Deciphering the gut microbiome of grass carp through multi-omics approach

Ming Li[1], Hui Liang[1], Hongwei Yang[1], Qianwen Ding[2], Rui Xia[1], Jie Chen[1], Wenhao Zhou[1], Yalin Yang[2], Zhen Zhang[2], Yuanyuan Yao[2], Chao Ran[2*] and Zhigang Zhou[1*]

Abstract

Background Aquaculture plays an important role in global protein supplies and food security. The ban on antibiotics as feed additive proposes urgent need to develop alternatives. Gut microbiota plays important roles in the metabolism and immunity of fish and has the potential to give rise to novel solutions for challenges confronted by fish culture. However, our understanding of fish gut microbiome is still lacking.

Results We identified 575,856 non-redundant genes by metagenomic sequencing of the intestinal content samples of grass carp. Taxonomic and functional annotation of the gene catalogue revealed specificity of the gut microbiome of grass carp compared with mammals. Co-occurrence analysis indicated exclusive relations between the genera belonging to *Proteobacteria* and *Fusobacteria/Firmicutes/Bacteroidetes*, suggesting two independent ecological groups of the microbiota. The association pattern of *Proteobacteria* with the gene expression modules of fish gut and the liver was consistently opposite to that of *Fusobacteria*, *Firmicutes*, and *Bacteroidetes*, implying differential functionality of *Proteobacteria* and *Fusobacteria/Firmicutes/Bacteroidetes*. Therefore, the two ecological groups were considered as two functional groups, i.e., Functional Group 1: *Proteobacteria* and Functional Group 2: *Fusobacteria/Firmicutes/Bacteroidetes*. Further analysis revealed that the two functional groups differ in genetic capacity for carbohydrate utilization, virulence factors, and antibiotic resistance. Finally, we proposed that the ratio of "Functional Group 2/Functional Group 1" can be used as a biomarker that efficiently reflects the structural and functional characteristics of the microbiota of grass carp.

Conclusions The gene catalogue is an important resource for investigating the gut microbiome of grass carp. Multi-omics analysis provides insights into functional implications of the main phyla that comprise the fish microbiota and shed lights on targets for microbiota regulation.

Keywords Gut microbiota, Gene modules, Gene catalogue, Grass crap

*Correspondence:
Chao Ran
ranchao@caas.cn
Zhigang Zhou
zhouzhigang03@caas.cn
[1] China-Norway Joint Lab On Fish Gastrointestinal Microbiota, Institute of Feed Research, Chinese Academy of Agricultural Sciences, Beijing 100081, China
[2] Biotechnology of the Ministry of Agriculture and Rural Affairs, Institute of Feed Research, Chinese Academy of Agricultural Sciences, Beijing 100081, China

Background

Fish consumption accounts for 1/6 of the world's animal protein intake (FAO, 2020). Due to the limited resources of capture fisheries, aquaculture has become the main way to improve the global supply of fish products [1, 2]. The limitations of production factors (land, feed, etc.) and aquaculture environmental stresses (pathogens, parasites, etc.) have resulted in continuous challenges to the high-efficiency and green development of aquaculture [3–5]. Moreover, the ban on antibiotics as feed additive proposes urgent need to develop alternatives.

Studies of fish microbiome have the potential to give rise to novel solutions to challenges confronted by the aquaculture industry [6]. The important role of the commensal microbiota in immune homeostasis, disease, and health has been demonstrated in humans [7–9], and a comprehensive gut microbiome gene catalogue was established [10, 11]. Studies in fish have also found that commensal microbiota plays important roles for host metabolism and immunity [12]. Regarding metagenomic study, a gene catalogue of the gut microbiome of zebrafish was constructed using metagenomics, including 1,569,102 non-redundant genes, which provides resources for gut microbiome-related research of this model animal [13]. Metagenomic study of salmonid-related *Mycoplasma* species revealed adaption to salmonid host and specific functions to benefit the host, such as biosynthesis of essential amino acids and metabolism of B vitamins [14–16]. However, as farmed animals, the metagenomic studies in fish lagged behind those in livestock and poultry. For instance, a metagenomic study in goat constructed 719 high-quality metagenome-assembled genomes (MAGs) and revealed their functions in the production of short-chain fatty acids (SCFAS) [17]. Microbial function was discovered for fiber digestion in the rumen, and a potential cross talk between microbiome and host cells was confirmed in dairy cows by metagenomic sequencing [18]. A gut microbial gene catalogue was constructed in chicken, and metagenomic analysis provided insights into the growth-promoting effect of *Macleaya cordata* extract (MCE) [19]. In contrast, the microbial gene catalogue has not yet been constructed in economic fish species and functions of main phyla in the aspect of host-interaction remains unclear. Weighted gene co-expression network analysis (WGCNA) has been used in correlations between host gene sets and factors (such as phenotypic traits and environmental factors) [20], which helps to identify key relationships between gene co-expression modules and microbial taxa [21].

Diet is a key factor that influences the structure and function of the gut microbiota [22]. Feeds in aquaculture have been shifting from animal proteins derived from marine resources to plant proteins [4]. Replacement of animal proteins by plant proteins is common in aquaculture, giving rise to formulations with differential percentage of animal versus plant protein sources. Thus, carnivorous (animal protein dominated), omnivorous (relatively balanced in animal and plant proteins), and herbivorous (plant protein dominated) diets have become distinctive features of feeds in aquaculture, and microbial alteration associated with the three dietary types is representative when investigating the structural and functional characteristics of fish microbiota.

Grass carp (*Ctenopharyngodon idella*), belonging to the Cyprinidae family, is the most important freshwater farmed fish species in China. At present, the production value of grass carp ranks the first among cultured freshwater fish globally. In this study, we investigate the structural and functional characteristics of the gut microbiome of grass carp. Metagenome sequencing of microbiota was conducted in fish fed carnivorous, omnivorous, and herbivorous diets. The gene catalogue was established, and the functional implications of key microbial taxa in the aspect of host interaction were evaluated by multi-omics approach.

Materials and methods

Experimental diets

According to the nutritional requirements of NRC (2011) and Wang et al. [12], the experimental formulations of different diets for grass carp were designed with equal nitrogen and lipid levels (Table 1). In brief, soy protein concentrate and wheat gluten protein were the protein sources for the herbivorous diet (HD), and casein and gelatin were the protein sources for the carnivorous diet (CD). Equal proportions of protein sources in the carnivorous and herbivorous diets make up the omnivorous diet (OD). Microcrystalline cellulose has been added as a dietary fiber for herbivorous fish. Soybean oil is applied as a fat source and adapted to the high-fat level feeding levels in aquaculture.

Experimental design and sample collection

Grass carp with an initial weight of about 20 ± 0.28 g were selected and maintained in the culture system for 3 weeks. Before the formal experiment, the fish were treated with the CD diet for 1 week. Then, 108 fish of the same-sized fish were randomly allocated to three tanks and fed three times a day by satiety feeding, with one of the three diets CD/OD/HD for 3 weeks (Supplementary Fig. 1A). During the experiment, the water temperature was maintained at 26 ± 2 °C. At weeks 1 and 3, 18 fish were randomly selected from each dietary group to form six biological replicates, with each replicate consisting of a mixture of three fish to ensure sufficient intestinal content samples for shotgun sequencing. The intestinal contents and tissues were sampled 4–6 h after feeding. Specifically, after anesthesia with MS-222, the hindgut was quickly stripped to collect the contents. In addition, the corresponding hindgut and liver tissues were collected and stored. All samples were quickly frozen directly in liquid nitrogen and finally transferred to -80 °C.

Table 1 Experimental formulations for different diets of grass carp

Ingredient (g/kg diet)	Carnivorous diet	Omnivorous diet	Herbivorous diet
Soybean protein concentrate	0	210.32	420.64
Vital gluten	0	44	88
Casein	304	152	0
Gelatin	76	38	0
Vital gluten	310	265	220
Soybean oil	84.2	82.7	81.1
Lysine	0	3.69	7.37
Methionine	7.24	9.245	11.25
VC phosphate	1	1	1
Vitamin premix[a]	2	2	2
Mineral premix[b]	2	2	2
Monocalcium phosphate	20	20	20
Choline chloride	2	2	2
Microcrystalline cellulose	80	80	80
Zeolite powder	111.56	88.045	64.64
Total	1000	1000	1000
Conventional nutrients			
Crude protein (%)	35.27	35.25	35.95
Crude lipid (%)	8.47	8.27	8.64
Crude ash (%)	4.72	4.06	3.94
Dry matter (%)	95.28	95.94	96.06

[a] Containing the following (g/kg vitamin premix): thiamine, 0.438; riboflavin, 0.632; pyridoxine, 0.908; d-pantothenic acid, 1.724; nicotinic acid, 4.583; biotin, 0.211; folic acid, 0.549; vitamin B12, 0.001; inositol, 21.053; menadione sodium bisulfite, 0.889; retinyl acetate, 0.677; cholecalciferol, 0.116

[b] Containing the following (g/kg mineral premix): $CoCl_2.6H_2O$, 0.074; $CuSO_4.5H_2O$, 2.5; $FeSO_4.7H_2O$, 73.2; NaCl, 40.0; $MgSO_4.7H_2O$, 284.0; $MnSO_4.H_2O$, 6.50; KI, 0.68; Na_2SeO_3, 0.10; $ZnSO_4.7H_2O$, 131.93; cellulose, 501.09

DNA extraction and metagenome sequencing

Total DNA was extracted from all intestinal content samples using the CTAB method. The integrity of the extracted DNA was measured by electrophoresis on a 1% agarose gel, and the concentration of DNA was determined using Qubit (Qubit™ dsDNA HS Assay Kit, Invitrogen, USA). High-quality metagenomic libraries (average size of DNA constructs 420–580 bp) were then constructed according to the manufacturer's instructions of the library preparation kit (VAHTS® Universal Plus DNA Library Pren Kit for Illumina). Sequencing of metagenomic libraries was performed on the Illumina NovaSeq 6000 sequencing technology platform (Illumina, San Diego, CA, USA) in paired-end 150-bp mode (PE150).

RNA extraction, transcriptome sequencing, and data analysis

RNA of intestinal and liver samples was extracted according to the operating instructions of TRIzol reagent (Invitrogen, USA). Subsequently, nucleic acid concentration was determined in NanoDrop 2000 (Thermo Fisher Scientific, Waltman, MA, USA), and integrity testing was performed using Agilent Bioanalyzer 2100 system, Hieff NGS® Ultima Dual-mode mRNA Library Prep Kit for Illumina® (cat. no.13533ES96) was used for library construction. The libraries were tested by Qsep-400 method and then sequenced via the Illumina NovaSeq 6000 technology platform (150 bp paired-end) using the NovaSeq 6000 S4 Reagent Kit (Illumina, San Diego, CA, USA). A total of 466.18 Gb Clean Data was obtained from 72 transcriptome sequencing samples. After the clean data were mapped onto the reference genome (http://www.ncgr.ac.cn/grasscarp/files/C_idella_female_scaffolds.fasta.v1.gz) [23] by HISAT2 (version 2.0.4,–dta -p 6 –max-intronlen 5,000,000) [24], the transcript was assembled and using String-Tie (Version V1.3.4d,–merge -F 0.1 -T 0.1) [25]. Gene expression was quantified by FPKM method [26]. Fold-Change > 1.5 and p-value < 0.05 were considered to be differentially expressed genes (DEG) using edgeR (version3.8.6) [27]. Enrichment analysis of KEGG pathways was performed using clusterProfiler (version 3.10.1), and q-value < 0.05 was considered to be a significant enrichment. Default parameters were used for unlisted parameters.

