frontiers in PHYSIOLOGY

Aquatic Physiology

Impact of a novel protein meal on the gastrointestinal microbiota and host transcriptome of larval zebrafish Danio rerio

Eugene Rurangwa, Detmer Sipkema, Jeroen Kals, Menno ter_Veld, Maria Forlenza, Gianina M. Bacanu, Hauke Smidt and Arjan P. Palstra

Journal Name:	Frontiers in Physiology
ISSN:	1664-042X
Article type:	Original Research Article
Received on:	05 Dec 2014
Accepted on:	14 Apr 2015
Provisional PDF published on:	14 Apr 2015
Frontiers website link:	www.frontiersin.org
Citation:	Rurangwa E, Sipkema D, Kals J, Ter_veld M, Forlenza M, Bacanu GM, Smidt H and Palstra AP(2015) Impact of a novel protein meal on the gastrointestinal microbiota and host transcriptome of larval zebrafish Danio rerio. *Front. Physiol.* 6:133. doi:10.3389/fphys.2015.00133

Impact of a novel protein meal on the gastrointestinal microbiota and the host transcriptome of larval zebrafish *Danio rerio*

Eugene Rurangwa[1#], Detmer Sipkema[2#], Jeroen Kals[1], Menno ter Veld[3], Maria Forlenza[4], Gianina M. Bacanu[2], Hauke Smidt[2], Arjan P. Palstra[1]*

[1] Institute for Marine Resources and Ecosystem Studies (IMARES), Wageningen University and Research Centre, Yerseke, the Netherlands
[2] Laboratory of Microbiology, Wageningen University, Wageningen, the Netherlands
[3] Aquaculture and Fisheries Group, Wageningen University, Wageningen, the Netherlands
[4] Cell Biology and Immunology Group, Wageningen University, Wageningen, the Netherlands

* Correspondence:
Dr. Arjan P. Palstra
The Institute for Marine Resources and Ecosystem Studies (IMARES)
Wageningen University and Research Centre
Korringaweg 5
4401 NT Yerseke
The Netherlands
arjan.palstra@wur.nl

[#] These authors contributed equally.

Running head: Gastrointestinal microbiota and host transcriptome of larval zebrafish

Abstract

Larval zebrafish was subjected to a methodological exploration of the gastrointestinal microbiota and transcriptome. Assessed was the impact of two dietary inclusion levels of a novel protein meal (NPM) of animal origin (ragworm *Nereis virens*) on the gastrointestinal tract (GIT). Microbial development was assessed over the first 21 days post egg fertilisation (dpf) through 16S rRNA gene-based microbial composition profiling by pyrosequencing. Differentially expressed genes in the GIT were demonstrated at 21 dpf by whole transcriptome sequencing (mRNAseq). Larval zebrafish showed rapid temporal changes in microbial colonization but domination occurred by one to three bacterial species generally belonging to *Proteobacteria* and *Firmicutes*. The high iron content of NPM may have led to an increased relative abundance of bacteria that were related to potential pathogens and bacteria with an increased iron metabolism. Functional classification of the 328 differentially expressed genes indicated that the GIT of larvae fed at higher NPM level was more active in *transmembrane ion transport* and *protein synthesis*. mRNAseq analysis did not reveal a major activation of genes involved in the immune response or indicating differences in iron uptake and homeostasis in zebrafish fed at the high inclusion level of NPM.

Keywords: zebrafish nutrition, 16S rRNA-based microbial composition, pyrosequencing, mRNA sequencing, gastrointestinal tract transcriptome, iron metabolism, aquaculture

1. Introduction

The diet has profound effects on the microbial composition and on the nutrient uptake by the enterocytes in the GIT. Moreover, the diet has effects on the interactions between host and

microbes, aspects of which can be very specific (Rawls et al., 2004) but which is also surprisingly conserved across all vertebrates (Rawls et al., 2006). During long-lasting interactions, coevolution between hosts and microbes has resulted in a microbial ecosystem that is monitored and controlled by the host while the microbiota influence their host to maintain a stable niche for its continued presence (Neish, 2009).

Commensal microbial communities play an important role in the host's GIT development, nutrition and protection against pathogens (Verschuere et al., 2000; Bates et al., 2006; Nayak, 2010; Ringø et al., 2010). In the absence of microbes (germ-free fish), specific aspects of GIT differentiation and functions are arrested or altered (Bates et al., 2006). GIT microbiota are involved in the host's feed digestion and physiological processes by producing vitamins, digestive enzymes, amino acids, essential growth factors and metabolites (Nayak, 2010). They affect a wide range of biological processes including nutrient processing and absorption, regulation of intestinal glycan expression, development of the mucosal immune system and fortification of the innate immune defences, angiogenesis, and epithelial renewal (reviewed by Kanther and Rawls, 2010, and Rawls et al., 2004).

Zebrafish offers interesting features as model organism to study the nutritional impact of alternative protein sources on the GIT functions, development of the microbial community and host–microbe interactions by combining several molecular based approaches. Key features of the zebrafish model include a characterized genome, a wide variety of molecular and bioinformatic tools and a well-characterized rapid embryonic development (Westerfield, 1993; Kimmel et al., 1995; Ulloa et al., 2011). With these advantages related to the use of zebrafish as an experimental fish model, nutritional research in aquaculture can be conducted at reduced cost, time and space needed in research facilities (Gomez-Requeni et al., 2010; Ribas and Piferrer, 2013). Zebrafish thereby offers an opportunity to gain mechanistic insights but, as any model, does not replace the commercial species of interest that has its own GIT characteristics.

When kept at 28 degrees, zebrafish larvae hatch from their chorions within 3 days post-fertilization (dpf), and the mouth opens around 74 hours post-fertilization (hpf). The GIT is colonized by microbiota from the environment after hatching within 12–24 h, concurrent with digestive tract differentiation (Hansen and Olafsen, 1999; Bates et al., 2006; Rawls et al., 2007; Kanther and Rawls, 2007; Nayak, 2010). By 4 dpf, within a day after mouth opening, the digestive tract is colonized by a small number of bacteria and their number increases after swallowing has started (Bates et al., 2006). At 5 dpf, the GIT is fully functional from a nutritional point of view when lipid and protein macromolecule uptake is apparent (Farber et al., 2001; Wallace et al., 2005), a regular pattern of spontaneous movement is visible and exogenous feeding commences (Holmberg et al., 2004). The yolk is largely absorbed and GIT morphogenesis has proceeded to a stage that supports feeding and digestion (Farber et al., 2001; Rawls et al., 2004).

This study aimed to determine the impact of a novel protein meal (NPM) on the gastrointestinal microbiota and the host transcriptome of larval zebrafish. The NPM that was tested is of animal origin (ragworm *Nereis virens*) and has been demonstrated to be beneficial for hematocrit levels and general physiological performance as suggested by improved growth in common sole *Solea solea* when compared to fish fed with commercial pelleted feeds (Kals, 2014). Therewith it potentially has considerable importance for aquaculture nutrition. The present study was undertaken to assess the impact of two dietary inclusion levels of the NPM on the GIT of the developing zebrafish, specifically on 1) microbial development over the first 21 dpf, and 2) molecular differentiation in physiological processes in the host by differentially expressed genes at 21 dpf as determined by mRNAseq. These investigations are performed with an unbiased approach.

Pathogens residing in the GIT are known to be stimulated in their pathogenic potential by increased iron availability (Kortman et al., 2012). As feed ingredients of animal origin are

expected to be rich iron sources, a bias was introduced to particularly assess (1) changes in the abundance of potential pathogens and consequences for the expression of GIT genes involved in immune response, and (2) changes in the abundance of bacteria with increased iron metabolism and consequences for the expression of GIT genes involved in iron uptake and homeostasis. It is hypothesized that increasing the dietary inclusion level of the NPM will lead to (1) a higher abundance of potential pathogens and bacteria with an increased iron metabolism, and (2) differential expression of genes indicating an activation of the immune response, lower iron uptake and increased attention for maintaining homeostasis.

2. Methods

2.1 Zebrafish husbandry

Zebrafish (*Danio rerio* Hamilton 1822) embryos were obtained from breeders of the Zod2F7 strain. The ancestral diet consisted of live nauplii of *Artemia* (brine shrimp) and commercial flake diet for ornamental fish (Tetra). At 5 dpf, 1,800 larvae from a single batch were randomly distributed over 6 experimental aquaria (each 6 L with 50 individuals per liter) installed in a thermo-regulated water bath and with individual inflow water connected to a flow-through system. Larvae were reared at a photoperiod of 14/10 h light/dark and under optimal water quality conditions (Temperature 25.9 ± 0.3 ^{o}C; pH 8.1 ± 0.1; dissolved oxygen 7.6 ± 0.3 mg/L; ammonium, nitrogen and nitrate null; nitrite concentration 0.00-0.05 mg/L). Larvae were fed *Paramecium* (diet p) at 4 and 5 dpf, then gradually weaned to experimental diets between 6 and 9 dpf (diets pB and pE; Fig. 1) and, from 10 dpf onwards, fed with experimental diets (diets B and E) until 21 dpf. During the transitional feeding period from live to inert dry feed, larvae were fed a daily ration of paramecium decreasing with 20% per day. Before live prey had disappeared in the tanks, feeding was completed with an increasing amount of experimental dry feeds.

2.2 Experimental diets

The experimental diets were isonitrogenous, isoenergetic, equal in amino acids composition, calcium and phosphates, but differed in concentration of the NPM ragworm (*Nereis virens*) meal (Seabait Ltd, Woodhorn Village, UK). Diet B (10% NPM) and E (75% NPM) were fed by hand till satiation 4-7 times per day. The dry micro-particulate lyophilised diets were prepared through cold extrusion (Research Diet Services, Wijk bij Duurstede, the Netherlands). 200 µm micro-particulates were fed between 6 and 10 dpf and 200-300 µm between 11 and 20 dpf for both experimental diets. The crude composition (Table 1) was analysed at Nutrilab bv (Giessen, the Netherlands) and the iron content was analysed at the Chemical Biological Soil Laboratory (Wageningen, the Netherlands) using Inductively Coupled Plasma Atomic Emission Spectroscopy.

2.3 Larval sample collection and storage

Larvae were sacrificed by an overdose of the anaesthetic 1.0% tricaine methane sulfonate buffered with 1.5% NaHCO3. Triplicate pools of ten larvae per diet were collected in sterile condition at 5, 7, 14 and 21 dpf. External surfaces of larvae were disinfected by rinsing with 70% ethanol for 2 min and then several times with sterile filtered (0.2 micron) Milli-Q water. In addition, at 21 dpf, extra triplicate pools of ten larvae per diet were collected for GIT sampling. Larvae were anaesthetized, disinfected, and dissected on ice in sterile conditions using flamed instruments between two different samples. Whole larvae were kept frozen in sterile Eppendorff tubes at -20°C directly upon sampling and were then stored at -80°C, while extra GIT samples

from day 21 were stored in RNAlater (Applied Biosystems, Nieuwerkerk a/d lJssel, the Netherlands) at -20°C.

2.4 Microbiological analyses: DNA isolation

Microbial DNA was isolated using the protocol described by Roeselers et al. (2011) with some modifications: Ten larvae were combined in 2.0 ml screw-cap tubes containing 0.1mm Zirconia/silica beads and 2.5 mm Glass beads (Biospec Products). 800 µl 120mM Na-phosphate buffer (pH 8.0) and 400 µl of lysis solution containing 10% sodium dodecyl sulfate, 0.5M Tris-HCl (pH 8.0) and 0.1M NaCl was added homogenisation in a Mini-beadbeater (Biospec Products) for 6 min at 5500 rpm. The supernatants were transferred to new tubes and lysozyme was added to a final concentration of 10 mg/ml followed by incubation at 42°C for 30 min. Ammonium acetate (7.5M) was added to the supernatant (2:5 v/v) and samples were incubated at -20°C for 5 min. Samples were centrifuged for 5 min at 12,000 g and the supernatants were transferred to new tubes. DNA was precipitated at room-temperature with isopropyl alcohol (500 µl) and pelleted by centrifugation at 12,000 rpm for 30 min at 4°C. Pellets were washed with -20°C 70% ethanol and air-dried for 45 min before resuspension in 50 µl nuclease free water.

