

## Response of gut microbiota to feed-borne bacteria depends on fish growth rate: a snapshot survey of farmed juvenile *Takifugu obscurus*

Xingkun Jin<sup>1</sup>  Ziwei Chen,<sup>1</sup> Yan Shi,<sup>1</sup>  
Jian-Fang Gui<sup>1,2</sup> and Zhe Zhao<sup>1</sup> 

<sup>1</sup>Department of Marine Biology, College of Oceanography, Hohai University, Nanjing, 210098, China.

<sup>2</sup>State Key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, The Innovation Academy of Seed Design, Chinese Academy of Sciences, Wuhan, 430072, China.

### Summary

**Environmental bacteria have a great impact on fish gut microbiota, yet little is known as to where fish acquire their gut symbionts, and how gut microbiota response to the disturbance from environmental bacteria. Through the integrative analysis by community profiling and source tracking, we show that feed-associated bacteria can impose a strong disturbance upon the hindgut microbiota of cultured fugu. Consequently, marked alterations in the composition and function of gut microbiota in slow growth fugu were observed, implying a reduced stability upon bacterial disturbance from feed. Moreover, quantitative ecological analyses indicated that homogeneous selection and dispersal limitation largely contribute to the community stability and partial variations among hosts in the context of lower degree of disturbance. While the disturbance peaked, variable selection leads to an augmented interaction within gut microbiota, entailing community unstability and shift. Our findings emphasized the intricate linkage between feed and gut microbiota and highlighted the importance of resolving the feed source signal before the conclusion of comparative analysis of**

**microbiota can be drawn. Our results provide a deeper insight into aquaculture of fugu and other economically important fishes and have further implications for an improved understanding of host-microbe interactions in the vertebrate gastrointestinal tract.**

### Introduction

Microbiome can directly or indirectly affect the host's physiological functions such as metabolism, immune defence, growth and development, as evidenced by increasingly large body of microbiome researches (Berg *et al.*, 2020). Integrative conceptual frame of holobiont had been proposed, in which animals and plants are no longer seen as autonomous entities but rather as biological networks comprised with both host and its associated microbiome (Zilber-Rosenberg and Rosenberg, 2008; Bordenstein and Theis, 2015; Theis, *et al.*, 2016). Understanding the factors that underlie changes in such interactive network is essential and could provide further insights into how the normal balance between host and microbiome (eubiosis) can be maintained or intervened if disrupted (dysbiosis) (Lozupone, *et al.*, 2012; Xiong, *et al.*, 2017; Berg, *et al.*, 2020). Systematically surveys of a broad range of environments and host species using the contemporary high-throughput sequencing technologies have unravelled the structural and functional diversity of complex microbial communities and enabled in-depth studies of microbiome-associated ecology, physiology and molecular function (Sunagawa, *et al.*, 2015; Lloyd-Price, *et al.*, 2017; Hatzenpichler, *et al.*, 2020).

Exploring the composition and function of gut microbiome on fish have evoked increased research interest, not only for its representativeness as the most diverse vertebrate lineage, they also have a great economic significance, particularly in aquaculture (Perry, *et al.*, 2020). The development of high-potential microbiome-based innovations offers opportunities for disease control and health promotion in aquaculture practice (Chao, *et al.*; de Bruijn, *et al.*, 2017). Therefore, there is an urgent need of establishing an optimized microbial management strategy to achieve a more sustainable aquaculture (Lieke, *et al.*, 2020). As a prerequisite, understanding

Received 1 September, 2020; revised 7 December, 2020; accepted 12 December, 2020.

*Microbial Biotechnology* (2022) 15(2), 683–702  
doi:10.1111/1751-7915.13741

**Funding Information** This work was supported by the National Key Research and Development Program of China (2018YFD0900200); the National Natural Science Foundation of China (31872597); the Fundamental Research Funds for the Central Universities from China (B200202141); and the Earmarked Fund for Jiangsu Agricultural Industry Technology System (JATS[2020]465).