Metagenomic assembly and non-redundant gene catalogue construction

The raw tags were filtered using Trimmomatic (version 0.33) to obtain high-quality sequencing data (CleanTags) with the parameters (LEADING:3 TRAILING:3 SLIDINGWINDOW:50:20 MINLEN:100), and the bowtie2 (version 2.2.4) was used to perform sequence alignment with the host genome removing host contamination. Metagenome assembly was performed using the software MEGAHIT (Version 1.1.2), and contig sequences shorter than 300 bp were filtered [28]. The QUAST (Version 2.3) software was used to evaluate the assembly results [29]. MetaGeneMark (Version 3.26) software was used to identify coding regions in the genome using default parameters [30]. Redundancy was removed using MMseqs2 software (Version 12-113e3) using a similarity threshold set to 95% and a coverage threshold set to 90% to construct non-redundant gene catalogue [31, 32].

Taxonomic profiling, functional annotation, and microbial data analysis

The protein sequences from the non-redundant gene set were aligned to the NCBI-nr database (2019–03) to obtain annotated taxonomic information using Diamond software (version 0.9.24) [33] with a threshold of e-value < 1e-05. The NR that could not be classified to any taxa were defined as unknown taxa. In addition, Eukaryota and Metazoa with very low abundance were excluded. Protein sequences from non-redundant gene catalogue were aligned to the KEGG (Kyoto Encyclopedia of Genes and Genomes) database [34] using Diamond software (version 0.9.24). The threshold was e-value < 1e-05, and if there was more than one match hit, the best match was selected as the annotation for that sequence. Similarly, the protein sequence of NR was aligned with eggNOG (version4.0) [35] using Diamond software (version 0.9.24). Carbohydrate-activated enzymes (CAZys) were annotated by aligning the protein sequence of NR to the dbCAN database (HMMdb V8) [36] using HMMER (version3.0). The antibiotic resistance genes (ARGs) were annotated by alignment with the Comprehensive Antibiotic Resistance Database (CARD) [37] using RGI (version4.2.2). Protein sequences of NR were annotated by the virulence factor annotation database (VFDB) [38] using BLAST (version2.2.31 +) software [39].

Abundance calculations for sequencing were referred to the previous method [40]. Based on taxonomic, KO, and CAZys annotations, the relative abundance of phylum, family, genus, species, KO, and CAZys was calculated by summing up the abundance of genes belonging to each category [19]. In addition, the abundance of virulence factor and antibiotic resistance genes in each functional group was calculated from the number of related genes normalized to the number of total NCBI nr-annotated genes in the functional group. The ratio of Functional Group 2/Functional Group 1 = the sum of the relative abundance of *Fusobacteria*, *Firmicutes* and *Bacteroidetes* in the sample/the relative abundance of *Proteobacteria* in this sample. Based on the R language (v3.1.1) and python2, alpha-diversity (including Ace, Chao1, Shannon, Simpson index) analysis was applied via the picante package (v1.8.2), principal coordinate analysis (PCoA) applied to assess the beta diversity of gut bacterial communities based on python2 (cogent, v1.5.3), the vegan package (v2.3–0) was applied for the analysis of PERMANOVA/ANOSIM, and the pheatmap package (v1.0.2) was applied for the plotting of heat maps. Venn diagram (v1.6.9) was applied for the analysis of vennd, mothur (v1.22.2) was applied for the analysis of rarefaction curve, and matplotlib (v1.5.1) was applied for the analysis of taxonomic composition.

Metagenome-assembled genomes

Metagenomic binning was applied to assembled data by using three different algorithms: MetaBAT2 (version2.12) [41], MaxBin (version2.2.6) [42], and CONOCOCT (version1.0.0) [43]. Subsequently, the software DAS_Tool (version1.1.2 −search_engine diamond −write_bins 1 −score_threshold 0) was used to integrate the results of the different metagenomic binning software [44]. The integrated metagenomic bins (or metagenome-assembled genomes, MAGs) were evaluated using checkM (version1.1.3, default parameters) [45] software. High-quality bins were defined by selecting completeness ≥ 80 contamination ≤ 10 [46]. MAGs were clustered into species-level genomic bins (SGBs) using dRep (version3.0.3) [47] with a threshold of 95% ANI. Finally, the GTDB-Tk (v1.2.0) software was used to annotate bins for taxonomic classification by reference to the Genome Taxonomy Database (GTDB) [48]. SGBs containing at least one MAGs in the GTDB were considered to be known SGBs, otherwise, the SGB was considered as unknown [49].

Weighted gene co-expression network analysis

We performed weighted gene co-expression network analysis (WGCNA) using WGCNA package [50] on the platform BMKCloud (www.biocloud.net). Firstly, the expression of all genes in the gut and the liver was used to construct gene co-expression modules. Secondly, Pearson correlations between module eigengenes (MEs) and ecological groups were calculated.

Co-occurrence network analysis

The relative abundance of the core microbiota was applied to construct a matrix of correlations, followed by random matrix theory to determine the threshold of

correlation. The correlation network data was visualized through the igraph (version 1.2.7) package.

Statistical analysis

The unpaired *t*-test was used to assess the differences between the two groups, and the Tukey's multiple comparisons test was used to analyze the three or more groups. The plots and diagrams were displayed by ggplot2 (2.2.1) using R language. GraphPad Prism (version 8.0) software was applied for graphing.

Results

Establishment and assessment of grass carp microbiome gene catalogues

Shotgun metagenomics yielded 476,855,093 valid reads (Table S1). A total of 2,099,912 contigs were assembled after quality control and de-hosting, and 575,856 non-redundant (NR) genes were identified with an average length of 1.65 kb (Table S2). Rarefaction analysis of all samples showed curves approaching saturation (at sequences number of 372,000–5308000) (Supplementary Fig. 1B). A total of 374,843 NR genes (65.09%) can be blasted to the NCBI-nr database (Table S3). A total of 64.88% of the NR genes could be taxonomically classified. Among them, 97.28% were assigned to bacteria, with the remaining genes being assigned to viruses (1.41%), fungi (1.42%), or archaea (0.17%) (Fig. 1A). At the phylum level, most of the annotated genes (53.74%) belonged to *Proteobacteria*, followed by *Bacteroidetes* (16.99%), *Firmicutes* (13.81%), and *Fusobacteria* (6.86%) (Fig. 1B). Our further analysis revealed that 48.71% of the NR genes were assigned to bacterial genera, predominantly in *Aeromonas* (17.04%), *Bacteroides* (6.94%), *Shewanella* (5.33%), and *Cetobacterium* (5.10%) (Fig. 1C).

Using KEGG and eggNOG for function classification, 187,972 (32.64%) and 329,336 (57.19%) NR genes were annotated with KEGG orthologous groups (KOs) (Table S4) and eggNOG orthologous groups (OGs) (Table S5). The proportion of genes annotated to KOs or OGs is lower than that in mammals [51], reflecting that gut commensal bacteria were less studied in fish. The KEGG profiles showed similarities in gut microbial functions of grass carp compared with human and pig. However, the genes for carbohydrate metabolism, amino acid metabolism, nucleotide metabolism, and energy metabolism are more abundant in human and pig, while the genes for cell motility and cellular community are more abundant in the fish microbiota (Fig. 1D) [19]. These results suggest a deviation of the gut microbial function of fish compared with mammals, with a general lower genetic capacity for
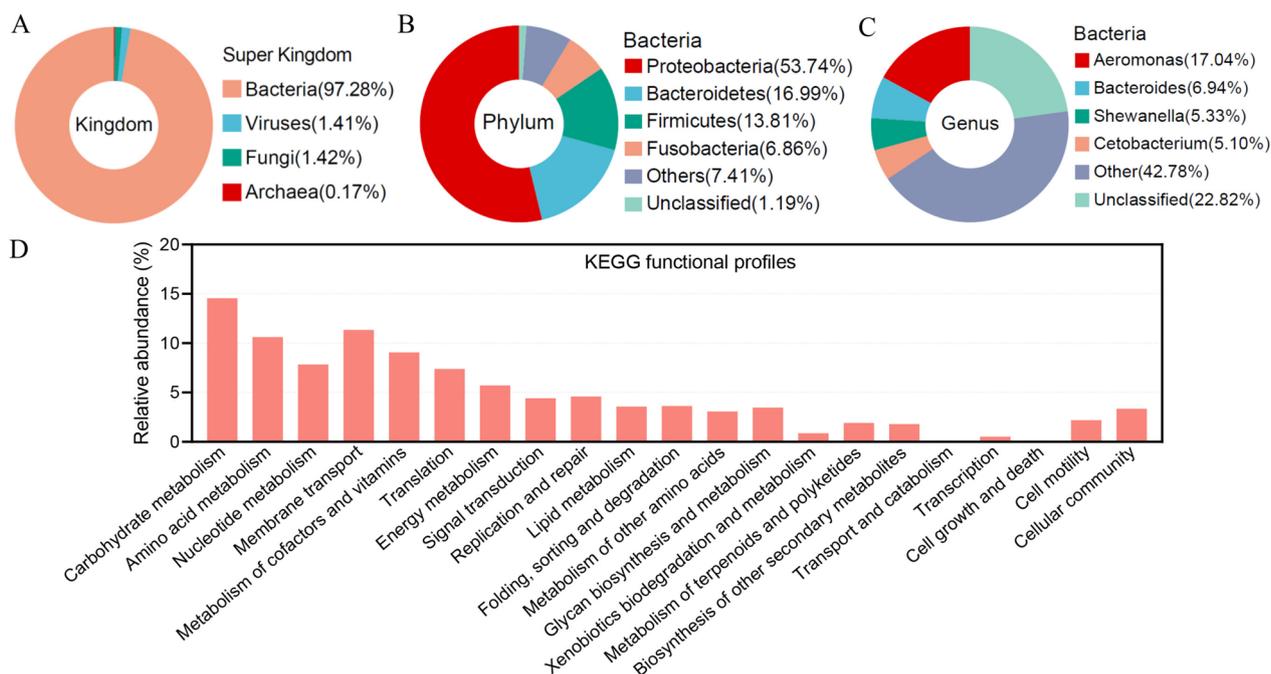
Fig. 1 Establishment and analysis of grass carp gut microbiome gene catalogues. **A** Taxonomic annotation of the grass carp gut gene catalog at the super kingdom level. Only non-redundant genes that could be classified from the NCBI nr database were involved in the analysis. **B** and **C** The NR genes assigned to bacteria were taxonomically annotated at the phylum (**B**) and genus (**C**) level. **D** KEGG functional files of the grass carp gut gene catalogue. Genes without functional annotation were excluded, and relative gene abundance analysis was performed at the second level of the KEGG

nutrient metabolism but higher capacity for microbe-microbe interactions.