2.5 Microbiological analyses: RT-PCR

RT-PCR targeting 16S rRNA was performed with the primers 27F and 1492R (Lane, 1991) for total RNA extracted from two pools of GIT for both diet B and diet E (for RNA extraction procedure see section "*mRNAseq: Total RNA isolation*"). The RT reaction (20 µl) contained 50 mM Tris-HCl, 75 mM KCl, 3 mM $MgCl_2$, 5 µM DTT, 0.5 mM of each dNTP, 2 pmol of primers 27F and 1492R, 200 U of of SuperScript™ Reverse Transcriptase (Invitrogen), 40 U of RNasin® Plus RNase Inhibitor (Promega) and 1 µg of RNA extracted from 10 pooled larvae or GITs for each of the two diets. Reactions were incubated at 55 °C for 60 minutes, followed by 15 minutes at 70 °C to denature the reverse-transcriptase.

2.6 Microbiological analyses: PCR and Sequencing

For 16S rRNA gene-based microbial composition profiling, barcoded amplicons from the V1-V2 region of 16S rRNA genes were generated from all DNA and reverse transcribed RNA samples by PCR using the 27F-DegS primer that was appended with the titanium sequencing adaptor A and an 8 nt sample-specific barcode at the 5' end, and an equimolar mix of two reverse primers (338R I and II), that carried the titanium adaptor B at the 5' end [58].

PCRs were performed using a thermocycler GS0001 (Gene Technologies, Braintree, U.K.) in a total volume of 100 µl containing 1× HF buffer (Finnzymes, Vantaa, Finland), 2 µl 10 mM (each nucleotide) PCR Grade Nucleotide Mix (Roche, Diagnostics GmbH, Mannheim, Germany), 2 units of Phusion® Hot Start II High-Fidelity DNA polymerase, 500 nM of a forward and reverse primer mix (Biolegio BV, Nijmegen, the Netherlands), and 0.2-0.4 ng/µl of template DNA (or cDNA). The amplification program consisted of an initial denaturation at 98°C for 30 s, 35 cycles of: denaturation at 98°C for 10 s, annealing at 56°C for 20 s and elongation at 72°C for 20 s, and a final extension at 72°C for 10 min. PCR products were purified with the High Pure Cleanup Micro Kit (Roche) using 10 µl nuclease-free water for elution, and quantified using a NanoDrop ND-1000 spectrophotometer. Purified PCR products were mixed in equimolar amounts and run on an agarose gel, followed by excision and purification by the DNA gel extraction kit (Millipore, Billerica, MA, USA). Purified amplicon pools were pyrosequenced using a Genome Sequencer FLX in combination with titanium chemistry (GATC-Biotech, Konstanz, Germany). Pyrosequencing data were deposited at the European Bioinformatics Institute in the sequence read archive under study accession number

PRJEB4784 and sample accession numbers ERS362581 – ERS362592 and ERS362595 - ERS362598.

2.7 Microbiological analyses: Sequence analysis

Pyrosequencing data were analysed using the QIIME 1.5.0 pipeline (Caporaso et al., 2010). Low quality sequences were removed using default parameters. Operational taxonomic units (OTUs) were identified at the 97% identity level. Representative sequences from the OTUs were aligned using PyNAST (DeSantis et al., 2006). The taxonomic affiliation of each OTU was determined using the RDP Classifier at a confidence threshold of 80% against the 12_10 Greengenes core set (Wang et al., 2007). Possible chimeric OTUs were identified using QIIME's ChimeraSlayer and removed from the initially generated OTU list, producing a final set of non-chimeric OTUs.

2.8 Microbiological analyses: Statistical analysis

OTU singletons and OTUs related to chloroplasts were removed prior to analyses. The relationship between microbial community composition, diet and time was analyzed by canonical correspondence analysis (CCA) using CANOCO 5 (Ter Braak, C.J.F., Šmilauer, P. *Canoco Reference Manual And User's Guide: Software For Ordination, Version 5.0*. Ithaca: Microcomputer Power, USA. pp. 496 (2012)). Rare OTUs were down-weighted using the default option. The different diets (p, pB, pE, B and E) were tested for significant contribution to the explanation of the variation in the OTU distribution with the Monte Carlo permutation test associated with the forward selection subroutine. The OTUs that contributed most to different microbial profiles between diet B and E were calculated using SIMPER in the software package PRIMER 6 v6.1.9 (PRIMER-E Ltd, Plymouth, UK) using normalised OTU tables (square root) of day 14 and day 21.

2.9 mRNAseq: Total RNA isolation

GITs of two pools of ten larvae per diet which had been stored in RNAlater were lysed in QIAzol Lysis Reagent, a Qiagen TissueRuptor was used to cut up the tissue samples and RNA was extracted using the Qiagen miRNeasy Mini Kit according to the manufacturer's description (Qiagen Benelux BV, Venlo, the Netherlands). RNA was eluted in 50 µl and quantified by Nanodrop (Thermo Fisher Scientific, Amsterdam, the Netherlands). Integrity of the RNA was confirmed using an Agilent bioanalyzer2100.

2.10 mRNAseq: Library preparation and sequencing

For each sample a RNA-seq library was prepared with the Illumina Truseq mRNASeq Sample Preparation Kit according to the manufacturer's description (Illumina, San Diego CA, USA). Each library was sequenced twice in a paired-end sequencing run with a read length of 50 nucleotides on a Illumina HiSeq2000 with version 2 sequencing chemistry. For each library approximately 20 to 30 million read pairs were obtained.

2.11 mRNAseq: Data analysis

Raw reads were quality trimmed using the quality_trim module in the CLCBio assembly cell version 4.01. Reads were mapped to the annotated cDNA's in the ZV9 zebrafish genome assembly using the ref_assembly_short module in the CLCBio assembly cell version 4.01. The data were converted to a table using the assembly_table module in the CLCBio assembly cell version 4.01. A custom perl script was used to convert this table to a tab separated value table.

This table was used in R package DESeq v1.0.6 to analyse expression in the different samples (Anders and Huber, 2010). Raw RNA-seq data (reads) have been submitted to the NCBI project data archive under Bioproject number 229446 (Biosample numbers SRS506058 for B1, SRS506087 for B2, SRS506089 for E1, SRS506092 for E2).
Gene expression of differentially expressed genes at $P<0.05$, both up- or down-regulated, was functionally characterised and classified using DAVID 6.7 (The Database for Annotation, Visualization and Integrated Discovery, Huang et al., 2009ab).

2.12 Ethics

All experiments were performed in accordance with relevant guidelines and regulations. Protocols used complied with the current laws of the Netherlands and were approved by the Animal Experimental Committee (DEC) of the Wageningen UR in Lelystad (The Netherlands) under number 2011102.

3. Results

3.1 Impact of diets on GIT microbial community

The development of the microbial community was assessed over the first 21 days post egg fertilisation (dpf). At a confidence threshold of 80%, 97,675 out of 97,894 qualified non-chimeric reads could be assigned to a known phylum using the RDP classifier. Qualified non-chimeric read numbers ranged from 1,740 to 10,685 reads per sample (average: 6,118 reads per sample), and the rarefaction curves showed that samples were sufficiently deep sequenced to discuss similarity and differences for the more abundant OTUs (supplementary Fig. S1).
All 16S rRNA sequences found in zebrafish 5 dpf were classified as either *Proteobacteria* (48%) or *Firmicutes* (52%). The bacterial composition changed at 7 dpf with the appearance of *Actinobacteria* for both diets tested (paramecium plus diet B: pB, and paramecium plus diet E: pE) and *Bacteroidetes* for zebrafish fed with pE (Fig. 1). This was followed by a further increase of diversity at 14 dpf for both diets (B and E without paramecium). The trend of diversifying microbiota was halted at 21 dpf due to the increased dominance of *Proteobacteria*, especially on diet B where they made up for 99% of all reads. Although for day 21, we only had one zebrafish larvae-derived sample per diet, 16S rRNA gene analysis based on RNA-extracted from 4 additional samples (2 for diet B and 2 for diet E) from 21 dpf confirmed the relative abundance of 99% at this day.

The bacterial colonisation of zebrafish and the impact of diet on the colonisation were analysed at the approximate species level (97% identity based on rRNA gene sequence) by canonical correspondence analysis (CCA). The first two CCA axes had eigenvalues of 0.89 and 0.80, respectively and explained 27% of the variation in species data and 72% of the variation in the microbial taxa-time interactions (Fig. 2). The microbiota changed dramatically during the first 21 dpf and the high impact of time ($P=0.002$) masked treatment (diet) effects. Therefore, the impact of diet on zebrafish-associated microbiota was based on the zebrafish gut-derived RNA samples on day 21 for diet B and diet E for which replicate samples were available. These RNA-based samples were similar to the DNA-based samples for 21 dpf with respect to their microbial profiles at the OTU-level (Fig. S2), which indicates that the approach used to obtain 16S rRNA gene amplicons (whole fish + DNA extraction vs. fish gut + RNA extraction and reverse transcription) did not have a major impact on the data obtained. The dominant early bacterial colonisers of zebrafish included members of the *Clostridia* (*Firmicutes*) and *Procabacteriaceae*, *Trabulsiella*, and *Xanthomonadaceae* (all *Proteobacteria*) (Fig. 3). These OTUs were mostly absent at day 7 and were replaced by OTUs most closely related to *Propionibacterium acnes*

(*Actinobacteria*), *Rhodanobacter* and *Rhizobium* (both *Proteobacteria*). *Propionibacterium acnes* was still highly abundant at day 14 as was *Rhodanobacter* albeit at a lower percentage. Populations within the *Rhodobacteraceae*, and *Methylobacteraceae* (both *Proteobacteria*) were newly appearing dominant OTUs at day 14. Although OTUs most closely related to *Propionibacterium acnes* and *Rhodobacteraceae* were still found at low levels at day 21, those falling within *Rhodanobacter* and *Methylobacteraceae* had disappeared. Instead, populations within the *Comamonadaceae*, *Aeromonadaceae*, *Acidovorax* and *Pseudomonas* (all *Proteobacteria*) dominated the zebrafish-associated microbiota.

The OTUs that contributed most (1% or more) to the differences in microbial profiles in zebrafish fed with diet B or diet E at day 21 were identified by SIMPER (Table 2). OTUs with a higher relative abundance for diet E that mainly contributed to the differences in microbial profiles between the diets were OTU45 (*Plesiomonas shigelloides*), OTU286 (*Acidovorax* sp.), OTU1407 (family *Aeromonadaceae*), OTU 441 (*Trabulsiella* sp.), OTU552 (order *Legionellales*), OTU1459 (*Novospirillum itersonii*), OTU1414 (*Rheinheimera* sp.), OTU832 (*Propionibacterium acnes*), OTU463 (family *Alcaligenaceae*), OTU16 (*Burkholderia* sp.) and OTU 1182 (*Halomonas* sp.). Other OTUs that contributed less to the difference observed between diet E and B, but which were completely absent in zebrafish fed with diet B were OTU1278 (*Achromobacter* sp.), OTU568 (*Janibacter* sp.) and OTU1260 (*Cupriavidus* sp.). The OTUs with a higher relative abundance with diet B were *Pseudomonas* spp. (OTU1214 and OTU1306), which were among the most dominant OTUs found in zebrafish at day 21 and had an even higher relative abundance for diet B than for diet E.

3.2 mRNAseq gene expression analyses

In order to assess potential differences in how the zebrafish host responds to the different dietary treatments, intestinal tissue-associated gene expression was measured at 21 dpf. Reads were mapped to the 27,882 annotated cDNA's of the ZV9 zebrafish genome assembly that were used for further analysis. In total, 328 genes were differentially expressed corresponding to 1.18% of the total number of genes (Table S1). Of these, expression of 214 genes was up-regulated and expression of 114 genes was down-regulated in larvae fed with diet E vs B. Among these, expression of 16 differentially expressed genes were detected exclusively for larvae fed with diet E and were below detection thresholds for larvae of diet B (fold change - fc "inf"), whereas for 10 other differentially expressed genes the opposite was true (fc "0").

Of the total number of 27,882 genes that were analysed, 19,990 genes could be converted to a DAVID id, and of the total number of 328 differentially regulated genes, 264 genes could be converted as such. Genes with a DAVID id were used for unbiased functional annotation.

