De Santis C, Crampton V, Bicskei B & Tocher DR (2015) Replacement of dietary soy- with air classified faba bean protein concentrate alters the hepatic transcriptome in Atlantic salmon (Salmo salar) parr, *Comparative Biochemistry and Physiology - Part D: Genomics and Proteomics*, 16, pp. 48-58.

DOI: 10.1016/j.cbd.2015.07.005

Replacement of dietary soy- with air classified faba bean protein concentrate alters the hepatic transcriptome in Atlantic salmon (*Salmo salar*) parr

Christian De Santis[1*], Viv O. Crampton[2], Beatrix Bicskei[1] and Douglas R. Tocher[1]

[1] *Institute of Aquaculture, University of Stirling, Stirling FK9 4LA, Scotland UK*

[2] *EWOS Innovation, 4335 Dirdal, Norway*

Running Title: Liver nutrigenomics in salmon fed vegetable proteins

ms. has 45 pages, 1 tables, 7 figures, 6 supplementary files

[*]Corresponding author:
Dr. Christian De Santis
Institute of Aquaculture, School of Natural Sciences, University of Stirling
Stirling FK9 4LA, Scotland, UK
Tel: +44 1786 467993
Christian.desantis@stir.ac.uk

Funding Source: Technology Strategy Board (Innovate UK)

Keywords: dietary protein replacement, liver, transcriptome, Atlantic salmon, *Salmo salar*, faba bean, *Vicia faba*

Reviewer's Login details to Array Express data:
Username: Reviewer_E-MTAB-2878
Password: hbhrmgda

Abstract

The production of carnivorous fish such as Atlantic salmon (*Salmo salar*) is dependent on the availability of high quality proteins for feed formulations. For a number of nutritional, strategic and economic reasons, the use of plant proteins has steadily increased over the years, however a major limitation is associated with the presence of anti-nutritional factors and the nutritional profile of the protein concentrate. Investigating novel raw materials involves understanding the physiological consequences associated with the dietary inclusion of protein concetrates. The primary aim of the present study was to assess the metabolic response of salmon to increasing inclusion of air-classified faba bean protein concentrate (BPC) in feeds as a replacement for soy protein concentrate (SPC). Specifically, we tested treatments with identical contents of fishmeal (222.4 g kg^{-1}) and progressively higher inclusion of BPC (0 g kg^{-1}, 111.8 g kg^{-1}, 223.6 g kg^{-1}, 335.4 g kg^{-1}, 447.2 g kg^{-1}) substituting SPC. This study demonstrated a dose-dependent metabolic response to a plant ingredient and was the first to compare the nutrigenomic transcriptional responses after substitution of terrestrial feed ingredients such as BPC and SPC without withdrawal of marine ingredients. It was found that after eight weeks a major physiological response in liver was only evident above 335.4 g kg^{-1} BPC and included decreased expression of metabolic pathways, and increased expression of genes regulating transcription and translation processes and the innate immune response. Furthermore, we showed that the nutritional stress caused by BPC resembled, at least at hepatic transcriptional level, that caused by soybean meal (included as a positive control in our experimental design). The outcomes of the present study suggested that Atlantic salmon parr might efficiently utilize moderate substitution of dietary SPC with BPC, with the optimum inclusion level being around 120 g kg^{-1}in the type of feeds tested here.

1. Introduction

The sustainable and profitable production of farmed fish is increasingly dependent on their capacity to perform under dietary regimes based on plant ingredients. Carnivorous fish species such as Atlantic salmon (*Salmo salar*) are in high demand due in part to their beneficial effects on human health as well as the eating quality of their flesh (Whelton et al., 2004, Johnston et al., 2006). To satisfy the increasing demand, high quality protein sources are required for the formulation of the feeds. The stagnation of wild fisheries and the limited availability of marine products such as fishmeal (FM) (FAO 2008-2015), traditionally the ideal protein source for farmed fish (National Research Council, 2011), has increased pressure for the introduction of alternative raw materials as dietary source of proteins (Gatlin et al., 2007). In recent years, considerable research has addressed the performance of fish utilizing alternative feed materials partially or completely substituting FM. Thus far, plant meals and proteins such as soy have been the most economically viable alternative raw materials (Gatlin et al., 2007). As a result of ongoing research, modern commercial feeds for Atlantic salmon utilize significant inclusion levels of alcohol-extracted soy protein concentrate (SPC) as the predominant substitute for FM with, to a lesser extent, a range of other plant sources (Ytrestøyl et al., 2014). However, the use of SPC as the major alternative to FM has raised economic (price of soy fluctuates due to high demand from a number of industries), environmental (over-exploitation of land for farming), strategic (over-reliance on a single ingredient for feed manufacture) and nutritional (less balanced composition compared with mixed sources) concerns, prompting continuing research towards the development of new alternative raw materials to be used in combination with others as protein concentrate (Burr et al., 2012, De Santis et al., 2015a, Zhang et al., 2012).

To be nutritionally attractive for commercial use, candidate alternative feedstuffs for salmon feeds must have relatively high protein content (48 %-80 %, ideally higher than 60 %) with a reasonable balance of essential amino acids, and have low levels of carbohydrates (e.g. fibre and starch) and antinutritional factors (ANFs) (Gatlin et al., 2007). In addition, good alternative feed ingredients must be accessible in terms of availability, price and sustainability. Many of these characteristics are partially dependent on the processing of the raw material that can greatly affect the resulting feedstuff both nutritionally and economically. In this respect, protein concentrates from faba bean (*Vicia faba*) (BPC) possess a number of favorable characteristics. Recently, we investigated the performance of Atlantic salmon fed BPC produced by air-classification, a simple and inexpensive process (De Santis

et al., 2015a). Using a mixture model approach the effects of 16 different combinations of FM, SPC and BPC on growth and health performance of salmon were assessed using parr in freshwater. The screening study aimed to identify an appropriate and more confined range of replacement levels to be applied in seawater studies involving larger fish. It was demonstrated that salmon efficiently utilized BPC at inclusion levels ranging from 50 to 200 g kg^{-1}, partially replacing SPC and/or FM, resulting in increased growth, protein content, fat content and ash. Evidence of detrimental effects on gut health, commonly observed in response to dietary inclusion of some vegetable proteins such as soybean meal (SBM) (Baeverfjord and Krogdahl, 1996, Kortner et al., 2012 and Urán et al., 2008), were not observed at low levels of inclusions (De Santis et al., 2015a). In contrast, it was shown that a high inclusion level of BPC (447.2 g kg^{-1}) caused mild gut inflammation, comparable but not as severe as that caused by SBM in post-smolt, seawater adapted salmon. The most important outcome of that study was that a superior performance was observed in response to mixed ingredients with the optimum formulation being 200.8 g kg^{-1} FM, 268.9 g kg^{-1} SPC, 117.4 g kg^{-1} BPC, providing a strong basis for continuing research on BPC utilization in salmon (De Santis et al., 2015a).

Nutrigenomics is a powerful approach to determine detailed metabolic responses (Mutch et al., 2005). Recently, nutrigenomics has been used as a tool to study the response of fish to vegetable dietary proteins, primarily focusing on hepatic or intestinal profiles (De Santis et al., 2015b, Kortner et al., 2012, Overturf et al., 2012, Panserat et al., 2009, Skugor et al., 2011 and Tacchi et al., 2012,). Specifically in salmon, two studies have reported the hepatic transcriptional signatures underlying a SBM-induced nutritional stress (De Santis et al., 2015b and Skugor et al., 2011). Further studies are however required to elucidate, understand and discriminate the general and specific molecular mechanisms underlying utilization of terrestrial proteins in salmon and fish in general. In this context, the present study aimed to provide insights into the metabolic responses of salmon parr to the utilization of air classified BPC as an alternative source of dietary protein.

The overall aim of the present study was to determine and compare hepatic transcriptomes in Atlantic salmon fed increasing levels of BPC as a substitute for dietary SPC. It is important to emphasize the rationale behind the feeds tested in this study. The experimental feeds used (B0, B20, B40, B60, B80) contained the same level of FM and varying levels of two vegetable proteins: SPC, a refined protein concentrate obtained by aqueous alcohol extraction of soybean, widely established as a dietary ingredient of farmed Atlantic salmon (Ytrestøyl et al., 2014) and BPC, a protein concentrate from faba bean produced with the a dry processing

method (air-classification). In addition, a feed formulated with high levels of SBM (360 g kg^{-1}) was included as positive control to benchmark detrimental effects associated with the plant material and affected by the processing method. Specifically, the objectives of this study were to *a*) establish if the mild effects on gut metabolism, health and impaired growth observed after high inclusion of BPC (i.e. 447.2 g kg^{-1},) was reflected in the alteration of hepatic metabolism, perhaps similar to that observed with high inclusion of SBM (positive control for nutritional stress); *b*) provide metabolic evidence to determine the maximum level of BPC inclusion that is efficiently utilized by salmon; and *c*) understand the metabolic processes underlying the improved growth performance observed previously with low/moderate BPC inclusion (De Santis et al., 2015a) by studying the response to lower inclusion levels of BPC (i.e. 111.8 g kg^{-1}, 223.6 g kg^{-1}). A well-described and validated custom-made Atlantic salmon 44K oligo microarray was utilized for the nutrigenomic profiling. The present study demonstrated a dose-dependent metabolic response to a plant ingredient and represents the first report in fish where the transcriptional response to three terrestrial feed ingredients (BPC, SPC and SBM) is compared.

2. Materials and Methods

2.1. Nutritional trial and experimental treatments

The nutritional trial, including full experimental design and diet formulations is described in detail elsewhere (De Santis et al., 2015a). Briefly, the trial was conducted in the freshwater facilities of EWOS Innovation (Dirdal, Norway) using a farmed population of Atlantic salmon parr of average initial weight of around 1.5g. Fish were acclimatized for two weeks before application of the experimental feeds, which were then fed to quadruplicate tanks. All feeds were formulated to meet the nutritional requirement of salmon (National Research Council, 2011) and to have the same protein, lipid and energy content. After eight weeks of feeding, liver was dissected from 24 individuals per dietary treatment (6 per tank), immediately placed in RNA Later (Life Technologies, Paisley, UK) and processed as per the manufacturer's instructions before being stored at -20°C prior to analyses. For hepatic transcriptional profiling a subset of five dietary treatments was chosen to span the most heterogeneous range of growth and health performance and allow to directly comparing SPC and BPC. The treatments had identical contents of FM (222.4 g kg^{-1}) and progressively higher inclusion of BPC (0 g kg^{-1}, 111.8 g kg^{-1}, 223.6 g kg^{-1}, 335.4 g kg^{-1}, 447.2 g kg^{-1}) substituting SPC referred to as diets B0, B20, B40, B60, B80 respectively (Table 1). In

addition, a feed formulated with 440 g kg^{-1} FM and 360 g kg^{-1} SBM was also analysed as a positive reference. Since knowledge on the hepatic transcriptomic response of Atlantic salmon parr to plant proteins is limited, we used the positive control to benchmark and define transcriptional profiles that could be supposedly associated with detrimental effects of the plant. Inclusion levels of SBM similar to those used in this study are in fact well documented to induce enteropathy in adult salmon (Baeverfjord and Krogdahl, 1996, Krogdahl et al., 2010, Urán et al., 2008, Urán et al., 2009). The feeds analysed in this study corresponded to 20:80:00 (B0), 20:40:40 (B20), 20:60:20 (B40), 20:20:80 (B60), 20:00:80 (B80) and HiSBM (SBM) from our previous nutritional trial (De Santis et al., 2015a) and were renamed for clarity of presentation and understanding.

2.2. Transcriptome analysis

Transcriptomic analysis was conducted using custom-made 4 x 44K Atlantic salmon oligo microarray slides (Agilent Technologies, Wokingham, UK; ArrayExpress accession no. A-MEXP-2065) described in detail previously (Tacchi et al., 2011). The array design and laboratory procedures utilized have been widely used and validated in several previous studies (Betancor et al., 2015a, Betancor et al., 2015b, Bicskei et al., 2014, De Santis et al., 2015b, Morais et al., 2012a, Morais et al., 2012b and Tacchi et al., 2012). The full laboratory protocol and pipeline for bioinformatics analysis are reported in detail in De Santis et al. (2015b). Briefly, equal amounts of RNA from three livers of fish from the same tank were extracted individually using TRI Reagent (Sigma-Aldrich, Dorset, UK), were pooled together and analyzed as a single biological replicate, thus providing 2 experimental samples per tank and 8 replicates per dietary treatment. The same RNA pools were used both for transcriptomic analyses and subsequent RT-qPCR validation. While known to reduce biological variance, within-tank pooling was deliberately chosen as a strategy to maximize the informative power of each biological replicate analyzed in this study and justified by having a non-limiting number of individual samples per experimental unit. The resulting RNA samples were amplified using TargetAmp™ 1-Round Aminoallyl-aRNA Amplification Kit, (Epicentre Technologies Corporation, Madison, Wisconsin, USA) following recommended procedures. Aminoallyl-amplified RNA (aRNA) samples were labelled with Cy3 dye (GE HealthCare Life Sciences, Buckinghamshire, UK) while a pool of all aRNA samples was labelled with Cy5 dye (GE HealthCare Life Sciences) and was used as a common reference. A dual-label common reference design was adopted, where equal amounts of each individual aRNA sample and the common reference pool were hybridized to

one array. Throughout the experiment samples were always randomized, avoiding samples from the same treatment being overrepresented in a particular batch in order to avoid unintentional biases. Details of microarray experiment have been submitted to ArrayExpress under accession number E-MTAB-2878.

Data analysis was performed using R v.3.0.1 and Bioconductor v.2.13 (Gentleman et al. 2004 and R Core Team. 2013). Quality control, data pre-processing and identification of differentially expressed features/genes were conducted using the package limma (Smyth, 2005). Features consistently expressed just above background noise (defined as those features whose intensity was lower than 5th percentile of the distribution in 75% or more of the analysed samples) were removed. Multiple testing correction (False Discovery Rate) was used for differential expression analysis (Benjamini and Hochberg, 1995). Features of the array were annotated using BLAST 2.2.29+ (blastx) against the entire non-redundant protein database as well as using the KEGG Automatic Annotation Server to obtain functional annotations (Altschul et al. 1990 and Moriya et al. 2007). A total of 89.6 % of all probes were returned with a BLAST annotation (annotation date Dec 2014) with e-value < 0.001, while 59 % of probes were returned with a functional annotation (KEGG identifier) using the KAAS server. Features representing the same target gene as implied from KEGG annotation were reduced into a unique value obtained by selecting the feature with the highest F-value calculated on all contrasts. A new dataset was therefore generated for further analyses where each gene was represented by a single feature only. Selecting a subset of features resulted in a dataset of 6740 annotated features targeting unique genes.

