a *q*-value < 0.1  
236 were considered significant, where the *q*-value represented the *p*-value adjusted for false  
237 discovery rate. KEGG classification was used for these analyses and all figures were  
238 produced using the software package *ggplot2*.

239

## 240 **3. Results**

### 241 *3.1 Growth performance*

242 There were no significant differences in recorded mortality between treatments. Dietary SBM  
243 inclusion of 300 g kg<sup>-1</sup> (i.e. feed S30) induced statistically significant growth retardation in  
244 salmon in terms of both final weight and specific growth rate (SGR) compared to fish fed all  
245 other treatments (Table 2). The SGR of salmon fed SBM at an inclusion level of 200 g kg<sup>-1</sup>  
246 was significantly lower than fish fed S0 or S10, but significantly higher than fish fed S30,  
247 although final weights were not different to fish fed the diets with lower inclusion of SBM.  
248 Based on these data, transcriptome analyses were performed on tissue samples from fish fed  
249 diets S0 and S30 to guarantee the largest difference between samples and increase the  
250 possibility to detect differential expression.

### 251 *3.2 Transcriptome analysis: Overview*

252 Hierarchical clustering performed on the normalized gene expression data indicated the  
253 existence of a clear structure in the dataset with four main clusters (Fig. 1). A clear separation  
254 was evident between the liver and the intestine forming two independent clusters fully  
255 supported by bootstrap values (100). Within each tissue cluster, the dietary treatments S0 and  
256 S30 also induced clearly identifiable differential gene expression responses. The liver  
257 demonstrated a weaker, but significant, bootstrap support (63-80), compared with the  
258 response of the intestine that was fully supported by statistical bootstraps (100), indicating a  
259 less defined response in the liver compared with that of the intestine.

260 The analysis of differential expression computed by limma reflected the structure observed by  
261 hierarchical clustering, indicating a less pronounced, but still evident response in liver, where  
262 133 genes were differentially expressed ( $p < 0.01$ , absolute FC  $> 1.3$ ), compared to a marked  
263 response in the intestine where a total of 1918 transcripts were differentially expressed  
264 ( $p < 0.01$ , absolute FC  $> 1.3$ ). The difference between the two tissues was again apparent in

265 that the differentially expressed transcripts of the liver showed a generally decreased (fold  
266 change) and less significant ( $p$  value) response compared with that of the intestine (Fig. 2a, b).  
267 Due to the minor transcriptional effect observed in the liver associated with primary  
268 metabolic role, the use of multiple testing corrections (Benjamini-Hochberg) on this tissue  
269 returned only 11 genes below the adjusted cutoff  $p$  value of 0.05 and proved to be over-  
270 conservative. Applying multiple testing correction (adj.  $p < 0.05$ ) in the intestine resulted in  
271 2664 differentially expressed genes. For consistency between the tissues we therefore elected  
272 to consider all genes with a  $p < 0.01$  and absolute FC  $> 1.3$  to be truly differentially  
273 expressed.

274 A total of 44 genes were significantly affected in both tissues, while the remaining 88 and  
275 1873 genes were represented exclusively as tissue-specific responses in the liver and intestine  
276 respectively. Of the 44 genes affected in both liver and intestine, 31 were regulated in the  
277 same direction (i.e. either up- or down-regulated in both tissues) possibly indicating a  
278 systemic response, while 13 were regulated in opposite directions (Table 3). Expression was  
279 always calculated relative to the S0 treatment. Amongst the genes whose transcript abundance  
280 increased in both liver and intestine, there was an over-representation of ribosomal subunits  
281 (L5, L21, L11, L7, S3A); enzymes involved in ribosome biogenesis such as ribonucleases  
282 P/MRP protein subunit RPP25 and casein kinase II subunit alpha; proteins participating in  
283 RNA processing such as translation initiation factor 5, cleavage and polyadenylation  
284 specificity factor subunit 5 and the spliceosome component U5 snRNP protein and factors  
285 engaged in protein processing such as ER degradation enhancer, mannosidase alpha 3 (Table  
286 3). In addition, other important genes that concurrently increased in both tissues included bile-  
287 salt stimulated lipase, complement component C7, sulfinoalanine decarboxylase and an  
288 enzyme involved in fatty acid  $\beta$ -oxidation, carnitine O-palmitoyltransferase 1. In contrast,  
289 down-regulated genes coded for proteins involved in metabolic processes such as lipid  
290 metabolism (glycerol kinase, elongation of very long chain fatty acids protein 5, very long

291 chain-3 hydroxyacyl-dehydratase), sugar metabolism (GDP-L-fucose synthase) and glycan  
292 metabolism (alpha-1,6-mannosyl-glycoprotein beta-1,2-N-acetylglucosaminyltransferase).

293 The following sections (sections 3.3 and 3.4) report on the tissue-specific responses. Gene-set  
294 enrichment analysis was utilized as a tool to identify mechanistic changes of groups of genes.  
295 However, while the limited number of differentially expressed genes identified in the liver  
296 also allowed a more in-depth analysis on individual genes, this approach was not possible  
297 with the distal intestine without introducing bias in selecting the gene list. Thus, in the distal  
298 intestine only gene-set enrichment analysis was applied representing a more robust approach  
299 compared to individual gene analysis. The liver was found to be only marginally affected,  
300 whereas a large number of gene-sets were significantly affected ( $q < 0.1$ ) in the intestine,  
301 clearly reflecting the trend identified above targeting differentially expressed genes.

### 302 *3.3 Liver*

303 Compared with the distal intestine, the liver was partially affected by diet S30 (Fig. 2) with  
304 only 88 genes uniquely differentially expressed compared to fish fed diet S0, although a set of  
305 key pathways were significantly affected (Fig. 3). The *2d* test was able to capture the most  
306 significantly perturbed pathways possessing genes both up and down-regulated. The most  
307 affected pathway in this category was the complement and coagulation cascades. A closer  
308 view on the complement cascade is shown in Fig. 4. Other affected processes identified by the  
309 *2d* test included lipid metabolism; in particular lipid digestion and absorption, regulation of  
310 steroid biosynthesis and the PPAR signalling pathway, amino acid metabolism; principally  
311 glycine/serine/threonine and alanine/aspartate/glutamate metabolism and metabolism of  
312 vitamins (retinol) including vitamin digestion and absorption. In contrast, the *1d* GAGE test  
313 was able to address expression changes in genes moving in the same direction, either up or  
314 down-regulated, and was particularly helpful in identifying the overall trend of expression  
315 within a specific pathway. This analysis indicated an up-regulation of the PI3K-Akt signalling

316 pathway and a down regulation of oxidative phosphorylation, proteasome apparatus, protein  
317 export and terpenoid backbone biosynthesis.

318 In Table 4, a selection of the genes differentially expressed only in the liver is listed. To  
319 minimize bias, we included most of the genes involved in the function of interest (the full list  
320 included 40 of the 88 affected genes). These included the up-regulation of several genes  
321 coding for pancreatic enzymes such as trypsin, carboxypeptidase B, etc. and bile salt-  
322 stimulated lipase, and key genes participating in energy metabolism for instance, 2-  
323 oxoglutarate dehydrogenase and succinate dehydrogenase (citric acid cycle). Generally,  
324 down-regulated genes were shown to relate to amino acid, lipid and sterol metabolism (Table  
325 4). Interestingly, ribosomal genes that were affected only in the liver were all down-regulated,  
326 in contrast with those affected in both tissues that were up-regulated. Finally, genes  
327 contributing to the innate immune response were also significantly altered in liver, where the  
328 complement cascade pathway possessed the highest significance score. Genes of this pathway  
329 that were significantly up-regulated included complement components 3 (C3), that was only  
330 affected in the liver ( $p < 0.001$ , Table 4), and C7 that was up-regulated in both tissues in  
331 tandem (Table 3). In addition, other genes of the complement cascade such as C5, C6,  
332 complement component receptor type 2, the mannan-binding lectin serine protease 1  
333 (MASP1) and the CD59 antigen were also perturbed to a lesser extent ( $p < 0.05$ ) and are  
334 shown in Fig. 4.

### 335 *3.4 Distal Intestine*

336 The distal intestine showed a pronounced global response, involving gene-sets regulating a  
337 number of processes (Fig. 4). This response at a gene-set level is summarized by the major  
338 biological processes affected in the intestine identified in both GAGE *1d* and *2d* analyses and  
339 included primarily metabolic (e.g. lipid, amino acid and energy metabolism), organismal



340 systems (e.g. digestion and absorption, immune and endocrine system) and other cellular  
341 processes including pathways involved in the protein synthesis.