Functional annotation clustering revealed that the differentially expressed genes represented 27 annotation clusters. 24 terms representing 7 of these clusters were significantly enriched at $P<0.05$ (Table S2). 52 terms were not clustered. The functional annotation clusters were associated with *ribosome components and activity* (enrichment score - es 14.61) and *transport* (es 1.4) as the dominant clusters. Other clusters involved *other glycan degradation* (es 1.25); *glycosaminoglycan metabolic process* (es 1.13); *extracellular matrix structural constituent* (es 0.66); *keratin type-1* (es 0.61) and *hydrogen ion transmembrane transporter activity* (es 0.60). The functional annotation chart revealed 65 records of which 40 were significantly enriched (Table S3). In addition to many terms that were associated with *ribosome*, *transport* and the other mentioned clusters, enriched records were associated with: *interferon-induced 6-16*, *interferon binding* and *interferon receptor activity*; *glycoside hydrolase* and *oxidative phosphorylation*.

Gene functional classification showed the presence of 7 gene groups as determined by 73 differentially expressed genes (Table 3; S4) agreeing with the existence of 7 enriched functionally annotated clusters. Also here the dominant gene groups represented *ribosome components and activity* and *transport*. The gene group *ribosome components and activity* consisted of 27 genes that were all significantly up-regulated at fc 1.47-1.85 in larvae fed with diet E vs those fed with diet B (75% vs 10% NPM). Also in the gene group *transport* all four genes were up-regulated at fc 1.56 up to 24.20 for *solute carrier family 12 (sodium/chloride transporters), member 3*. Other groups involved *WD40 repeats acting as protein-protein interaction sites*; *nucleotide binding*; *transcription*; *metal binding/zinc fingers* and *membrane*. In gene groups *WD40 repeats* (4 genes) and *metal binding/zinc fingers* (13 genes) all genes were up-regulated while groups *nucleotide binding*, *transcription* and *membrane* contained both genes that were up-regulated as well as genes that were down-regulated. Among them were genes that were specific for larvae fed with diet E (*zgc:110560* or *hypothetical protein LOC100150958*; *similar to Serine/threonine-protein kinase Pim-3* or *si:dkey-108d22.5*; *forkhead box G1* and *mediator of RNA polymerase II transcription subunit 11*) and genes specific for larvae fed with diet B (*zgc:172065* or *hypothetical LOC100001153*).

4. Discussion

This study represents a methodological exploration assessing the development of the gastrointestinal microbiota through 16S rRNA gene-based microbial composition profiling by pyrosequencing and the functional response of the transcriptome by mRNAseq. Both approaches provided complementary information on the nutritional impact of a novel protein source relevant for aquaculture. The impact of the novel protein meal concerned the impact on the GIT microbiota and, either directly or indirectly through the microbiota, on the host transcriptome.

The two experimental diets especially differed in the level of iron (Table 1). This difference originated from the difference in NPM as it is a rich iron source, consisting mostly of heme iron. Feed ingredients of animal origin, like the NPM in this study, are rich in iron. Iron can be taken up in two forms; as heme (e.g. hemoglobin), and as non-heme (e.g. iron sulfate). The uptake of heme differs greatly from the absorption of inorganic iron as the uptake of heme iron is not physiologically regulated and independent of the intestinal pH in contrast to the uptake of inorganic iron (Kraemer and Zimmermann, 2007). Uptake of inorganic iron is more complex and requires reduction of Fe^{3+} to Fe^{2+} which, in turn, requires an acidic environment as provided by the gut in monogastric animals. Despite very few mechanistic studies of piscine intestinal iron uptake, zebrafish is supposed to take up iron from the diet in the intestinal enterocytes not any different than by the mechanism that all vertebrates apply (reviewed by Bury et al., 2003). Iron homeostasis is crucial since, in excess, iron can be detrimental to health because of its production of oxygen free radicals, and, when too low, loss of energy due to the decrease of haemoglobin concentrations and cytochrome capacity in aerobic metabolism. In diets B and E, calculated iron content was 312 and 1486 mg kg^{-1}, respectively (Table 1). These levels exceed the daily iron requirements in fish that ranges between 30 and 170 mg kg^{-1} DM food (Watanabe et al., 1997). Also experimental studies on dietary iron supplementation in fish report on findings that indicate that such levels have already reached a plateau for physiological effects (tilapia: Shiau and Su, 2003; rainbow trout: Carriquiriborde et al., 2004). So most probably any difference in GIT response to the experimental diets does not relate to an iron shortage in fish fed with diet B but may reflect effects of iron overloading, especially in fish fed with diet E.

Of 97,894 representative sequences of GIT microbiota in the present study, the dominant phylum was *Proteobacteria*. Other main phyla encountered and ranked according to average relative abundance were *Firmicutes*, *Actinobacteria* and *Bacteroidetes*. The rapid temporal

changes in GIT microbiota make it difficult to compare our results directly to other studies that have been done on the identification of zebrafish GIT microbiota (Bates et al., 2006; Roeselers et al., 2011; Semanova et al., 2012; Lan and Love, 2012). However, our data are similar to published data for two important aspects: (i) larval zebrafish GITs are dominated by one to three bacterial species, (ii) these dominant species generally belong to *Proteobacteria* and *Firmicutes*.

Firstly, many larval and juvenile animals still have immature gut microbiota that are not yet fully functional and may be dominated by a few early colonisers. For human gut microbiota it has been shown that infant gut bacterial species have faster growth rates than adult gut bacterial species, which favours early colonisation (De Muinck et al., 2013). With respect to the second point, dominant bacteria in the GITs of fish juveniles have been identified as *Pseudomonas* (*Proteobacteria*) in zebrafish and salmon (Bates et al., 2006; Navarette et al., 2009), and an unidentified gammaproteobacterium in juvenile pinfish (Givens, 2012). Similarly, the GIT microbiota of juvenile Siberian sturgeon was shown to be mono-dominated, only by *Cetobacterium somerae* (*Fusobacteria*) (Geraylou et al., 2013). Thereby it should be noted that the herbivorous pinfish and the carnivorous salmon and sturgeon have stomachs and GIT morphologies that are different from the omnivorous stomachless zebrafish.

At 5 dpf, zebrafish GIT in our study was dominated by members of the genus *Clostridium* (*Firmicutes*) with 40 ± 11% of the reads and the family *Procabacteriaceae* (*Proteobacteria*) with 22 ± 21% of the reads. *Clostridium* is a well-known inhabitant of the animal gut. Although the genus *Clostridium* has been related to animal diseases, commensal *Clostridium* spp. are dominant players in the maintenance of gut homeostasis in man and other animals including many fish species (Sullam et al., 2012; Lopetuso et al., 2013). *Clostridium* spp. were previously found to be abundant in the GIT of fed juvenile zebrafish (Semanova et al., 2012), but absent or not abundant in starved juveniles and in the adult zebrafish GIT (Roeselers et al., 2011; Semanova et al., 2012), which may indicate that it is an early coloniser of the zebrafish GIT that disappears with gut maturation. This is also confirmed by our data as *Clostridium* spp. were detected only in a few samples at later time points. Little information exists on the family *Procabacteriaceae* (*Proteobacteria*) and they are not frequently encountered in animal gut. Candidatus *Procabacter acanthamoebae* was identified as an *Acanthamoeba* endosymbiont (Horn et al., 2002). *Acanthamoeba* is related to a number of animal diseases (Paterson et al., 2011), however, no visible signs of distress were recorded at 5 dpf and onwards no more reads were detected that were affiliated to *Procabacteriaceae*. The zebrafish-associated microbiota shifted remarkably between 5 and 7 dpf (Fig. 2) and only few of the OTUs present at 5 dpf were also found at 7 dpf. One OTU that was recovered at both days was a *Rhodanobacter* sp. (*Proteobacteria*) that accounted for 37 ± 34% of the reads at 7 dpf. *Rhodanobacter* spp. are typically known for their potential for partial or complete denitrification (Kostka et al., 2012) and may have been derived from the fish tank water filtration system. The other dominant OTU at 7dpf was closely related to *Propionibacterium acnes* (*Actinobacteria*) and this OTU remained traceable at 14 and 21 dpf albeit at lower relative abundance. A large number of new OTUs was found at 14 dpf of which many were lost again at 21 dpf. However, some, such as a number of OTUs belonging to the family *Rhodobacteraceae*, appeared at 14 dpf and remained. *Rhodobacteraceae* are commonly found in the aquatic habitat, but are not typical gut-associated bacteria (Elifantz et al., 2013). Despite the large fluctuations of zebrafish-associated microbiota over time, it is apparent that the inter-individual variation within time points decreases (Fig. 2, Fig. S2), which indicates that a more stable and homogenous microbiota becomes associated with the zebrafish population at 21 dpf. In addition, the dominant OTU at 21 dpf, a *Pseudomonas* sp., is in line with previous studies in zebrafish (Bates et al., 2006; Roeselers et al., 2011; Semanova et al., 2012; Lan and Love, 2012), which could indicate gut maturation.

The OTUs that were present at a higher relative abundance in fish fed with the experimental diet E as compared to animals fed with the control diet B and contributed most to the difference between gut microbiota for the different diets can roughly be divided into three groups based on comparison to their near neighbours: (1) potential pathogens, (2) bacteria with an increased iron metabolism, and (3) common aquatic bacteria. From the first group, *Plesiomonas shigelloides* (OTU45) contributed most to the difference in microbiota between diet E and diet B. It is an emerging pathogen that is widespread in the aquatic environment and has been related to gastrointestinal infections and other diseases in a wide range of animal hosts including fish (Chen et al., 2013; Joh et al., 2013). Also the family *Aeromonadaceae* (to which OTU1407 belongs) harbours many fish pathogens that are associated to gastroenteritis and wound infections (Tomás, 2012). The genus *Trabulsiella* was proposed in 1991 (McWhorter et al., 1991) as a genus that is highly related to pathogenic *Salmonella* sp. Currently two species belonging to this genus have been described and were isolated from human and termite gut. Although *T. guamensis* can occur in human diarrheal stools, there still is no evidence that it actually causes diarrhea (McWhorter et al., 1991; Chou et al., 2007). The order *Legionellales* (OTU 552) comprises the families *Legionellaceae* and *Coxiellaceae* that are both known to represent common animal pathogens (Garrity et al., 2005). The genus *Burkholderia* (OTU16) represents both pathogenic (animals and plants) and non-pathogenic species (Estrada-de los Santos et al., 2013). *Propionibacterium acnes* (OTU832) and *Cupriavidus* (OTU1260) spp. are commensal inhabitants of the skin and GIT of animals, but are also related to infections, especially in immuno-compromised individuals (Perry and Lambert, 2006; Balada-Llasat et al., 2010). In addition, some *Cupriavidus* spp., such as *C. gilardii* and *C. metallidurans* are particularly resistant to high metal concentrations (Kirsten et al., 2011). The *Halomonas* sp. (OTU1182) that was markedly increased in zebrafish fed at high NPM inclusion level shared 96 to 98% percent identity with *Halomonas titanicae* strains SSA831, SSA728 and SSA637 based on the 16S rRNA gene. *Halomonas titanicae* was isolated from corroded parts of the RMS Titanic wreck and possesses an unusual high number of iron reductases, iron uptake regulators, ferrochelase, iron transporters, and iron-binding periplasmic protein-encoding genes (Sánchez-Porro et al., 2013). In addition, also *Novospirillum itersonii* (OTU1459), previously named *Aquaspirillum itersonii*, is best known for its iron reduction capacities (Dailey and Lascelles, 1977) and may have been selected for by the high iron content of diet E. In future studies, a tank water control should be included, in this case to be able to confirm that these iron reducers originated from the tank water.

In this study, we have applied mRNAseq in an unbiased approach to investigate the molecular differentiation of physiological processes in the GIT as indicated by differentially expressed genes. In other recent studies that we found in literature, only whole-body mRNAseq was performed at such young stages of developing zebrafish, or microarray studies specifically on the GIT. Note that because in this study the whole GIT was analyzed, any variation in physiological processes occurring along the GIT was thereby discarded (Clements et al., 2014). Although zebrafish belongs to the cyprinid family, a family of fishes that that do not possess a stomach, also stomachless fish show regional differentiation in GIT function (German, 2009) and microbial communities (Clements et al., 2014). In our study only approximately 1% of the total number of genes was differentially expressed. Functional classification of genes revealed that by far the most dominant gene groups represented *ribosome components and activity* and *transport* that were enriched in their expression in the larvae fed at high inclusion levels vs. those fed at low inclusion levels of the NPM. These gene groups included 23 ribosomal proteins and several solute carrier families of sodium, potassium, chloride, dicarboxylate, and aminoacid and glucose transporters. These data would suggest that the GIT of larvae fed at higher inclusion of the protein meal is much more active in transmembrane ion transport and protein synthesis, perhaps for making the machinery to perform this transport.