2.3 Data mining

Similarity between treatments. Overall similarity between experimental treatments was estimated by association between gene expression profiles analyzed using the limma function *genas* (subset = Fpval). This analysis determined the correlation based on the log transformed fold change (log_2FC) on a feature-by-feature basis. The algorithm also allowed discriminating true biological correlation (ρbiol) from technical correlation (ρtech) (Majewski et al., 2010).

Overview of differential expression. All figures based on differentially expressed genes were plotted using the R package ggplot2 (Wickham, 2009). For figures involving functional information, the KEGG database was used as the preferred classification system.

Gene-Set Enrichment Analysis (GSEA). Unique annotated sequences were analyzed using the R function gage of the software package gage [Generally Applicable Gene-set Enrichment, (Luo et al., 2009)] to identify mechanistic changes as suggested by coordinated expression

changes in gene-sets. For completeness, two types of test were performed: 1 direction (*1d*), testing all genes in a gene-set moving towards the same direction; and 2 directions (*2d*), testing genes in a gene-set that move towards both directions. Gene-sets with a *q*-value < 0.01 were considered significant, where the *q*-value represented the *p*-value adjusted for false discovery rate (Benjamini and Hochberg, 1995). KEGG classification was used for these analyses and all figures were produced using the software package ggplot2.

2.4. RT-qPCR validation

Validation of microarray expression data was performed by reverse transcriptase RT-qPCR. A total of six targets were analyzed, including a reference and five target features. EF1a (Table 2) was used for normalization and was selected based on stability across the analyzed treatments from a number of candidate reference genes (data not reported). The expression of the target genes was normalized using the delta-delta Ct approach (Pfaffl, 2001). Target genes for validation investigated in this study were selected based on various criteria from the microarray dataset (e.g. large fold change, p value < 0.0001, 0.01 < p value < 0.05, etc.). Primers for the target genes were designed using the program PerlPrimer (Marshall, 2004) either to overlap the probe sequence or, where not possible, in proximity of it to ensure sufficient validation by amplification of the same target sequence. Protocols for reverse transcription and qPCR were described in detail previously (Bicskei et al, 2014). Primers utilized for validation and results are provided in Supplementary Table 1 and Supplementary Fig. 1, respectively.

3. Results

3.1 Somatic, biochemical and histological data

Results of somatic, biochemical and histological analyses were presented and discussed in detail previously (De Santis et al., 2015a) and a summary is provided in Fig. 1. Briefly, for the treatments analysed in this study statistical modeling of growth data indicated that maximum weight gain over the 8-weeks experimental period (~ 10.5 g) was achieved in response to feed B20, whereas feeds B0 and B40 had a marginally lower effect on weight gain (~ 10.2 g). At BPC inclusion levels higher than 223.6 g kg^{-1} (B40) fish performance, measured in terms of weight gain, deteriorated resulting in ~ 9.1 g in response to B60 and ~ 7.3 g in response to B80. Protein and oil content varied marginally within a range of 0.5% with the highest protein and fat content generally resulting in response to B20 and B40. For

enteritis assessment, only scores for supernuclear vacuoles (SNV) and goblet cells (GC) in the posterior intestine resulted significantly different between dietary treatments. Fish fed with SBM feed had the highest scores followed by those fed B80. Overall, fish fed B0, B20, B40 and B60 had GC and SNV scores similar to fish fed with a negative control diet (Fig. 1). Notably, fish fed the SBM positive control developed an marginally inflamed posterior intestine (only GC and SNV) but did not show significant growth retardation. Further insights into these results can be found in De Santis et al. (2015a).

3.2 Overview of transcriptomic analysis

Hepatic transcriptome analysis indicated that there was a substantially different response in salmon when only SPC (B0) or BPC (B80) was included as a dietary source of protein. Specifically, the response to the two extreme treatments B0 and B80 differed by 2692 genes that were expressed differently (p value < 0.05). The salmon hepatic transcriptome responded in a dose-dependent manner with an evident correlation between inclusions of BPC and number of differentially expressed genes (i.e. $DEG_{B0\text{-}B20}$ = 6, $DEG_{B0\text{-}B40}$ = 295, $DEG_{B0\text{-}B60}$ = 1503, $DEG_{B0\text{-}B80}$ = 2692) (Supplementary Fig. 1). The majority of expression differences were relatively low (i.e. absolute fold change < 2) with the largest changes observed in the contrast B0-B80. The biological correlation (ρbiol), statistically determined on expression values (LogFC), revealed that moderate inclusions of BPC (i.e. B20 and B40) did not significantly affect the salmon hepatic transcriptome ($\rho biol_{B0\text{-}B20}$ = 0.961, $\rho biol_{B0\text{-}B40}$ = 0.897) and differences emerged above BPC inclusions of 335.4 g kg^{-1} ($\rho biol_{B0\text{-}B60}$ = 0.639, $\rho biol_{B0\text{-}B80}$ = 0.052) (Fig. 2). By comparison of the treatments with the hepatic response to the SBM treatment it was evident that B80 and to a lesser extent B60 were more similar to SBM than treatments with lower levels of BPC (i.e. B0, B20 and B40) (Supplementary Fig 2).

3.3 Gene Set Enrichment Analyses (GSEA)

To elucidate the mechanistic changes that occurred in salmon liver in response to the substitution of dietary SPC with BPC, GSEA was performed using gage. Results of GSEA are shown in Fig. 3 (*1d*) and Fig. 4 (*2d*), while details on specific contrasts are reported in Supplementary Tables 2 & 3. The result of the *1d* analysis was informative in that it indicated the overall direction of change of the genes underlying each gene-set (Fig. 3). The overall trend and similarity between treatments was also immediately evident from this analysis and confirmed the results mentioned above indicating that the responses to diets B20 and B40 were similar to that of B0, whereas B60 and B80 differed significantly. GSEA also confirmed

that the profile of fish fed diets B60 and B80 were closely comparable to the dietary transcriptional signatures observed in response to a SBM-induced nutritional stress, used as a positive control in the trial. The stress signature, identifiable both in response to SBM and B60/B80, included *a*) the down-regulation of all metabolic pathways including two pathways of the endocrine and digestive system (PPAR signaling pathway, vitamin absorption); *b*) the up-regulation of most pathways involved in genetic information processing or protein synthesis (e.g. ribosome biogenesis, protein export, RNA degradation, RNA transport, etc.); *c*) up-regulation of a number of organismal system pathways, which included immune system (cytosolic-DNA sensing pathways, natural killer cell cytotoxicity and Fc epsilon RI signaling pathway) and digestive functions (pancreatic secretion, protein digestion and absorption); and *d*) a marked and significant down-regulation of proteasome-related genes and oxidative phosphorylation. Two further notable features that emerged through GSEA were a set of pathways only down-regulated by high levels of BPC and not by low BPC levels or SBM, including glycolysis/gluconeogenesis, amino acid metabolism (phenylalanine, tryptophan, valine-leucine-isoleucine degradation, glycine-serine-threonine), PPAR signaling pathway and vitamin digestion and absorption, and a group of gene-sets consistently down-regulated in response to diet B40 that included ribosome, protein processing in the endoplasmatic reticulum, protein export, N-glycan biosynthesis and terpenoid backbone biosynthesis (Fig. 3).

In contrast to the *1d* analysis, the *2d* analysis shown in Fig. 4 assumed that genes underlying a certain gene-set changed in both directions. This analysis is particularly effective to identify differential expression in metabolic pathways that often include more complex gene networks (Luo et al., 2009). This test revealed that at least two pathways (steroid biosynthesis and PPAR signaling pathways) were affected in all treatments including BPC (see Supplementary Table 3 for details). In addition, steroid biosynthesis was also the most affected *2d* pathway (highest magnitude and lowest *q* value) in response to high inclusion of BPC (BPC60 and BPC80) as well as SBM, followed by glutathione metabolism. The latter was not altered when dietary BPC was included in small proportions. Lipid digestion and absorption and biosynthesis of unsaturated fatty acids also clustered with these pathways suggesting that lipid metabolism was a particularly affected process. Finally, a large group of metabolic pathways was significantly affected by diet B80, and to a slightly lesser extent also by diets B60 and SBM, including glycolysis/gluconeogenesis, citrate cycle, pentose phosphate pathway, pyruvate metabolism and other metabolic pathways (Fig. 4). As for the *1d* analysis, a set of pathways primarily affected by high BPC, but not by SBM, was identified through

the *2d* test and included amino acid metabolism (tyrosine and beta-alanine), digestive functions (carbohydrate digestion and absorption and pancreatic secretion) and the immune response (complement and coagulation cascades).

3.4 BPC-specific effects

Through the GSEA described in section 3.2, a set of pathways specifically affected by high dietary inclusions of BPC, but not SBM, was identified. These pathways included those regulating digestive functions (vitamins and carbohydrates digestion and absorption, pancreatic secretion), amino acid metabolism (phenylalanine, tyrosine, glycine-serine-threonine, tryptophan, etc.) and the complement and coagulation cascades. To identify the specific mechanisms affected by BPC and not SBM, the genes underlying these processes were analyzed in detail. Genes (by *process*) specifically affected by high levels of BPC included: *digestive processes*) apolipoprotein A1 and B, membrane transporters such as the MFS transporter SP and OPA families, and sodium/potassium-transporting ATPase (subunits alpha and beta) and transcobalamin-2 (Fig. 5); *amino acid metabolism*) genes regulating glycine-serine-threonine metabolism (i.e. glycine N-methyltransferase, serine-glyoxylate transaminase, glycine hydroxymethyltransferase, dimethylglycine dehydrogenase, etc.), phenylalanine metabolism (phenylalanine-4-hydroxylase and 4-hydroxyphenylpyruvate dioxygenase) and tryptophan metabolism (tryptophan 2,3-dioxygenase) (Fig. 6); *complement and coagulation cascades*) complement components (C4, C5, CR3b/4b, C1, C8) and coagulation factors (F2, F5, F8, F13) (Fig. 7).

3.5 RT-qPCR validation of array data

Five features were analyzed by RT-qPCR in a subset of three treatments (B0, B40, B80) to validate the data obtained with the oligo microarray. Five targets were considered sufficient as the custom-made 44K salmon array has been successfully validated in several previous studies (Betancor et al., 2015b, Bicskei et al. 2014, Carmichael et al. 2013 and Morais et al. 2011). RT-qPCR data were comparable to that from the array showing identical direction of changes with comparable fold-changes, treatment variance and *p* values (Supplementary Fig. 3).

4. Discussion

The overarching aim of this study was to gain insight into the consequences of the utilization of BPC as a dietary protein source in Atlantic salmon and it was the first investigation of its kind to study the metabolic response to this protein source. Commercially used SPC was substituted with increasing levels of BPC produced using relatively simple and inexpensive air-classification technology. To benchmark transcriptional profiles supposedly associated with detrimental effects of plant material, a feed including high levels of SBM and known to induce enteritis and nutritional stress was used as a positive control of degraded fish welfare. The metabolic status of fish was determined by analysis of the hepatic transcriptome since the liver is arguably the most important metabolic organ playing key roles in support of digestion, protein and hormone synthesis and detoxification. Specifically, we sought to: *a*) describe and establish differences in the hepatic transcriptome between high inclusion of BPC (i.e. 447.2 g kg^{-1}), where some detrimental effects on gut health and growth were observed, and the SBM reference that caused some clinical sign of enteritis in the gut but not growth retardation (De Santis et al., 2015a); *b*) provide metabolic evidence to determine the maximum level of BPC inclusion that is efficiently used by Atlantic salmon; and *c*) describe the metabolic responses, where possible, underlying the improved performance observed with moderate inclusion of BPC (i.e. 111.8 g kg^{-1}, 223.6 g kg^{-1}). It is important to emphasize that the experimental feeds used in this study contained the same level of FM and that the SBM positive control was formulated with a higher level of FM than the remaining treatments. Therefore, the results discussed herein are solely associated with properties of the plant protein sources and not with a withdrawal of marine ingredients.

The physiological response to dietary SBM is the best characterized nutritional stress caused by a plant protein source in fish and is supported by histological, biochemical and transcriptional data in tissues such as the intestine and more recently, liver (De Santis et al., 2015b, Kortner et al. 2012, Kortner et al., 2013, Overturf et al., 2012, Skugor et al., 2011 and Tacchi et al., 2012). For this reason, a SBM-rich feed was chosen as the positive reference to induce degraded fish welfare in this study. Histological analysis of the posterior intestine aiming to assess the presence of gut inflammation revealed that while salmon parr fed the positive reference were the most affected amongst all treatments, they only developed a mild gut inflammation primarily involving a reduction of supernuclear vacuoles and increased number of goblet cells. Despite the high level of SBM used in the diet (360 g kg^{-1}), the detrimental effects obtained were only partially comparable to those well documented in studies performed with post-smolt, seawater adapted salmon (Baeverfjord and Krogdahl, 1996, Urán et al., 2008, Urán et al., 2009). Our results were in line with those recently

reported in Atlantic salmon parr suggesting that early life stages of salmon might be either more tolerant to inclusions of plant material in the feed or having a less pronounced immune response in the gut due to a partially developed immune system (Sahlmann et al. 2015). Despite the response in the gut, however, a significant alteration of the hepatic transcriptome was observed in comparison to the control feed B0 containing only SPC. In addition, we demonstrated that the hepatic transcriptional response of salmon fed very high inclusion of BPC was largely consistent with that caused by SBM. These results were informative considering that SBM and SPC are obtained from the same seed through different processing. The dietary transcriptional signatures underlying the nutritional stress (very high BPC- or SBM-induced) included the down-regulation in liver of metabolic pathways including energy metabolism, digestive functions (except pancreatic secretion) and proteasome-related genes. Up-regulated processes included protein synthesis as well as pathways associated with the immune response. Interestingly, this signature was consistent with that observed in salmon smolts fed high levels of SBM in saltwater (De Santis et al. 2015b, Skugor et al. 2011), although contrarily to the intestinal inflammatory response that appeared to be only marginal, the response of the liver seemed more pronounced. This might suggest that in early life stages such as parr the liver might play a more central role in facing the nutritional challenge of plant proteins, an observation that supports other studies in parr showing that this organ is one of the first to develop (Sahlmann et al. 2015).