342 For clarity of presentation and due to the number of pathways identified in the intestine,  
343 pathways were separated into “highly significantly different” ( $q < 0.0001$ ) and “significantly  
344 different” ( $0.1 < q < 0.0001$ ), however this should not minimize the importance of pathways  
345 with lower  $q$  values. Among the highly significantly regulated gene-sets identified through the  
346  $2d$  test ( $q < 0.0001$ ) were pathways associated with phagocytosis (cellular organelles  
347 lysosome and phagosome), lipid and vitamin digestion and absorption pathways, PPAR  
348 signalling pathway of the endocrine system, antigen processing and presentation (immune  
349 system) and key metabolic pathways including glycan degradation, amino sugar and  
350 nucleotide sugar degradation, retinol metabolism, arginine/proline metabolism,  
351 alanine/aspartate/glutamate metabolism and glutathione metabolism (Fig. 5a). Only one  
352 immune pathway, antigen processing and presentation, was identified by the  $2d$  test with a  
353 highly significant  $q$  value. A further 22 pathways, mainly metabolic ( $n= 15$ ), were  
354 differentially expressed below the  $q = 0.0001$  cut-off (Fig. 5a).

355 Interestingly, only two pathways classified as “genetic information processing” were  
356 identified by the  $2d$  test and included the proteasome complex and aminoacyl-tRNA  
357 biosynthesis. Notably, the  $1d$  test revealed that this category of pathways was significantly  
358 up-regulated, with most pathways exhibiting a highly significant  $q$  value ( $< 0.0001$ ) (e.g.  
359 ribosome, proteasome, spliceosome, RNA transport, etc., Fig. 5b). Furthermore, other gene-  
360 sets with increased expression in fish fed the S30 treatment comprised those regulating  
361 functions such as cell growth and death, cell cycle, apoptosis and MAPK signalling (Fig. 5b).  
362 An overall increased expression was also observed in a number of immune pathways. The  
363 immune response involved up-regulation of the TNF signalling pathway ( $p < 0.0001$ ), NOD-  
364 like receptor interaction, NF-kB signalling pathway, cytosolic DNA sensing pathway, Jak-

365 STAT signalling pathway, cytokine-cytokine receptor interaction and T-cell receptor  
366 signalling pathway (Fig. 5b).

367 .

368 A number of gene-sets that tested significant in the *2d* test were also significant with the *1d*  
369 test, particularly amongst the down-regulated pathways (Fig. 5b). These included gene-sets  
370 centred on metabolic pathways and processes such as digestion and absorption of proteins,  
371 lipids, vitamins and minerals, indicating the genes underlying these processes were primarily  
372 decreased in response to diet S30. Finally, the *1d* test revealed that proteins associated with  
373 the lysosome and peroxisome organelles were also reduced (Fig. 5b).

#### 374 **4. Discussion**

375 The present study reports a molecular snapshot reflecting the hepatic and intestinal (distal)  
376 physiology of Atlantic salmon exposed to a plant-based nutritional stress. The results herein  
377 were produced utilizing an unbiased global analysis approach to identify mechanistic changes  
378 of gene expression with the precise aim of avoiding the bias and restrictions often associated  
379 with the selection of limited gene lists. The study aimed to provide a platform to screen the  
380 most perturbed processes and identify molecular markers useful for phenotyping and  
381 discriminating influential processes affected in response to diets formulated with high levels  
382 of plant-based ingredients. The complete lists of differentially expressed genes identified are  
383 publicly available and can be consulted for further information on Array Express (accession  
384 number E-MTAB-3291).

385 The nutritional trial revealed that dietary SBM reduced growth performance in salmon when  
386 included at levels over 200 g kg<sup>-1</sup>. It was not within the scope of the study to determine in  
387 detail the performance of salmon in response to dietary SBM inclusion, since this had already  
388 been investigated previously and the effects on growth were in agreement with those observed

389 in the present trial (growth data reviewed in Collins et al., 2013, health effects reviewed in  
390 Krogdahl et al., 2010). The aim of this nutritional trial was solely to reproduce the expected  
391 effects and provide an experimental model of nutritional stress that would satisfy the  
392 objectives of the present gene expression study. The experimental model included two lower  
393 inclusion levels ( $100 \text{ g kg}^{-1}$  and  $200 \text{ g kg}^{-1}$ ) that did not show consistent and/or significant  
394 growth retardation compared with the control diet (S0). However, the results could provide  
395 experimental material to test the potential molecular biomarkers and hypothesis developed in  
396 the current study in dietary treatments that have only “subclinical” effects on performance.

397 The detected differential transcriptomic response observed in salmon fed  $300 \text{ g kg}^{-1}$  SBM  
398 compared to FM control was greater in the distal intestine compared to the liver. These results  
399 were expected due to the inflammation of the distal intestine and consequent immune  
400 response that is generally observed after feeding similar levels of dietary SBM (Baeverfjord  
401 and Krogdahl, 1996, Bakke-McKellep et al., 2007, Urán et al., 2008, Urán et al., 2009). Since  
402 not directly affected by the contact with the feed itself and hence the ANFs, the liver response  
403 was different and possibly accessory to that of the intestine. While in the former we observed  
404 processes related to the immune response, cell growth and regeneration and evidence of  
405 impairment of metabolic functions, the liver showed activation of mainly metabolic pathways  
406 and an immune response that reflected the inflammation observed locally in the distal  
407 intestine.

408 The role played by the liver in the physiological response to plant-based nutritional stress is  
409 not fully understood. Kortner and colleagues (2013) reported an increase of hepatic  
410 cholesterol metabolism suggesting it was a compensatory response to reduced dietary uptake  
411 through endogenous cholesterol biosynthesis. However, this was in contrast to results from  
412 the present study where all genes of lipid and sterol metabolism were down regulated in liver  
413 with the exception of cholest-5-ene-3beta,7alpha-diol 3beta-dehydrogenase (up-regulated),

414 which is involved in the initial stages of bile acid synthesis from cholesterol (Schwarz et al.,  
415 2000). The discrepancy observed might indicate that a different response is to be expected  
416 depending upon SBM dietary inclusion level and, most importantly, the severity of the  
417 immune response. Whichever the case, it is evident that the cholesterol pathway and its  
418 underlying genes might represent a reliable marker describing the nutritional status of fish and  
419 should be further investigated.

420 A well-known characteristic and important limitation of plant ingredients is the presence of  
421 protease inhibitors, proteins involved in the plant's defence mechanisms against herbivores  
422 (Hartl et al., 2011). Protease inhibitors are tightly linked to proteins and might consequently  
423 be co-extracted with the protein fraction during the production of concentrates. In the gut  
424 lumen, protease inhibitors impact on feed digestibility supposedly by interfering with  
425 endogenous proteases such as trypsins, chymotrypsins, etc., which in turn can result in over-  
426 secretion of these enzymes from pancreatic tissue and consequent loss of endogenous  
427 resources (e.g. sulphur-rich amino acids methionine and cysteine), eventually leading to  
428 decreased growth (Chikwati et al., 2012, Sarwar-Gilani et al., 2012). In salmon liver, we  
429 detected a significantly increased expression of genes coding for pancreatic enzymes such as  
430 proteases (trypsin, chymotrypsin, etc.) and lipases. This was not surprising considering that  
431 fish liver contains several cell types including pancreatic-like cells (Rust, 2002). These results  
432 may resemble the "hyperactive pancreas" syndrome observed in mammals (Sarwar-Gilani et  
433 al., 2012) and encourage further investigation of these genes as biomarkers for protein  
434 digestibility.

435 At a metabolic level, a number of key hepatic genes involved in amino acid metabolism were  
436 down-regulated including those coding for enzymes catalysing the degradation of essential  
437 amino acids such as threonine (glycine C-acetyltransferase), tryptophan (indoleamine 2,3-  
438 dioxygenase) and methionine (S-adenosylmethionine synthetase) which suggested a

439 compensatory response to preserve these important nutrients. There was also up-regulation in  
440 liver of two genes fundamental to the citric acid cycle (2-oxoglutarate dehydrogenase and  
441 succinate dehydrogenase), which suggested an activation of energy metabolism. Interestingly,  
442 the up-regulation of the carnitine O-palmitoyltransferase gene, involved in the  $\beta$ -oxidation of  
443 fatty acids via the synthesis of acyl-CoA, possibly reflected the undernourished condition of  
444 fish fed  $300 \text{ g kg}^{-1}$  SBM, with the liver approaching a ketotic state. This is consistent with a  
445 previous study demonstrating a number of similarities in the hepatic response of Atlantic  
446 salmon to restricted feeding and feed containing extracted SBM (Skugor et al., 2011).

447 Expression data from the liver suggested that this organ also participated in the innate  
448 immune response at a systemic level. This was evident from the up-regulation of several  
449 genes of the complement cascade, which plays a crucial role in mediating non-specific  
450 defence against pathogens in the blood, and the liver is the major site of production of  
451 complement proteins (Jain et al., 2014). The up-regulation of the MASP1 genes suggested  
452 that activation of the cascade possibly occurred through the lectin pathway, which is  
453 stimulated through recognition of carbohydrates or glycoproteins present on invading  
454 pathogens or at the lesion. Notably, CD59 antigen abundance decreased, suggesting a  
455 stimulation of the formation of the membrane attack complex that is formed as an end result  
456 of the complement cascade. Activation of proteins of the complement cascade in the liver was  
457 not reported previously in Atlantic salmon fed a lower level ( $200 \text{ g kg}^{-1}$ ) of SBM (Skuggor et  
458 al., 2011). Since the microarray utilized in the earlier study (NCBI GEO Omnibus GPL6154)  
459 contained probes targeting several complement component proteins, it can be hypothesized  
460 that the activation of this pathway may occur at SBM inclusions above  $200 \text{ g kg}^{-1}$  and that  
461 below this level the process remains more localized to the intestine.