Among the individual genes, we have found 10 uniquely expressed genes for fish fed at low inclusion levels and 16 genes for fish fed at high inclusion levels. The 10 uniquely expressed genes for fish fed at low inclusion levels were all uncharacterised genes except for one: *secretogranin V (7B2 protein)*, a gene required for the production of an active Proprotein convertase 2 (PC2) enzyme (also known as prohormone convertase 2 or neuroendocrine convertase 2 enzyme) that is responsible for the first step in the maturation of many neuroendocrine peptides from their precursors, such as the conversion of proinsulin to insulin intermediates (Mbikay et al., 2001; Portela-Gomes et al., 2008). The 16 uniquely expressed genes for fish fed at high inclusion levels were all, except for four genes, characterized. Among them were *cadherin 16, KSP-cadherin,* a calcium-dependent, membrane-associated glycoprotein, and *claudin 19*, involved in magnesium transport. These genes also have a clear role in *transport* and so has the highest up-regulated expressed gene in fish fed at high inclusion level at fc 768: *stanniocalcin 1, like*. Stanniocalcin 1 is involved in calcium homeostasis. It has been found to reduce Ca^{2+} uptake via the inhibition of epithelial Ca^{2+} channel mRNA expression in zebrafish embryos (Tseng et al., 2009). As such it would fit well with a role in the GIT, with the dominant gene group *transport* and with other strongly up-regulated genes expressing channels and transporters (*purinergic receptor P2X, ligand-gated ion channel, 3b* at fc 102; *solute carrier family 12 (sodium/chloride transporters), member 3* at fc 24; *chloride channel accessory 2* at fc 4.93; *slc12a10.3 solute carrier family 12 (sodium/potassium/chloride transporters), member 10.3* at fc 3.24; *solute carrier family 16 (monocarboxylic acid transporters), member 9a* at fc 2.86; *solute carrier family 12, member 10.1* at fc 2.48; *solute carrier family 25, member 38a* at fc 2.06; *solute carrier family 2 (facilitated glucose transporter), member 13* at fc 2.06). However, such high level of difference in expression is often indicative for immune-related genes. Stanniocalcin 1 also has an immune-related function. It is an inhibitor of macrophage chemotaxis and chemokinesis (Kanellis et al., 2004) and modulates transendothelial migration of leukocytes (Chakraborty et al., 2007) in humans. Thus a role in modulating the immune/inflammatory response could be expected. There are more signs for an immune response in fish fed at high NPM levels given the roles of other up-regulated immune-related genes such as *interlectin 2* at fc 8.64; *radical S-adenosyl methionine domain containing 2* at fc 6.29 (see also later); *ISG15 ubiquitin-like modifier* at fc 4.99 and *B-cell CLL/lymphoma 6a (zinc finger protein 51)* at fc 2.83. Rawls et al. (2004) performed DNA microarray comparisons of gene expression in the digestive tracts of 6 dpf zebrafish and revealed that 212 genes were regulated by the microbiota, including genes involved in innate immune responses. Thus we cannot rule out that fish fed at high NPM levels display an gastrointestinal immune response.

Because the two experimental diets especially differed in the level of iron, in a biased approach, we have analyzed the expression profiles of genes functionally involved in iron uptake and homeostasis. Twenty seven genes were identified as involved in iron homeostasis but were non-differentially expressed at a fold change 0.27 – 1.54 (Table S5). Among them were genes encoding for *ferritin*, *transferrin receptors*, *hephaestin*, *ferrochelatase*, an *iron-responsive element binding protein*, an *iron-regulated transporter* and *ceruloplasmin*; but also *heme oxygenase*, *heme binding protein*, a *heme transporter*, and finally *hepcidin* and the *interleukin 6 receptor*. In a parallel study we have investigated the effects on adult zebrafish fed with the same experimental diets for one month (Palstra et al., unpublished data). In a biased approach, quantitative real-time PCR was performed on individual GIT and liver of these fish. Here we did find significant differential expression of several of these genes: GIT expression of marker gene *hepcidin antimicrobial peptide 1* (*hamp1*) was significantly higher, and of *hephaestin-like 1* significantly lower in fish fed at high vs. low iron level. Liver expression of marker genes *transferrin a* and *hamp1* was significantly higher, and of *ferritin heavy polypeptide 1a* significantly lower in fish fed at high vs. low iron level. These expression profiles, supported by data on body composition, suggest that in adult fish fed at higher iron level, less dietary iron uptake occurs, less iron is released in the circulation, and less iron is taken up and stored in the

570 liver. This suggests a metabolic defense mechanism against iron overload. Indeed, metal
571 absorption is lower when metal concentrations are elevated (reviewed by Karasov and Douglas,
572 2013). However, as based on the absence of differential expression of such genes for the larval
573 zebrafish in this study, we cannot conclude that major changes occur in iron uptake and
574 homeostasis during the earliest stages of development. Although diets especially differed in iron
575 content, no data have been collected in this study that show that this difference also leads to a
576 difference in iron availability to the GIT. The competitory activity of the microbiota may result
577 in an alteration of the iron availability for the gastrointestinal functions. Some genes that were
578 differentially expressed may have a relation with iron homeostasis. Among them was *wdr45 like*
579 that was up-regulated at fold change 1.6 in larvae fed at higher iron level. Wdr45 is associated
580 with human brain iron accumulation (Haack et al., 2012). *Radical S-adenosyl methionine domain*
581 *containing 2* (*rsad2*) is an interferon-inducible iron-sulfur cluster-binding antiviral protein that
582 was up-regulated at fold change 6.29 at higher iron levels. Other genes that may be involved
583 could be many of the unknown differentially expressed genes belonging to the cluster metal
584 binding/zinc fingers.
585
586 The developing gastrointestinal microbiota of larval zebrafish showed rapid temporal
587 changes until a suspected stable and mature state at 21 dpf. At all times, the larval zebrafish GITs
588 were dominated by one to three bacterial species generally belonging to *Proteobacteria* and
589 *Firmicutes*. The OTUs that contributed most to the difference between gastrointestinal
590 microbiota for the different diets represented common aquatic bacteria but also bacteria related
591 to potential pathogens and bacteria with an increased iron metabolism. As for the gastrointestinal
592 transcriptome at 21 dpf, the GIT of larvae fed at higher NPM inclusion is more active in
593 transmembrane ion transport and protein synthesis. Although some indications existed,
594 transcriptomic analysis did not reveal signs for the occurrence of a major immune/inflammatory
595 activation and/or iron overload response.
596
597 The gained insights on the impact of the NPM on larval zebrafish GIT microbiology and
598 physiology are valuable information for fish specifically, and vertebrates in general. Caution is
599 required with the transfer of knowledge towards commercially produced species in aquaculture,
600 particularly because many of them are carnivorous. Carnivorous fishes like salmon (Navarette et
601 al., 2009) and sole (Martin-Antonio et al., 2007; Tapia-Paniagua et al., 2010) possess a
602 specialized GIT region with an acidic environment, or a stomach, while the omnivorous cyprinid
603 zebrafish is stomachless. Such interspecific differences in GIT morphology have important
604 consequences for the GIT microbial composition and physiology (Clements et al., 2014), and
605 thus for the dietary impact of the NPM.
606
607
608 **Acknowledgements**
609
610 The authors wish to thank W. Nusselder, S. Visser and T. Wiegers (Aquatic Research
611 Facility of Wageningen University, Wageningen, the Netherlands) for feeding the fish during the
612 experiment and assistance with sampling, M. Naim (Wageningen University, Wageningen, the
613 Netherlands) for performing the reverse transcription reaction with the RNA samples and
614 preparing those samples for pyrosequencing and H. Jansen (ZF-screens B.V., Leiden, the
615 Netherlands) for performing mRNAseq and providing an initial data analysis. The research was
616 supported by a grant from the Dutch Ministry of Economic Affairs, Agriculture and Innovation
617 (project “Marine Aquaculture: Customized Nutrition”, KB-14-005-014 to A.P. Palstra).
618
619
620 **Author Contributions**
621

Conceived and designed the experiments: ER, DS, JK, APP. Performed the experiments: ER, DS, MtV, MF, GMB. Analyzed the data: ER, DS, APP. Wrote the paper: ER, DS, JK, HS, APP.

Competing Financial Interests statement

The authors declare no competing financial interests.

References

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biology* **11**: R106.

Balada-Llasat, J.M., Elkins, C., Swyers, L., Bannerman, T., Pancholi, P. (2010). Pseudo-outbreak of *Cupriavidus pauculus* infection at an outpatient clinic related to rinsing culturette swabs in tap water. *J. Clin. Microbiol.* **48**: 2645-2647.

Bates, J.M., Mittge, E., Kuhlman, J., Baden, K.N., Cheesman, S.E. & Guillemin, K. (2006). Distinct signals from the microbiota promote different aspects of zebrafish gut differentiation. *Dev. Biol.* **297**: 374-386.

Bury, N.R., Walker, P.A., Glover, C.N. (2003). Nutritive metal uptake in teleost fish. *J. Exp. Biol.* **206**: 11-23.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Meth.* **7**: 335-336.

Carriquiriborde, P, Handy, R.D., Davies, S.J. (2004). Physiological modulation of iron metabolism in rainbow trout (*Oncorhynchus mykiss*) fed low and high iron diets. *J. Exp. Biol.* **207**: 75-86.

Chakraborty, A., Brooks, H., Zhang, P., Smith, W., McReynolds, M.R., Hoying, et al. (2007). Stanniocalcin-1 regulates endothelial gene expression and modulates transendothelial migration of leukocytes. *Am. J. Physiol. Renal. Physiol.* **292**: F895-904.

Chen, X., Chen, Y., Yang, Q., Kong, H., Yu, F., Han, D., et al. (2013). *Plesiomonas shigelloides* infection in Southeast China. *PloS ONE* **8**: e77877.

Chou, J.H., Chen, W.M., Arun, A.B., Young, C.C. (2007). *Trabulsiella odontotermitis* sp. nov., isolated from the gut of the termite *Odontotermes formosanus* Shiraki. *Int. J. Syst. Evol. Microbiol.* **57**: 696-700.

Clements, K.D., Angert, E.R., Montgomery, W.L., Choat, J.H. (2014). Intestinal microbiota in fishes: what's known and what's not. *Mol. Ecol.* **23**: 1891–1898.

Dailey, H.A., Lascelles, J. (1977). Reduction of iron and synthesis of protoheme by *Spirillum itersonii* and other organisms. *J. Bacteriol.* **129**: 815-820.

De Muinck, E.J., Stenseth, N.C., Sachse, D., Vander Roost, J., Rønningen, K.S., Rudi, K., et al. (2013). Context-dependent competition in a model gut bacterial community. *PLoS ONE* **8**: e67210.

DeSantis, T.Z., Hugenholtz, P., Keller, K., Brodie, E.L., Larsen, N., Piceno, Y.M., et al. (2006). NAST: a multiple sequence alignment server for comparative analysis of 16S rRNA genes. *Nucleic Acids Res.* **34**: W394–W399.

Elifantz, H., Horn, G., Ayon, M., Cohen, Y., Minz, D. (2013). *Rhodobacteraceae* are the key members of the microbial community of the initial biofilm formed in Eastern Mediterannean coastal seawater. *FEMS Microbiol. Ecol.* **85**: 348-357.

Estrada-de los Santos, P., Vinuesa, P., Martinez-Aguilar, L., Hirsch, A.M., Caballero-Mellado, J. (2013). Phylogenetic analysis of *Burkholderia* species by multilocus sequence analysis. *Curr. Microbiol.* **67**: 51-60.