The majority of transcriptional changes observed in the liver in response to B80 (and partially to B60) were common to those developing in response to SBM, thus suggesting that these changes are likely to be associated with the processing of the plant rather than the nutritional profile of soy and faba bean. The processing method is in fact critical in determining the level of ANFs remaining in protein concentrates (Krogdahl et al., 2010). This hypothesis is supported by the fact that the control diet B0 contained more soy-derived protein (from SPC) and less marine proteins (FM) than the positive control SBM and if detrimental effects were associated with the nutritional profile of the soybean, these should be more evident in the experimental feed B0. Thus, SBM had a nutritional profile different to any other experimental feed tested in this study due to the significantly higher content of FM (440 g kg^{-1} vs 222.4 g kg^{-1}), arguably the best protein source for fish. The effects of ANFs have been widely investigated and can have a significant impact on digestive capacity, metabolism and health of farmed animals in general (Krogdahl et al., 2010). For example, the glycoprotein conglycinin, a seed storage protein, can induce oxidative damage and inflammation in fish resulting in an impaired endogenous antioxidant response as well as negatively affecting

hepatic lipid metabolism (Yamazaki et al., 2012 and Zhang et al., 2013). Similarly, the heat-stable glycosides, saponins, can impact lipid metabolism, negatively affecting plasma cholesterol levels, bile salt concentrations and hepatic bile synthesis in salmon and other organisms (Francis et al., 2001 and Gu et al., 2014), and trigger the response of components of the innate immune response in fish (Fuentes-Appelgren et al., 2014). Finally, trypsin inhibitors may increase pancreatic hyperactivity with over secretion of digestive enzymes such as chymotrypsins, elastases and carboxypeptidases into the intestinal lumen with subsequent loss of resources (i.e. sulphur-rich amino acids) in the faeces and as a result lead to growth suppression and also affect protein and amino acid digestibility (Rosewicz, 1989, Sarwar Gilani et al., 2012).

Evidence of impaired metabolism in the liver was observed both in response to high BPC and SBM, including a significant decrease of lipid, carbohydrate, energy and amino acid metabolism. Liver is arguably the most active metabolic tissue and, as such, it is rich in mitochondria to sustain its metabolic functions. One hypothesis to explain the observed reduced metabolism is that high levels of dietary BPC or SBM caused a degree of oxidative stress resulting in impaired mitochondrial function. This was supported by the fact that beans, and faba beans in particular, are potentially potent oxidative stressors due to the presence of vicin, a toxic alkaloid glycoside (Lattanzio et al., 1983). The oxidative potential of vicin is well known in humans, where consumption of faba beans in individuals with glucose-6 phosphate dehydrogenase deficiency causes the disease favism associated with the inability of erythrocytes to tolerate vicin-induced oxidative stress (Cappellini and Fiorelli, 2008). Activation of glutathione metabolism indicated by GSEA supported this hypothesis. In fact, glutathione plays an important role as a cofactor of enzymes with antioxidant functions protecting the cell from oxidative damage (Pompella et al., 2003). A second hypothesis to explain reduced metabolic functions is reduced nutrient uptake from the intestine due either to effects such as the inhibition of lipid absorption by saponins (Gu et al., 2014) or by directly affecting digestion itself (digestibility). Notably, in this study genes coding for pancreatic enzymes were up-regulated after feeding high levels of BPC or SBM in salmon parr, a trend also observed in post-smolt salmon (De Santis et al., 2015b). The increased levels of these digestive enzymes in salmon liver might resemble the condition known as “hyperactive pancreas” described in mammalian models where increased synthesis of pancreatic enzymes is observed after exposure to plant-born protease inhibitors (De Santis et al., 2015b and Sarwar Gilani et al., 2012). These hypotheses are not mutually exclusive, however a

significant reduction of nutrient uptake seems less likely at least in the case of fish fed SBM, since degraded growth was not observed within the timeframe of the experiment.

In contrast to genes regulating metabolic functions that showed consistently decreased expression, a number of other genes showed the opposite trend following the nutritional stress (B80, SBM). These included mainly transcripts coding for factors involved in protein synthesis such as, for example, the biogenesis of ribosomes, RNA transport and degradation, protein export and protein processing in the endoplasmatic reticulum, indicating an increased potential for net protein synthesis and export. This observation, together with data indicating reduced metabolic functions, might suggest that high dietary BPC as well as SBM may induce an investment of energy in other directions, for example towards the synthesis of proteins involved in detoxification or immune responses. The liver receives most of its blood supply through the portal vein that connects this organ to the intestine. The blood flowing to the liver is rich in nutrients that are digested in the intestinal lumen and absorbed by the enterocytes, but it may also contain bacterial products, toxins and food antigens. The ingress of unwanted products may be facilitated by enterocyte disruption caused by the local inflammation generally observed in response to plant material and also detected after feeding high levels of BPC and SBM in salmon parr (De Santis et al., 2015b).

There is evidence that liver plays an important role in the immune response towards pathogens (Gao et al, 2008). The results of this study suggest that only a marginal inflammatory response was seen at intestinal level, whereas in the liver we identified at least three interrelated mechanisms that might indicate an activation of the first line of defense against pathogens. These included the up-regulation of the cytosolic DNA sensing pathway, a family of pattern recognition receptors that are responsible for the activation of the response against the invading pathogen (Takaoka and Taniguchi, 2008). The activation of pattern recognition receptors is responsible, amongst other, of triggering an extracellular response that involves the recruitment of natural killer cells for the elimination of the infected cells. Genes involved in the natural killer cell-mediated cytotoxicity were also up regulated. One last important mechanism suggested by our results was the possible involvement of the complement cascade, which can occur through three different biochemical pathways: classical, alternative and lectin pathways (Degn and Thiel, 2013). The classical pathway requires antigen-antibody complexes for activation and is mediated by complement component 1 protein complex (C1q, C1s, C1r). The expression of genes coding for elements of the classical pathway (e.g. C1-subcomponents) were down-regulated by the use of high levels of BPC but not when low levels or SBM were used. The other two routes leading to the

common terminal pathway are the alternative and lectin pathways, which are mediated by complement component 3 and mannan-binding serine protease-I, respectively (Janeway, 2001 and Janeway and Travers, 1994). Mediators of the alternative and lectin pathways (e.g. C3 and MASP-1) were significantly up-regulated in response to high BPC and, to a lesser extent, SBM. The up-regulation of genes regulating the lectin pathway was also observed in the liver of salmon smolts fed high levels of dietary SBM (De Santis et al., 2015b), whereas in zebrafish it was showed that dietary supplementation of ANFs such as saponin resulted in an increased innate immune response as indicated by increased number of granulocytes associated with the digestive tract as well as higher expression of genes in the larvae known to regulate this response including C3 (Fuentes-Appelgren et al., 2014).

The hepatic transcriptional response of salmon to diets B60 and B80 was in several aspects comparable to that observed in response to the SBM control, suggesting that high levels of these two protein concentrates may induce similar detrimental effects on fish. However, three major processes were identified in which salmon responded differently. These processes could mirror the slower growth rate observed in B60 and B80 compared with SBM, in contrast with the remaining processes that should relate to the specific response to ANFs. Understanding these differences could help clarify the issues of utilization of BPC as a dietary protein source. The first major difference concerned amino acid metabolism, where three genes including dimethylglycine dehydrogenase, glycine N-methyltransferase and betaine-homocysteine S-methyltransferase were down-regulated in response to high BPC but not SBM and they participate in the synthesis of glycine from choline. Soybean is a rich source of choline in the form of phosphatidylcholine and, although most of this is concentrated in the oil fraction, a proportion of choline still remains in the protein concentrate (Menten et al., 1997). In addition, BPC-related down-regulation of genes coding for catabolic enzymes regulating tryptophan (tryptophane 2,3 dioxygenase) and phenylalanine (phenylalanine 4-hydroxilase and 4-hydroxyphenylpyruvate dioxygenase) degradation was observed. This might be a response to different dietary levels of the aromatic amino acids supplied by BPC compared to SPC and SBM. For example, the level of tryptophan in BPC used in this trial was 40 % lower than that measured in SPC (De Santis et al., 2015a). The differential expression of these catabolic genes might explain the utilization of the resources, however are unlikely to result in major differences at phenotypic level. In fact, while not every amino acid was balanced in the diet, they all met the nutritional requirement of salmon (De Santis et al., 2015a). Finally, differences were also identified in the expression of apolipoprotein-B and to a lesser extent apolipoprotein-A1, involved in the formation of

chylomicrons and consequently in the transport of vitamins and triacylglycerol, as well as in the expression of transcobalamin-2 necessary for the uptake of vitamin B_{12}. These differences however were only noted above BPC inclusion of 335.4 g kg^{-1} and not below that level when a more equal contribution of SPC and BPC was provided in the diet.

Testing novel raw material such as plant protein can be challenging. The presence of xenobiotics, together with the fact that these foreign compounds may have individual as well as synergistic actions on the physiological response encourage studies to provide an overall view of the physiological state of the animal. The present investigation provided a profile of the hepatic transcriptome response to dietary BPC in salmon. The overall metabolic status of salmon was generally not significantly altered by inclusion of BPC up to a level of 223.6 g kg^{-1} (B40) and feeds below these inclusions resulted in a similar response. The molecular outcome of the present study supported our previous results describing the effects on growth performance and health indicating that Atlantic salmon could tolerate moderate substitution of dietary SPC with BPC, with the optimum being 117.4 g kg^{-1} (De Santis et al., 2015a). In addition, it was shown that BPC inclusions above 335.4 g kg^{-1} (B60) triggered a transcriptional response that suggested ongoing nutritional stress detrimental for fish health and welfare, similar to that observed at maximum inclusions of BPC (447.2 g kg^{-1}) and in response to 360 g kg^{-1} SBM. Our previous study showed that diet B60 did not result in gut inflammation and resulted in only a minor loss of performance, however data from the present study demonstrated that this level of substitution exceeded the amount that can be efficiently metabolized and utilized by salmon at the hepatic level a particularly useful information considering that post-smolt, seawater adapted salmon might be more susceptible to these level of inclusion compared to parr. Future studies should therefore focus on the use of dietary BPC in seawater smolts not exceeding inclusion levels of 223.6 g kg^{-1} and include a wider range of sampling times and tissues to fully understand the utilization of plant proteins in salmon. Further, the possibility that salmon could develop detrimental effects when fed a diet with BPC inclusion level below 223.6 g kg^{-1} over longer period of feeding should be investigated.

Availability of supporting data

The data set supporting the results of this article is available in the ArrayExpress public repository with the accession number E-MTAB-2878.

Competing interests

563 The authors declare that they have no competing interests.

564

565 **Authors' contributions**

566 The microarray study was conceived and designed by CDS and DRT. CDS performed,
567 interpreted and evaluated all molecular and bioinformatic analyses. VC designed and
568 managed all aspects of the salmon feeding trial. BB assisted in the molecular analyses. The
569 manuscript was written by CDS and DRT. All authors read and approved the final
570 manuscript.

571

572 **Acknowledgements**

573 This research and CdS were funded by the UK Technology Strategy Board (TSB) Sustainable
574 Protein Production initiative (project no.101096). We gratefully acknowledge Dr. Laura
575 Martinez-Rubio for the collection of tissue samples and Dr. John B. Taggart for endless
576 advice and assistance with laboratory procedures.

577

578 **References**

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

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

584 Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful
585 approach to multiple testing. J. R. Stat. Soc. Ser. B Stat. Methodol. 57, pp. 289-300.

586 Betancor, M.B., Sprague, M., Sayanova, O., Usher, S., Campbell, P.J., Napier, J.A., Caballero, M.J.,
587 Tocher, D.R., 2015a. Evaluation of a high-EPA oil from transgenic Camelina sativa in feeds for
588 Atlantic salmon (*Salmo salar* L.): Effects on tissue fatty acid composition, histology and gene
589 expression. Aquaculture 444, 1-12.

590 Betancor, M.B., Sprague, M., Usher, S., Sayanova, O., Campbell, P.J., Napier, J.A., Tocher, D.R.,
591 2015b. A nutritionally-enhanced oil from transgenic Camelina sativa effectively replaces fish oil
592 as a source of eicosapentaenoic acid for fish. Scientific Reports 5, 8104.

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

596 Burr, G.S., Wolters, W.R., Barrows, F.T., Hardy, R.W., 2012. Replacing fishmeal with blends of
597 alternative proteins on growth performance of rainbow trout (*Oncorhynchus mykiss*), and early or
598 late stage juvenile Atlantic salmon (*Salmo salar*). Aquaculture 334–337, 110-116.

599 Cappellini, M., Fiorelli, G., 2008. Glucose-6-phosphate dehydrogenase deficiency. The Lancet 371,
600 64-74.

601 Carmichael, S., Bron, J., Taggart, J., Ireland, J., Bekaert, M., Burgess, S., Skuce, P., Nisbet, A.,
602 Gharbi, K., Sturm, A., 2013. Salmon lice (*Lepeophtheirus salmonis*) showing varying emamectin
603 benzoate susceptibilities differ in neuronal acetylcholine receptor and GABA-gated chloride
604 channel mRNA expression. BMC Genomics 14, 408.

605 De Santis, C., Ruohonen, K., Tocher, D.R., Martin, S.A.M., Król, E., Secombes, C.J., Bell, J.G., El-
606 Mowafi, A., Crampton, V.O., 2015a. Atlantic salmon (*Salmo salar*) parr as a model to predict the
607 optimum inclusion of air classified faba bean protein concentrate in feeds for seawater salmon.
608 Aquaculture 444, 70-78.

609 De Santis, C., Bartie, K.L., Olsen, R.E., Taggart, J.B., Tocher, D.R., 2015b. Nutrigenomic profiling of
610 transcriptional processes affected in liver and distal intestine in response to a soybean meal-
611 induced nutritional stress in Atlantic salmon (*Salmo salar*). Comp. Biochem. Phys. D 15, 1-11.

612 Degn, S.E., Thiel, S., 2013. Humoral Pattern Recognition and the Complement System. Scand. J.
613 Immunol. 78, 181-193.

614 FAO 2008-2015. Fisheries and Aquaculture topics. The State of World Fisheries and Aquaculture
615 (SOFIA). Topics Fact Sheets. Text by Jean- Francois Pulvenis. In: FAO Fisheries and
616 Aquaculture Department [online]. Rome. Updated 19 May 2015. [Cited 7 July 2015].
617 http://www.fao.org/fishery/sofia/en.