Reconstruction of microbial genomes from metagenomic sequencing data has been reported in humans [52], ruminants [46], and pigs [51, 53] by using metagenomic binning methods. We constructed metagenome-assembled genomes (MAGs) based on metagenomic data of grass carp and assembled a total of 129 high-quality MAGs (completeness > 80% and contamination < 10%) (Table S6). The 129 MAGs were further organized into species-level genome bins (SGBs) by average nucleotide identity (ANI) threshold of 95%. This resulted in a total of 18 SGBs. Fifteen SGBs were without any publicly annotatable genomes and were defined as unknown SGBs (uSGB). Subsequent taxonomic annotation by the Genome Taxonomy Database Toolkit (GTDB-Tk) revealed that the 18 SGBs are mainly from *Proteobacteria*, *Fusobacteria*, *Bacteroidetes*, and *Firmicutes* (Table S7).

The ecological interactions of gut microbiota of grass carp

The top 30 genera contributed more than 80% to the total abundance of the microbiota (Supplementary Fig. 2A and B), while the VENN analysis showed that 23 genera were shared among different dietary groups and time points (Supplementary Fig. 2C). Further analysis among the taxonomically annotated genera revealed that shared genera contributed 79.4–89.20% of the total abundance (Fig. 2A), indicating that shared genera dominated the microbial community and thus can be considered as core genera. Co-occurrence network analysis was performed to further explore the ecological interactions between the core genera. At week 1, co-exclusive patterns were observed between the core genera belonging to *Proteobacteria* (except *Acinetobacter*) and *Fusobacteria*, *Bacteroidetes*, or *Firmicutes* (Fig. 2B). Similarly, the genera of *Proteobacteria* (except *Photobacterium*, *Acinetobacter*, *Pseudomonas*) showed exclusive relations with one or more of the genera of *Bacteroidetes*, *Firmicutes* (except *Enterococcus*), or *Fusobacteria* at week 3 (Fig. 2C). Similarly, analysis at the phylum level showed that there were mutually exclusive patterns between *Proteobacteria* and *Fusobacteria*, *Firmicutes*, or *Bacteroidetes* (Fig. 2D and E). Thus, the results indicated mutually exclusive relations between the core genera belonging to *Proteobacteria* and those belonging to *Fusobacteria*, *Firmicutes*, and *Bacteroidetes*, suggesting two independent ecological groups of the intestinal microbiota of grass carp: *Proteobacteria* and *Fusobacteria/Firmicutes/Bacteroidetes*.

Gut microbiota was associated with host gene modules

In order to explore the functional characteristics of the gut microbiota, correlation analysis of the gut microbiota and host gene modules was conducted. WGCNA analysis revealed 24 and 31 gene modules of gut expressed genes at weeks 1 and 3, respectively (Supplementary Fig. 2D and F). In the liver, 31 and 25 gene modules were obtained at weeks 1 and 3, respectively (Supplementary Fig. 2E and G).

We further investigated the association of the two ecological groups (*Proteobacteria* and *Fusobacteria/Firmicutes/Bacteroidetes*) with gene modules (M) of host by Pearson analysis. Gene modules were differentiated by colors (Fig. 3). The association pattern of *Proteobacteria* with the gene modules of gut and the liver at week 1 was consistently opposite to that of *Fusobacteria*, *Firmicutes*, and *Bacteroidetes* (Fig. 3A and B), and the opposite association was also observed at week 3 (Fig. 3C and D), indicating differential functionality of *Proteobacteria* and *Fusobacteria/Firmicutes/Bacteroidetes* in terms of their interaction with fish host. Therefore, the ecological groups can be considered as two functional groups, i.e., Functional Group 1: *Proteobacteria* and Functional Group 2: *Fusobacteria/Firmicutes/Bacteroidetes*.

KEGG analysis revealed that gut gene modules significantly associated with the microbiota mainly included functions of recognition of microorganisms by the immune system (e.g., Toll-like receptor signaling pathway, NOD-like receptor signaling pathway, and RIG-I-like receptor signaling pathway) and the downstream cellular processes (e.g., apoptosis, phagosome, cell adhesion molecules, and ferroptosis) at week 1 (Fig. 3E). In the liver, gene modules with significant association with microbiota involved functions of nutrient metabolism (e.g., propanoate metabolism, fatty acid metabolism, fatty acid biosynthesis, and biosynthesis of unsaturated fatty acids), the immune system (e.g., NOD-like receptor signaling pathway and RIG-I-like receptor signaling pathway), and the endocrine systems (e.g., PPAR signaling pathway and adipocytokine signaling pathway) (Fig. 3E). Similarly, the gut gene modules included functions of the immune system (e.g., Toll-like receptor signaling pathway, NOD-like receptor signaling pathway and RIG-I-like receptor signaling pathway), nutrient metabolism pathways (e.g., fatty acid biosynthesis, fatty acid metabolism and fat digestion, and absorption), and the endocrine system pathway (PPAR signaling pathway) at week 3 (Fig. 3F), while in the liver, the gene modules involved nutritional metabolic pathways (e.g., purine metabolism, pyrimidine metabolism, and primary bile acid biosynthesis) and endocrine systems involved in metabolism (insulin signaling pathway, PPAR signaling pathway, and adipocytokine signaling pathway) (Fig. 3F). Taken together, the two functional groups were associated with host nutrient metabolism and immunity.
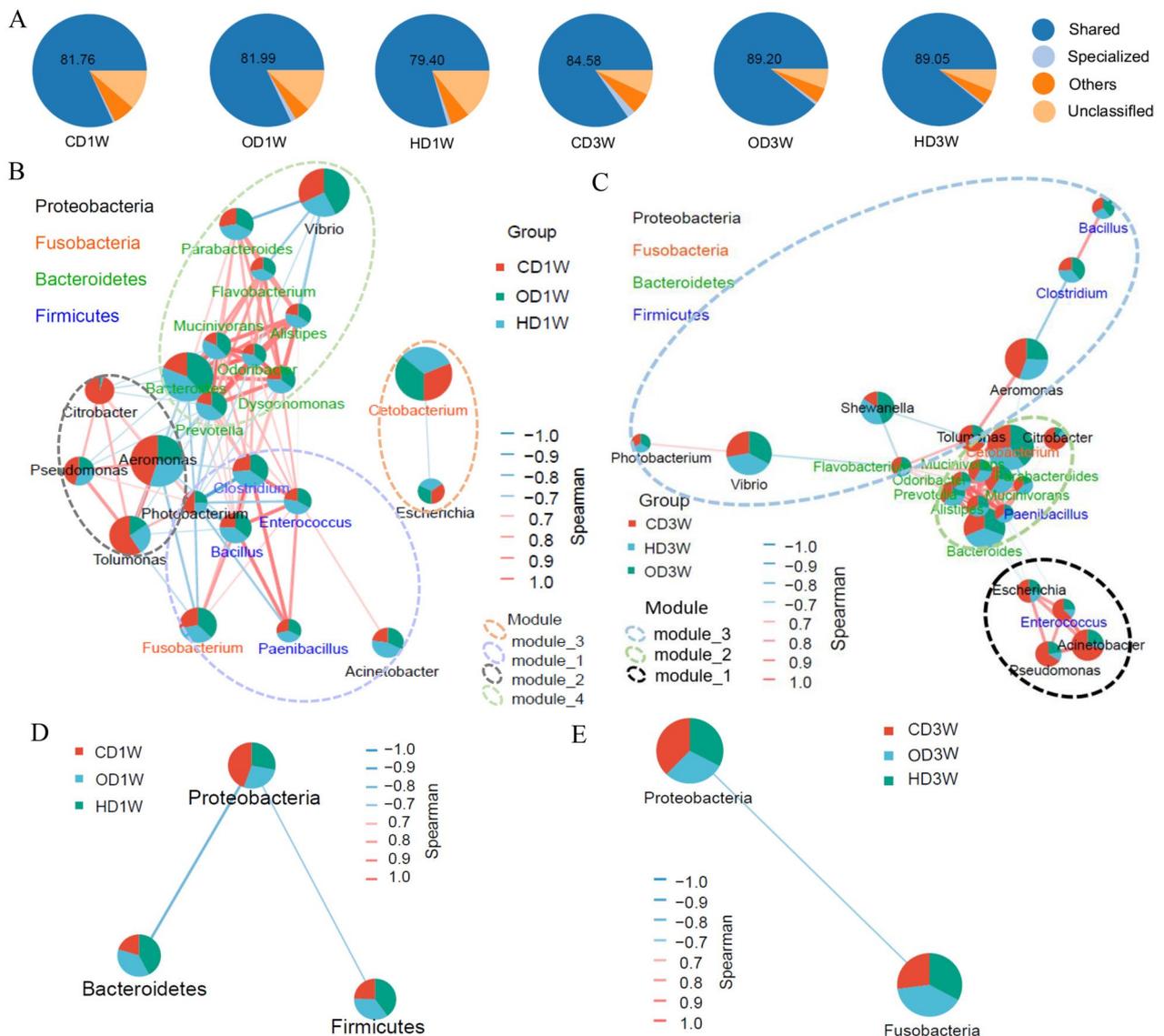
Fig. 2 Co-occurrence network analysis of the core genera of grass carp gut microbiota. **A** The relative abundance of shared genera in the microbiota. **B** and **C** Co-occurrence network of core genera at weeks 1 and 3. **D** and **E** Co-occurrence network of microbiota at the phylum level at weeks 1 and 3. The connection strength threshold is Spearman's r > 0.5, and correlation is considered as significant when *p*-value < 0.05. The size of the connection point represents the relative abundance of a specific microbe. Red lines indicate co-occurrence. Green lines indicate co-exclusion, and the thickness of the line shows the strength of the correlation. The dotted ellipse indicates a specific modular unit

The two functional groups differ in genetic capacity for carbohydrate utilization, virulence factors, and antibiotic resistance

The gut microbiota utilizes carbohydrates by carbohydrate-active enzymes (CAZy) and produces SCFAs, which are beneficial to the host in both nutritional and immune aspects [19, 54, 55]. In contrast, virulence factors and antibiotic resistance are generally negative factors that may exert detrimental effect [56–58]. We therefore analyzed the genes encoding CAZy, virulence

factors, and antibiotic resistance harbored by the two functional groups. The results showed that members of Functional Group 2 enriched CAZy genes encoding enzymes degrading arabinoxylan, pectin, mucin, inulin, and cellulose compared with Functional Group 1, and only the starch-related CAZy gene family was enriched in Functional Group 1 (Fig. 4A and B).