462 The gene expression response detected in the distal intestine was consistent with the results of  
463 recent studies investigating SBM-induced enteritis (Sahlmann et al., 2013) and that caused by

464 pea protein concentrate supplemented with soyasaponin (Kortner et al., 2012), and  
465 significantly overlapped with the immune response developed during human inflammatory  
466 bowel disease (Maloy and Powrie, 2011). The present results confirmed, at an individual gene  
467 level (dataset available as supplementary material in Array Express), the majority of the  
468 immune-related responses of chronically inflamed distal intestine in salmon reported  
469 previously (Kortner et al., 2012, Skugor et al., 2011). These included *a*) the activation of T-  
470 cell mediated processes via the up-regulation of the CD86 antigen, CTLA4 IL-18 and IL-22  
471 and increased expression of T-cell receptors, which was also confirmed by the up-regulation  
472 of the T-cell receptor signalling pathway identified by gene-set analysis; *b*) activation of  
473 TNF- and NF- $\kappa$ B-mediated response and up-regulation of components of the respiratory burst  
474 complex through the up-regulation of the TNF signalling pathway, a critical trigger for the  
475 release of pro-inflammatory cytokines and also the activation of a number of intracellular  
476 pathways eventually leading to apoptosis and cell survival (Chu, 2013); and finally *c*) up-  
477 regulation of anti-inflammatory proteins, notably annexin A1, that was found to increase by  
478 approximately 10-fold in agreement with previous studies (Kortner et al., 2012). All these  
479 genes represent an informative pool of potential molecular markers of the immune response in  
480 the salmon intestine. More focused immunological analyses of intestinal transcriptome  
481 profiling during the development of enteritis have been reported previously (Kortner et al.,  
482 2012, Sahlmann et al., 2013). In addition to these results, we reported a notable intestinal up-  
483 regulation of a set of pathways involved in protein synthesis (ribosome, proteasome, tRNA  
484 biosynthesis, RNA transport, etc.) and cell proliferation (cell cycle, apoptosis, etc.) that  
485 implied active regeneration of the damaged intestinal tissue.

486 Metabolic pathways and other cellular and physiological processes such as digestion and  
487 absorption of nutrients (proteins, lipid, vitamins and minerals) were also significantly affected  
488 in the distal intestine. A down-regulation of virtually all metabolic pathways following the  
489 SBM-induced nutritional stress was observed, consistent with studies indicating that distal

490 intestine undergoing severe inflammation was characterized by “tissue malfunction” (Kortner  
491 et al., 2012, Sahlmann et al., 2013). Studies from other fish species revealed that impaired  
492 metabolism could also develop independently from the immune response (Murray et al.,  
493 2010), which suggested that some of the effects on metabolism might be possibly associated  
494 with the nutritional properties of SBM or plant products in general rather than tissue  
495 malfunction *per se*. To better identify and characterise these processes it will be necessary to  
496 analyse dietary treatments inducing a lower (subclinical) level of stress that does not directly  
497 affect the functionality of the intestine (e.g. 100 g kg<sup>-1</sup>). Marked down-regulation of genes  
498 involved in cellular processes associated with lysosomes, peroxisomes and, to a lesser extent,  
499 phagosomes, whose primary role is the degradation of various cellular metabolites, as well as  
500 the digestion of invading pathogens, was possibly also linked to the tissue impairment.

501 Amongst the processes affected, lipid and in particular sterol metabolism in the distal intestine  
502 appeared to be altered by the use of plant proteins. We observed a down-regulation of the  
503 genes involved in several pathways of lipid metabolism, digestion and absorption, primary  
504 bile acid biosynthesis and secretion. This was in agreement with previous observations in  
505 salmon fed 200 g kg<sup>-1</sup> SBM where processes of lipid and bile uptake were impaired (Kortner  
506 et al., 2013). It is unclear whether impaired lipid metabolism in the distal intestine was a  
507 consequence of the tissue damage or if it was affected by the presence of certain ANFs such  
508 as, for example, soyasaponins, that interfere with the normal absorption of these nutrients and  
509 are known to cause hypocholesterolemia (Francis et al., 2001, Francis et al., 2002, Gu et al.,  
510 2014, Kortner et al., 2013).

511 In conclusion, the present study contributes towards the development of a nutrigenomic  
512 model describing molecular signatures observed following nutritional stress induced by  
513 dietary plant proteins. Our results build on recent nutrigenomic findings in Atlantic salmon  
514 fed pea protein concentrate and on those of salmon undergoing restricted/SBM feeding

515 towards the identification of diagnostic markers in the intestine. More importantly, the  
516 analyses of the liver transcriptome provided a platform for selection of responses that could  
517 be more general and related to the overall performance, and hence have the potential as  
518 predictive markers for monitoring health, welfare and performance traits. This study is a  
519 comprehensive report on transcriptome profiles of distal intestine and liver, highlighting the  
520 role of the latter tissue in fish undergoing SBM-induced nutritional stress.

## 521 **Acknowledgements**

522 KLB and this study were funded by the European Commission FP7 Combination of  
523 Collaborative projects and Coordination and Support Actions (CP-CSA) Project No. 262336,  
524 Aquaculture infrastructures for excellence in European fish research (AQUAEXCEL). The  
525 authors also acknowledge the support of the MASTS pooling initiative (The Marine Alliance  
526 for Science and Technology for Scotland) in the completion of this study. MASTS is funded  
527 by the Scottish Funding Council (grant reference HR09011) and contributing Institutions.

## 528 **References**

529 Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment  
530 search tool. *J. Mol. Biol.* 215, 403-410.

531 Azaza, M.S., Wassim, K., Mensi, F., Abdelmouleh, A., Brini, B., Kraïem, M.M., 2009.  
532 Evaluation of faba beans (*Vicia faba* L. var. *minuta*) as a replacement for soybean meal in  
533 practical diets of juvenile Nile tilapia *Oreochromis niloticus*. *Aquaculture* 287, 174-179.

534 Baeverfjord, G., Krogdahl, A., 1996. Development and regression of soybean meal induced  
535 enteritis in Atlantic salmon, *Salmo salar* L., distal intestine: a comparison with the intestines  
536 of fasted fish. *J. Fish Dis.* 19, 375-387.



537 Bakke-McKellep, A.M., Frøystad, M.K., Lilleeng, E., Dapra, F., Refstie, S., Krogdahl, Å,  
538 Landsverk, T., 2007. Response to soy: T-cell-like reactivity in the intestine of Atlantic  
539 salmon, *Salmo salar* L. J. Fish Dis. 30, 13-25.

540 Bicskei, B., Bron, J., Glover, K., Taggart, J., 2014. A comparison of gene transcription  
541 profiles of domesticated and wild Atlantic salmon (*Salmo salar* L.) at early life stages, reared  
542 under controlled conditions. BMC Genomics 15, 884.

543 Booth, M.A., Allan, G.L., Frances, J., Parkinson, S., 2001. Replacement of fish meal in diets  
544 for Australian silver perch, *Bidyanus bidyanus*: IV. Effects of dehulling and protein  
545 concentration on digestibility of grain legumes. Aquaculture 196, 67-85.

546 Buttle, L.G., Burrells, A.C., Good, J.E., Williams, P.D., Southgate, P.J., Burrells, C., 2001.  
547 The binding of soybean agglutinin (SBA) to the intestinal epithelium of Atlantic salmon,  
548 *Salmo salar* and Rainbow trout, *Oncorhynchus mykiss*, fed high levels of soybean meal. Vet.  
549 Immunol. Immunopathol. 80, 237-244.

550 Chikwati, E.M., Venold, F.F., Penn, M.H., Rohloff, J., Refstie, S., Guttvik, A., Hillestad, M.,  
551 Krogdahl, Å, 2012. Interaction of soyasaponins with plant ingredients in diets for Atlantic  
552 salmon, *Salmo salar* L. Br. J. Nutr. 107, 1570-1590.

553 Chomczynski, P., Mackey, K., 1995. Short technical reports. Modification of the TRI reagent  
554 procedure for isolation of RNA from polysaccharide- and proteoglycan-rich sources.  
555 Biotechniques 19, 942-945.

556 Chu, W., 2013. Tumor necrosis factor. Cancer Lett. 328, 222-225.

557 Collins, S.A., Øverland, M., Skrede, A., Drew, M.D., 2013. Effect of plant protein sources on  
558 growth rate in salmonids: Meta-analysis of dietary inclusion of soybean, pea and  
559 canola/rapeseed meals and protein concentrates. Aquaculture 400–401, 85-100.