672 Farber, S.A., Pack, M., Ho, S.-Y., Johnson, I.D., Wagner, D.S., Dosch, R., et al. (2001). Genetic
673 Analysis of Digestive Physiology Using Fluorescent Phospholipid Reporters. *Science* **292**:
674 1385-1388.
675 Garrity, G.M., Bell, J.A., Lilburn, T. (2005). "Order VI. *Legionellales* ord. nov.," in *Bergey's*
676 *Manual of Systematic Bacteriology,* eds. Brenner, D.J., Krieg, N.R., Staley, J.J., Garrity,
677 G.M. (New York: Springer), 210-236.
678 Geraylou, Z., Souffreau, C., Rurangwa, E., Maes, G.E., Spanier, K.I., Courtin, C.M., et al.
679 (2013). Prebiotic effects of arabinoxylan oligosaccharides on juvenile Siberian sturgeon
680 (*Acipenser baerii*) with emphasis on the modulation of the gut microbiota using 454
681 pyrosequencing. *FEMS Microbiol. Ecol.* **86**: 257-371.
682 German, D.P. (2009). Do herbivorous minnows have "plug-flow reactor" guts? Evidence from
683 digestive enzyme activities, gastrointestinal fermentation, and luminal nutrient
684 concentrations. *J. Comp. Physiol. B* **179**: 759-771.
685 Givens, C.E. (2012). *A fish tale: comparison of the gut microbiome of 15 fish species and the*
686 *influence of diet and temperature on its composition*. PhD thesis University of Georgia,
687 USA. https://getd.libs.uga.edu/pdfs/givens_carrie_e_201212_phd.pdf.
688 Gomez-Requeni, P., Conceicao, L.E.C., Olderbakk Jordal, A.E., Ronnestad, I. (2010) A
689 reference growth curve for nutrition experiments in zebrafish (*Danio rerio*) and changes in
690 whole body proteome during development. *Fish Physiol. Biochem.* **36**: 1199-1215.
691 Guss, A.M., Roeselers, G., Newton, I.L.G., Young, C.R., Klepac-Ceraj, V., Lory, S., et al.
692 (2011). Phylogenetic and metabolic diversity of bacteria associated with cystic fibrosis.
693 *ISME J.* **5**: 20-29.
694 Haack, T.B., Hogarth, P., Kruer, M.C., Gregory, A., Wieland, T., Schwarzmayr, T. et al. (2012).
695 Exome sequencing reveals de novo *WDR45* mutations causing a phenotypically distinct, X-
696 linked dominant form of NBIA. *Am. J. Hum. Genet.* **91**: 1144-1149.
697 Hansen, G.A., and Olafsen, J.A. (1999). Bacterial interactions in early life stages of marine cold
698 water fish. *Microb. Ecol.* **38**: 1-26.
699 Holmberg, A., Schwerte, T., Pelster, B., Holmgren, S. (2004). Ontogeny of the gut motility
700 control system in zebrafish *Danio rerio* embryos and larvae. *J. Exp. Biol.* **207**: 4085-4094.
701 Horn, M., Fritsche, T.R., Linner, T., Gautom, R.K., Harzenetter, M.D., Wagner, M. (2002).
702 Obligate bacterial endosymbionts of *Acanthamoeba* spp. related to the *β-proteobacteria*:
703 proposal of '*Candidatus* Procabacter acanthamoebae' gen. nov. sp. nov. *Int. J. Syst. Evol.*
704 *Microbiol.* **52**: 599-605.
705 Huang, D.W., Sherman, B.T., Lempicki, R.A. (2009a). Systematic and integrative analysis of
706 large gene lists using DAVID Bioinformatics Resources. *Nature Protoc.* **4**: 44-57.
707 Huang, D.W., Sherman, B.T., Lempicki, R.A. (2009b). Bioinformatics enrichment tools: paths
708 toward the comprehensive functional analysis of large genes. *Nucleic Acids Res.* **37**: 1-13.
709 Joh, S.J., Ahn, E.H., Lee, H.J., Shin, G.W., Kwon, J.H., Park, C.G. (2013). Bacterial pathogens
710 and flora isolated from farm-cultured eels (*Anguilla japonica*) and their environmental
711 waters in Korean eel farms. *Vet. Microbiol.* **163**: 190-195.
712 Kals, J. (2014). *Alleviation of anaemic growth suppression in fish*. International patent
713 WO/2014/032837.
714 Kanellis, J., Bick, R., Garcia, G., Truong, L., Tsao, C.C., Etemadmoghadam, D. et al. (2004).
715 Stanniocalcin-1, an inhibitor of macrophage chemotaxis and chemokinesis. *Am. J. Physiol.*
716 *Renal. Physiol.* **286**: F356-F362.
717 Kanther, M., and Rawls, J.F. (2010). Host–microbe interactions in the developing zebrafish.
718 *Curr. Opin. Immunol.* **22**: 10–19.
719 Karasov, W.H., Douglas, A.E. (2013). Comparative Digestive Physiology. *Compr. Physiol.* **3**:
720 741-783.
721 Kimmel, C.B., Ballard, W.W., Kimmel, S.R., Ullmann, B. & Schilling, T.F. (1995). Stages of
722 embryonic development of the zebrafish. *Dev. Dyn.* **203**: 253–310.

723 Kirsten, A., Herzberg, M., Voigt, A., Seravalli, A. Grass, G., Scherer, J., et al. (2011).
724 Contributions of five secondary metal uptake systems to metal homeostasis of *Cupriavidus*
725 *metallidurans* CH34. *J. Bacteriol.* **193**: 4652-4663.
726 Kortman, G.A., Boleij, A., Swinkels, D.W., Tjalsma, H. (2012). Iron availability increases the
727 pathogenic potential of *Salmonella typhimurium* and other enteric pathogens at the intestinal
728 epithelial interface. *PLoS ONE* **7**: e29968.
729 Kostka, J.E., Green, S.J., Rishiswar, L., Prakash, O., Katz, L.S., Mariño-Ramirez, L., et al.
730 (2012). Genome sequences for six *Rhodanobacter* strains, isolated from soils and the
731 terrestrial subsurface, with variable denitrification capacities. *J. Bacteriol.* **194**: 4461-4462.
732 Kraemer, K., and Zimmermann, M.B. (2007). *Nutritional Anemia*. Basel: Sight and Life press.
733 Lan C.-C., and Love, D.R. (2012). Molecular Characterisation of Bacterial Community
734 Structure along the Intestinal Tract of Zebrafish (*Danio rerio*): A Pilot Study. *ISRN*
735 *Microbiol.* 2012: Article ID 590385, Doi: 10.5402/2012/590385.
736 Lane, D.J. (1991). "16S/23S rRNA sequencing," in: *Nucleic acid techniques in bacterial*
737 *systematics*, eds. Stackebrandt, E., Goodfellows, M. (New York: Wiley), 115–175.
738 Lopetuso, L.R., Scaldaferri, F., Petito, V., Gasbarrini, A. (2013). Commensal Clostridia: leading
739 players in the maintenance of gut homeostasis. *Gut Pathogens* **5**: 23.
740 Martin-Antonio, B., Manchado, M, Infante, C., Zerolo, R., Labella, A., Alonso, C., et al. (2007).
741 Intestinal microbiota variation in Senegalese sole (*Solea senegalensis*) under different
742 feeding regimes. *Aquac. Res.* **38**: 1213-1222.
743 Mbikay, M., Seidah, N.G., Chrétien, M. (2001). Neuroendocrine secretory protein 7B2:
744 structure, expression and functions. *Biochem. J.* **357**: 329–42.
745 McWhorter, A.C., Haddock, R.L., Nocon, F.A., Steigerwalt, A.G., Brenner, D.J., Aleksić, S. et
746 al. (1991). *Trabulsiella guamensis*, a new genus and species of the family
747 *Enterobacteriaceae* that resembles *Salmonella* subgroups 4 and 5. *J. Clin. Microbiol.* **29**:
748 1480-1485.
749 Navarette, P., Espejo, R.T., Romero, J. (2009). Molecular analysis of microbiota along the
750 digestive tract of juvenile Atlantic salmon (*Salmo salar* L.). *Microb Ecol* **57**: 550-561.
751 Nayak, S.K. (2010) Probiotics and immunity: A fish perspective. *Fish Shellfish Immun.* **29**: 2–
752 14.
753 Neish, A.S. (2009). Microbes in gastrointestinal health and disease. *Gastroenterology* **136**: 65–
754 80.
755 Paterson, G.N., Rittig, M., Siddiqui, R., Khan, N.A. (2011). Is *Acanthamoeba* pathogenecity
756 associated with intracellular bacteria? *Exp. Parasitol.* **129**: 207-210.
757 Perry, A.L., and Lambert, P.A. (2006). *Propionibacterium acnes*. *Lett. Appl. Microbiol.* **42**:
758 185-188.
759 Portela-Gomes, G.M., Grimelius, L., Stridsberg, M. Prohormone convertases 1/3, 2, furin and
760 protein 7B2 (Secretogranin V) in endocrine cells of the human pancreas. *Regul. Peptides*
761 (2008) **146**: 117–124.
762 Rawls, J.F., Mahowald, M.A., Goodman, A.L., Trent, C.M. & Gordon, J.I. (2007). In vivo
763 imaging and genetic analysis link bacterial motility and symbiosis in the zebrafish gut. *Proc.*
764 *Natl. Acad. Sci.* **104**: 7622-7627.
765 Rawls, J.F., Mahowald, M.A., Ley, R.E., Gordon, J.I. (2006). Reciprocal gut microbiota
766 transplants from zebrafish and mice to germ-free recipients reveal host habitat selection.
767 *Cell* **127**: 423–433.
768 Rawls, J.F., Samuel, B.S. & Gordon, J.I. (2004). Gnotobiotic zebrafish reveal evolutionarily
769 conserved responses to gut microbiota. *Proc. Natl. Acad. Sci.* **101**: 4596-4601.
770 Ribas, L., and Piferrer, F. (2013). The zebrafish (*Danio rerio*) as a model organism, with
771 emphasis on applications for finfish aquaculture research. *Rev. Aquacult.* **5**: 1–32.
772 Ringø, E., Olsen R.E., Gifstad, T.Ø., Dalmo, R.A., Amlund, H., Hemre G.-I., et al. (2010).
773 Prebiotics in aquaculture: a review. *Aquacult. Nutr.* **16**: 117–136.

774 Roeselers, G., Mittge, E.K., Stephens, W.Z., Parichy, D.M., Cavanaugh, C.M., Guillemin, K., et
775 al. (2011). Evidence for a core gut microbiota in the zebrafish. *ISME J.* **5**: 1595–1608.
776 Sánchez-Porro, C., De la Haba, R.R., Cruz-Hernandez, N., Gonzalez, J.M., Reyes-Guirao, C.,
777 Navarro-Sampedro, L., et al. (2013). Draft genome of the marine Gammaproteobacterium
778 *Halomonas titanicae*. *Genome Announc.* **1**: e00083-13.
779 Semanova, I., Carten, J.D., Stombaugh, J., Mackey, L.C., Knight, R., Farber, S.A., et al. (2012).
780 Microbiota regulate intestinal absorbtion and metabolism of fatty acids in the zebrafish. *Cell*
781 *Host Microbe* **12**: 277-288.
782 Shiau, S-Y., and Su, L-W. (2003). Ferric citrate is half as effective as ferrous sulfate in meeting
783 the iron requirement of juvenile Tilapia, *Oreochromis niloticus* × *O. aureus*. *J. Nutr.* **133**:
784 483-488.
785 Sullam, K.E., Essinger, S.D., Lozupone, C.A., O'Connor, M.P., Rosen, G.L., Knight, R., et al.
786 (2012). Environmental and ecological factors that shape the gut bacterial communities of
787 fish: a meta-analysis. *Mol. Ecol.* **21**: 3363-78.
788 Tapia-Paniagua, S.T., Chabrillón, M., Díaz-Rosales, P., García de la Banda, I., Lobo, C.,
789 Balebona, Ma.C., et al. (2010). Intestinal microbiota diversity of the flat fish *Solea*
790 *senegalensis* (Kaup, 1858) following probiotic administration. *Microb. Ecol.* **60**: 310–319.
791 Tomás, J.M. (2012). The main *Aeromonas* pathogenic factors. *ISRN Microbiol.* 25621.
792 Tseng, D-Y., Chou, M-Y., Tseng, Y-C., Hsiao, C-D., Huang, C-J., Kaneko, T., et al. (2009).
793 Effects of stanniocalcin 1 on calcium uptake in zebrafish (*Danio rerio*) embryo. *Am. J.*
794 *Physiol. Regul. Integr. Comp. Physiol.* **296**: R549–R557.
795 Ulloa, P.E., Iturra, P., Neira, R., Araneda, C. (2011). Zebrafish as a model organism for
796 nutrition and growth: towards comparative studies of nutritional genomics applied to
797 aquacultured fishes. *Rev. Fish Biol. Fisheries* **21**: 649-666.
798 Verschuere, L., Rombaut, G., Sorgeloos, P., Verstraete W. (2000). Probiotic bacteria as
799 biological control agents in aquaculture. *Microbiol. Mol. Biol. Rev.* **64**: 655-671.
800 Wallace, K.N., Akhter, S., Smith, E.M., Lorent, K., Pack, M. (2005). Intestinal growth and
801 differentiation in zebrafish. *Mech. Dev.* **1222**: 157-173.
802 Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, JR. (2007). Naive Bayesian classifier for rapid
803 assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.*
804 **73**: 5261–5267.
805 Watanabe, T., Kiron, V., Satoh, S. (1997). Trace minerals in fish nutrition. *Aquaculture* **151**:
806 185-207.
807 Westerfield, M. (1993). *The Zebrafish Book: A Guide For The Laboratory Use Of Zebrafish*
808 *(Brachydanio rerio)*. Eugene: University of Oregon Press, Eugene.