Furthermore, members of Functional Group 1 encoded more virulence factor (VF) genes compared with Functional Group 2 (Fig. 4C). In particular, VF
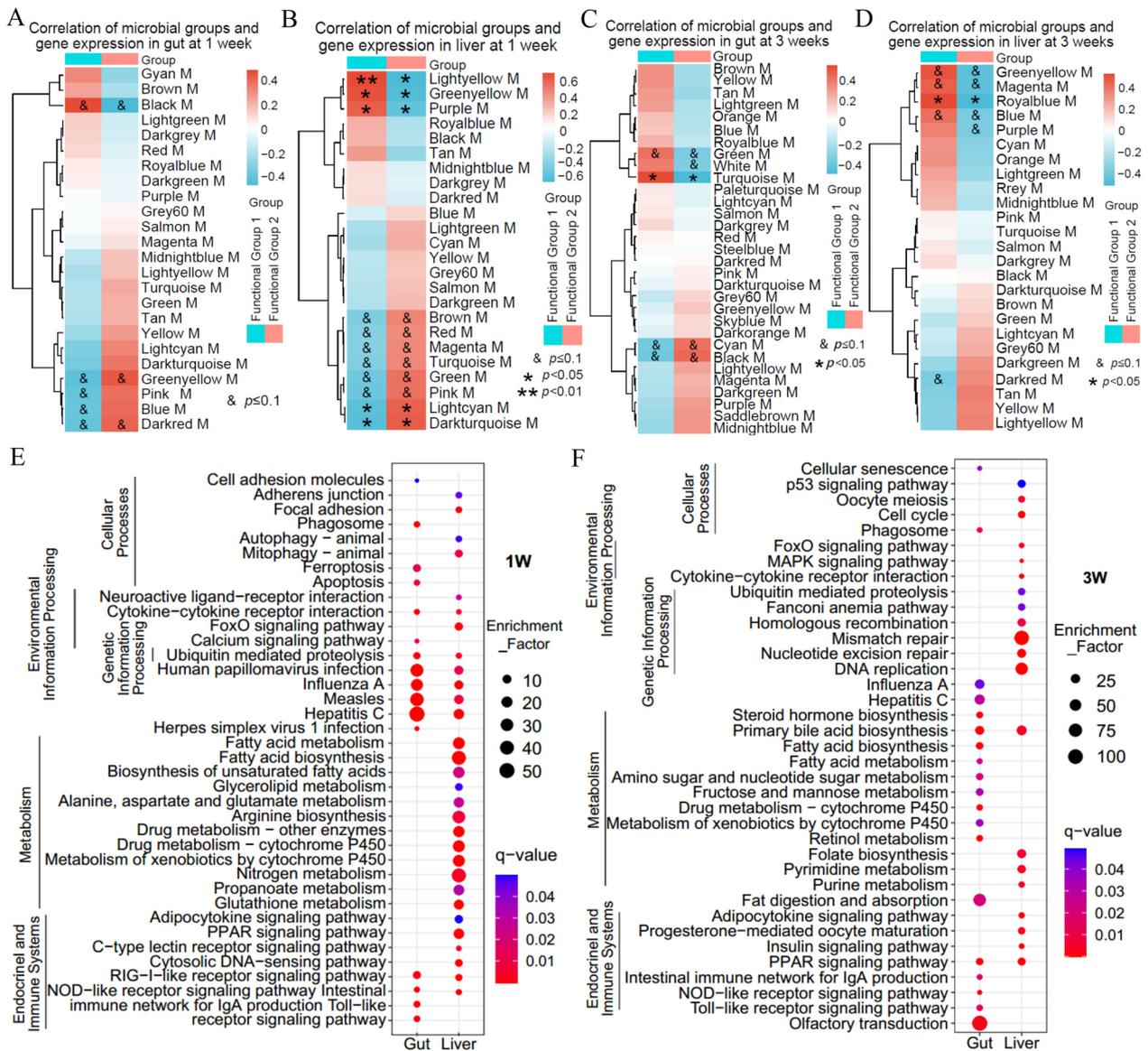
Fig. 3 Gut microbiota is associated with host gene modules. **A**, **B**, **C**, and **D** Weighted gene co-expression network analyses (WGCNA) was performed to identify co-expressed gene modules (M). Gene modules are clusters of highly interconnected genes, which are designated by color codes (such as black M). The association of the two ecological groups with host gene modules was investigated by Pearson analysis. **A** and **B** represent association with gut and liver gene modules at week 1, and **C** and **D** represent association with gut and liver gene modules at week 3. Heatmaps show the correlation between gene modules and the two ecological groups of microbiomes. Clustering was performed by a complete clustering method using Euclidean distances. KEGG pathways enriched in the gene modules correlated with the functional groups (**E** and **F**). Enriched bubble plots have been shown for pathways with *q*-values < 0.05. Five or six biological replicates were included during analysis, \$*p* < 0.1; *p < 0.05; **p < 0.01; ***p < 0.001

genes involved in antiphagocytosis, endotoxin, and manganese are enriched in Functional Group 2, while Functional Group 1 enriched VF genes across 12 different VF classes (Fig. 4D), suggesting that Functional Group 1 may exert more negative phenotypes in the host-microbes interaction. In terms of antibiotic resistance genes (ARG), the overall number of ARGs is higher in Functional Group 1 versus Functional Group 2 (Fig. 4E). Functional Group 2 encoded more resistance genes against oxazolidinone, macrolide, rifamycin, bicyclomycin, lincosamide, and streptogramin, while Functional Group 1 enriched ARGs for resistance to 11 antibiotic classes (Fig. 4F).
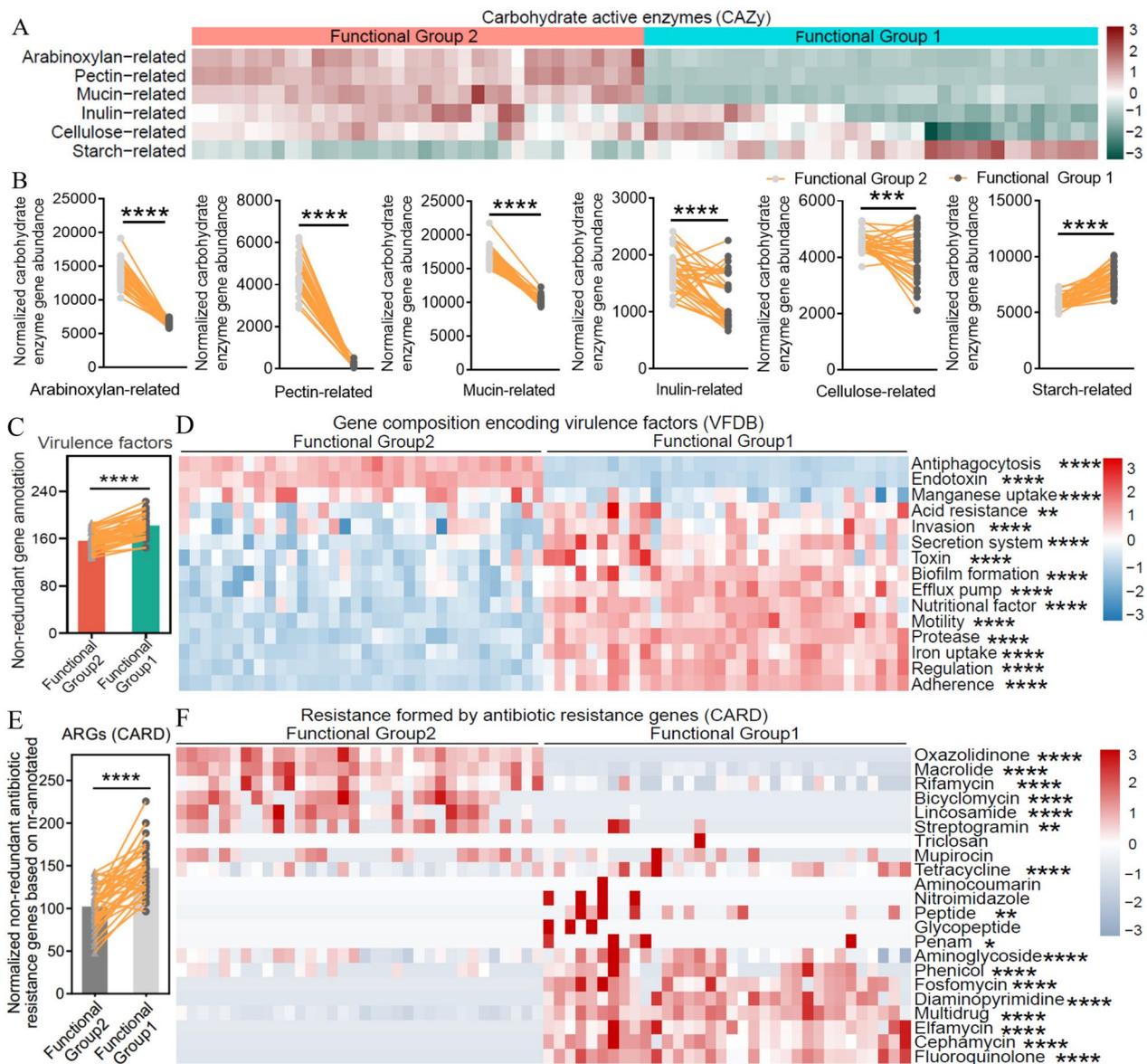
Fig. 4 The two functional groups differ in genetic capacity for carbohydrate utilization, virulence factors, and antibiotic resistance. **A** and **B** Differences in genetic capacity for carbohydrate substrate utilization. **A** Heat map showing the proportion of each carbohydrate-active enzyme (CAZy) category in different functional groups. **B** Abundance of genes involved in carbohydrate substrate utilization. (The abundance of genes annotated to different types of carbohydrates was normalized to the total carbohydrate-active enzyme in each sample. The abundance of total carbohydrate-active enzyme genes in different functional groups was adjusted to 100,000.) Arabinoxylan-related CAZy families, CE1, CE2, CE4, CE6, CE7, GH10, GH11, GH115, GH43, GH51, GH67, GH3, and GH5; pectin-related CAZy families, CE12, CE8, GH28, PL1, and PL9; mucin-related CAZy families, GH1, GH2, GH3, GH4, GH18, GH19, GH20, GH29, GH33, GH38, GH58, GH79, GH84, GH85, GH88, GH89, GH92, GH95, GH98, GH99, GH101, GH105, GH109, GH110, GH113, PL6, PL8, PL12, PL13, and PL21; inulin-related CAZy families, GH32 and GH91; cellulose-related CAZy families, GH1, GH44, GH48, GH8, GH9, GH3, and GH5; starch-related CAZy families, GH13, GH31, and GH97. **C** Number of genes annotated to virulence factors in the two functional groups. **D** The abundance of genes encoding virulence factors in the two functional groups. Genes for each VF class were normalized to the number of total NCBI nr-annotated genes in each functional group. **E** The number of antibiotic resistance genes in the two functional groups. Antibiotic resistance genes were normalized to the total NCBI nr-annotated genes in each functional group (adjusted to 100,000). **F** The abundance of ARGs in the two functional groups. Genes for each resistance class were normalized to the number of total NCBI nr-annotated genes in each functional group. Thirty-four biological replicates were included in each functional group during analysis. The Mann–Whitney test was used to analyze differences between functional groups, $*p < 0.05$; $**p < 0.01$; $***p < 0.001$; $****p < 0.0001$
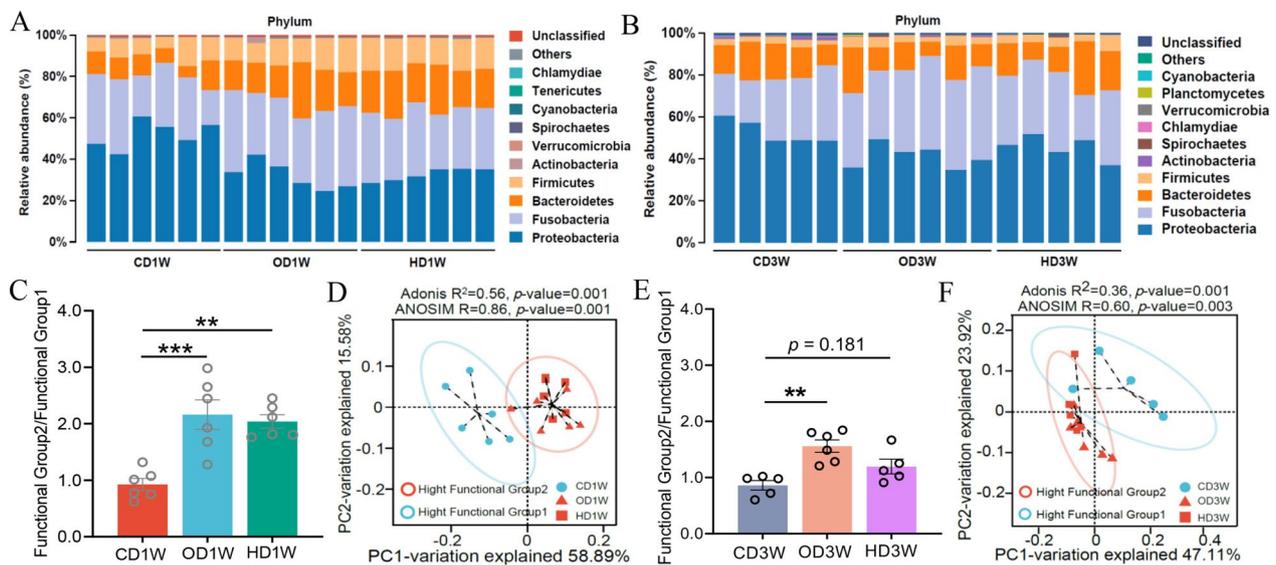
The ratio of "Functional Group 2/Functional Group 1" reflects the structural and functional characteristics of the microbiota