560 Francis, G., Kerem, Z., Makkar, H.P.S., Becker, K., 2002. The biological action of saponins  
561 in animal systems: a review. Br. J. Nutr. 88, 587-605.

562 Francis, G., Makkar, H.P.S., Becker, K., 2001. Antinutritional factors present in plant-derived  
563 alternate fish feed ingredients and their effects in fish. Aquaculture 199, 197-227.

564 Gatlin, D.M., Barrows, F.T., Brown, P., Dabrowski, K., Gaylord, T.G., Hardy, R.W., Herman,  
565 E., Hu, G., Krogdahl, Å, Nelson, R., Overturf, K., Rust, M., Sealey, W., Skonberg, D., J  
566 Souza, E., Stone, D., Wilson, R., Wurtele, E., 2007. Expanding the utilization of sustainable  
567 plant products in aquafeeds: a review. Aquacult. Res. 38, 551-579.

568 Gentleman, R., Carey, V., Bates, D., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier,  
569 L., Ge, Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F.,  
570 Li, C., Maechler, M., Rossini, A., Sawitzki, G., Smith, C., Smyth, G., Tierney, L., Yang, J.,  
571 Zhang, J., 2004. Bioconductor: open software development for computational biology and  
572 bioinformatics. Genome Biol. 5, R80.

573 Glencross, B., Evans, D., Dods, K., McCafferty, P., Hawkins, W., Maas, R., Sipsas, S., 2005.  
574 Evaluation of the digestible value of lupin and soybean protein concentrates and isolates when  
575 fed to rainbow trout, *Oncorhynchus mykiss*, using either stripping or settlement faecal  
576 collection methods. Aquaculture 245, 211-220.

577 Gu, M., Kortner, T.M., Penn, M., Hansen, A.K., Krogdahl, Å, 2014. Effects of dietary plant  
578 meal and soya-saponin supplementation on intestinal and hepatic lipid droplet accumulation  
579 and lipoprotein and sterol metabolism in Atlantic salmon (*Salmo salar* L.). Br. J. Nutr. 111,  
580 432-444.

581 Hartl, M., Giri, A.P., Kaur, H., Baldwin, I.T., 2011. The multiple functions of plant serine  
582 protease inhibitors: Defense against herbivores and beyond. *Plant Signal. Behav.* 6, 1009-  
583 1011.

584 Hartviksen, M., Bakke, A., Vecino, J., Ringø, E., Krogdahl, Å, 2014. Evaluation of the effect  
585 of commercially available plant and animal protein sources in diets for Atlantic salmon  
586 (*Salmo salar* L.): digestive and metabolic investigations. *Fish Physiol. Biochem.* 40, 1621-37.

587 Jain, U., Otle, A.R., Van Limbergen, J., Stadnyk, A.W., 2014. The complement system in  
588 inflammatory bowel disease. *Inflamm. Bowel Dis.* 20, 1628-37.

589 Kaushik, S.J., Covès, D., Dutto, G., Blanc, D., 2004. Almost total replacement of fish meal by  
590 plant protein sources in the diet of a marine teleost, the European seabass, *Dicentrarchus*  
591 *labrax*. *Aquaculture* 230, 391-404.

592 Kortner, T.M., Skugor, S., Penn, M.H., Mydland, L.T., Djordjevic, B., Hillestad, M.,  
593 Krasnov, A., Krogdahl, A., 2012. Dietary soyasaponin supplementation to pea protein  
594 concentrate reveals nutrigenomic interactions underlying enteropathy in Atlantic salmon  
595 (*Salmo salar*). *BMC Vet. Res.* 8, 101.

596 Kortner, T.M., Gu, J., Krogdahl, Å, Bakke, A.M., 2013. Transcriptional regulation of  
597 cholesterol and bile acid metabolism after dietary soyabean meal treatment in Atlantic salmon  
598 (*Salmo salar* L.). *Br. J. Nutr.* 109, 593-604.

599 Krogdahl, Å, Penn, M., Thorsen, J., Refstie, S., Bakke, A.M., 2010. Important antinutrients in  
600 plant feedstuffs for aquaculture: an update on recent findings regarding responses in  
601 salmonids. *Aquacult. Res.* 41, 333-344.

602 Luo, W., Brouwer, C., 2013. Pathview: an R/Bioconductor package for pathway-based data  
603 integration and visualization. *Bioinformatics* 29, 1830-1831.

604 Luo, W., Friedman, M.S., Shedden, K., Hankenson, K.D., Woolf, P.J., 2009. GAGE:  
605 generally applicable gene set enrichment for pathway analysis. *BMC Bioinformatics* 10, 161.

606 Maloy, K.J., Powrie, F., 2011. Intestinal homeostasis and its breakdown in inflammatory  
607 bowel disease. *Nature* 474, 298-306.

608 Martinez-Rubio, L., Morais, S., Evensen, Ø, Wadsworth, S., Ruohonen, K., Vecino, J.L.G.,  
609 Bell, J.G., Tocher, D.R., 2012. Functional feeds reduce heart inflammation and pathology in  
610 Atlantic salmon (*Salmo salar* L.) following experimental challenge with Atlantic Salmon  
611 Reovirus (ASRV). *PLoS ONE* 7, e40266.

612 Morais, S., Silva, T., Cordeiro, O., Rodrigues, P., Guy, D.R., Bron, J.E., Taggart, J.B., Bell,  
613 J.G., Tocher, D.R., 2012a. Effects of genotype and dietary fish oil replacement with vegetable  
614 oil on the intestinal transcriptome and proteome of Atlantic salmon (*Salmo salar*). *BMC*  
615 *Genomics* 13, 448.

616 Morais, S., Taggart, J., Guy, D., Bell, J., Tocher, D., 2012b. Hepatic transcriptome analysis of  
617 inter-family variability in flesh n-3 long-chain polyunsaturated fatty acid content in Atlantic  
618 salmon. *BMC Genomics* 13, 410.

619 Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A.C., Kanehisa, M., 2007. KAAS: an automatic  
620 genome annotation and pathway reconstruction server. *Nucleic Acids Res.* 35, W182-5.

621 Muller, M., Kersten, S., 2003. Nutrigenomics: goals and strategies. *Nat. Rev. Genet.* 4, 315-  
622 322.

623 Murray, H.M., Lall, S.P., Rajaselvam, R., Boutilier, L.A., Blanchard, B., Flight, R.M.,  
624 Colombo, S., Mohindra, V., Douglas, S.E., 2010. A nutrigenomic analysis of intestinal  
625 response to partial soybean meal replacement in diets for juvenile Atlantic halibut,  
626 *Hippoglossus hippoglossus*, L. *Aquaculture* 298, 282-293.

627 National Research Council (NRC), 2011. Nutrient Requirements of Fish and Shrimp. National  
628 Academy Press, Washington, DC.

629 Øverland, M., Sørensen, M., Storebakken, T., Penn, M., Krogdahl, Å, Skrede, A., 2009. Pea  
630 protein concentrate substituting fish meal or soybean meal in diets for Atlantic salmon (*Salmo*  
631 *salar*) - Effect on growth performance, nutrient digestibility, carcass composition, gut health,  
632 and physical feed quality. *Aquaculture* 288, 305-311.

633 Panserat, S., Hortopan, G.A., Plagnes-Juan, E., Kolditz, C., Lansard, M., Skiba-Cassy, S.,  
634 Esquerré, D., Geurden, I., Médale, F., Kaushik, S., Corraze, G., 2009. Differential gene  
635 expression after total replacement of dietary fish meal and fish oil by plant products in  
636 rainbow trout (*Oncorhynchus mykiss*) liver. *Aquaculture* 294, 123-131.

637 R Core Team, 2013. R: A language and environment for statistical computing. R Foundation  
638 for Statistical Computing, Vienna, Austria. (<http://www.R-project.org/>)

639 Ritchie, M.E., Silver, J., Oshlack, A., Holmes, M., Diyagama, D., Holloway, A., Smyth, G.K.,  
640 2007. A comparison of background correction methods for two-colour microarrays.  
641 *Bioinformatics* 23, 2700-2707.

642 Russell, S., Meadows, L.A., Russell, R.R., 2009. *Microarray Technology in Practice*.  
643 Elsevier, San Diego (USA).

644 Rust, M.B., 2002. Nutritional Physiology. In: Halver, J.E., Hardy, R.W. (Eds.), *Fish Nutrition*.  
645 Academic Press, New York, pp. 367-452.

646 Sahlmann, C., Sutherland, B.J.G., Kortner, T.M., Koop, B.F., Krogdahl, Å, Bakke, A.M.,  
647 2013. Early response of gene expression in the distal intestine of Atlantic salmon (*Salmo*  
648 *salar* L.) during the development of soybean meal induced enteritis. *Fish Shellfish Immunol.*  
649 34, 599-609.