Figure legends

Fig. 1: Taxonomic distribution of bacterial 16S rRNA gene reads retrieved from zebrafish raised with different diets during the first 21 dpf. Only phyla that represent more than 1% of the reads in at least one of the samples are shown and chloroplasts-affiliated reads were removed prior to analysis. The pie diagrams shown are averages of replicate samples with the number of samples and total number of reads below each pie chart. Not all replicate samples that were initially taken resulted in successful DNA extraction due to the small sample size. The step-wise decrease of *Paramecium* in the diet is indicated below the pie charts, with the percentage of *Paramecium* in yellow and the percentage of Diet B/E in green.

Fig. 2: CCA-ordination plot of the zebrafish microbiota. The red triangles represent the centroids of the datasets belonging to different time points indicated with the number in red. Each data point refers to DNA extracted from 10 pooled zebrafish. Sample names are build up as follows: ZF= zebrafish; dpf (5, 7, 14 or 21); diet (p, pB, pE, B, E); replicate (1, 2).

Fig. 3: Heatmap of the operational taxonomic units (OTUs) (97% similarity) that represented more than 1% of the reads in at least one of the zebrafish samples. These OTUs represented 85-100% of the reads in the different samples. Relative abundance of OTUs is marked according to the legend in the figure. Samples are organised per day and according to diet B or E within day 14 and day 21. OTUs were classified up to the phylum (p), class (c), order (o), family (f), genus (g) or species (s) level. “Bact” refers to the phylum *Bacteroidetes*, “Planct” refers to the phylum *Planctomycetes*. Sample names are built up as follows: ZF= zebrafish; dpf (5, 7, 14 or 21); diet (p, pB, pE, B, E); replicate (1, 2); (R) if a sample is derived from RNA.

833 **Tables**

834 **Table 1: Experimental diets.**

	Dietary treatment	
Amount of novel protein meal (%)	**10**[l]	**75**[l]
Code	**B**	**E**
Ingredients in %		
Novel protein meal[a]	10.00	75.00
Pea protein[b]	21.88	7.94
Casein[c]	19.42	3.66
Soy Protein Concentrate[d]	15.20	0.00
Fish Oil[e]	10.24	0.00
Diamol[f]	8.33	0.00
Sugar[g]	1.32	0.00
Lime[h]	0.20	0.00
Wheat gluten[i]	5.00	5.00
Binder 1[j]	2.00	2.00
Binder 2[k]	2.00	2.00
Salt[l]	2.00	2.00
Binder 3[m]	1.00	1.00
Mineral and vitamin premix[†]	1.36	1.36
Betaine[n]	0.05	0.05
Calculated (composition)	**B**	**E**
DM ($g.kg^{-1}$)	921.7	941.9
ASH ($g.kg-^{1}$)	144.5	157.4
CP (g.kg-1)	533.2	533.1
EE ($g.kg^{-1}$)	134.8	134.8
Ca ($g.kg^{-1}$)	2.4	2.3
P ($g.kg^{-1}$)	5.6	6.1
GE	21.4	21.2
CP/GE	25.9	26.2
Iron ($mg.kg^{-1}$)	312	1486
Calculated amino acids ($g.kg^{-1}$)		
Lysine[*]	37.1	34.1
Methionine[*]	10.3	10.3
Cysteine[**]	5.6	5.7
Threonine[*]	21.2	19.7
Tryptophan[*]	6.3	6.1
Isoleucine[*]	26.0	21.8
Arginine[*]	32.9	32.8
Phenylalanine[*]	27.6	22.0
Histidine[*]	14.5	13.4
Leucine[*]	45.2	37.0
Tyrosine[**]	21.6	19.1
Valine[*]	29.7	25.5

Alanine	22.2		34.1	
Asparagine	50.8		44.2	
Glutamate	106.0		78.7	
Glycine	18.9		26.3	
Proline	40.1		38.3	
Serine	27.0		21.0	
* essential, ** conditionally essential				
Analysed (composition)				
Size of feed	**200µm**	**200-300µm**	**200µm**	**200-300µm**
DM ($g.kg^{-1}$)	960	956	961	948
ASH ($g.kg^{-1}$.dm)	164	159	185	184
CP ($g.kg^{-1}$.dm)	573	586	556	572
EE ($g.kg^{-1}$.dm)	122	120	103	156
GE	20.9	21.0	20.0	21.3
CP/GE	27.4	27.9	27.8	26.9

[1]Percentage of novel protein meal: Recipes are isoenergetic. Composition of diet B and E are equal in macronutrients, amino acids, calcium and phosphates. The novel protein meal contains 17 percent of fat (ether extract) of which the composition is comparable to that of fish oil as the novel protein meal is made of a marine invertebrate. Calculated omega 3 content of diet B and E were 21.3 and 19.5 $g.kg^{-1}.dm^{-1}$ respectively. [a] Ingredient is not specified because of confidentiality reasons of ongoing research; [b] Roquette Freres, Lestrem, France; [c] Acid casein 30/60 mesh, Lactalis, Bourgbarré, France; [d] Soycomil R ADM Eurpoort BV, the Netherlands; [e] Coppens International, the Netherlands; [f] Damolin A/S, Hamburg, Germany; [g] Melis Suikerunie, Dinteloord, the Netherlands; [h] Inducal 250, Sibelco/Ankerpoort, Maastricht, the Netherlands; [i] Gluvital 21000, Cargill, Bergen op Zoom, the Netherlands; [j] Binder1, [k] Binder2 and [m] Binder3: Ingredients are not specified because of confidentiality reasons of ongoing research; [l] Animalfeed salt, Kloek zout, the Netherlands, [n] Betafin, Danisco Animal Nutrition Marlborough UK; [†] vitamins (mg or IU kg^{-1} diet) include: vitamin A (retinyl acetate), 2.4 mg, 8000 IU; vitamin D3 (cholecalciferol), 0.04 mg, 1700 IU; vitamin K3 (menadione sodium bisulfite), 10 mg; vitamin B1 (thiamine), 8 mg; vitamin B2 (riboflavin), 20 mg; vitamin B6, [2] vitamin B_{12} (cyanocobalamin) 0.02mg (pyridoxine hydrochloride), 10 mg; folic acid, 6 mg; biotin, 0.7 mg; inositol, 300 mg; niacin, 70 mg; pantothenic acid, 30 mg, choline, 1500 mg; vitamin C, 500 mg; vitamin E, 300 mg; Minerals (g or mg kg^{-1} diet): Mn (manganese oxide), 20 mg; I (potassium iodide), 1.5 mg; Cu (copper sulphate), 5mg; Co (cobalt sulphate), 0.1 mg; Mg (magnesium sulphate), 500 mg; Zn (zinc oxide), 30 mg; Se (sodium selenite), 0.3 mg; Fe (Iron Sulfate), 60 mg; Calcium carbonate, 2150 mg; Dicalcium phosphate, 5000 mg; Potasium Chloride, 1000mg; Antixoidant BHT (E300-321), 100 mg; Anti-fungal Calcium propionate, 1000 mg. Abbreviations: DM= dry matter; CP= crude protein; EE= ether extract; C= calcium; P= phosphate; GE= gross energy.

Table 2: Bacterial OTUs that contribute most to the difference between diet E and diet B at day 21.

#OTU ID	ZF21BR [%]	ZF21ER [%]	ZF21ER-ZF21BR	Av.Diss	Diss/SD	Contrib [%]	Cum [%]	taxon ID
11	0.76	3.24	2.48	4.2	3.69	10.95	10.95	(s) *Plesiomonas shigelloides*
1214	2.62	2.06	-0.56	2.9	1.33	7.55	18.5	(s) *Pseudomonas alcaligenes*
286	1.13	2.43	1.3	2.28	1.16	5.93	24.43	(g) *Acidovorax*
1407	1.66	1.96	0.3	2.22	1.68	5.8	30.23	(f) *Aeromonadaceae*
1306	8.68	7.96	-0.72	1.52	1.49	3.96	34.19	(g) *Pseudomonas*
1394	0.99	0.45	-0.54	0.96	1.3	2.49	36.69	(f) *Aeromonadaceae*
1392	2.22	2.06	-0.16	0.94	1.46	2.45	39.13	(f) *Comamonadaceae*
1538	0.51	0.37	-0.14	0.88	2.37	2.28	41.41	(g) *Vogesella*
441	0.05	0.48	0.43	0.74	2.32	1.93	43.34	(g) *Trabulsiella*
552	0.34	0.42	0.08	0.73	3.68	1.89	45.23	(o) *Legionellales*
1459	0.21	0.42	0.21	0.56	1.42	1.46	46.69	(s) *Novospirillum itersonii*
1381	0.29	0.14	-0.15	0.51	1.27	1.32	48.01	(f) *Comamonadaceae*
1414	0.5	0.58	0.08	0.5	5.57	1.32	49.33	(g) *Rheinheimera*
370	0.37	0.12	-0.25	0.5	1.27	1.31	50.64	(g) *Pseudomonas*
832	0.08	0.37	0.29	0.5	1.26	1.3	51.94	(s) *Propionibacterium acnes*
422	0.27	0	-0.27	0.46	1.75	1.2	53.14	(c) *Gammaproteobacteria*
463	0	0.26	0.26	0.45	6.56	1.18	54.32	(f) *Alcaligenaceae*
1086	0.25	0.07	-0.18	0.42	1.09	1.1	55.42	(f) *Rhodobacteraceae*
8	0	0.24	0.24	0.42	3.77	1.09	56.51	(f) *Sphingobacteriaceae*
859	0.32	0.14	-0.18	0.4	1.4	1.05	57.56	(o) *Aeromonadales*
16	0.13	0.36	0.23	0.39	5.99	1.02	58.58	(g) *Burkholderia*
1182	0.15	0.37	0.22	0.38	4.28	1	59.58	(g) *Halomonas*
1127	0.31	0.1	-0.21	0.37	1.73	0.96	60.54	(p) *Cyanobacteria*
576	0.2	0	-0.2	0.34	0.86	0.9	61.43	(f) *Comamonadaceae*
1278	0	0.2	0.2	0.34	14.24	0.89	62.33	(g) *Achromobacter*
1167	0.18	0	-0.18	0.32	3.31	0.82	63.15	(f) *Pseudomonadaceae*
568	0	0.18	0.18	0.31	0.87	0.82	63.97	(g) *Janibacter*
1117	0.18	0	-0.18	0.31	0.86	0.82	64.79	(p) *Proteobacteria*
152	0.05	0.18	0.13	0.3	1.13	0.79	65.58	(c) *Gammaproteobacteria*
1260	0	0.17	0.17	0.3	3.77	0.77	66.35	(g) *Cupriavidus*
1016	0.05	0.21	0.16	0.29	1.53	0.75	67.1	(f) *Halomonadaceae*
1015	0.17	0	-0.17	0.28	21.34	0.74	67.84	(f) *Comamonadaceae*
1133	0.17	0	-0.17	0.28	21.34	0.74	68.57	(f) *Pseudomonadaceae*
414	0.07	0.17	0.1	0.28	1.23	0.74	69.31	(c) *Gammaproteobacteria*
288	0.16	0.07	-0.09	0.27	1.17	0.72	70.03	(f) *Pseudomonadaceae*

Only OTUs that cumulatively contribute to 50% of the difference between bacterial profiles between diet B and E are shown based on SIMPER analysis. Average relative abundances of square root-transformed data for diet B and E is indicated in the columns ZF21B and ZF21E, respectively. "Av.Diss." indicates the average dissimilarity between the diets for each OTU; "Diss/SD" indicates the dissimilarity divided by the standard deviation; "Contrib" is the relative contribution to the difference between diet B and E; and "Cum" represents the cumulative relative contribution to the difference starting from the top with the OTUs that contribute most to the difference. OTUs in grey refer to OTUs that were more abundant for diet E. OTUs were classified up to the phylum (p), class (c), order (o), family (f), genus (g) or species (s) level.