Dietary ingredients are key factors influencing the gut microbiota [22]. We evaluated the effect of diets on the microbial composition of grass carp. Rarefaction curve analysis of the samples indicated saturation of sequencing depth (Supplementary Fig. 3A and C). At week 1, the relative abundance of *Proteobacteria* decreased in OD and HD groups compared with CD group, while the abundance of *Firmicutes* and *Bacteroidetes* increased (Fig. 5A). Similarly, the abundance of *Proteobacteria* decreased, and *Fusobacteria* increased in OD and HD groups compared with CD group at week 3 (Fig. 5B). Diets had no significant influence on the α-diversity of the microbiota including ACE, Chao1, Shannon, and Simpson indices (Supplementary Fig. 3E and F). Principal coordinate analysis (PCoA) revealed a significant alteration of the microbiota due to dietary groups at both weeks 1 and 3. The microbiota of CD clustered alone, while the microbiota of OD and HD clustered together (Supplementary Fig. 3G, H, I, K, and L).

Considering the ecological and functional difference of Functional Group 1 and Functional Group 2, the ratio of the abundance of Functional Group 2 and Functional Group 1, designated as "Functional Group 2/Functional Group 1," was calculated to evaluate the structural and functional characteristics of the microbiota. At week 1, "Functional Group 2/Functional Group

1" was significantly higher in OD/HD diets versus CD diet (Fig. 5C), and a similar trend in "Functional Group 2/Functional Group 1" was observed among the three dietary groups at week 3 (Fig. 5E). PCoA analysis showed that the microbial structure associated with different diets was well explained by "Functional Group2/Functional Group 1," with the microbiota of high ratio (OD/HD) deviating from those of low ratio (CD) at both week 1 (Fig. 5D) and week 3 (Fig. 5F). Thus, the ratio of "Functional Group2/Functional Group 1" can be used as a parameter to evaluate the structural characteristics of the gut microbiota of grass carp.

Furthermore, PCoA analysis showed a robust separation between the microbiota of high ratio groups (OD/HD) and low ratio CD group in terms of the abundance of genes encoding carbohydrate-coding enzymes (ANOSIM, $R=0.85$, $p=0.001$; $R=0.51$, $p=0.007$), and the separation was efficient at both weeks 1 and 3 (Fig. 6A and B). Compared with low ratio microbiota (CD), high ratio microbiota (OD/HD) enriched CAZy genes for arabinoxylan, pectin, and cellulose utilization (Fig. 6C and D). On top of it, the microbiota with high ratio harbored more abundance of the gene encoding the key enzyme (FTHFS) for acetate production. A similar trend was observed for butyrate producing key enzymes (Buk, AtoA/D), while no obvious difference was observed for the propionate producing gene (PcoAt) between the high ratio (OD/HD) and low ratio (CD) groups (Fig. 6E). Together, these results indicated that the microbiota of high ratio groups

Fig. 5 The ratio of "Functional Group 2/Functional Group 1" reflects the structural difference of microbiota associated with different dietary groups. **A** and **B** The relative abundance of top ten phyla of the gut microbiota at weeks 1 and 3. **C** and **E** Ratio of "Functional Group 2/Functional Group 1" at weeks 1 and 3. **D** and **F** Principal coordinate analysis (PCoA) of the microbiota associated with different dietary groups by Bray–Curtis's distance. Samples with "high Functional Group 2" or "high Functional Group 1" were marked with different colors. The dotted ellipse borders represent the 95% confidence interval. Data were expressed as the mean ± SEM (*n* = 5 or 6 biological replicates), **$p < 0.01$; ***$p < 0.001$
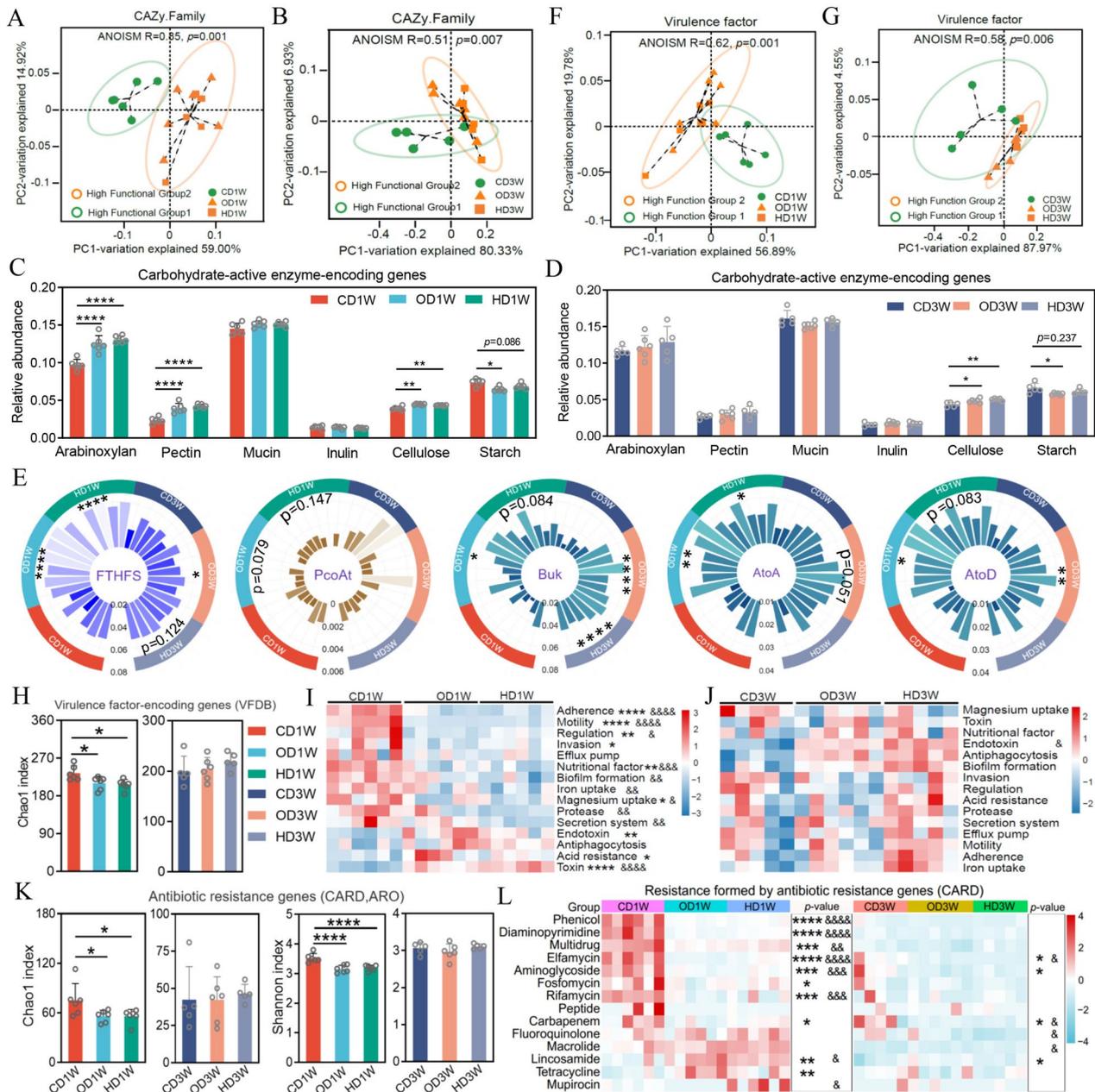
Fig. 6 The ratio of "Functional Group 2/Functional Group 1" reflects the functional characteristics of the microbiota. **A** and **B** PCoA analysis of the abundance of CAZy genes in different dietary groups. Samples with "high Functional Group 2" or "high Functional Group 1" were marked with different colors. **C** and **D** Relative abundance of CAZy genes for a particular substrate in different dietary groups. **E** The relative abundance of key enzyme genes associated with the production of acetate (FTHFS: acetyl-formyltetrahydrofolate synthase/formate-tetrahydrofolate ligase), propionate (PcoAt: propionyl-CoA:succinate-CoA transferase/propionate CoA-transferase), and butyrate (Buk, butyrate kinase; AtoA, acetoacetate CoA transferase beta subunit; and AtoD, acetoacetate CoA transferase alpha subunit) in different dietary groups based on KEGG orthologous groups. **F** and **G** PCoA analysis of genes encoding virulence factors in different dietary groups. Samples with "high Functional Group 2" or "high Functional Group 1" were marked with different colors. **H** Chao1 index analysis of virulence factor gene abundance. **I** and **J** The relative abundance of each class of virulence factor genes in different dietary groups at week 1 (**I**) and week 3 (**J**). **K** The richness and diversity of antibiotic resistance genes were analyzed by Chao1 and Shannon indices, respectively. **L** The relative abundance of each class of antibiotic resistance genes in different dietary groups. ARGs with relative abundance below 1 in 100,000 were excluded in the analysis

had higher functional capability for carbohydrate utilization and SCFAs production.