650 Sarwar Gilani, G., Wu Xiao, C., Cockell, K.A., 2012. Impact of antinutritional factors in food  
651 proteins on the digestibility of protein and the bioavailability of amino acids and on protein  
652 quality. Br. J. Nutr. 108, S315-S332.

653 Schwarz, M., Wright, A.C., Davis, D.L., Nazer, H., Björkhem, I., Russell, D.W., 2000. The  
654 bile acid synthetic gene 3beta-hydroxy-Delta(5)-C(27)-steroid oxidoreductase is mutated in  
655 progressive intrahepatic cholestasis. J. Clin. Invest. 106, 1175-1184.

656 Silver, J.D., Ritchie, M.E., Smyth, G.K., 2009. Microarray background correction: maximum  
657 likelihood estimation for the normal-exponential convolution. Biostatistics 10, 352-363.

658 Skugor, S., Grisdale-Helland, B., Refstie, S., Afanasyev, S., Vielma, J., Krasnov, A., 2011.  
659 Gene expression responses to restricted feeding and extracted soybean meal in Atlantic  
660 salmon (*Salmo salar* L.). Aquacult. Nutr. 17, 505-517.

661 Smyth, G.K., 2004. Linear models and empirical bayes methods for assessing differential  
662 expression in microarray experiments. Stat. Appl. Genet. Mol. Biol. 3, Article3.

663 Smyth, G.K., Speed, T., 2003. Normalization of cDNA microarray data. Methods 31, 265-  
664 273.

665 Suzuki, R., Shimodaira, H., 2011. pvclust: Hierarchical Clustering with P-Values via  
666 Multiscale Bootstrap Resampling (<http://cran.r-project.org/web/packages/pvclust/>)

667 Tacchi, L., Bron, J.E., Taggart, J.B., Secombes, C.J., Bickerdike, R., Adler, M.A., Takle, H.,  
668 Martin, S.A.M., 2011. Multiple tissue transcriptomic responses to *Piscirickettsia salmonis* in  
669 Atlantic salmon (*Salmo salar*). Physiol. Genom. 43, 1241-1254.

670 Tacchi, L., Secombes, C., Bickerdike, R., Adler, M., Venegas, C., Takle, H., Martin, S., 2012.  
671 Transcriptomic and physiological responses to fishmeal substitution with plant proteins in  
672 formulated feed in farmed Atlantic salmon (*Salmo salar*). BMC Genomics 13, 363.

673 The World Bank, 2013. Fish to 2030: Prospects for Fisheries and Aquaculture. (www.  
674 worldbank.org)

675 Urán, P.A., Schrama, J.W., Rombout, J.H.W.M., Taverne-Thiele, J.J., Obach, A., Koppe, W.,  
676 Verreth, J.A.J., 2009. Time-related changes of the intestinal morphology of Atlantic salmon,  
677 *Salmo salar* L., at two different soybean meal inclusion levels. J. Fish Dis. 32, 733-744.

678 Urán, P.A., Schrama, J.W., Rombout, J.H.W.M., Obach, A., Jensen, L., Koppe, W., Verreth,  
679 J.A.J., 2008. Soybean meal-induced enteritis in Atlantic salmon (*Salmo salar* L.) at different  
680 temperatures. Aquacult. Nutr. 14, 324-330.

681 Wacyk, J., Powell, M., Rodnick, K., Overturf, K., Hill, R.A., Hardy, R., 2012. Dietary protein  
682 source significantly alters growth performance, plasma variables and hepatic gene expression  
683 in rainbow trout (*Oncorhynchus mykiss*) fed amino acid balanced diets. Aquaculture 356–357,  
684 223-234.

685 Wickham, H., 2009. H. Wickham. ggplot2: elegant graphics for data analysis. Springer New  
686 York, 2009.

687

688 **Figure Legends**

689 **Figure 1.** Hierarchical clustering of gene expression profiles in liver and intestine. Two types  
690 of  $p$ -values are provided: AU (Approximately Unbiased) is computed by multiscale bootstrap  
691 resampling (1000 rep) and reported in red and BP (Bootstrap Probability) is computed by  
692 normal bootstrap resampling and reported in green.

693 **Figure 2.** Histograms plotting the results of differential expression analysis in liver and  
694 intestine of salmon fed diet S30 compared to fish fed diet S0. Figure 2a plots the number of  
695 differentially expressed genes based on  $p$  value and indicates that a larger number of genes  
696 with a  $p < 0.01$  (the dashed line depict the  $p$  value = 0.01) were present in the intestine  
697 compared to the liver. Figure 2b only plots genes with  $p < 0.01$ . FC is fold change.

698 **Figure 3.** Bubble graph plotting GAGE results from the  $2d$  test (a) and  $1d$  test (b) in the liver.  
699 The x-axis plots the fold change (mean statistic) while the y-axis plots the  $\text{Log}_{10}$  transformed  
700  $q$ -value. The size of the bubbles is proportional to the number of genes used by the algorithm  
701 to test the gene-set. Colours of bubbles refer to the biological process (KEGG classification).  
702 Figures are not on the same scale.

703 **Figure 4** Complement cascade generated with the R package *pathview* of affected genes in  
704 the liver (Luo and Brouwer, 2013). Red arrows indicate activation, blue arrow inhibition and  
705 dotted arrows indirect effect. Highlighted in blue are the names of the differentially expressed  
706 genes and the corresponding KEGG identifier.

707 **Figure 5.** Bubble graph plotting GAGE results from the  $2d$  test (a) and  $1d$  test (b) in the  
708 intestine. The x-axis plots the fold change (mean statistic) while the y-axis plots the  $\text{Log}_{10}$   
709 transformed  $q$ -value. The size of the bubbles is proportional to the number of genes applied  
710 by the algorithm to test the gene-set. Colours of bubbles refer to the biological process type  
711 (KEGG classification). The dashed line denotes the cut-off of “highly significant gene-sets”  
712 ( $q < 0.0001$ ). Figures are not on the same scale.



713 **Table 1.** Formulation (g / Kg) and analyzed proximate compositions  
 714 of the experimental diets

	S0	S10	S20	S30
Fishmeal <sup>1</sup>	420.10	368.00	344.00	291.00
Hi-Pro Soymeal <sup>2</sup>	0.00	100.00	200.00	299.50
Corn gluten 60	77.00	77.00	33.00	23.00
Sunflower cake	125.00	104.70	95.80	70.60
Horsebeans	160.00	135.60	105.30	85.30
Fish oil std 18	154.00	158.00	164.00	170.00
Rapeseed oil	40.00	40.00	40.00	40.00
Amino acids	5.70	6.40	7.30	12.10
Vitamins and Minerals	31.90	31.90	31.90	31.90
<b>Proximate composition</b>				
Protein - crude (%)	44.06	44.18	43.83	43.59
Fat - crude (%)	25.97	25.96	25.96	26.02
Ash (%)	7.28	7.47	7.75	7.42
Energy - gross (MJ/kg)	23.65	23.78	23.70	23.75

715

716 <sup>1</sup>NA LT70; <sup>2</sup>Soya cake 48 Hi-Pro (solvent extracted soybean meal cake)

717 All values are represented as g kg<sup>-1</sup> unless otherwise stated.

718

719

720 **Table 2.** Growth performance and somatic indexes ( $\pm$ SD)

<b>Feed</b>	<b>Initial weight</b>	<b>Final weight</b>	<b>SGR</b>
S0	175 $\pm$ 27	424 $\pm$ 78 <sup>a</sup>	1.03 $\pm$ 0.22 <sup>a</sup>
S10	170 $\pm$ 25	422 $\pm$ 76 <sup>a</sup>	1.03 $\pm$ 0.21 <sup>a</sup>
S20	176 $\pm$ 28	405 $\pm$ 67 <sup>a</sup>	0.96 $\pm$ 0.16 <sup>b</sup>
S30	175 $\pm$ 25	378 $\pm$ 67 <sup>b</sup>	0.87 $\pm$ 0.18 <sup>c</sup>

721

722 Number of fish in each triplicate tank was 45. Weights are in g. Calculations of SGR (specific  
 723 growth rate) were based on growth of individual fish. Values within a column not sharing  
 724 superscript letters are significantly different ( $p < 0.05$ ). Growth data was compared using a  
 725 one-way ANOVA for statistical significance.

726

**Table 3.** Forty-four genes differentially expressed ( $p < 0.01$ ) in both liver and intestine of fish fed S30 compared with S0. LogFC is the log<sub>2</sub> transformed fold change.