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

620 Fuentes-Appelgren, P., Opazo, R., Barros, L., Feijoó, C.G., Urzúa, V., Romero, J., 2014. Effect of the
621 dietary inclusion of soybean components on the innate immune system in zebrafish. Zebrafish
622 11(1), 41-49.

623 Gao, B., Jeong, W., Tian, Z., 2008. Liver: An organ with predominant innate immunity. Hepatology
624 47, 729-736.

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

629 Gentleman, R., Carey, V., Bates, D., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge,
630 Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F., Li, C.,
631 Maechler, M., Rossini, A., Sawitzki, G., Smith, C., Smyth, G., Tierney, L., Yang, J., Zhang, J.,
632 2004. Bioconductor: open software development for computational biology and bioinformatics.
633 Genome Biol. 5, R80.

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

637 Janeway, C., 2001. Immunobiology five. Garland Pub.

638 Janeway, C., Travers, P., 1994. Immunobiology: The Immune System in Health and Disease. Current
639 Biology Limited.

640 Johnston, I.A., Li, X., Vieira, V.L.A., Nickell, D., Dingwall, A., Alderson, R., Campbell, P.,
641 Bickerdike, R., 2006. Muscle and flesh quality traits in wild and farmed Atlantic salmon.
642 Aquaculture 256, 323-336.

643 Kortner, T.M., Skugor, S., Penn, M.H., Mydland, L.T., Djordjevic, B., Hillestad, M., Krasnov, A.,
644 Krogdahl, A., 2012. Dietary soyasaponin supplementation to pea protein concentrate reveals
645 nutrigenomic interactions underlying enteropathy in Atlantic salmon (*Salmo salar*). BMC Vet.
646 Res. 8, 101-6148-8-101.

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

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

653 Lattanzio, V., Bianco, V., Crivelli, G., Miccolis, V., 1983. Variability of amino acids, protein, vicine
654 and convicine in *Vicia faba* (L) cultivars. J. Food Sci. 48, 992-993.

655 Luo, W., Friedman, M.S., Shedden, K., Hankenson, K.D., Woolf, P.J., 2009. GAGE: generally
656 applicable gene set enrichment for pathway analysis. BMC Bioinf. 10, 161-2105-10-161.

657 Majewski, I.J., Ritchie, M.E., Phipson, B., Corbin, J., Pakusch, M., Ebert, A., Busslinger, M., Koseki,
658 H., Hu, Y., Smyth, G.K., Alexander, W.S., Hilton, D.J., Blewitt, M.E., 2010. Opposing roles of
659 polycomb repressive complexes in hematopoietic stem and progenitor cells. Blood 116, 731-739.

660 Marshall, O.J., 2004. PerlPrimer: cross-platform, graphical primer design for standard, bisulphite and
661 real-time PCR. Bioinformatics 20, 2471-2472.

662 Menten, J., Pesti, G., Bakalli, R., 1997. A new method for determining the availability of choline in
663 soybean meal. Poultry Science 76, 1292-1297.

664 Morais, S., Silva, T., Cordeiro, O., Rodrigues, P., Guy, D.R., Bron, J.E., Taggart, J.B., Bell, J.G.,
665 Tocher, D.R., 2012a. Effects of genotype and dietary fish oil replacement with vegetable oil on
666 the intestinal transcriptome and proteome of Atlantic salmon (*Salmo salar*). BMC Genomics 13,
667 448.

668 Morais, S., Taggart, J.B., Guy, D.R., Bell, J.G., Tocher, D.R., 2012b. Hepatic transcriptome analysis
669 of inter-family variability in flesh n-3 long-chain polyunsaturated fatty acid content in Atlantic
670 salmon. BMC Genomics 13, 410-2164-13-410.

671 Morais, S., Pratoomyot, J., Taggart, J., Bron, J., Guy, D., Bell, J.G., Tocher, D., 2011. Genotype-
672 specific responses in Atlantic salmon (*Salmo salar*) subject to dietary fish oil replacement by
673 vegetable oil: a liver transcriptomic analysis. BMC Genomics 12, 255.

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

676 Mutch, D.M., Wahli, W., Williamson, G., 2005. Nutrigenomics and nutrigenetics: the emerging faces
677 of nutrition. The FASEB Journal 19, 1602-1616.

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

680 Overturf, K., Vallejo, R., Palti, Y., Barrows, F., Parsons, J., 2012. Microarray analysis of differential
681 utilization of plant-based diets by rainbow trout. Aquacult. Int. 20, 213-232.

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

686 Pfaffl, M.W., 2001. A new mathematical model for relative quantification in real-time RT-PCR.
687 Nucleic Acids Res. 29, e45.

688 Pompella, A., Visvikis, A., Paolicchi, A., Tata, V.D., Casini, A.F., 2003. The changing faces of
689 glutathione, a cellular protagonist. Biochem. Pharmacol. 66, 1499-1503.

690 R Core Team, 2013. R: A language and environment for statistical computing. R Foundation for
691 Statistical Computing, Vienna, Austria. http://www. R-project. org/.

692 Rosewicz, S.L., 1989. Pancreatic digestive enzyme gene expression: effects of CCK and soybean
693 trypsin inhibitor. Am. J. Physiol. Gastr. L. 256, G733-G738.

694 Sahlmann, C., Gu, J., Kortner, T.M., Lein, I., Krogdahl, Å, Bakke, A.M., 2015. Ontogeny of the
695 Digestive System of Atlantic Salmon (*Salmo salar* L.) and Effects of Soybean Meal from Start-
696 Feeding. PLoS ONE 10, e0124179.

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

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

703 Smyth, G.K., 2005. Limma: linear models for microarray data. In: Gentleman, R., Carey, V., Dudoit,
704 S., Irizarry, R., Huber, W. (Eds.), Bioinformatics and Computational Biology Solutions using R
705 and Bioconductor. Springer, New York, pp. 397-420.

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

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

712 Takaoka, A., Taniguchi, T., 2008. Cytosolic DNA recognition for triggering innate immune
713 responses. Adv. Drug Deliv. Rev. 60, 847-857.

714 Urán, P.A., Schrama, J.W., Rombout, J.H.W.M., Obach, A., Jensen, L., Koppe, W., Verreth, J.A.J.,
715 2008. Soybean meal-induced enteritis in Atlantic salmon (*Salmo salar* L.) at different
716 temperatures. Aquacult. Nutr. 14, 324-330.

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

720 Warnes, G.R., Bolker, B., Bonebakker, L., Gentleman, R., Liaw, W.H.A., Lumley, T., Maechler, M.,
721 Magnusson, A., Moeller, S., Schwartz, M., Venables, B., 2013. gplots: Various R programming
722 tools for plotting data.

723 Whelton, S.P., He, J., Whelton, P.K., Muntner, P., 2004. Meta-Analysis of observational studies on
724 fish intake and coronary heart disease. Am. J. Cardiol. 93, 1119-1123.

725 Wickham, H., 2009. ggplot2: elegant graphics for data analysis.

726 Yamazaki, T., Kishimoto, K., Miura, S., Ezaki, O., 2012. Dietary β-conglycinin prevents fatty liver
727 induced by a high-fat diet by a decrease in peroxisome proliferator-activated receptor γ2 protein.
728 J. Nutr. Biochem. 23, 123-132.

729 Ytrestøyl, T., Synnøve., T., Åsgård, A.T., 2014. Resource utilisation of Norwegian salmon farming in
730 2012 and 2013. NOFIMA (2014). http://nofima.no/en/pub/1151443/

731 Zhang, J., Guo, L., Feng, L., Jiang, W., Kuang, S., Liu, Y., Hu, K., Jiang, J., Li, S., Tang, L., Zhou,
732 X., 2013. Soybean β-conglycinin induces inflammation and oxidation and causes dysfunction of
733 intestinal digestion and absorption in fish. PLoS ONE 8, e58115.

734 Zhang, Y., Øverland, M., Shearer, K.D., Sørensen, M., Mydland, L.T., Storebakken, T., 2012.
735 Optimizing plant protein combinations in fish meal-free diets for rainbow trout (Oncorhynchus
736 mykiss) by a mixture model. Aquaculture 360–361, 25-36.

Figure Legends

Figure 1. **Summary of somatic, histological and biochemical data extracted from De Santis et al., (2015a)**. From top left moving clockwise: a) contour plot showing weight gain (g fish) over eight weeks, "X" indicate feeds tested in the original study while red dots indicate feeds considered in the present study (B0, B20, B40, B60, B80; see Table 1 for formulations). The contour lines join points of equal growth. Lighter colours denote regions with higher weight gain; b) Modelled effect of feeds on supernuclear vacuoles (SNV) and the goblet cells (GC) in the posterior intestine. Mean and 95% confidence interval are shown. Positive number means a more severe effect. Blue dotted line denotes the mean score assigned to the positive control for enteritis while red dotted line that of fish fed the negative control; c) Contour plot showing the whole body protein content of fish; d) Contour plot showing the whole body oil content of fish. For further information on these results the reader is referred to the original publication (De Santis et al., 2015a).

Figure 2. Association of gene expression profiles as analysed by *genas* (limma package). Green ellipses denote the biological correlation (ρbiol) compared to the technical correlation (ρtech) represented in purple as calculated by the algorithm. Dots represent gene expression $\log_2$ transformed fold change, reported relatively to the common reference pool.

Figure 3. Heatmap comparing results of GSEA (*1d*). The figure was generated using heatmap.2 (package gplots) (Warnes et al. 2013). Hierarchical clustering (Euclidean distance) was performed on individual expression values (i.e. "statistics") calculated using the gage algorithm. All expression values are reported compared to B0. For clarity, only pathways that were significantly different ($q < 0.01$) in at least one contrast are plotted. All replicates analyzed are shown (BPC, n = 8; SBM, n = 6), represented by an individual square.

Figure 4. Heatmap comparing results of GSEA (*2d*). The figure was generated using heatmap.2 (package gplots) (Warnes et al. 2013). Hierarchical clustering (Euclidean distance) was performed on individual expression values (i.e. "statistics") calculated using the gage algorithm. All expression values are reported compared to B0. For clarity, only pathways that were significantly different ($q < 0.01$) in at least one contrast are plotted. All replicates analyzed are shown (BPC, n = 8; SBM, n = 6), represented by a square.

Figure 5. **Expression of genes regulating digestive functions**. Heatmap plotting Log_2-transformed fold change expression values (relatively to B0) of genes regulating pancreatic secretion, vitamin digestion and absorption and carbohydrate digestion and absorption. Genes were selected according to KEGG classification. For clarity, only the most variable genes

were selected (Interquartile range > $Log_2(1.3)$). Note: Fig. 5, Fig. 6 and Fig. 7 are not on the same scale. Genes highlighted in yellow were only affected by high concentration of BPC but not SBM.

Figure 6. **Expression of genes regulating amino acid metabolism**. Heatmap plotting Log_2-transformed fold change expression values (relative to B0) of genes regulating amino acid metabolism. Genes were selected according to KEGG classification. For clarity, only the most variable genes were selected (Interquartile range > $Log_2(1.2)$). Note: Fig. 5, Fig. 6 and Fig. 7 are not on the same scale. Genes highlighted in yellow were only affected by high concentration of BPC but not SBM.

Figure 7. **Expression of genes regulating complement and coagulation cascades**. Heatmap plotting Log_2-transformed fold change expression values (relative to B0) of genes regulating the complement and coagulation cascades. Genes were selected according to KEGG classification. For clarity, only the most variable genes were selected (Interquartile range > $Log_2(1.15)$). Note: Fig. 5, Fig. 6 and Fig. 7 are not on the same scale. Genes highlighted in yellow were only affected by high concentration of BPC but not SBM.

Legend to Supplementary Figures

Supplementary Figure 1. Plot of genes that resulted differentially expressed ($p < 0.01$, absolute fold change > 1.5) for the contrasts B80 vs B0, SBM vs B0 and SBM vs B80. Hierarchical clustering (Euclidean distance) was performed on individual expression values. The figure was generated using heatmap.2 (package gplots) (Warnes et al. 2013).

Supplementary Figure 2. Pairwise comparison of number of unique differentially expressed genes (DEG) in hepatic transcriptomes between fish fed the different dietary treatments. Features targeting the same gene were excluded hence the number represents features targeting unique genes affected by each dietary treatments.

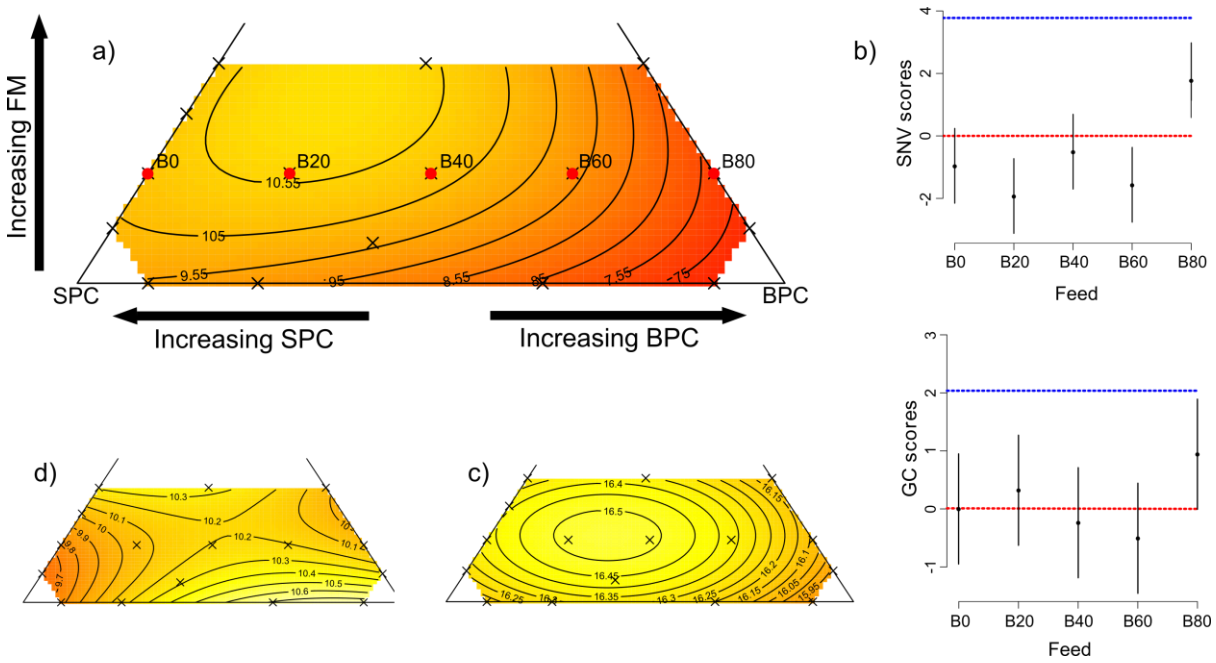
Supplementary Figure 3. Validation of the expression of five probes analysed using array technology and RT-qPCR.