The expression of GPR43 in gut was higher in high Functional Group 2 diets (OD/HD) compared with Functional Group 1 group (CD) (Supplementary Fig. 4A). In the gut, acetyl-CoA production from pyruvate, amino acids (leucine, isoleucine), and fatty acids was all suppressed in dietary groups with high Functional Group 2 (OD/HD) versus high Functional Group 1 group (CD) (Supplementary Fig. 4B). On the other hand, the expression of ACSS1_2, the key enzyme for acetate assimilation, was upregulated. Similarly, high Functional Group 2 groups (OD/HD) featured reduced acetyl-CoA production from pyruvate, amino acids (leucine, valine, isoleucine), and fatty acids in the liver when compared with high Functional Group 1 group (CD), and the expression of ACSS1_2 was enhanced (Supplementary Fig. 4A). In both gut and the liver (Supplementary Fig. 4A and B), the fatty acids synthesis from acetyl-CoA was enhanced in high Functional Group 2 dietary groups (OD/HD) versus CD group, indicating that the exogenous acetate influenced lipid metabolism. Together, these results indicate that the differential SCFAs production capability of the microbiota with different "Functional Group 2/Functional Group 1" ratio affected host metabolism, and acetate was the main SCFA that exert the influence.

Further analysis of the virulence factors revealed significant structural differences in the genes encoding the virulence factors for dietary groups with different ratio of "Functional Group2/Functional Group 1" (ANOSIM, $R = 0.62$, $p = 0.001$; $R = 0.58$, $p = 0.006$) (Fig. 6F and G). High ratio groups encoded lower abundance of virulence factor genes compared with low ratio group (Fig. 6 H, I, and J). This trend was more obvious at week 1 compared with week 3, which accorded with the reduced difference in the ratio among groups. Regarding antibiotic resistance, high ratio groups (OD/HD) had lower abundance and diversity of ARGs compared with low ratio group (CD) (Fig. 6K and L).

The diversity and composition of the gut microbiota were associated with the diversity and composition of DNA viruses

The nonredundant gene catalogue was annotated through the nr database for the DNA virome, and rarefaction curves indicated that the shotgun metagenomics sequencing depth was adequate to capture the DNA viruses (Supplementary Fig. 5A and B). The analysis revealed that the intestinal DNA virome was dominated by *Caudovirales*, which consisted of *Myoviridae*, *Siphoviridae*, and *Podoviridae* (Supplementary Fig. 5C and D, Table S12). Compared to the CD group, the relative abundance of *Myoviridae* increased, and *Siphoviridae*

decreased in the OD/HD group at both weeks 1 and 3 (Supplementary Fig. 5C and D). At the genus level, the gut DNA virome mainly consisted of *Spn3virus*, *Vhmlvirus*, *Phikzvirus*, *Eah2virus*, and *Machinavirus* (Table S12). Compared to CD, the relative abundance of *Spn3virus*, *Vhmlvirus*, *Machinavirus*, and *Eah2virus* increased, and *Phikzvirus* decreased in the OD/HD dietary groups (Supplementary Fig. 5E and F). Interestingly, the Chao1 and Shannon indexes of the microbiota were strongly positive correlated with DNA virome (Fig. 7A and B; Supplementary Fig. 6A, B, and C). A correlation study of beta diversity showed a strong positive correlation between the structural characteristics of the gut microbial community and DNA viruses (Fig. 7C and D), which was stable at both weeks 1 and 3 (Supplementary Fig. 6D). More specifically, *Aeromonas* and *Tolumonas* showed negative correlations with phage, while *Bacteroides*, *Fusobacterium*, *Clostridium*, and *Bacillus* exhibited positive correlations at week 1, including *Spn3virus*, *Vhmlvirus*, *Phikzvirus*, *Machinavirus* (Fig. 7E). At week 3, *Vibrio*, *Aeromonas*, *Shewanella*, *Acinetobacter*, and *Tolumonas* were found to correlate with phages (Fig. 7F). Therefore, the diversity and structure of gut microbiota were significantly correlated with DNA viruses (mainly phages) in grass carp. Similar results were reported in human [59], indicating a conserved correlation of commensal bacteria and phages.

Discussion

The gene catalogue of grass carp reported in this study comprises 575,856 NR genes, which can serve as a reference for further metagenomic studies of this important economic fish species. To our knowledge, this is the first fish gut microbial gene catalogue for economic fish species. Despite large compositional difference, functional analysis revealed shared functions of fish microbiota with those of mammals, indicating functional redundancy of the microbiota. Nevertheless, compared to mammals, the microbiota of grass carp featured lower abundance of genes related to nutrient metabolism, suggesting lower contribution of the microbiota to host metabolism. It will be interesting to evaluate whether this is a common feature of fish microbiota, which awaits gut metagenomic analysis of more fish species.

While *Firmicutes* constitutes the dominant phylum of fish microbiome in some studies [60, 61], fish microbiota comprises high abundance of *Proteobacteria* in most cases [62], which is also true in this study on grass carp. *Proteobacteria* is known to include many pathogenic or opportunistically pathogenic genera/species and is generally regarded as negative components in the commensal microbiota [63]. However, it is a paradox that this "negative" phylum often occupies high abundance in fish
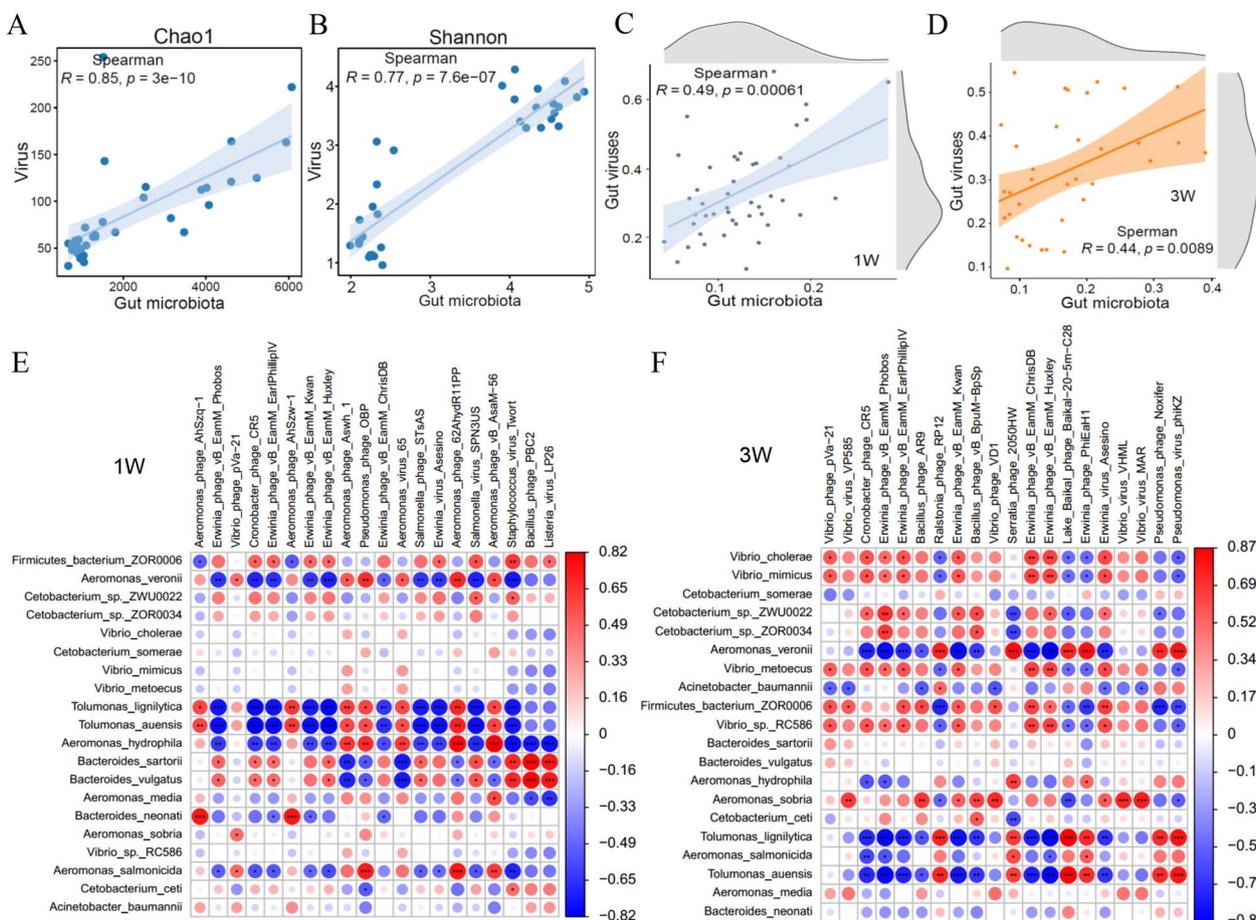
Fig. 7 Diversity and composition of DNA viruses were associated with the diversity and composition of the gut microbiota. **A** and **B** Spearman correlation analysis between gut DNA viral and microbiota alpha diversity at week**s** 1 and 3. **C** and **D** Correlation analysis of beta diversity of intestinal DNA viruses and microbiota (**E** and **F**). The heat map shows the correlation between intestinal DNA viruses (mainly phages) and microbiota at the species level. Correlation was conducted by Spearman's analysis. Red represents positive correlation, blue represents negative correlation**,** and the shade of the color indicates the value of the correlation coefficient

microbiota, and the functionality of *Proteobacteria* in fish has never been assessed in a systematic way. The results in our study indicate that the functional implication of *Proteobacteria* is generally negative, as they harbor less genes for carbohydrate degradation and SCFAs production while encoding more virulence factor and antibiotic resistance genes.