KOID	Description	LIVER		INTESTINE	
		LogFC	P.Value	LogFC	P.Value
K02377	GDP-L-fucose synthase	-1.01	0.0028	-4.15	< 0.0001
K15985	cAMP-dependent protein kinase inhibitor alpha	-0.68	0.0013	-1.22	< 0.0001
K00736	alpha-1,6-mannosyl-glycoprotein beta-1,2-N-acetylglucosaminyltransferase	-0.54	0.0035	-1.17	< 0.0001
K00864	glycerol kinase	-0.53	0.0050	-1.10	< 0.0001
K03940	NADH dehydrogenase (ubiquinone) Fe-S protein 7	-0.53	0.0076	-0.75	0.0004
K11251	histone H2A	-0.53	0.0066	-0.79	0.0002
K10244	elongation of very long chain fatty acids protein 5	-0.52	0.0039	-1.12	< 0.0001
K10703	very-long-chain (3R)-3-hydroxyacyl-dehydratase	-0.50	0.0013	-1.22	< 0.0001
K00522	ferritin heavy chain	-0.40	0.0063	-0.85	< 0.0001
K04644	clathrin light chain A	-0.37	0.0031	-0.51	0.0001
K11428	histone-lysine N-methyltransferase SETD8	-0.30	0.0034	-0.28	0.0057
K06109	Ras-related protein Rab-13	-0.29	0.0074	-0.46	< 0.0001
K02977	small subunit ribosomal protein S27Ae	-0.23	0.0099	-0.24	0.0064
K09850	Ras association domain-containing protein 1	0.26	0.0084	0.27	0.0063
K14525	ribonucleases P/MRP protein subunit RPP25	0.33	0.0023	0.55	< 0.0001
K12859	U5 snRNP protein, DIM1 family	0.34	0.0090	0.40	0.0024
K02932	large subunit ribosomal protein L5e	0.35	0.0091	0.59	< 0.0001
K02889	large subunit ribosomal protein L21e	0.35	0.0049	0.57	< 0.0001
K03262	translation initiation factor 5	0.36	0.0019	0.35	0.0027
K02868	large subunit ribosomal protein L11e	0.38	0.0016	0.51	< 0.0001
K02984	small subunit ribosomal protein S3Ae	0.39	0.0051	0.53	0.0003
K00940	nucleoside-diphosphate kinase	0.40	0.0064	0.60	0.0001
K12298	bile salt-stimulated lipase	0.41	0.0004	0.30	0.0058
K03996	complement component 7	0.41	0.0061	0.52	0.0009
K10086	ER degradation enhancer, mannosidase alpha 3	0.43	0.0076	0.43	0.0072
K03097	casein kinase II subunit alpha	0.47	0.0007	0.58	< 0.0001
K02936	large subunit ribosomal protein L7Ae	0.52	0.0003	0.70	< 0.0001
K08765	carnitine O-palmitoyltransferase 1	0.65	0.0023	0.74	0.0007
K09414	heat shock transcription factor 1	0.77	< 0.0001	0.45	0.0038
K14397	cleavage and polyadenylation specificity factor subunit 5	0.90	0.0002	0.96	< 0.0001
K01594	sulfinioalanine decarboxylase	1.03	0.0060	2.56	< 0.0001
K11262	acetyl-CoA carboxylase / biotin carboxylase	-0.84	0.0015	0.70	0.0068
K00222	delta14-sterol reductase	-0.61	0.0028	0.64	0.0017
K01507	inorganic pyrophosphatase	-0.53	0.0042	1.08	< 0.0001
K01875	seryl-tRNA synthetase	-0.47	0.0043	0.82	< 0.0001
K04440	c-Jun N-terminal kinase	-0.25	0.0062	0.30	0.0015
K10577	ubiquitin-conjugating enzyme E2 I	0.32	0.0057	-0.39	0.0011
K07292	hepatocyte nuclear factor 4-alpha	0.36	0.0035	-0.51	0.0001
K10574	ubiquitin-conjugating enzyme E2 B	0.38	0.0059	-0.48	0.0008
K06624	cyclin-dependent kinase inhibitor 1B	0.55	0.0034	-0.81	< 0.0001
K08550	estrogen receptor alpha	0.63	0.0082	-0.77	0.0015
K12408	cholest-5-ene-3beta,7alpha-diol 3beta-dehydrogenase	0.68	0.0055	-1.40	< 0.0001
K01103	6-phosphofructo-2-kinase / fructose-2,6-bisphosphatase	0.75	0.0004	-2.99	< 0.0001
K04630	guanine nucleotide-binding protein G(i) subunit alpha	1.52	0.0018	-4.41	< 0.0001

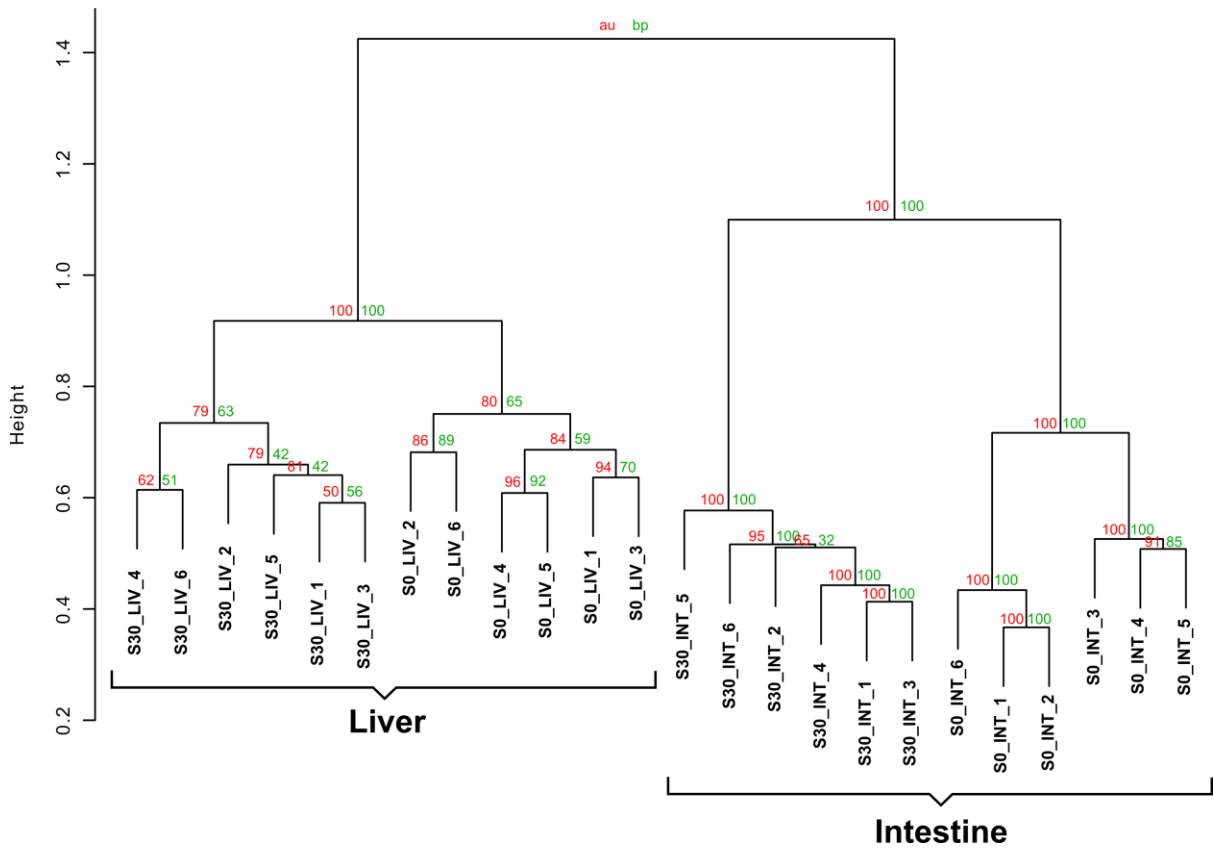
Up-regulated genes are highlighted in red, while down-regulated genes in green. Colour intensity is relative to the magnitude of change. All expression values are reported as Log<sub>2</sub> fold change (LogFC) relative to the control treatment S0. KOID = KEGG identifier.

**Table 4.** Selection of genes differentially expressed only in the liver of fish fed S30 compared with S0. LogFC is the log<sub>2</sub> transformed fold change.