© 2021 The Authors. *Microbial Biotechnology* published by Society for Applied Microbiology and John Wiley & Sons Ltd.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

the ecological processes and mechanisms in controlling the assemblage of fish-associated microbiome becomes critical (De Schryver and Vadstein, 2014; Zhang and Gui, 2018).

Fish gut can be considered as a dynamic ecosystem in which spatial-temporal disturbance in the microbial nutrient niches shape the bacterial community (Ringø, *et al.*, 2016; Pereira and Berry, 2017). The potential niche space in fish gut environment can be determined by the host in multiple ways (Wang, *et al.*, 2018), yet it also can be largely determined by the diet as well as its associated bacteria (Parris, *et al.*, 2019; Wilkes Walburn, *et al.*, 2019; Abdul Razak and Scribner, 2020). Recent findings suggest that a limited numbers of bacterial taxa are shared inter-individually from a given host population and comprise a core gut microbiota (Roeselers, *et al.*, 2011; Wong, *et al.*, 2013; Riiser, *et al.*, 2018). Fish-associated microbiota can be potentially acquired from a variety of sources with which organisms come into intimate biological interactions (Wang, *et al.*, 2018).

*Takifugu* (aka fugu), a genus of pufferfish (*Tetraodontidae*), is better known to be able to accumulate a potent neurotoxin, tetrodotoxin, and possess high toxicity (Zhang, *et al.*, 2020b). As a vertebrate, fugu displays several peculiarities in its genome, such as shortened intergenic and intronic sequences devoid of repetitive elements (Brenner, *et al.*, 1993; Aparicio, *et al.*, 2002), thereby makes it distinctive animal model for genome evolution study (Van de Peer, 2004e Peer, 2004). Moreover, the artificially reared fugu becomes non-toxic and is considered a delicacy and commercially farmed widely across many East Asian countries including Japan, Korea and China (Wang, *et al.*, 2019; Gao, *et al.*, 2020). In 2016, China had deregulated the fugu farming industry, and large-scale aquaculture of two fugu species *T. rubripes* and *T. obscurus* were licensed. With the market demand continues to grow, the aquaculture fugu production is progressively increased (Wang, *et al.*, 2019; Zhang, *et al.*, 2020b). Therefore, understanding how and determining the extent to which the rearing environments effect fugu health and production is needed.

As a demersal fish, fugu live on or near the bottom of water, foraging food by scraping them off the surface (Okabe, *et al.*, 2019). Meanwhile, faeces from farmed fugu can also play a part of bacterial enrichment in rearing water and sediment as seen in other fish species (Muziasari, *et al.*, 2017; Gołaś, *et al.*, 2019; Zhang, *et al.*, 2020a). Such dynamic reciprocal interactions might be expected to lead to a closer association between the gut of farmed fugu and its rearing environment to some extent, that is distinctions in the gut bacterial communities mirrored the environments and *vice versa*. Especially while considering the abundant bacterial species in the surrounding environments as a

potential reservoir (Giatsis, *et al.*, 2015; Eichmiller, *et al.*, 2016; Parris, *et al.*, 2019). On the other hand, unlike the pond water and sediment, the feed and gut are inherently linked by the high nutrient niche, which might impose a strong environmental selection of microbiota with functional traits linked to copiotroph (De Schryver and Vadstein, 2014; Troussellier *et al.*, 2017). Despite the evidences for the broad existence of core gut microbiota, accumulating studies had also showed substantial inter-individual variation in the gut microbial communities for both cultured (Li, *et al.*, 2017; Perry, *et al.*, 2020) and wild-captured fish species (Star, *et al.*, 2013; Egerton, *et al.*, 2018; Riiser, *et al.*, 2020). Several factors can contribute to such variability, including environmental filtering either by the host, for example variations between regions of the gastrointestinal tract, state of health and nutrition; or by the rearing environments such as diet (Wong, *et al.*, 2013; Clements, *et al.*, 2014). Such variations should be resolved before the conclusion of comparative analysis of microbiota can be drawn.