872 **Table 3: Functional gene groups and their differentially expressed genes.**
873

Gene Group 1	***ribosome components and activity***		
Enrichment Score: 14.61			
ENSEMBL_GENE_ID	**Gene Name**	**fc**	**pval**
ENSDARG00000034291	ribosomal protein L37; hypothetical LOC100000999	1.85	0.002
ENSDARG00000043509	ribosomal protein L11	1.80	0.003
ENSDARG00000051783	ribosomal protein, large, P0	1.70	0.006
ENSDARG00000077291	ribosomal protein S2	1.69	0.006
ENSDARG00000030602	ribosomal protein S19	1.67	0.008
ENSDARG00000009285	ribosomal protein L15	1.64	0.011
ENSDARG00000034897	ribosomal protein S10	1.64	0.012
ENSDARG00000035871	ribosomal protein L30	1.62	0.013
ENSDARG00000042389	zgc:171772	1.58	0.019
ENSDARG00000036875	ribosomal protein S12	1.58	0.018
ENSDARG00000070849	ribosomal protein S15	1.57	0.021
ENSDARG00000046119	ribosomal protein S3	1.55	0.023
ENSDARG00000030408	ribosomal protein S26, like	1.54	0.027
ENSDARG00000037071	ribosomal protein S26	1.54	0.028
ENSDARG00000057556	zgc:65996	1.53	0.028
ENSDARG00000053058	ribosomal protein S11	1.51	0.033
ENSDARG00000035692	ribosomal protein S3A	1.51	0.034
ENSDARG00000055475	zgc:73262	1.50	0.043
ENSDARG00000020197	ribosomal protein L5a	1.50	0.037
ENSDARG00000011201	ribosomal protein, large P2, like	1.50	0.044
ENSDARG00000053457	similar to ribosomal protein L23; ribosomal protein L23	1.49	0.043
ENSDARG00000037350	similar to ribosomal protein L9; ribosomal protein L9	1.49	0.040
ENSDARG00000041435	ubiquitin A-52 residue ribosomal protein fusion product 1	1.49	0.043
ENSDARG00000023298	zgc:109888	1.49	0.042
ENSDARG00000014867	ribosomal protein L8	1.48	0.041
ENSDARG00000015490	ribosomal protein L24	1.48	0.050
ENSDARG00000025073	ribosomal protein L18a	1.47	0.044
ENSDARG00000013012	ribosomal protein L36	1.47	0.045

Gene Group 2	***transport***		
Enrichment Score: 1.31			
ENSEMBL_GENE_ID	**Gene Name**	**fc**	**pval**
ENSDARG00000013855	solute carrier family 12 (sodium/chloride transporters), member 3	24.20	0.000
ENSDARG00000055253	similar to Solute carrier family 12 member 3 (Thiazide-sensitive sodium-chloride cotransporter) (Na-Cl symporter); slc12a10.3 solute carrier family 12 (sodium/potassium/chloride transporters), member 10.3	3.24	0.020
ENSDARG00000013743	solute carrier family 12 (sodium/potassium/chloride transporters), member 10.1; solute carrier family 12, member 2-like	2.48	0.033
ENSDARG00000053853	solute carrier family 13 (sodium-dependent dicarboxylate transporter), member 2	1.56	0.024

Gene Group 3	***WD40 repeats acting as protein-protein interaction sites***		
Enrichment Score: 0.22			
ENSEMBL_GENE_ID	**Gene Name**	**fc**	**pval**
ENSDARG00000075883	BUB3 budding uninhibited by benzimidazoles 3 homolog (yeast)	1.97	0.027
ENSDARG00000045019	zgc:85939	1.63	0.026
ENSDARG00000021557	wdr45 like	1.60	0.026
ENSDARG00000041619	guanine nucleotide binding protein (G protein), beta polypeptide 2-like 1	1.48	0.043

Gene Group 4 ***nucleotide binding***
Enrichment Score: 0.09

ENSEMBL_GENE_ID	Gene Name	fc	pval
ENSDARG00000055385	zgc:110560; hypothetical protein LOC100150958	Inf	0.021
ENSDARG00000074873	similar to Serine/threonine-protein kinase Pim-3; si:dkey-108d22.5	Inf	0.001
ENSDARG00000052900	zgc:153642	7.94	0.000
ENSDARG00000002210	zgc:92836	2.29	0.005
ENSDARG00000041411	RAD51 homolog (RecA homolog, E. coli) (S. cerevisiae)	2.05	0.040
ENSDARG00000002344	zgc:55461; zgc:123194; zgc:153264; zgc:123292; tubulin, beta 2c; zgc:153426	1.61	0.019
ENSDARG00000031164	tubulin, alpha 8 like 2	1.49	0.040
ENSDARG00000040984	heat shock protein 13	0.56	0.038
ENSDARG00000056443	zgc:152753	0.41	0.038
ENSDARG00000015134	similar to calcium/calmodulin-dependent protein kinase kinase 1, alpha; zgc:194737	0.27	0.049
ENSDARG00000035178	guanine nucleotide binding protein (G protein), alpha 14	0.19	0.035
ENSDARG00000014373	vasa homolog	0.18	0.022
ENSDARG00000030644	guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 3	0.07	0.049
ENSDARG00000078093	zgc:172065; hypothetical LOC100001153	0.00	0.012

Gene Group 5 ***transcription***
Enrichment Score: 0.02

ENSEMBL_GENE_ID	Gene Name	fc	pval
ENSDARG00000070769	forkhead box G1	Inf	0.042
ENSDARG00000069337	Mediator of RNA polymerase II transcription subunit 11	Inf	0.000
ENSDARG00000079406	homeo box C11a; homeo box C11b	14.81	0.041
ENSDARG00000058133	forkhead box D1	2.22	0.024
ENSDARG00000043531	v-jun sarcoma virus 17 oncogene homolog (avian)	1.81	0.033
ENSDARG00000040253	one cut domain, family member, like	0.48	0.024
ENSDARG00000056407	interferon regulatory factor 8	0.35	0.034

Gene Group 6 ***metal binding/zinc fingers***
Enrichment Score: 0.02

ENSEMBL_GENE_ID	Gene Name	fc	pval
ENSDARG00000056907	zgc:173949; hypothetical protein LOC100006493; similar to retinoblastoma-binding protein 6	26.61	0.026
ENSDARG00000070864	B-cell CLL/lymphoma 6 (zinc finger protein 51)	2.83	0.005
ENSDARG00000021677	similar to PHD finger protein 6	2.37	0.016
ENSDARG00000013279	zgc:153635	2.02	0.023
ENSDARG00000028476	zgc:65779; hypothetical LOC791614	1.90	0.003
ENSDARG00000001897	zgc:110815	1.76	0.032
ENSDARG00000043323	ligand of numb-protein X 1	1.73	0.041
ENSDARG00000071558	zgc:154176	1.71	0.012
ENSDARG00000008218	zgc:77303	1.69	0.029
ENSDARG00000038006	odd-skipped related 2 (Drosophila)	1.67	0.048
ENSDARG00000014794	zgc:92453	1.52	0.045
ENSDARG00000055475	zgc:73262	1.50	0.043
ENSDARG00000023298	zgc:109888	1.49	0.042

Gene Group 7 ***membrane***
Enrichment Score: 0.006

ENSEMBL_GENE_ID	Gene Name	fc	pval
ENSDARG00000044990	zgc:65811	2.54	0.000

ENSDARG00000076899	solute carrier family 2 (facilitated glucose transporter), member 13a	2.06	0.024
ENSDARG00000055307	synaptophysin-like 2a	0.63	0.043
ENSDARG00000019137	translocating chain-associating membrane protein 1	0.62	0.027
ENSDARG00000059824	hypothetical LOC564868; zgc:153102	0.44	0.019
ENSDARG00000027065	zgc:165543; similar to Sodium-coupled neutral amino acid transporter 3 (Na(+)-coupled neutral amino acid transporter 3) (System N amino acid transporter 1) (N-system amino acid transporter 1) (Solute carrier family 38 member 3)	0.34	0.008

874

875 Shown are each of the functional gene groups, enrichment score and name; the Ensemble gene
876 ID and gene name; fold change and P-values for differentially expressed genes ($P \leq 0.05$).

Figure 1.TIF

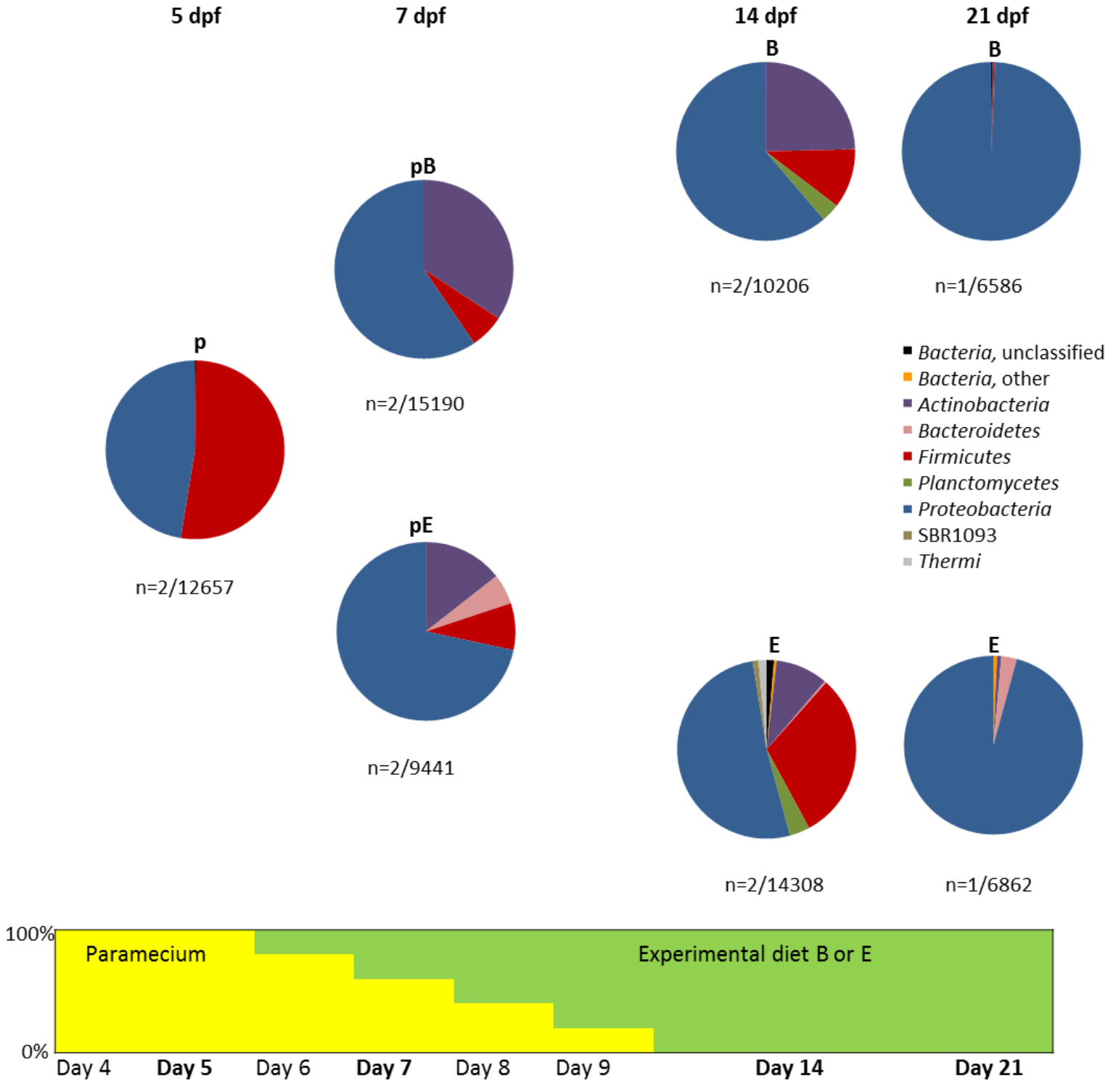

5 dpf
7 dpf
14 dpf
21 dpf
B
B
pB
p
pE
E
E
n=2/10206
n=1/6586
n=2/15190
n=2/12657
n=2/9441
n=2/14308
n=1/6862
Bacteria, unclassified
Bacteria, other
Actinobacteria
Bacteroidetes
Firmicutes
Planctomycetes
Proteobacteria
SBR1093
Thermi
100%
0%
Paramecium
Experimental diet B or E
Day 4
Day 5
Day 6
Day 7
Day 8
Day 9
Day 14
Day 21