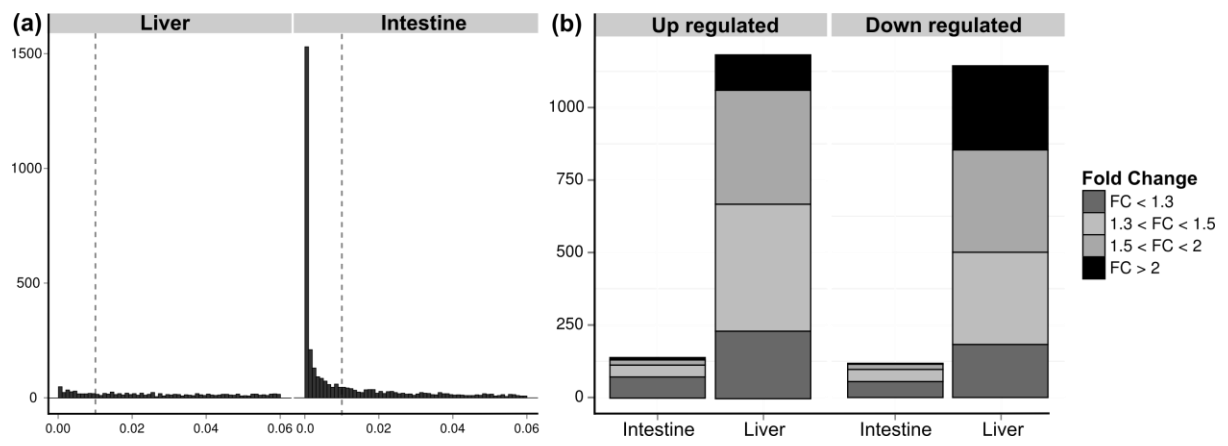
	KOID	Gene	LogFC	P Value
<b>Immune System</b>				
Complement Cascade	K03990	complement component 3	1.15	< 0.001
<b>Metabolism</b>				
Amino acid	K00639	glycine C-acetyltransferase	-0.49	< 0.001
Amino acid	K00789	S-adenosylmethionine synthetase	-0.85	0.003
Carbohydrate	K12047	maltase-glucoamylase	0.62	0.010
Energy	K00164	2-oxoglutarate dehydrogenase E1 component	0.61	< 0.001
Energy	K00234	succinate dehydrogenase (ubiquinone) flavoprotein subunit	0.41	0.003
Lipid	K01597	diphosphomevalonate decarboxylase	-0.63	< 0.001
Lipid	K10205	elongation of very long chain fatty acids protein 2	0.54	< 0.001
Lipid	K01830	prostaglandin-H2 D-isomerase	-0.65	0.001
Lipid	K07296	adiponectin	-0.89	0.002
Lipid	K01823	isopentenyl-diphosphate delta-isomerase	-0.63	0.005
Lipid	K00626	acetyl-CoA C-acetyltransferase	-0.49	0.005
Lipid	K00869	mevalonate kinase	-0.48	0.005
Lipid	K01852	lanosterol synthase	-0.41	0.006
Nucleotide	K03010	DNA-directed RNA polymerase II subunit RPB2	-0.43	0.001
Nucleotide	K01489	cytidine deaminase	-0.54	0.004
Vitamin	K01435	biotinidase	-0.86	< 0.001
Vitamin	K18271	retinol-binding protein 4	-0.90	0.009
<b>Cell Proliferation</b>				
Cell Cycle and Apoptosis	K04426	mitogen-activated protein kinase 5	0.44	< 0.001
Cell Cycle and Apoptosis	K10858	DNA mismatch repair protein PMS2	-0.44	< 0.001
Cell Cycle and Apoptosis	K04441	p38 MAP kinase	0.54	0.002
Cell Cycle and Apoptosis	K06245	laminin, beta 4	0.49	0.003
Cell Cycle and Apoptosis	K04436	mitogen-activated protein kinase 8 interacting protein 3	0.48	0.003
Cell Cycle and Apoptosis	K10567	endonuclease VIII-like 1	0.47	0.007
<b>Genetic Information processing</b>				
Proteasome	K05610	ubiquitin carboxyl-terminal hydrolase L5	-1.12	< 0.001
Protein Processing	K13431	signal recognition particle receptor subunit alpha	-0.59	< 0.001
Protein Processing	K07342	protein transport protein SEC61 subunit gamma and related proteins	-0.49	0.004
Ribosomal	K02891	large subunit ribosomal protein L22e	-0.62	< 0.001
Ribosomal	K02911	large subunit ribosomal protein L32	-0.52	< 0.001
Ribosomal	K17418	large subunit ribosomal protein L37, mitochondrial	-0.48	< 0.001
Ribosomal	K02950	small subunit ribosomal protein S12	-0.40	< 0.001
Ribosomal	K17433	large subunit ribosomal protein L52, mitochondrial	-0.52	0.002
Ribosomal	K17430	large subunit ribosomal protein L49, mitochondrial	-0.53	0.003
Ribosomal	K17428	large subunit ribosomal protein L47, mitochondrial	-0.44	0.003
Ribosomal	K17409	small subunit ribosomal protein S30, mitochondrial	-0.45	0.005
<b>Digestive system</b>				
Pancreatic	K01312	trypsin	1.11	< 0.001
Pancreatic	K01310	chymotrypsin	1.23	< 0.001
Pancreatic	K07886	Ras-related protein Rab-27B	0.58	< 0.001
Pancreatic	K01291	carboxypeptidase B	0.92	< 0.001
Pancreatic	K01346	pancreatic elastase II	0.58	0.001

Up-regulated genes are highlighted in red, while down-regulated genes in green. Colour intensity is relative to the magnitude of change. All expression values are reported as Log<sub>2</sub> fold change (LogFC) relative to the control treatment S0. KOID = KEGG identifier.

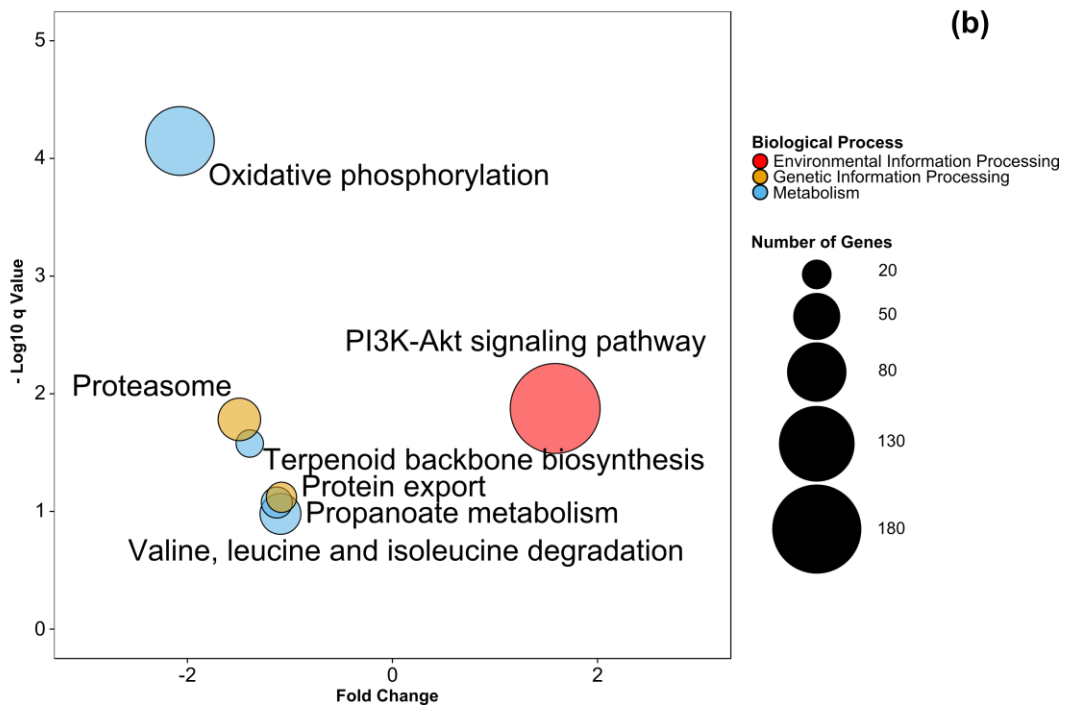
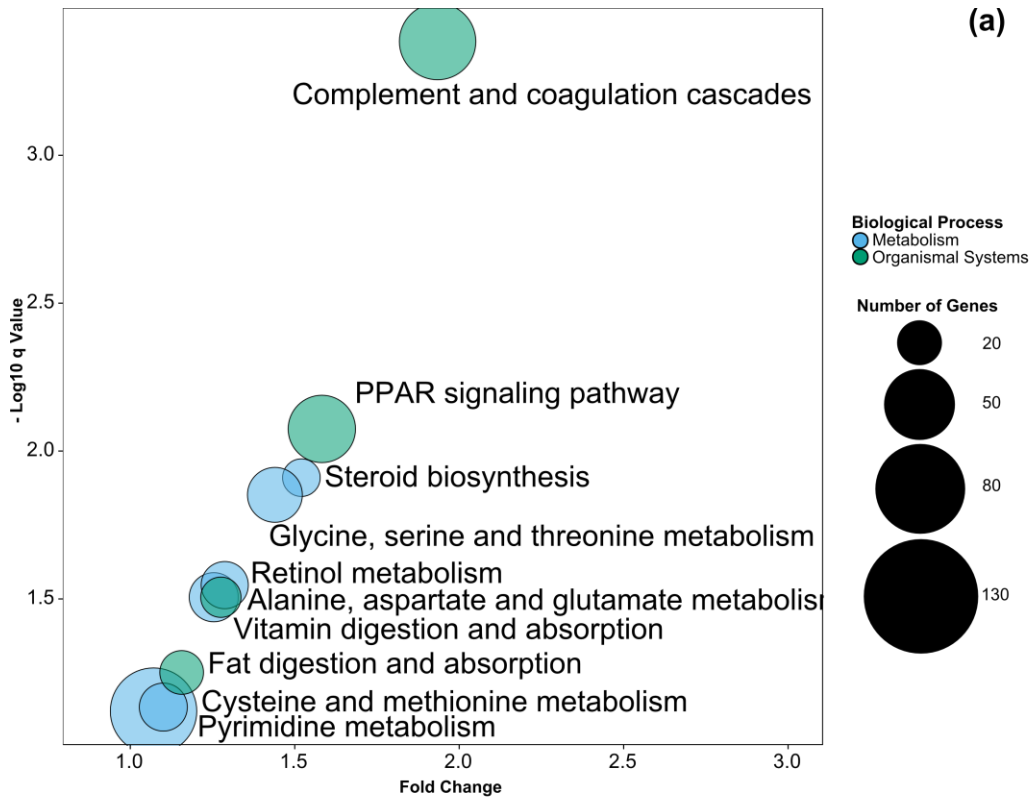
**Fig. 1**