We constructed the co-occurrence network of the core genera and observed that except for a few genera, members from *Proteobacteria* showed negative correlations with the genera belonging to *Firmicutes*, *Fusobacteria*, and *Bacteroidetes*. Therefore, *Proteobacteria* and *Firmicutes/Fusobacteria/Bacteroidetes* form two ecological groups. Further studies revealed consistent opposite association pattern of the two ecological groups with gene modules of host, which prompted us to define them as two functional groups. Consistent with their opposite association patters with the host, the two functional

groups showed difference in genetic capacity for carbohydrate utilization, SCFAs production, virulence factors, and antibiotic resistance. Considering their ecological and functional difference, we proposed that the ratio of "Functional Group 2/Functional Group 1" can be used to evaluate the structural and functional characteristics of grass carp gut microbiota. We found that the ratio was robust in reflecting overall compositional difference of the microbiota associated with different diets. Moreover, the functionality of the microbiota can also be reflected by the ratio. We also found that the ratio of "Functional Group 2/Functional Group 1" can also depict structural characteristics of gut microbiota of other fish species, including zebrafish and largemouth bass (Supplementary Figs. 7 and 8), suggesting potentially extensive application of the ratio in evaluating the gut microbiota of fish species. In mammals, the *Firmicutes/Bacteroidetes* ratio of intestinal microbiota reflects the degree of obesity

[64]. The ratio of "Functional Group 2/Functional Group 1" proposed in this study is a similar way of description of the microbiota. Our preliminary data suggest that the ratio reflects the overall health of fish, with higher ratio correlating with better health (Supplementary Fig. 9 and data not shown), but the specific functional implications of the ratio deserve further detailed studies. Notably, the two functional groups comprise major phyla of the gut microbiome of grass carp, and thus, the ratio is a relatively rough biomarker that depicts the characteristics of the microbiota. Further studies are warranted to investigate the ecological and functional interactions of the core subgroups within the two functional groups, which may give rise to more accurate microbiome signature as reported in human [58].

Metagenomic studies of fish microbiota have been rare. Recently, a series of studies investigated salmonid-related *Mycoplasma* species by metagenomic analysis. Firstly, the studies confirmed that the gut microbiota of Atlantic salmon is low in diversity and dominated by *Mycoplasma*. By constructing MAGs, the studies described the mutualistic relationship and co-diversification between commensal *Mycoplasma* and its salmonid host. Salmonid-related *Mycoplasma* can benefit the host through biosynthesis of essential amino acid and metabolism of B vitamins. Unlike that in salmon, the intestinal microbiota of grass carp showed relatively high diversity, with genera belonging to *Proteobacteria*, *Firmicutes*, *Fusobacteria*, and *Bacteroidetes* constituting considerable abundance in the microbiome. By metagenomic analysis, our study provided knowledge about the overall function of main taxonomic groups of the intestinal microbiota of grass carp. Further studies may uncover the specific functions of key commensal taxa through MAGs construction [14–16]. Studies have confirmed that the gut microbiota is able to ferment dietary fiber to produce short-chain fatty acids (e.g., acetate) that modulate lipid metabolism in the host [54, 65]. In our study, it was found that acetate metabolized by gut microbiota under group 1 diets (e.g., OD diets) was able to activate host GPR43 genes (especially intestinal) and participate in host fatty acid anabolism, which suggests conserved mechanism in the aspect of SCFA-mediated host-microbiota interactions in fish and mammals.

Enterotypes were first reported to exist in humans with robust clustering of gut microbial communities [66, 67]. Enterotype-like structures have also been reported in several animal studies, although their gut microbial composition differs from that of humans. In addition, there has been found to be a correlation between enterotype and host growth traits in pigs [68]. It should be noted that the microbiota variation of grass carp observed in this study is more like to be continuous, which is not a stratified variation that can be depicted by enterotypes. This might be due to that the influence of "artificial diets" on the intestinal microbiota of grass carp is too big, which overwhelmed the possible stratified variation of the microbiota in natural cases. The ratio of "Functional Group 2/Functional Group 1" can reflect the structural and functional features of fish microbiota, and the microbiota associated with different diets and time points can be roughly classified as "high ratio" and "low ratio." However, such classification does not fit the concept of enterotypes. Enterotypes that can robustly cluster the gut microbiota variation in fish deserves further investigation.

Conclusions

The gut microbial gene catalogue of grass carp extended our knowledge about the gut microbiome of this economically important fish species and provided resources for fish gut microbiome-related research. The taxonomic and functional annotation results reflected that the gut commensal bacteria were less studied in fish compared with mammals, highlighting the importance of more fundamental research in this field. We found that *Proteobacteria* is generally negatively correlated with the members of *Fusobacteria/Firmicutes/Bacteroidetes*, and the two groups showed differential functionality in terms of their interaction with fish host. Consistent with their differential functionality, the two functional groups differed in the genetic capacity for carbohydrate utilization, SCFAs production, virulence factors, and antibiotic resistance. Furthermore, the ratio of "Functional Group 2/Functional Group 1" can efficiently reflect the structural and functional characteristics of the gut microbiome of grass carp and could be used as a biomarker to assess the microbiota. Our results provide insights into functional implications of the main phyla that comprise the fish microbiota and shed lights on targets for microbiota regulation, which may promote the development of green inputs for aquaculture that derive from or target the gut microbiota.

Supplementary Information

Additional file 1: Supplementary Figure S1. Experimental design and rarefaction curve analysis of microbial non-redundant genes. **Supplementary Figure S2.** WGCNA analysis to reveal the association between host gene expression and gut microbiota. **Supplementary Figure S3.** Gut microbial composition and diversity of grass carp in different dietary groups. **Supplementary Figure S4.** Microbiota-derived acetate affected host metabolism. **Supplementary Figure S5.** Taxonomic composition of intestinal DNA viruses. **Supplementary Figure S6.** Diversity and composition of DNA viruses were associated with the diversity and composition of the gut microbiota. **Supplementary Figure S7.** The ratio of "Functional Group 2/Functional Group 1" reflects the structural characteristics of the

microbiota in omnivorous fish fed CD, OD and HD diets. **Supplementary Figure S8.** The ratio of "Functional Group 2/Functional Group 1" reflects the structural characteristics of the microbiota in carnivorous fish fed CD, OD and HD diets. **Supplementary Figure S9.** The ratio of "Functional Group 2/Functional Group 1" was associated with the host's inflammatory gene expression. **Supplementary Figure S10.**The diversity of gut DNA virus. **Supplementary Figure S11.** Two functional groups show negative correlation in relative abundance. **Supplementary Figure S12.** Virulence factors gene annotation of two functional groups.

Additional file 2: Table S1. Summary of metagenomic sequencing data. **Table S2.** The identification of Open Reading Fram (ORF) of the metagenomic. **Table S3.** Taxonomic and functional annotation of non-redundant genes. **Table S4.** KEGG annotation of non-redundant genes. **Table S5.** eggNOG annotation of non-redundant genes. **Table S6.** High-quality MAGs. **Table S7.** 18 SGBs were taxonomically annotated through the Genome Taxonomy Database Toolkit. **Table S8.** Microbiota-derived acetate affected gut metabolism. **Table S9.** Microbiota-derived acetate affected liver metabolism. **Table S10.** Metagenomes of identified predominant phyla, families, genera, and species of microbiota. **Table S11.** Metagenomes of identified predominant family, genus, and species of viruses. **Table S12.** Experimental formulations for different diets of Largemouth bass.

Availability of data and materials
The metagenomic data were deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA), with accession number PRJNA945347 (https://dataview.ncbi.nlm.nih.gov/object/PRJNA945347?reviewer=gadm9mf5jfj730s89ssjomd1ag). Raw data on the transcriptome of intestinal-liver tissues can be obtained from the NCBI database under BioProject number PRJNA945233 (https://dataview.ncbi.nlm.nih.gov/object/PRJNA945233?reviewer=roblhqfosr4sirq5p9bb6kiblo). The 16S rRNA sequence raw data from this study were deposited in SRA under BioProject number PRJNA945365 (https://dataview.ncbi.nlm.nih.gov/object/PRJNA945365?reviewer=raga6pq9te5raqohuerluo6i3j). One-hundred twenty-nine MAGs were deposited in Figshare (https://doi.org/10.6084/m9.figshare.22354015.v1).

Declarations

Ethics approval and consent to participate
All experimental procedures were performed in accordance with the "Guidelines for Experimental Animals" of the Ministry of Science and Technology (Beijing, China). The study was reviewed and approved by the Feed Research Institute of the Chinese Academy of Agricultural Sciences Animal Care Committee under the auspices of the China Council for Animal Care. All efforts were made to minimize suffering.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

References
1. Wenning R. The State of World Fisheries and Aquaculture (Sofia) 2020 report. Integr Environ Asses. 2020;16(5):800–1.
2. Gephart JA, Golden CD, Asche F, Belton B, Brugere C, Froehlich HE, Fry JP, Halpern BS, Hicks CC, Jones RC, et al. Scenarios for global aquaculture and its role in human nutrition. Rev Fisheries Sci Aquaculture. 2020;29(1):122–38.
3. Zhang W, Belton B, Edwards P, Henriksson PJG, Little DC, Newton R, Troell M. Aquaculture will continue to depend more on land than sea. Nature. 2022;603(7900):E2–4.
4. Naylor RL, Hardy RW, Buschmann AH, Bush SR, Cao L, Klinger DH, Little DC, Lubchenco J, Shumway SE, Troell M. A 20-year retrospective review of global aquaculture. Nature. 2021;591(7851):551–63.
5. Naylor RL, Kishore A, Sumaila UR, Issifu I, Hunter BP, Belton B, Bush SR, Cao L, Gelcich S, Gephart JA, et al. Blue food demand across geographic and temporal scales. Nat Commun. 2021;12(1):5413.
6. Limborg MT, Alberdi A, Kodama M, Roggenbuck M, Kristiansen K, Gilbert MTP. Applied hologenomics: feasibility and potential in aquaculture. Trends Biotechnol. 2018;36(3):252–64.
7. Brussow H. Problems with the concept of gut microbiota dysbiosis. Microb Biotechnol. 2020;13(2):423–34.
8. Byndloss MX, Pernitzsch SR, Baumler AJ. Healthy hosts rule within: ecological forces shaping the gut microbiota. Mucosal Immunol. 2018;11(5):1299–305.
9. Rejeski JJ, Wilson FM, Nagpal R, Yadav H, Weinberg RB. The impact of a Mediterranean diet on the gut microbiome in healthy human subjects: a pilot study. Digestion. 2022;103(2):133–40.
10. Human Microbiome Jumpstart Reference Strains C, Nelson KE, Weinstock GM, Highlander SK, Worley KC, Creasy HH, Wortman JR, Rusch DB, Mitreva M, Sodergren E, et al. A catalog of reference genomes from the human microbiome. Science. 2010;328(5981):994–9.
11. Zou YQ, Xue WB, Luo GW, Deng ZQ, Qin PP, Guo RJ, Sun HP, Xia Y, Liang SS, Dai Y, et al. 1,520 reference genomes from cultivated human gut bacteria enable functional microbiome analyses. Nat Biotechnol. 2019;37(2):179-+.
12. Wang AR, Zhang Z, Ding QW, Yang YL, Bindelle J, Ran C, Zhou ZG. Intestinal Cetobacterium and acetate modify glucose homeostasis via parasympathetic activation in zebrafish. Gut Microbes. 2021;13(1):1–15.
13. Gaulke CA, Beaver LM, Armour CR, Humphreys IR, Barton CL, Tanguay RL, Ho E, Sharpton TJ. An integrated gene catalog of the zebrafish gut microbiome reveals significant homology with mammalian microbiomes. bioRxiv 2020.
14. Rasmussen JA, Villumsen KR, Duchêne DA, Puetz LC, Delmont TO, Sveier H, Jørgensen LvG, Præbel K, Martin MD, Bojesen AM, et al. Genome-resolved metagenomics suggests a mutualistic relationship between mycoplasma and salmonid hosts. Commun Biol. 2021;4(1):579.
15. Rasmussen JA, Villumsen KR, Ernst M, Hansen M, Forberg T, Gopalakrishnan S, Gilbert MTP, Bojesen AM, Kristiansen K, Limborg MT. A multi-omics approach unravels metagenomic and metabolic alterations of a probiotic and synbiotic additive in rainbow trout (Oncorhynchus mykiss). Microbiome. 2022;10(1):21.
16. Rasmussen JA, Kiilerich P, Madhun AS, Waagbø R. Lock E-JR, Madsen L, Gilbert MTP, Kristiansen K, Limborg MT: Co-diversification of an intestinal mycoplasma and its salmonid host. ISME J. 2023;17(5):682–92.
17. Peng XF, Wilken S, Lankiewicz TS, Gilmore SP, Brown JL, Henske JK, Swift CL, Salamov A, Barry K, Grigoriev IV, et al. Genomic and functional analyses of fungal and bacterial consortia that enable lignocellulose breakdown in goat gut microbiomes. Nat Microbiol. 2021;6(4):499-+.
18. Xue MY, Wu JJ, Xie YY, Zhu SL, Zhong YF, Liu JX, Sun HZ. Investigation of fiber utilization in the rumen of dairy cows based on metagenome-assembled genomes and single-cell RNA sequencing. Microbiome. 2022;10(1):11.
19. Huang P, Zhang Y, Xiao K, Jiang F, Wang H, Tang D, Liu D, Liu B, Liu Y, He X, et al. The chicken gut metagenome and the modulatory effects of plant-derived benzylisoquinoline alkaloids. Microbiome. 2018;6(1):211.