799 **Table 1.** Formulations of experimental feeds

Ingredients	**B0**	**B20**	**B40**	**B60**	**B80**
LT FM	222.4	222.4	222.4	222.4	222.4
Selecta SPC 60	448.2	336.2	224.1	112.1	0.00
Fabaqua 62-65	0.00	111.8	223.6	335.4	447.2
Wheat gluten	80.0	80.0	80.0	80.0	80.0
Tapioca	76.1	76.3	76.6	76.7	77.0
Vitamin, Mineral and Pigments premixes	56.6	56.6	56.6	56.6	56.6
Synthetic amino acids	8.3	9.1	10.0	10.7	11.5
Fish oil	98.4	97.6	96.8	96.1	95.3
Lecithin source	10.0	10.0	10.0	10.0	10.0

800 Formulation of the experimental feeds used in the study. All values are represented as g kg^{-1}.
801 Further details on dietary formulations can be found in De Santis et al., 2015a)
802

803 **Figure 1**



804

805

Figure 2

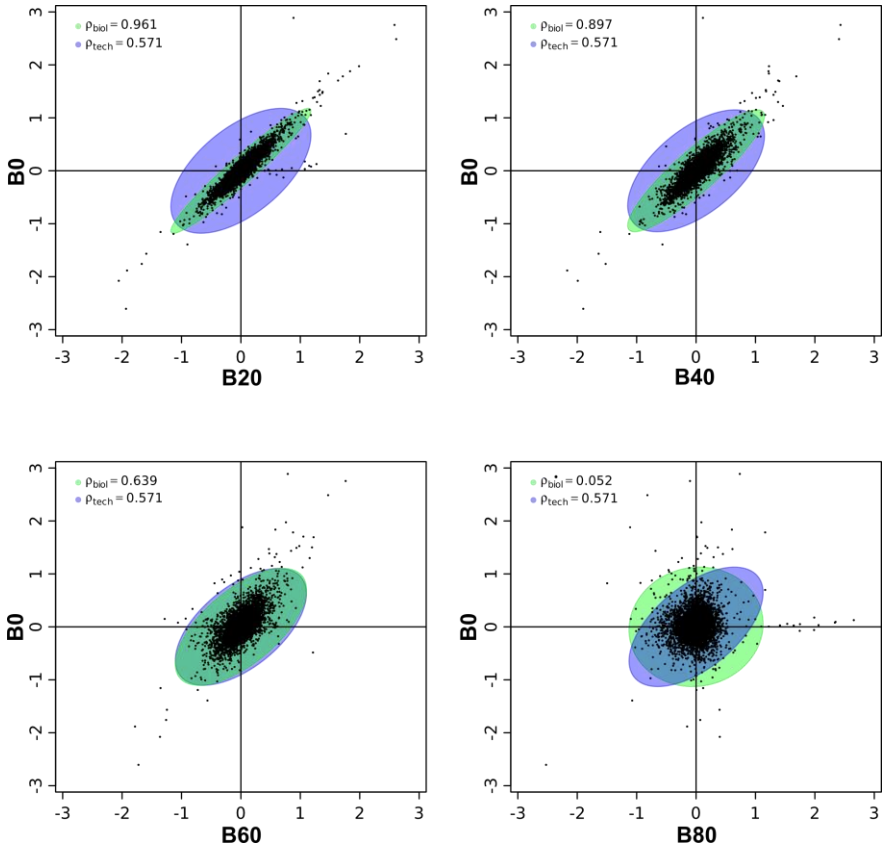

ρbiol = 0.961
ρtech = 0.571
B0
B20
ρbiol = 0.897
ρtech = 0.571
B0
B40
ρbiol = 0.639
ρtech = 0.571
B0
B60
ρbiol = 0.052
ρtech = 0.571
B0
B80

Figure 3


-4 -2 0 2 4
Fold Change
Genetic Information Processing
Metabolism
Organismal System
Environmental Information Processing
Cellular Process
Proteasome
Oxidative phosphorylation
Cytosolic DNA-sensing pathway
RNA polymerase
mRNA surveillance pathway
Natural killer cell mediated cytotoxicity
Fc epsilon RI signaling pathway
Neuroactive ligand-receptor interaction
Ubiquitin mediated proteolysis
Protein digestion and absorption
Pancreatic secretion
Ribosome biogenesis in eukaryotes
RNA transport
RNA degradation
Protein export
Protein processing in endoplasmic reticulum
Ribosome
N-Glycan biosynthesis
Terpenoid backbone biosynthesis
Aminoacyl-tRNA biosynthesis
PPAR signaling pathway
ABC transporters
Phenylalanine metabolism
Glycolysis / Gluconeogenesis
Tryptophan metabolism
Valine, leucine and isoleucine degradation
Glyoxylate and dicarboxylate metabolism
Glycine, serine and threonine metabolism
Vitamin digestion and absorption
Steroid biosynthesis
p53 signaling pathway
Mismatch repair
Glutathione metabolism
Citrate cycle (TCA cycle)
Fatty acid elongation
Biosynthesis of unsaturated fatty acids
Pentose phosphate pathway
Propanoate metabolism
Pyruvate metabolism
111.8 g kg⁻¹ BPC
223.6 g kg⁻¹ BPC
335.4 g kg⁻¹ BPC
447.2 g kg⁻¹ BPC
360 g kg⁻¹ SBM
B20
B40
B60
B80
SBM

Figure 4

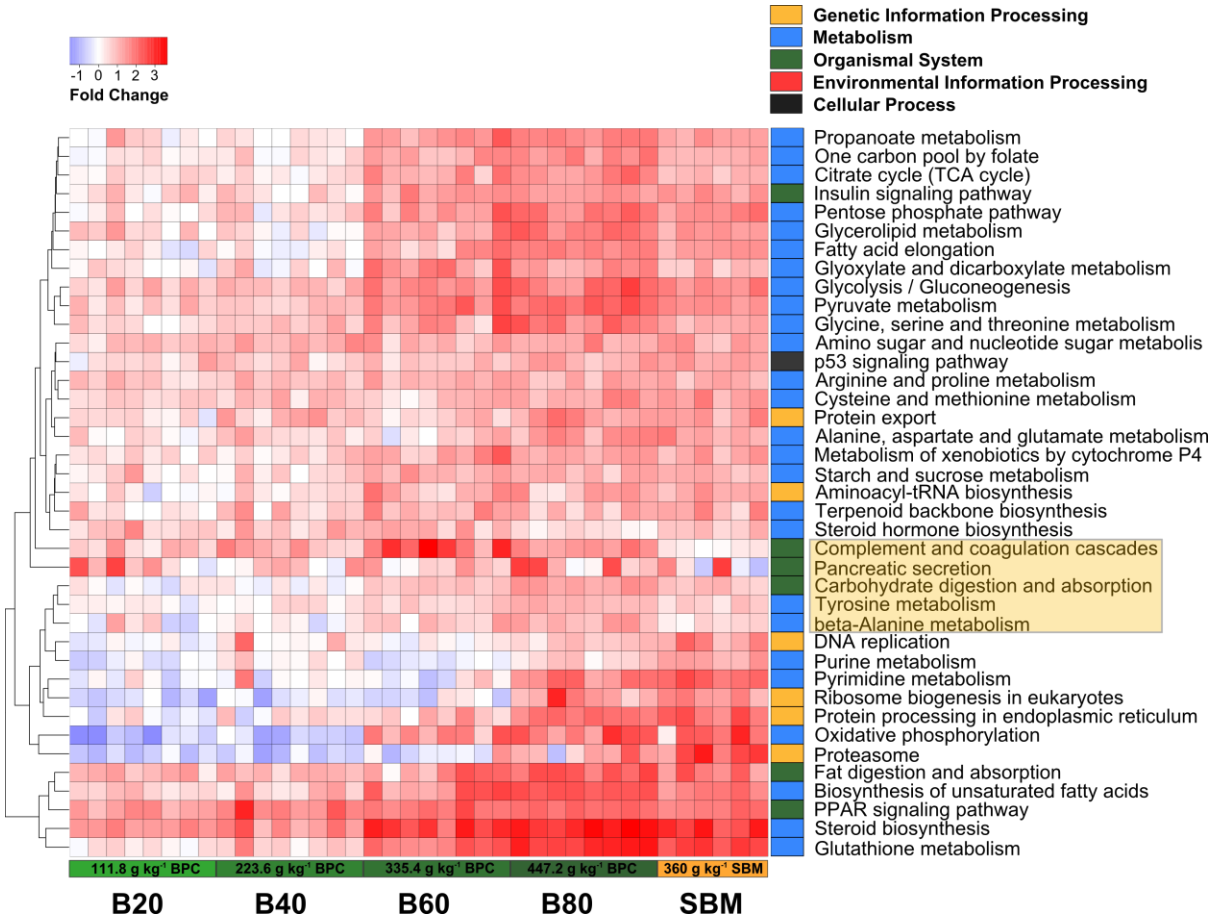

Genetic Information Processing
Metabolism
Organismal System
Environmental Information Processing
Cellular Process
-1 0 1 2 3
Fold Change
Propanoate metabolism
One carbon pool by folate
Citrate cycle (TCA cycle)
Insulin signaling pathway
Pentose phosphate pathway
Glycerolipid metabolism
Fatty acid elongation
Glyoxylate and dicarboxylate metabolism
Glycolysis / Gluconeogenesis
Pyruvate metabolism
Glycine, serine and threonine metabolism
Amino sugar and nucleotide sugar metabolis
p53 signaling pathway
Arginine and proline metabolism
Cysteine and methionine metabolism
Protein export
Alanine, aspartate and glutamate metabolism
Metabolism of xenobiotics by cytochrome P4
Starch and sucrose metabolism
Aminoacyl-tRNA biosynthesis
Terpenoid backbone biosynthesis
Steroid hormone biosynthesis
Complement and coagulation cascades
Pancreatic secretion
Carbohydrate digestion and absorption
Tyrosine metabolism
beta-Alanine metabolism
DNA replication
Purine metabolism
Pyrimidine metabolism
Ribosome biogenesis in eukaryotes
Protein processing in endoplasmic reticulum
Oxidative phosphorylation
Proteasome
Fat digestion and absorption
Biosynthesis of unsaturated fatty acids
PPAR signaling pathway
Steroid biosynthesis
Glutathione metabolism
111.8 g kg-1 BPC
223.6 g kg-1 BPC
335.4 g kg-1 BPC
447.2 g kg-1 BPC
360 g kg-1 SBM
B20
B40
B60
B80
SBM

Figure 5

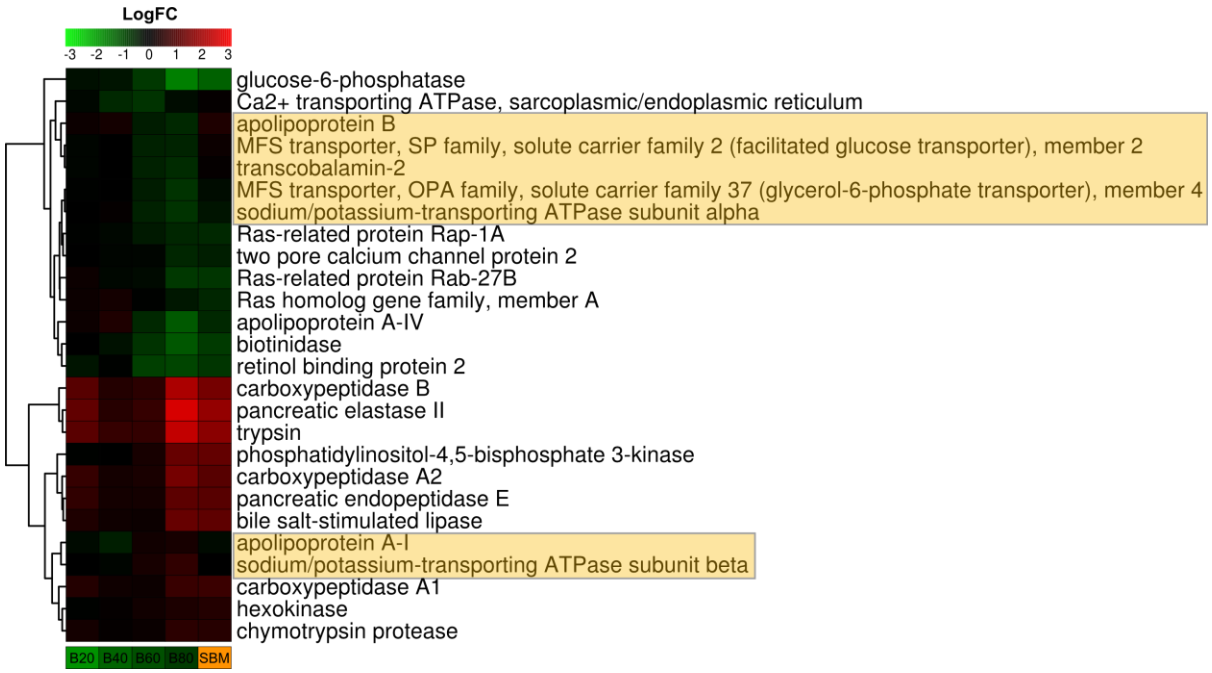

LogFC
-3 -2 -1 0 1 2 3
glucose-6-phosphatase
Ca2+ transporting ATPase, sarcoplasmic/endoplasmic reticulum
apolipoprotein B
MFS transporter, SP family, solute carrier family 2 (facilitated glucose transporter), member 2
transcobalamin-2
MFS transporter, OPA family, solute carrier family 37 (glycerol-6-phosphate transporter), member 4
sodium/potassium-transporting ATPase subunit alpha
Ras-related protein Rap-1A
two pore calcium channel protein 2
Ras-related protein Rab-27B
Ras homolog gene family, member A
apolipoprotein A-IV
biotinidase
retinol binding protein 2
carboxypeptidase B
pancreatic elastase II
trypsin
phosphatidylinositol-4,5-bisphosphate 3-kinase
carboxypeptidase A2
pancreatic endopeptidase E
bile salt-stimulated lipase
apolipoprotein A-I
sodium/potassium-transporting ATPase subunit beta
carboxypeptidase A1
hexokinase
chymotrypsin protease
B20 B40 B60 B80 SBM