Figure 2.TIF

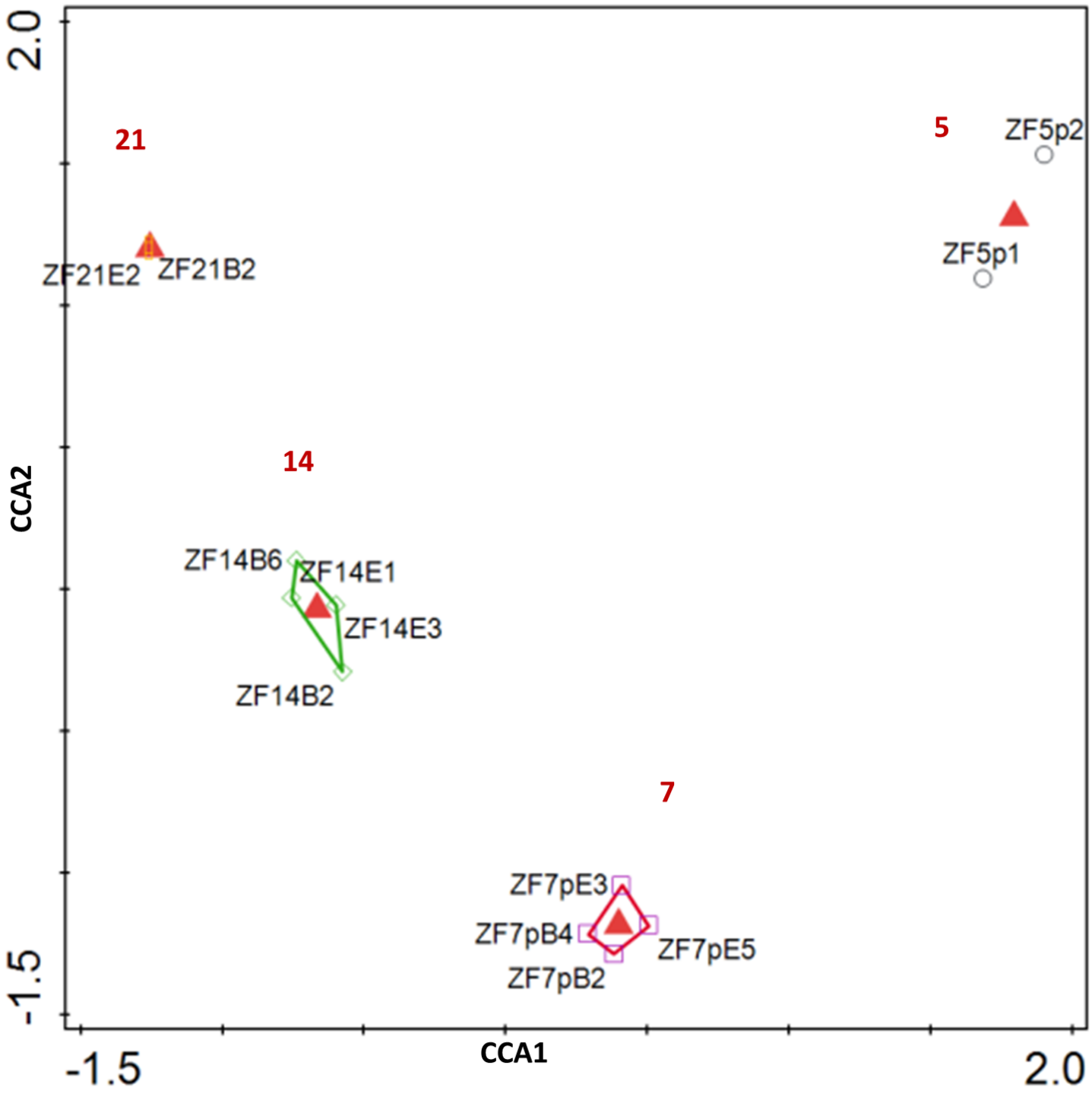
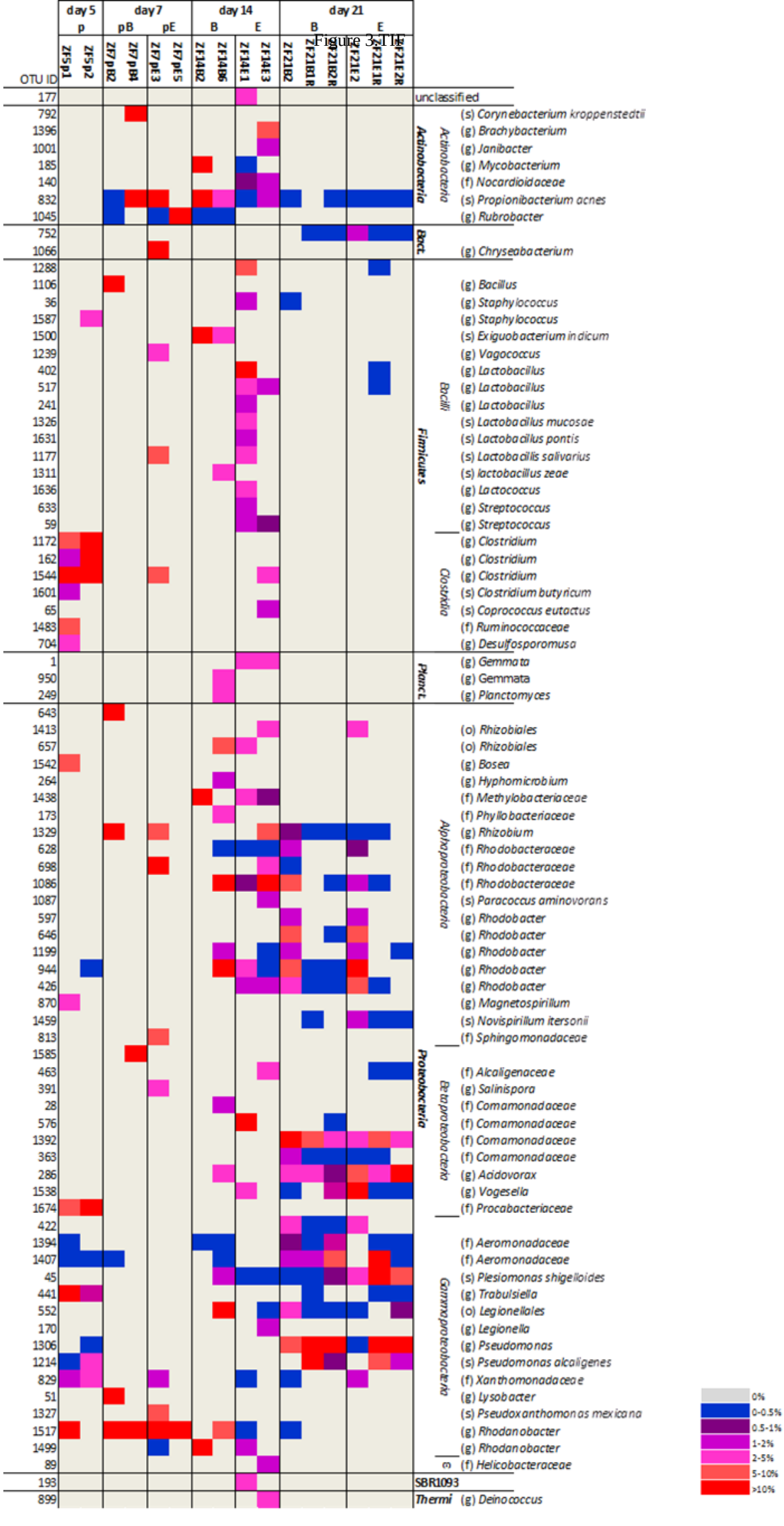

2.0
-1.5
CCA2
21
ZF21E2
ZF21B2
5
ZF5p2
ZF5p1
14
ZF14B6
ZF14E1
ZF14E3
ZF14B2
7
ZF7pE3
ZF7pB4
ZF7pE5
ZF7pB2
-1.5
CCA1
2.0

Figure 3. TIF


day 5
p
day 7
pB
pE
day 14
B
E
day 21
B
E
OTU ID
ZFS1
ZFS2
ZF7dB2
ZF7dB4
ZF7pE3
ZF7pE5
ZF14B2
ZF14B6
ZF14E1
ZF14E3
ZF21B2
ZF21B1R
ZF21B2R
ZF21E2
ZF21E1R
ZF21E2R
unclassified
(s) Corynebacterium kroppenstedtii
(g) Brachybacterium
(g) Janibacter
(g) Mycobacterium
(f) Nocardioidaceae
(s) Propionibacterium acnes
(g) Rubrobacter
Actinobacteria
Actinobacteria
Bact.
(g) Chryseobacterium
(g) Bacillus
(g) Staphylococcus
(g) Staphylococcus
(s) Exiguobacterium indicum
(g) Vagococcus
(g) Lactobacillus
(g) Lactobacillus
(g) Lactobacillus
(s) Lactobacillus mucosae
(s) Lactobacillus pontis
(s) Lactobacillis salivarius
(s) lactobacillus zeae
(g) Lactococcus
(g) Streptococcus
(g) Streptococcus
Bacilli
Firmicutes
(g) Clostridium
(g) Clostridium
(g) Clostridium
(s) Clostridium butyricum
(s) Coprococcus eutactus
(f) Ruminococcaceae
(g) Desulfosporomusa
Clostridia
(g) Gemmata
(g) Gemmata
(g) Planctomyces
Planct.
(o) Rhizobiales
(o) Rhizobiales
(g) Bosea
(g) Hyphomicrobium
(f) Methylobacteriaceae
(f) Phyllobacteriaceae
(g) Rhizobium
(f) Rhodobacteraceae
(f) Rhodobacteraceae
(f) Rhodobacteraceae
(s) Paracoccus aminovorans
(g) Rhodobacter
(g) Rhodobacter
(g) Rhodobacter
(g) Rhodobacter
(g) Rhodobacter
(g) Magnetospirillum
(s) Novispirillum itersonii
(f) Sphingomonadaceae
Alphaproteobacteria
(f) Alcaligenaceae
(g) Salinispora
(f) Comamonadaceae
(f) Comamonadaceae
(f) Comamonadaceae
(f) Comamonadaceae
(g) Acidovorax
(g) Vogesella
(f) Procabacteriaceae
Betaproteobacteria
Proteobacteria
(f) Aeromonadaceae
(f) Aeromonadaceae
(s) Plesiomonas shigelloides
(g) Trabulsiella
(o) Legionellales
(g) Legionella
(g) Pseudomonas
(s) Pseudomonas alcaligenes
(f) Xanthomonadaceae
(g) Lysobacter
(s) Pseudoxanthomonas mexicana
(g) Rhodanobacter
(g) Rhodanobacter
(f) Helicobacteraceae
Gammaproteobacteria
ε
SBR1093
Thermi
(g) Deinococcus
0%
0-0.5%
0.5-1%
1-2%
2-5%
5-10%
>10%