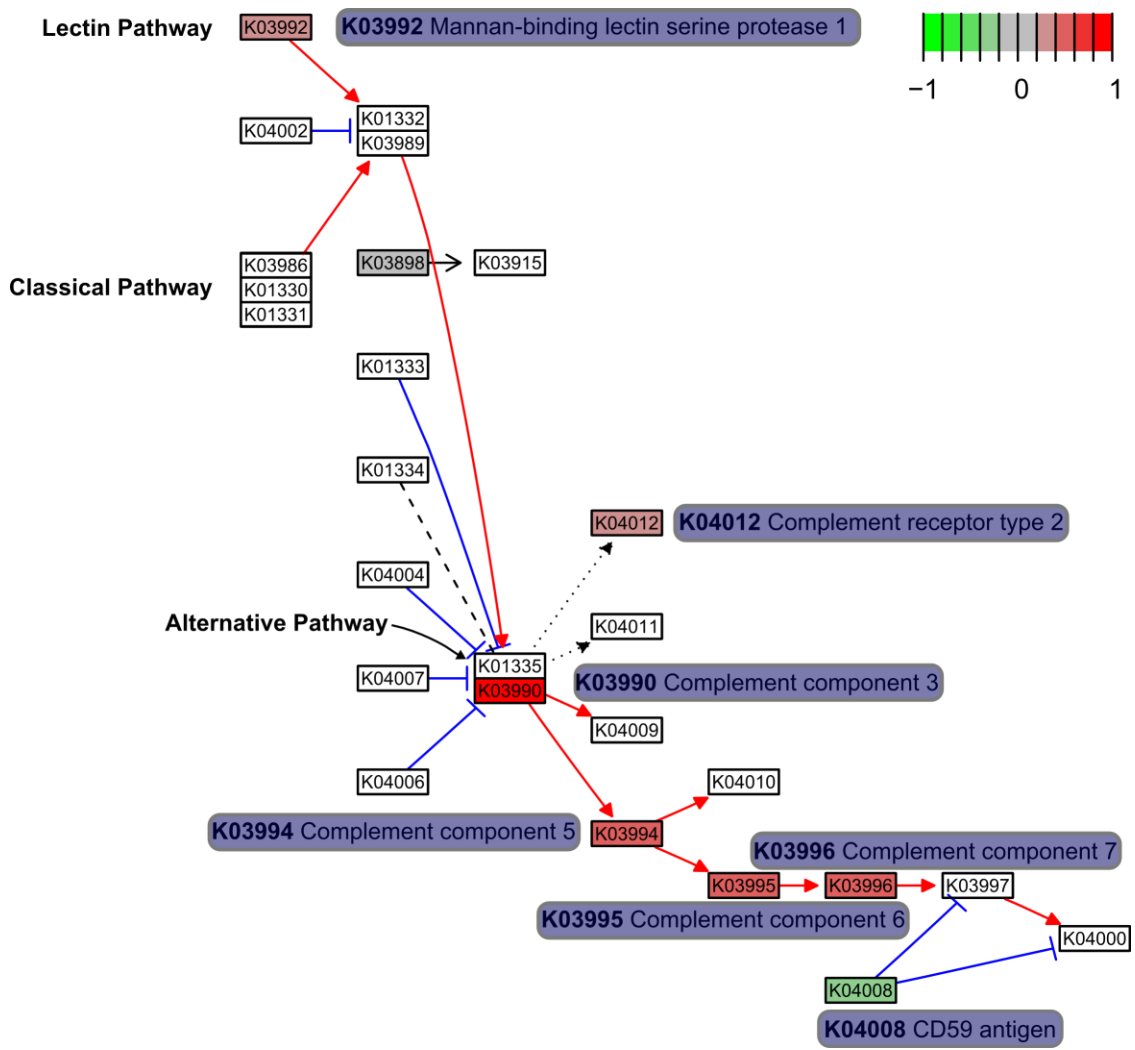
**Fig. 2**



**Fig. 3**

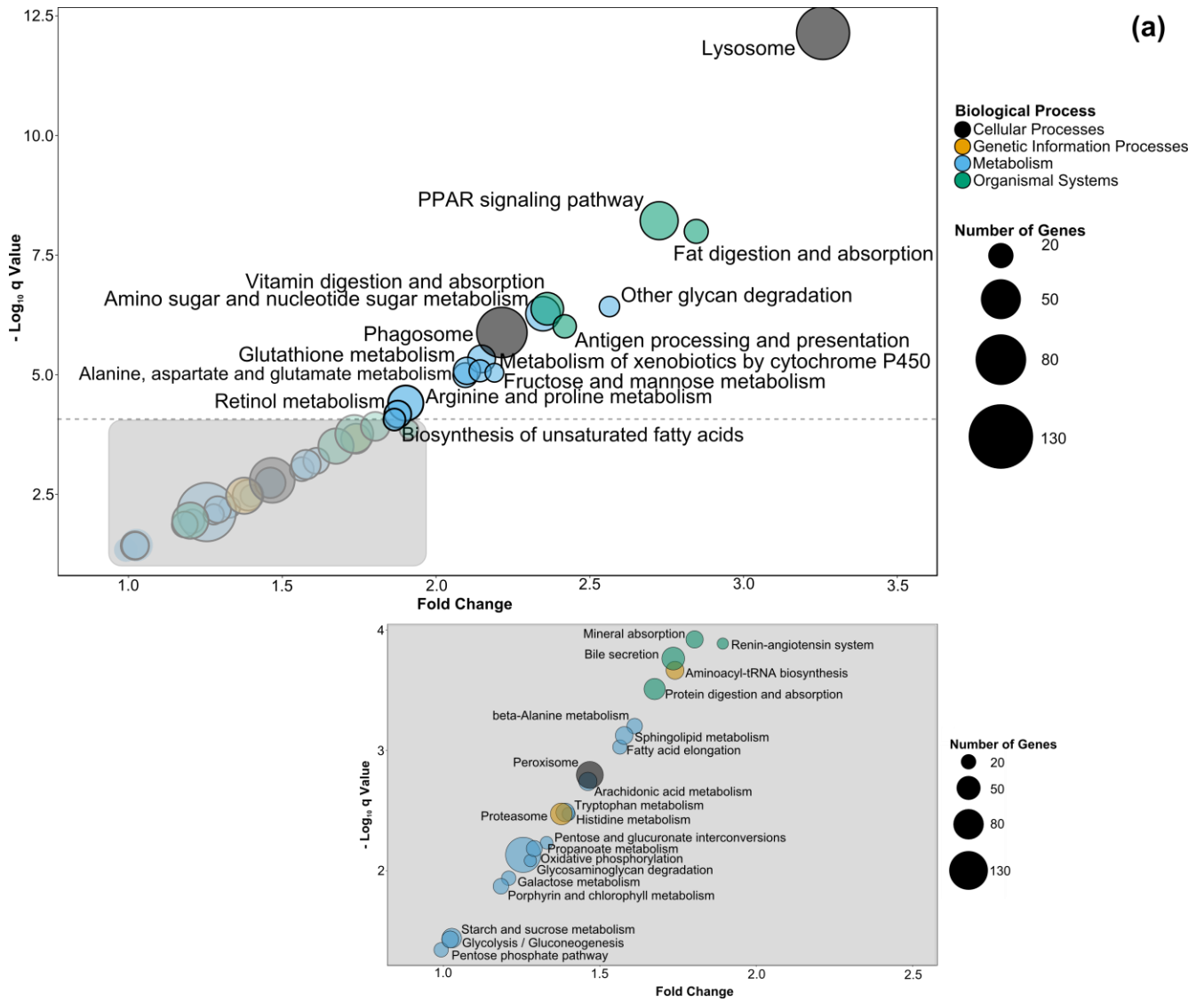


**Fig. 4**





**Fig 5**



(b)