Figure 6

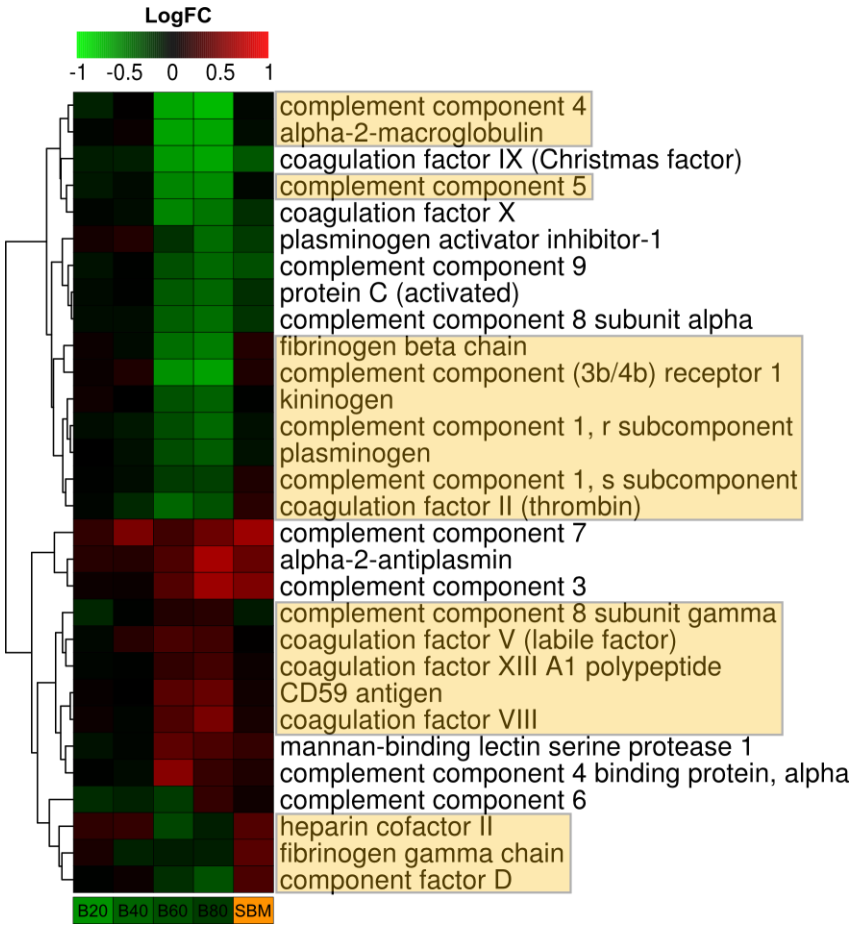
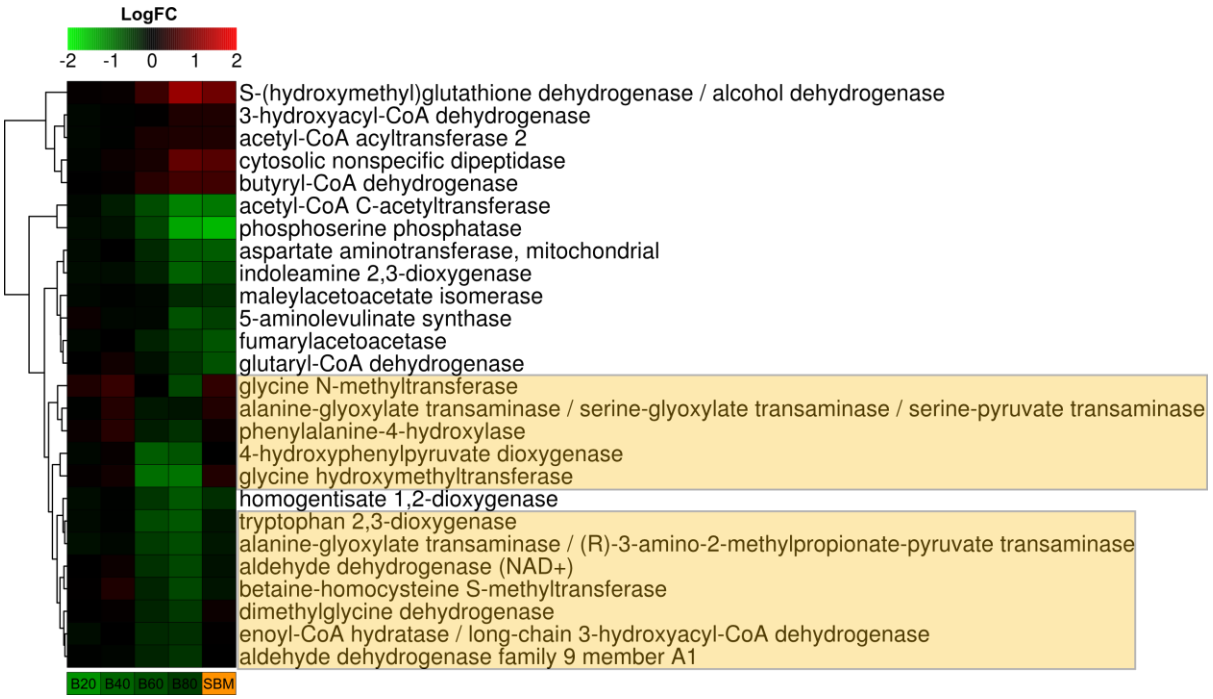
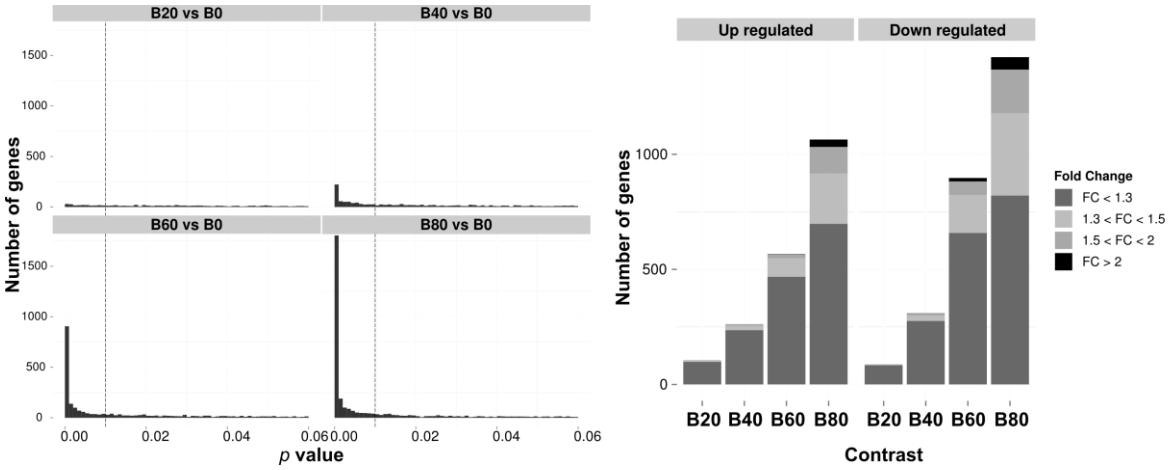

LogFC
-1 -0.5 0 0.5 1
complement component 4
alpha-2-macroglobulin
coagulation factor IX (Christmas factor)
complement component 5
coagulation factor X
plasminogen activator inhibitor-1
complement component 9
protein C (activated)
complement component 8 subunit alpha
fibrinogen beta chain
complement component (3b/4b) receptor 1
kininogen
complement component 1, r subcomponent
plasminogen
complement component 1, s subcomponent
coagulation factor II (thrombin)
complement component 7
alpha-2-antiplasmin
complement component 3
complement component 8 subunit gamma
coagulation factor V (labile factor)
coagulation factor XIII A1 polypeptide
CD59 antigen
coagulation factor VIII
mannan-binding lectin serine protease 1
complement component 4 binding protein, alpha
complement component 6
heparin cofactor II
fibrinogen gamma chain
component factor D
B20 B40 B60 B80 SBM

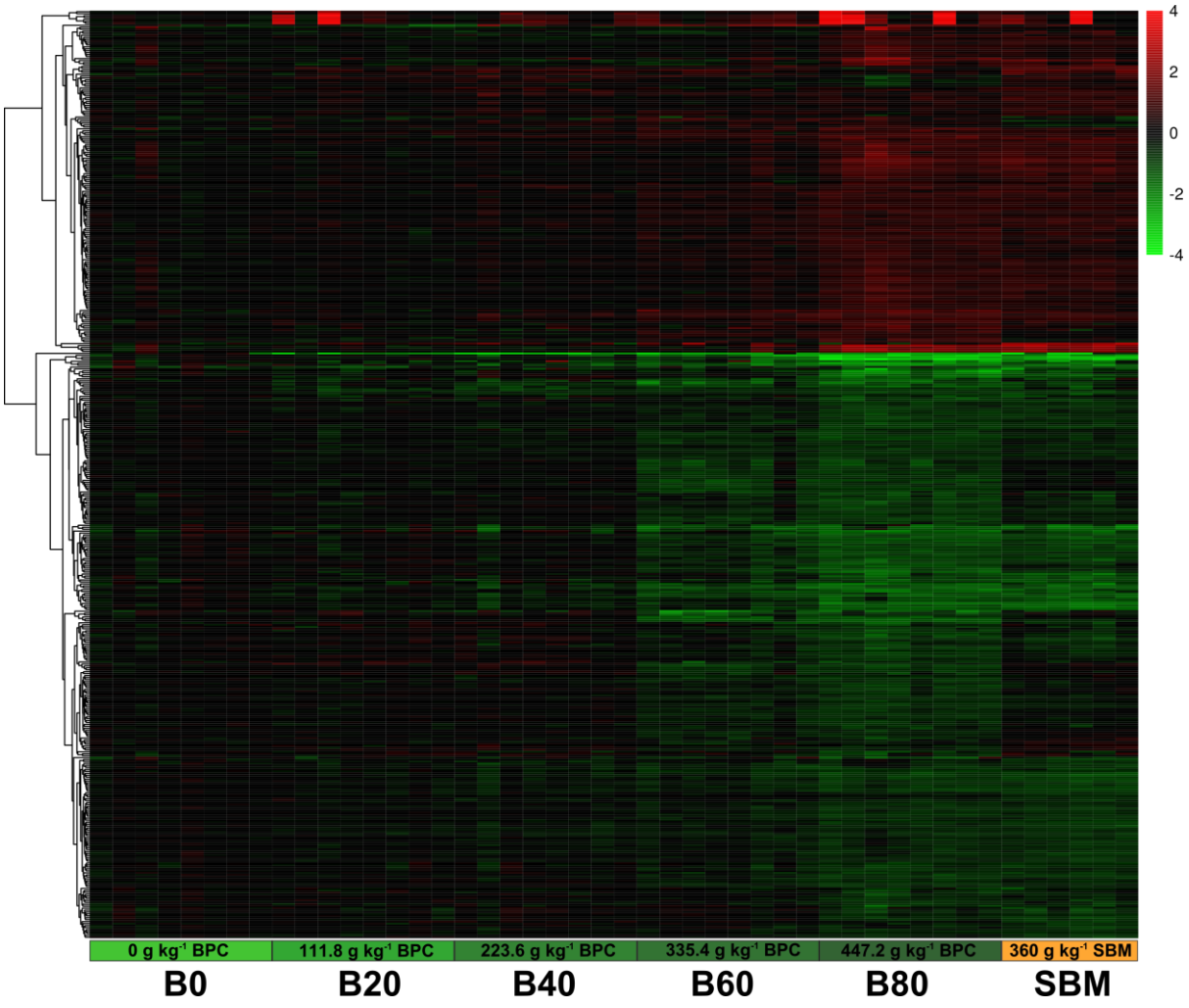
Figure 7


LogFC
-2
-1
0
1
2
S-(hydroxymethyl)glutathione dehydrogenase / alcohol dehydrogenase
3-hydroxyacyl-CoA dehydrogenase
acetyl-CoA acyltransferase 2
cytosolic nonspecific dipeptidase
butyryl-CoA dehydrogenase
acetyl-CoA C-acetyltransferase
phosphoserine phosphatase
aspartate aminotransferase, mitochondrial
indoleamine 2,3-dioxygenase
maleylacetoacetate isomerase
5-aminolevulinate synthase
fumarylacetoacetase
glutaryl-CoA dehydrogenase
glycine N-methyltransferase
alanine-glyoxylate transaminase / serine-glyoxylate transaminase / serine-pyruvate transaminase
phenylalanine-4-hydroxylase
4-hydroxyphenylpyruvate dioxygenase
glycine hydroxymethyltransferase
homogentisate 1,2-dioxygenase
tryptophan 2,3-dioxygenase
alanine-glyoxylate transaminase / (R)-3-amino-2-methylpropionate-pyruvate transaminase
aldehyde dehydrogenase (NAD+)
betaine-homocysteine S-methyltransferase
dimethylglycine dehydrogenase
enoyl-CoA hydratase / long-chain 3-hydroxyacyl-CoA dehydrogenase
aldehyde dehydrogenase family 9 member A1
B20
B40
B60
B80
SBM