20. Zhong C, Chen C, Gao X, Tan C, Bai H, Ning K. Multi-omics profiling reveals comprehensive microbe-plant-metabolite regulation patterns for medicinal plant Glycyrrhiza uralensis Fisch. Plant Biotechnol J. 2022;20(10):1874–87.

21. Gu W, Liu S, Chen L, Liu Y, Gu C, Ren HQ, Wu B. Single-cell RNA sequencing reveals size-dependent effects of polystyrene microplastics on immune and secretory cell populations from zebrafish intestines. Environ Sci Technol. 2020;54(6):3417–27.

22. Kolodziejczyk AA, Zheng D, Elinav E. Diet-microbiota interactions and personalized nutrition. Nat Rev Microbiol. 2019;17(12):742–53.

23. Wang Y, Lu Y, Zhang Y, Ning Z, Li Y, Zhao Q, Lu H, Huang R, Xia X, Feng Q, et al. The draft genome of the grass carp (Ctenopharyngodon idellus) provides insights into its evolution and vegetarian adaptation. Nat Genet. 2015;47(6):625–31.

24. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015;12(4):357–60.

25. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol. 2015;33(3):290–5.

26. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods. 2008;5(7):621–8.

27. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139–40.

28. Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics. 2015;31(10):1674–6.

29. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29(8):1072–5.

30. Zhu WH, Lomsadze A, Borodovsky M. Ab initio gene identification in metagenomic sequences. Nucleic Acids Res. 2010;38(12):e132.

31. Mirdita M, Steinegger M, Soding J. MMseqs2 desktop and local web server app for fast, interactive sequence searches. Bioinformatics. 2019;35(16):2856–8.

32. Steinegger M, Soding J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nat Biotechnol. 2017;35(11):1026–8.

33. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat Methods. 2015;12(1):59–60.

34. Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M. The KEGG resource for deciphering the genome. Nucleic Acids Res. 2004;32:D277–80.

35. Powell S, Forslund K, Szklarczyk D, Trachana K, Roth A, Huerta-Cepas J, Gabaldon T, Rattei T, Creevey C, Kuhn M, et al. eggNOG v4.0: nested orthology inference across 3686 organisms. Nucleic Acids Res. 2014;42(Database issue):D231-239.

36. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Res. 2009;37:D233–8.

37. Jia B, Raphenya AR, Alcock B, Waglechner N, Guo P, Tsang KK, Lago BA, Dave BM, Pereira S, Sharma AN, et al. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. Nucleic Acids Res. 2017;45(D1):D566–73.

38. Chen LH, Yang J, Yu J, Ya ZJ, Sun LL, Shen Y, Jin Q. VFDB: a reference database for bacterial virulence factors. Nucleic Acids Res. 2005;33:D325–8.

39. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215(3):403–10.

40. Qin J, Li Y, Cai Z, Li S, Zhu J, Zhang F, Liang S, Zhang W, Guan Y, Shen D, et al. A metagenome-wide association study of gut microbiota in type 2 diabetes. Nature. 2012;490(7418):55–60.

41. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, Wang Z. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ. 2019;7:e7359.

42. Wu YW, Tang YH, Tringe SG, Simmons BA, Singer SW. MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm. Microbiome. 2014;2:26.

43. Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, Lahti L, Loman NJ, Andersson AF, Quince C. Binning metagenomic contigs by coverage and composition. Nat Methods. 2014;11(11):1144–6.

44. Sieber CMK, Probst AJ, Sharrar A, Thomas BC, Hess M, Tringe SG, Banfield JF. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. Nature Microbiology. 2018;3(7):836-+.

45. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 2015;25(7):1043–55.

46. Stewart RD, Auffret MD, Warr A, Walker AW, Roehe R, Watson M. Compendium of 4,941 rumen metagenome-assembled genomes for rumen microbiome biology and enzyme discovery. Nat Biotechnol. 2019;37(8):953–61.

47. Olm MR, Brown CT, Brooks B, Banfield JF. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. Isme J. 2017;11(12):2864–8.

48. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics. 2019;36(6):1925–7.

49. Pasolli E, Asnicar F, Manara S, Zolfo M, Karcher N, Armanini F, Beghini F, Manghi P, Tett A, Ghensi P, et al. Extensive unexplored human microbiome diversity revealed by over 150,000 genomes from metagenomes spanning age, geography, and lifestyle. Cell. 2019;176(3):649-662 e620.

50. Langfelder P, Horvath S: WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559.

51. Chen C, Zhou Y, Fu H, Xiong X, Fang S, Jiang H, Wu J, Yang H, Gao J, Huang L. Expanded catalog of microbial genes and metagenome-assembled genomes from the pig gut microbiome. Nat Commun. 2021;12(1):1106.

52. Nayfach S, Shi ZJ, Seshadri R, Pollard KS, Kyrpides NC. New insights from uncultivated genomes of the global human gut microbiome. Nature. 2019;568(7753):505–10.

53. Wang W, Hu H, Zijlstra RT, Zheng J, Ganzle MG. Metagenomic reconstructions of gut microbial metabolism in weanling pigs. Microbiome. 2019;7(1):48.

54. Makki K, Deehan EC, Walter J, Backhed F. The impact of dietary fiber on gut microbiota in host health and disease. Cell Host Microbe. 2018;23(6):705–15.

55. Oliphant K, Allen-Vercoe E. Macronutrient metabolism by the human gut microbiome: major fermentation by-products and their impact on host health. Microbiome. 2019;7(1):91.

56. Sayers S, Li L, Ong E, Deng S, Fu G, Lin Y, Yang B, Zhang S, Fa Z, Zhao B, et al. Victors: a web-based knowledge base of virulence factors in human and animal pathogens. Nucleic Acids Res. 2019;47(D1):D693–700.

57. Zhang Z, Zhang Q, Wang T, Xu N, Lu T, Hong W, Penuelas J, Gillings M, Wang M, Gao W, et al. Assessment of global health risk of antibiotic resistance genes. Nat Commun. 2022;13(1):1553.

58. Wu G, Xu T, Zhao N, Lam YY, Ding X, Wei D, Fan J, Shi Y, Li X, Li M, et al. Two competing guilds as a core microbiome signature for health recovery. bioRxiv 2022:2022.2005.2002.490290.

59. Moreno-Gallego JL, Chou SP, Di Rienzi SC, Goodrich JK, Spector TD, Bell JT, Youngblut ND, Hewson I, Reyes A, Ley RE. Virome diversity correlates with intestinal microbiome diversity in adult monozygotic twins. Cell Host Microbe. 2019;25(2):261-272 e265.

60. Dhanasiri AKS, Jaramillo-Torres A, Chikwati EM, Forberg T, Krogdahl Å, Kortner TM. Effects of dietary supplementation with prebiotics and Pediococcus acidilactici on gut health, transcriptome, microbiota, and metabolome in Atlantic salmon (Salmo salar L.) after seawater transfer. Animal Microbiome. 2023;5(1):10.

61. Wang J, Jaramillo-Torres A, Li Y, Kortner TM, Gajardo K, Brevik ØJ, Jakobsen JV, Krogdahl Å. Microbiota in intestinal digesta of Atlantic salmon (Salmo salar), observed from late freshwater stage until one year in seawater, and effects of functional ingredients: a case study from a commercial sized research site in the Arctic region. Animal Microbiome. 2021;3(1):14.

62. Kostic AD, Howitt MR, Garrett WS. Exploring host-microbiota interactions in animal models and humans. Gene Dev. 2013;27(7):701–18.

63. Litvak Y, Byndloss MX, Baumler AJ. Colonocyte metabolism shapes the gut microbiota. Science. 2018;362(6418):eaat9076.

64. Ley RE, Turnbaugh PJ, Klein S, Gordon JI. Microbial ecology: human gut microbes associated with obesity. Nature. 2006;444(7122):1022–3.

65. Zhao S, Jang C, Liu J, Uehara K, Gilbert M, Izzo L, Zeng X, Trefely S, Fernandez S, Carrer A, et al. Dietary fructose feeds hepatic lipogenesis via microbiota-derived acetate. Nature. 2020;579(7800):586–91.

66. Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, Fernandes GR, Tap J, Bruls T, Batto JM, et al. Enterotypes of the human gut microbiome. Nature. 2011;473(7346):174–80.
67. Costea PI, Hildebrand F, Arumugam M, Backhed F, Blaser MJ, Bushman FD, de Vos WM, Ehrlich SD, Fraser CM, Hattori M, et al. Enterotypes in the landscape of gut microbial community composition. Nat Microbiol. 2018;3(1):8–16.
68. Ramayo-Caldas Y, Mach N, Lepage P, Levenez F, Denis C, Lemonnier G, Leplat JJ, Billon Y, Berri M, Dore J, et al. Phylogenetic network analysis applied to pig gut microbiota identifies an ecosystem structure linked with growth traits. Isme J. 2016;10(12):2973–7.

Publisher's Note