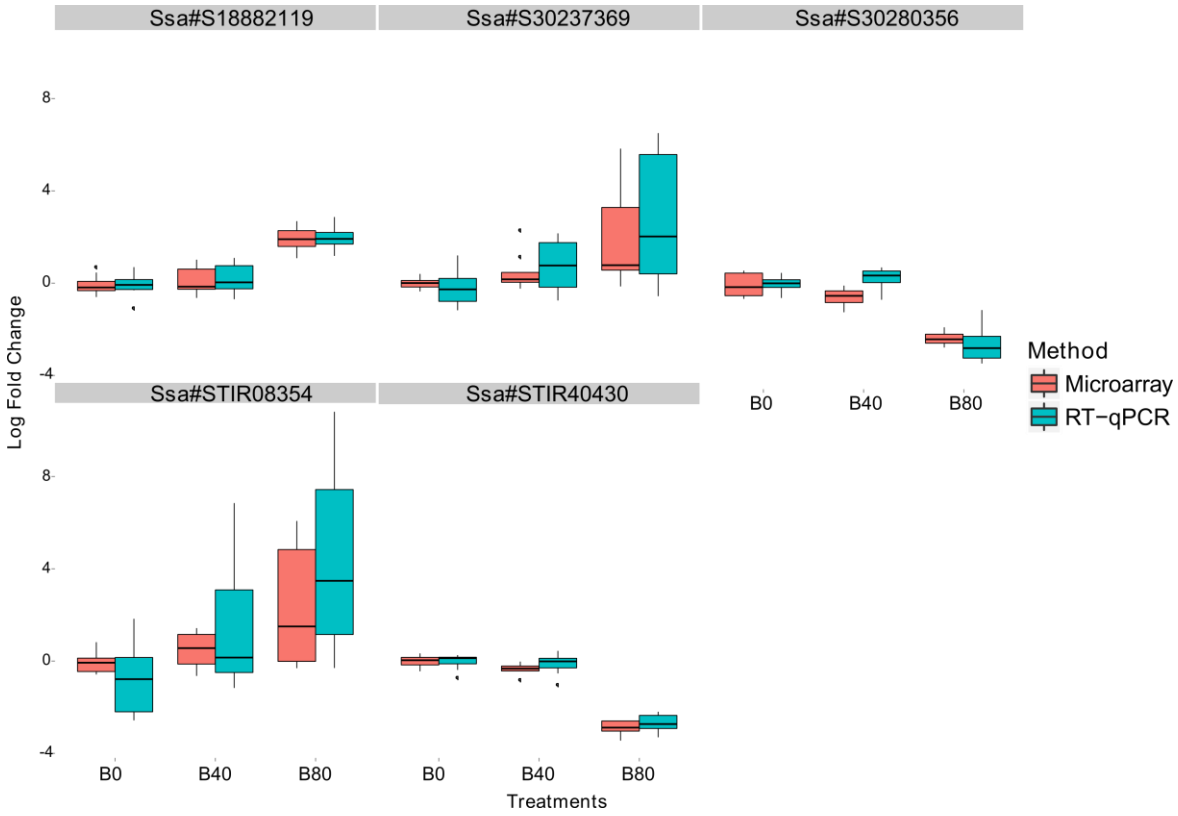
Supplementary Figure 1



Supplementary Figure 2



Supplementary Figure 3



833 **Supplementary Table 1.** RT-qPCR primers used for validation of microarray data.

Probe Name	Gene Description	Forward Primer	Reverse Primer	Efficiency
NA	Elongation Factor 1A	GCACCACGAGACCCTGGAAT	CACGTTGCCACGACGGATAT	101.4%
Ssa#S30280356	Acyl-CoA desaturase	CCCTAAACCACCGATGAGAC	CAAGAAGCACAAAGCAGTCC	101.6%
Ssa#STIR40430	elongation of very long chain fatty acids-like 6	CTGATGTTTCTTTGGCTCCC	CCCGCATTCTTCATAAGTACC	105.2%
Ssa#S18882119	Apolipoprotein F	TAGTACCGTTACAATAGTCCCTG	TCATCCCTCTTAGACCACCT	91.6%
Ssa#STIR08354	chymotrypsinogen b1	TCCCTGTCCATATTGCTATGTC	TGTCATGTTCAGTTCAGACCA	104.9%
Ssa#S30237369	glucagon I	AAGAAAGCAGAAAGCAACGG	GGCTTCCCTTCAACCTACAG	94.6%

834

835

836 **Supplementary Table 2**. Complete results of GSEA (*1d* test) of treatments compared to B0.

Gene Set	Biological Process	Class	Number of genes	Magnitude	q Value
B20 vs B0					
Ribosome	GIP	Translation	119	-1.37	0.017
Proteasome	GIP	Folding, sorting and degradation	41	-1.32	0.017
Protein digestion and absorption	Organismal Systems	Digestive system	43	1.33	0.030
B40 vs B0					
Ribosome	GIP	Translation	119	-2.57	0.000
Protein processing in endoplasmic reticulum	GIP	Folding, sorting and degradation	119	-1.96	0.000
N-Glycan biosynthesis	Metabolism	Glycan biosynthesis and metabolism	37	-1.58	0.001
Terpenoid backbone biosynthesis	Metabolism	Metabolism of terpenoids and polyketides	17	-1.56	0.001
Protein export	GIP	Folding, sorting and degradation	21	-1.49	0.001
Aminoacyl-tRNA biosynthesis	GIP	Translation	28	-1.33	0.005
Proteasome	GIP	Folding, sorting and degradation	41	-1.26	0.009
DNA replication	GIP	Replication and repair	33	-1.08	0.047
B60 vs B0					
Oxidative phosphorylation	Metabolism	Energy metabolism	106	-2.62	0.000
Aminoacyl-tRNA biosynthesis	GIP	Translation	28	-2.00	0.000
Propanoate metabolism	Metabolism	Carbohydrate metabolism	22	-1.74	0.000
Cardiac muscle contraction	Organismal Systems	Circulatory system	49	-1.67	0.000
Glyoxylate and dicarboxylate metabolism	Metabolism	Carbohydrate metabolism	20	-1.65	0.000
Biosynthesis of unsaturated fatty acids	Metabolism	Lipid metabolism	15	-1.69	0.000
Pentose phosphate pathway	Metabolism	Carbohydrate metabolism	18	-1.63	0.000
Proteasome	GIP	Folding, sorting and degradation	41	-1.57	0.000
Fatty acid elongation	Metabolism	Lipid metabolism	18	-1.55	0.000
Pyruvate metabolism	Metabolism	Carbohydrate metabolism	26	-1.55	0.000
Terpenoid backbone biosynthesis	Metabolism	Metabolism of terpenoids and polyketides	17	-1.54	0.000
Valine, leucine and isoleucine degradation	Metabolism	Amino acid metabolism	36	-1.45	0.000
Vitamin digestion and absorption	Organismal Systems	Digestive system	16	-1.36	0.001
Glycine, serine and threonine metabolism	Metabolism	Amino acid metabolism	31	-1.33	0.001
Carbon fixation pathways in prokaryotes	Metabolism	Energy metabolism	11	-1.30	0.003
Tryptophan metabolism	Metabolism	Amino acid metabolism	30	-1.16	0.007
Methane metabolism	Metabolism	Energy metabolism	14	-1.19	0.007
Citrate cycle (TCA cycle)	Metabolism	Carbohydrate metabolism	23	-1.17	0.007

Glutathione metabolism	Metabolism	Metabolism of other amino acids	23	-1.14	0.009
Ubiquitin mediated proteolysis	GIP	Folding, sorting and degradation	101	1.40	0.010
Neuroactive ligand-receptor interaction	EIP	Signaling molecules and interaction	158	1.28	0.012
SNARE interactions in vesicular transport	GIP	Folding, sorting and degradation	27	-1.06	0.015
Complement and coagulation cascades	Organismal Systems	Immune system	59	-1.04	0.019
Phenylalanine metabolism	Metabolism	Amino acid metabolism	12	-1.02	0.023
Glycerolipid metabolism	Metabolism	Lipid metabolism	26	-1.00	0.023
Carbon fixation in photosynthetic organisms	Metabolism	Energy metabolism	15	-1.01	0.023
Chemokine signaling pathway	Organismal Systems	Immune system	99	1.17	0.025
Progesterone-mediated oocyte maturation	Organismal Systems	Endocrine system	50	1.17	0.025
Natural killer cell mediated cytotoxicity	Organismal Systems	Immune system	48	1.12	0.034
Serotonergic synapse	Organismal Systems	Nervous system	58	1.08	0.035
Protein digestion and absorption	Organismal Systems	Digestive system	43	1.09	0.035
B80 vs B0					
Oxidative phosphorylation	Metabolism	Energy metabolism	106	-3.66	0.000
Ribosome biogenesis in eukaryotes	GIP	Translation	63	3.65	0.000
Proteasome	GIP	Folding, sorting and degradation	41	-2.86	0.000
Biosynthesis of unsaturated fatty acids	Metabolism	Lipid metabolism	15	-2.43	0.000
Pentose phosphate pathway	Metabolism	Carbohydrate metabolism	18	-2.22	0.000
Pyruvate metabolism	Metabolism	Carbohydrate metabolism	26	-2.18	0.000
Cardiac muscle contraction	Organismal Systems	Circulatory system	49	-2.10	0.000
Fatty acid elongation	Metabolism	Lipid metabolism	18	-2.14	0.000
Propanoate metabolism	Metabolism	Carbohydrate metabolism	22	-2.05	0.000
Citrate cycle (TCA cycle)	Metabolism	Carbohydrate metabolism	23	-1.89	0.000
Glutathione metabolism	Metabolism	Metabolism of other amino acids	23	-1.80	0.000
Carbon fixation pathways in prokaryotes	Metabolism	Energy metabolism	11	-1.84	0.000
Glyoxylate and dicarboxylate metabolism	Metabolism	Carbohydrate metabolism	20	-1.77	0.000
Glycine, serine and threonine metabolism	Metabolism	Amino acid metabolism	31	-1.66	0.000
PPAR signaling pathway	Organismal Systems	Endocrine system	44	-1.61	0.000
Protein export	GIP	Folding, sorting and degradation	21	1.80	0.000
Carbon fixation in photosynthetic organisms	Metabolism	Energy metabolism	15	-1.56	0.000
Valine, leucine and isoleucine degradation	Metabolism	Amino acid metabolism	36	-1.41	0.000
Phenylalanine metabolism	Metabolism	Amino acid metabolism	12	-1.45	0.000
RNA transport	GIP	Translation	115	1.55	0.001

Vitamin digestion and absorption	Organismal Systems	Digestive system	16	-1.40	0.001
RNA degradation	GIP	Folding, sorting and degradation	52	1.51	0.001
mRNA surveillance pathway	GIP	Translation	54	1.45	0.001
RNA polymerase	GIP	Transcription	28	1.40	0.002
Glycolysis / Gluconeogenesis	Metabolism	Carbohydrate metabolism	33	-1.27	0.002
Methane metabolism	Metabolism	Energy metabolism	14	-1.28	0.002
Protein digestion and absorption	Organismal Systems	Digestive system	43	1.35	0.004
Neuroactive ligand-receptor interaction	EIP	Signaling molecules and interaction	158	1.29	0.004
Cytosolic DNA-sensing pathway	Organismal Systems	Immune system	36	1.31	0.004
ABC transporters	EIP	Membrane transport	29	-1.18	0.005
Natural killer cell mediated cytotoxicity	Organismal Systems	Immune system	48	1.25	0.005
Pancreatic secretion	Organismal Systems	Digestive system	49	1.26	0.005
Nicotinate and nicotinamide metabolism	Metabolism	Metabolism of cofactors and vitamins	14	-1.16	0.006
Mismatch repair	GIP	Replication and repair	19	-1.14	0.007
p53 signaling pathway	Cellular Processes	Cell growth and death	45	-1.10	0.009
Aminoacyl-tRNA biosynthesis	GIP	Translation	28	-1.09	0.011
Leukocyte transendothelial migration	Organismal Systems	Immune system	65	-1.04	0.014
Mineral absorption	Organismal Systems	Digestive system	27	-1.03	0.015
Tryptophan metabolism	Metabolism	Amino acid metabolism	30	-1.02	0.015
Collecting duct acid secretion	Organismal Systems	Excretory system	15	-1.03	0.016
Cholinergic synapse	Organismal Systems	Nervous system	58	1.11	0.017
Protein processing in endoplasmic reticulum	GIP	Folding, sorting and degradation	119	1.10	0.018
Antigen processing and presentation	Organismal Systems	Immune system	32	1.08	0.019
Pyrimidine metabolism	Metabolism	Nucleotide metabolism	75	-0.95	0.025
Fc epsilon RI signaling pathway	Organismal Systems	Immune system	31	1.03	0.026
Serotonergic synapse	Organismal Systems	Nervous system	58	1.02	0.026
Steroid biosynthesis	Metabolism	Lipid metabolism	14	-0.97	0.027
Ribosome	GIP	Translation	119	1.01	0.028
DNA replication	GIP	Replication and repair	33	-0.93	0.028
Chemokine signaling pathway	Organismal Systems	Immune system	99	0.99	0.030
NOD-like receptor signaling pathway	Organismal Systems	Immune system	34	0.99	0.031
Porphyrin and chlorophyll metabolism	Metabolism	Metabolism of cofactors and vitamins	21	-0.88	0.044
Insulin signaling pathway	Organismal	Endocrine system	68	-0.86	0.045

	Systems				
SBM vs B0					
Oxidative phosphorylation	Metabolism	Energy metabolism	106	-4.33	0.000
Proteasome	GIP	Folding, sorting and degradation	41	-4.49	0.000
Ribosome biogenesis in eukaryotes	GIP	Translation	63	3.15	0.000
Glutathione metabolism	Metabolism	Metabolism of other amino acids	23	-2.27	0.000
Cardiac muscle contraction	Organismal Systems	Circulatory system	49	-1.90	0.000
Citrate cycle (TCA cycle)	Metabolism	Carbohydrate metabolism	23	-1.79	0.000
Biosynthesis of unsaturated fatty acids	Metabolism	Lipid metabolism	15	-1.79	0.000
Steroid biosynthesis	Metabolism	Lipid metabolism	14	-1.73	0.001
Carbon fixation pathways in prokaryotes	Metabolism	Energy metabolism	11	-1.70	0.001
Fatty acid elongation	Metabolism	Lipid metabolism	18	-1.58	0.002
Natural killer cell mediated cytotoxicity	Organismal Systems	Immune system	48	1.60	0.006
p53 signaling pathway	Cellular Processes	Cell growth and death	45	-1.41	0.006
Mismatch repair	GIP	Replication and repair	19	-1.41	0.006
Aminoacyl-tRNA biosynthesis	GIP	Translation	28	-1.40	0.006
RNA transport	GIP	Translation	115	1.50	0.006
Fc epsilon RI signaling pathway	Organismal Systems	Immune system	31	1.52	0.006
Pyruvate metabolism	Metabolism	Carbohydrate metabolism	26	-1.39	0.006
Protein processing in endoplasmic reticulum	GIP	Folding, sorting and degradation	119	1.46	0.008
Antigen processing and presentation	Organismal Systems	Immune system	32	1.41	0.010
mRNA surveillance pathway	GIP	Translation	54	1.39	0.010
Peroxisome	Cellular Processes	Transport and catabolism	64	1.35	0.013
Pentose phosphate pathway	Metabolism	Carbohydrate metabolism	18	-1.30	0.013
Collecting duct acid secretion	Organismal Systems	Excretory system	15	-1.28	0.016
Synaptic vesicle cycle	Organismal Systems	Nervous system	38	-1.23	0.017
DNA replication	GIP	Replication and repair	33	-1.22	0.018
Protein digestion and absorption	Organismal Systems	Digestive system	43	1.29	0.023
Alanine, aspartate and glutamate metabolism	Metabolism	Amino acid metabolism	25	1.23	0.031
Endocrine and other factor-regulated calcium reabsorption	Organismal Systems	Excretory system	24	-1.10	0.045
Pyrimidine metabolism	Metabolism	Nucleotide metabolism	75	-1.07	0.046
Fanconi anemia pathway	GIP	Replication and repair	43	-1.07	0.046
Mineral absorption	Organismal Systems	Digestive system	27	-1.06	0.049

837 Q Value is the p value adjusted for multiple corrections (Benjamini & Hochberg). Pathways are
838 ordered by decreasing significance.
839 Magnitude is the average of individual statistics and denotes how affected was the pathway.

Pathways, Biological Processes and Class are as per KEGG classification.
Number of Genes indicates the genes tested for the specific gene-set.
GIP is Genetic Information Processing
EIP is Environmental Information Processing

845 **Supplementary Table 3**. Results of GSEA (*2d* test) of treatments compared to B0.

Gene Set	Biological Process	Class	Number of genes	Magnitude	q Value
B20 vs B0					
Steroid biosynthesis	Metabolism	Lipid metabolism	14	1.54	0.005
PPAR signaling pathway	Organismal Systems	Endocrine system	44	1.36	0.010
B40 vs B0					
PPAR signaling pathway	Organismal Systems	Endocrine system	44	1.92	0.000
Steroid biosynthesis	Metabolism	Lipid metabolism	14	1.58	0.002
Amino sugar and nucleotide sugar metabolism	Metabolism	Carbohydrate metabolism	36	1.18	0.041
Complement and coagulation cascades	Organismal Systems	Immune system	59	1.13	0.046
B60 vs B0					
Steroid biosynthesis	Metabolism	Lipid metabolism	14	2.73	0.000
Complement and coagulation cascades	Organismal Systems	Immune system	59	2.43	0.000
PPAR signaling pathway	Organismal Systems	Endocrine system	44	1.99	0.000
Glutathione metabolism	Metabolism	Metabolism of other amino acids	23	1.99	0.000
Pyruvate metabolism	Metabolism	Carbohydrate metabolism	26	1.86	0.000
Methane metabolism	Metabolism	Energy metabolism	14	1.93	0.000
Glycolysis / Gluconeogenesis	Metabolism	Carbohydrate metabolism	33	1.74	0.000
Biosynthesis of unsaturated fatty acids	Metabolism	Lipid metabolism	15	1.83	0.000
Glyoxylate and dicarboxylate metabolism	Metabolism	Carbohydrate metabolism	20	1.69	0.000
Oxidative phosphorylation	Metabolism	Energy metabolism	106	1.57	0.000
Propanoate metabolism	Metabolism	Carbohydrate metabolism	22	1.54	0.000
Pentose phosphate pathway	Metabolism	Carbohydrate metabolism	18	1.53	0.000
Insulin signaling pathway	Organismal Systems	Endocrine system	68	1.47	0.000
Fat digestion and absorption	Organismal Systems	Digestive system	19	1.52	0.000
Glycine, serine and threonine metabolism	Metabolism	Amino acid metabolism	31	1.42	0.001
Citrate cycle (TCA cycle)	Metabolism	Carbohydrate metabolism	23	1.34	0.002
Carbon fixation pathways in prokaryotes	Metabolism	Energy metabolism	11	1.35	0.002
Metabolism of xenobiotics by cytochrome P450	Metabolism	Xenobiotics biodegradation and metabolism	12	1.35	0.002
Glycerolipid metabolism	Metabolism	Lipid metabolism	26	1.26	0.003
One carbon pool by folate	Metabolism	Metabolism of cofactors and vitamins	13	1.27	0.003
Fatty acid elongation	Metabolism	Lipid metabolism	18	1.24	0.004
Terpenoid backbone biosynthesis	Metabolism	Metabolism of terpenoids and polyketides	17	1.22	0.004
Starch and sucrose metabolism	Metabolism	Carbohydrate metabolism	23	1.19	0.005
Arginine and proline metabolism	Metabolism	Amino acid metabolism	39	1.14	0.006
Aminoacyl-tRNA biosynthesis	GIP	Translation	28	1.15	0.007
Carbon fixation in photosynthetic	Metabolism	Energy metabolism	15	1.12	0.009

organisms					
Steroid hormone biosynthesis	Metabolism	Lipid metabolism	24	1.08	0.010
Amino sugar and nucleotide sugar metabolism	Metabolism	Carbohydrate metabolism	36	1.07	0.010
Cardiac muscle contraction	Organismal Systems	Circulatory system	49	1.05	0.011
Vitamin digestion and absorption	Organismal Systems	Digestive system	16	1.06	0.011
Pentose and glucuronate interconversions	Metabolism	Carbohydrate metabolism	14	0.93	0.033
Phagosome	Cellular Processes	Transport and catabolism	79	0.88	0.038
Carbohydrate digestion and absorption	Organismal Systems	Digestive system	16	0.87	0.045
B80 vs B0					
Steroid biosynthesis	Metabolism	Lipid metabolism	14	3.16	0.000
Glutathione metabolism	Metabolism	Metabolism of other amino acids	23	2.98	0.000
Methane metabolism	Metabolism	Energy metabolism	14	2.34	0.000
Biosynthesis of unsaturated fatty acids	Metabolism	Lipid metabolism	15	2.39	0.000
Fat digestion and absorption	Organismal Systems	Digestive system	19	2.29	0.000
Oxidative phosphorylation	Metabolism	Energy metabolism	106	2.09	0.000
Pyruvate metabolism	Metabolism	Carbohydrate metabolism	26	2.15	0.000
PPAR signaling pathway	Organismal Systems	Endocrine system	44	2.09	0.000
Glycolysis / Gluconeogenesis	Metabolism	Carbohydrate metabolism	33	2.00	0.000
Protein processing in endoplasmic reticulum	GIP	Folding, sorting and degradation	119	1.94	0.000
Glycine, serine and threonine metabolism	Metabolism	Amino acid metabolism	31	1.93	0.000
Glycerolipid metabolism	Metabolism	Lipid metabolism	26	1.94	0.000
Pentose phosphate pathway	Metabolism	Carbohydrate metabolism	18	1.91	0.000
Pyrimidine metabolism	Metabolism	Nucleotide metabolism	75	1.79	0.000
Propanoate metabolism	Metabolism	Carbohydrate metabolism	22	1.77	0.000
Fatty acid elongation	Metabolism	Lipid metabolism	18	1.76	0.000
Citrate cycle (TCA cycle)	Metabolism	Carbohydrate metabolism	23	1.64	0.000
Carbon fixation pathways in prokaryotes	Metabolism	Energy metabolism	11	1.70	0.000
One carbon pool by folate	Metabolism	Metabolism of cofactors and vitamins	13	1.63	0.000
Carbon fixation in photosynthetic organisms	Metabolism	Energy metabolism	15	1.62	0.000
Complement and coagulation cascades	Organismal Systems	Immune system	59	1.53	0.000
Ribosome biogenesis in eukaryotes	GIP	Translation	63	1.46	0.000
Insulin signaling pathway	Organismal Systems	Endocrine system	68	1.44	0.000
Alanine, aspartate and glutamate metabolism	Metabolism	Amino acid metabolism	25	1.45	0.000
Glyoxylate and dicarboxylate metabolism	Metabolism	Carbohydrate metabolism	20	1.44	0.000
Protein export	GIP	Folding, sorting and degradation	21	1.45	0.000
Carbohydrate digestion and absorption	Organismal Systems	Digestive system	16	1.39	0.001

Pancreatic secretion	Organismal Systems	Digestive system	49	1.33	0.001
Cysteine and methionine metabolism	Metabolism	Amino acid metabolism	23	1.30	0.001
Metabolism of xenobiotics by cytochrome P450	Metabolism	Xenobiotics biodegradation and metabolism	12	1.29	0.002
Amino sugar and nucleotide sugar metabolism	Metabolism	Carbohydrate metabolism	36	1.22	0.002
p53 signaling pathway	Cellular Processes	Cell growth and death	45	1.19	0.003
beta-Alanine metabolism	Metabolism	Metabolism of other amino acids	21	1.16	0.004
Cardiac muscle contraction	Organismal Systems	Circulatory system	49	1.13	0.004
Arginine and proline metabolism	Metabolism	Amino acid metabolism	39	1.13	0.005
Aminoacyl-tRNA biosynthesis	GIP	Translation	28	1.12	0.005
Starch and sucrose metabolism	Metabolism	Carbohydrate metabolism	23	1.10	0.006
Tyrosine metabolism	Metabolism	Amino acid metabolism	24	1.07	0.008
Arachidonic acid metabolism	Metabolism	Lipid metabolism	24	1.03	0.010
Vitamin digestion and absorption	Organismal Systems	Digestive system	16	1.03	0.011
Galactose metabolism	Metabolism	Carbohydrate metabolism	18	1.02	0.012
Proteasome	GIP	Folding, sorting and degradation	41	1.02	0.012
Phagosome	Cellular Processes	Transport and catabolism	79	0.96	0.016
Terpenoid backbone biosynthesis	Metabolism	Metabolism of terpenoids and polyketides	17	0.97	0.017
Porphyrin and chlorophyll metabolism	Metabolism	Metabolism of cofactors and vitamins	21	0.91	0.027
Estrogen signaling pathway	Organismal Systems	Endocrine system	51	0.89	0.029
Pentose and glucuronate interconversions	Metabolism	Carbohydrate metabolism	14	0.90	0.029
Antigen processing and presentation	Organismal Systems	Immune system	32	0.88	0.030
Primary bile acid biosynthesis	Metabolism	Lipid metabolism	13	0.87	0.035
Lysosome	Cellular Processes	Transport and catabolism	89	0.80	0.048
SBM vs B0					
Steroid biosynthesis	Metabolism	Lipid metabolism	14	2.82	0.000
Glutathione metabolism	Metabolism	Metabolism of other amino acids	23	2.59	0.000
Proteasome	GIP	Folding, sorting and degradation	41	2.37	0.000
Pyrimidine metabolism	Metabolism	Nucleotide metabolism	75	2.16	0.000
Protein processing in endoplasmic reticulum	GIP	Folding, sorting and degradation	119	2.12	0.000
Oxidative phosphorylation	Metabolism	Energy metabolism	106	1.98	0.000
PPAR signaling pathway	Organismal Systems	Endocrine system	44	1.90	0.000
Fat digestion and absorption	Organismal Systems	Digestive system	19	1.79	0.000
Methane metabolism	Metabolism	Energy metabolism	14	1.77	0.000
Biosynthesis of unsaturated fatty acids	Metabolism	Lipid metabolism	15	1.79	0.000
Glycolysis / Gluconeogenesis	Metabolism	Carbohydrate metabolism	33	1.69	0.000

Pentose phosphate pathway	Metabolism	Carbohydrate metabolism	18	1.66	0.001
Pyruvate metabolism	Metabolism	Carbohydrate metabolism	26	1.59	0.001
Ribosome biogenesis in eukaryotes	GIP	Translation	63	1.54	0.001
DNA replication	GIP	Replication and repair	33	1.51	0.002
Insulin signaling pathway	Organismal Systems	Endocrine system	68	1.43	0.003
Cysteine and methionine metabolism	Metabolism	Amino acid metabolism	23	1.42	0.004
Glycerolipid metabolism	Metabolism	Lipid metabolism	26	1.41	0.004
Fatty acid elongation	Metabolism	Lipid metabolism	18	1.42	0.005
Protein export	GIP	Folding, sorting and degradation	21	1.38	0.005
p53 signaling pathway	Cellular Processes	Cell growth and death	45	1.33	0.006
Propanoate metabolism	Metabolism	Carbohydrate metabolism	22	1.33	0.007
Purine metabolism	Metabolism	Nucleotide metabolism	113	1.27	0.008
Alanine, aspartate and glutamate metabolism	Metabolism	Amino acid metabolism	25	1.29	0.008
Amino sugar and nucleotide sugar metabolism	Metabolism	Carbohydrate metabolism	36	1.28	0.008
Terpenoid backbone biosynthesis	Metabolism	Metabolism of terpenoids and polyketides	17	1.25	0.013
Aminoacyl-tRNA biosynthesis	GIP	Translation	28	1.23	0.013
Glycine, serine and threonine metabolism	Metabolism	Amino acid metabolism	31	1.17	0.017
Metabolism of xenobiotics by cytochrome P450	Metabolism	Xenobiotics biodegradation and metabolism	12	1.21	0.017
Carbon fixation in photosynthetic organisms	Metabolism	Energy metabolism	15	1.14	0.022
Carbon fixation pathways in prokaryotes	Metabolism	Energy metabolism	11	1.16	0.022
Folate biosynthesis	Metabolism	Metabolism of cofactors and vitamins	10	1.16	0.022
Arginine and proline metabolism	Metabolism	Amino acid metabolism	39	1.11	0.023
Primary bile acid biosynthesis	Metabolism	Lipid metabolism	13	1.13	0.024
One carbon pool by folate	Metabolism	Metabolism of cofactors and vitamins	13	1.12	0.024
Starch and sucrose metabolism	Metabolism	Carbohydrate metabolism	23	1.07	0.029
Various types of N-glycan biosynthesis	Metabolism	Glycan biosynthesis and metabolism	24	1.01	0.043
Citrate cycle (TCA cycle)	Metabolism	Carbohydrate metabolism	23	0.98	0.050

Q Value is the p value adjusted for multiple corrections (Benjamini & Hochberg). Pathways are ordered by decreasing significance.
Magnitude is the average of individual statistics and denotes how affected was the pathway.
Pathways, Biological Processes and Class are as per KEGG classification.
Number of Genes indicates the genes tested for the specific gene-set.
GIP is Genetic Information Processing