Environmental bacteria have a great impact on fish gut microbiota (Eichmiller, *et al.*, 2016; Muziasari, *et al.*, 2017; Lieke, *et al.*, 2020), yet little is known as to where fish acquire their gut symbionts, and to what extent gut microbiota response to it. Here we used 16S rRNA amplicon sequencing to characterize the hindgut microbiota from cultured obscure puffer (*T. obscurus*) in comparisons with its rearing environment. Juvenile offsprings from a full-sib family were reared in a closed indoor aquaculture facility with consistent pond management and husbandry condition as well as feeding procedure. Our findings illustrated a clear diversity–function relationships in fugu gut microbiota and emphasized the intricate linkage between feed and gut microbiota. These results provide a deeper insight into aquaculture of fugu and other economically important fishes and have further implications for an improved understanding of host–microbe interactions in the vertebrate gastrointestinal tract.

## Results

In this study, the microbiota from fugu gut and its rearing environments were profiled using 16S rRNA gene amplicon sequencing. Using Illumina MiSeq, we obtained a total of 2.6 million reads with an average length of 410 bp from 82 samples (fish = 61; feed = 3; water = 9; sediment = 9). After quality filtering and chimeras removal using USEARCH followed by taxonomy-based filtering of plastids (chloroplast) sequences via RDP classifier, a total of 2,289 exact sequence variants (ESVs) were clustered. A total of 2.2 million high-quality reads were retained (mappable to ESVs) with average 26,712 reads per sample (ranging from 7,244 to 66,203; Table S1). Further, data sets were rarefied at 7,000

sequences per sample, followed by coverage estimates by rarefaction curves to ensure a sufficient sequencing depth and to include all samples for downstream analysis (Fig. S1). The data set generated and analysed for this study is available in the NCBI sequencing reads archive (SRA), under BioProject accession number PRJNA658467.

#### *Comparative analysis of microbiota shows distinctions between fugu gut and rearing environment*

The calculated indices of Richness (Observed ESVs), Chao1, Shannon and Simpson, congruously, indicated that the alpha diversity of fugu gut microbiota was markedly lower comparing to that of rearing environment including feed, water and sediment (Fig. S2). In addition, statistical analysis of beta diversity based on Bray–Curtis distance showed that the pond ( $n = 3$ ) does have certain effect on community dissimilarity ( $R^2 = 0.11$ ;  $P = 0.001$ ; MeanSqs = 0.97; Table S2), thereby 'pond' was set as a strata. Further, the community composition was revealed to be significantly differed among sample groups including fugu gut (F), feed (Fd), sediment (S) and water (W) (PERMANOVA,  $R^2 = 0.24$ ;  $P = 0.001$ ; MeanSqs = 2.08; strata='pond'; Fig. 1A). In addition, the permutational analysis of multivariate dispersion (PERMDISP) results showed that the sample groups differed in homogeneities, and the intra-group beta dispersion was significantly lower in feed group than any other groups (Tukey HSD,  $P < 0.05$ , Table S2). This is in accordance with the fact that the commercial pelleted feed are technical replicates ( $n = 3$ ) (Fig. S3). Moreover, the abundance-weighted UniFrac distances of all intra-group pairs were significantly shorter than that of inter-group pairs ( $P < 0.001$ ; Mann–Whitney  $U$ -test), among others the fugu gut had the greatest within-group dispersions (F.F: range 0.069–0.996; coefficient of variation, 0.340) suggesting a much greater variability (Fig. 1B).

We further estimated the proportions of feed, sediment and water as a potential environmental bacteria 'source' for the gut microbiota (as 'sink') using SourceTracker. Results shown that the majority of bacterial proportions in fugu gut microbiota were, in fact, not identifiable from any given source ('unknown', 70.42%  $\pm$  30.87% s.d.), suggesting a higher degree of host specificity. Yet, a notable proportion was predicted from feed (29.19%  $\pm$  30.53% s.d.) (far left facet 'FishSink' in Fig. 1C and Table S3). Moreover, verifications using one each of the four sample groups as 'source' had confirmed the accuracy of the predictive measures in classifying among different groups, despite a minor proportion from an

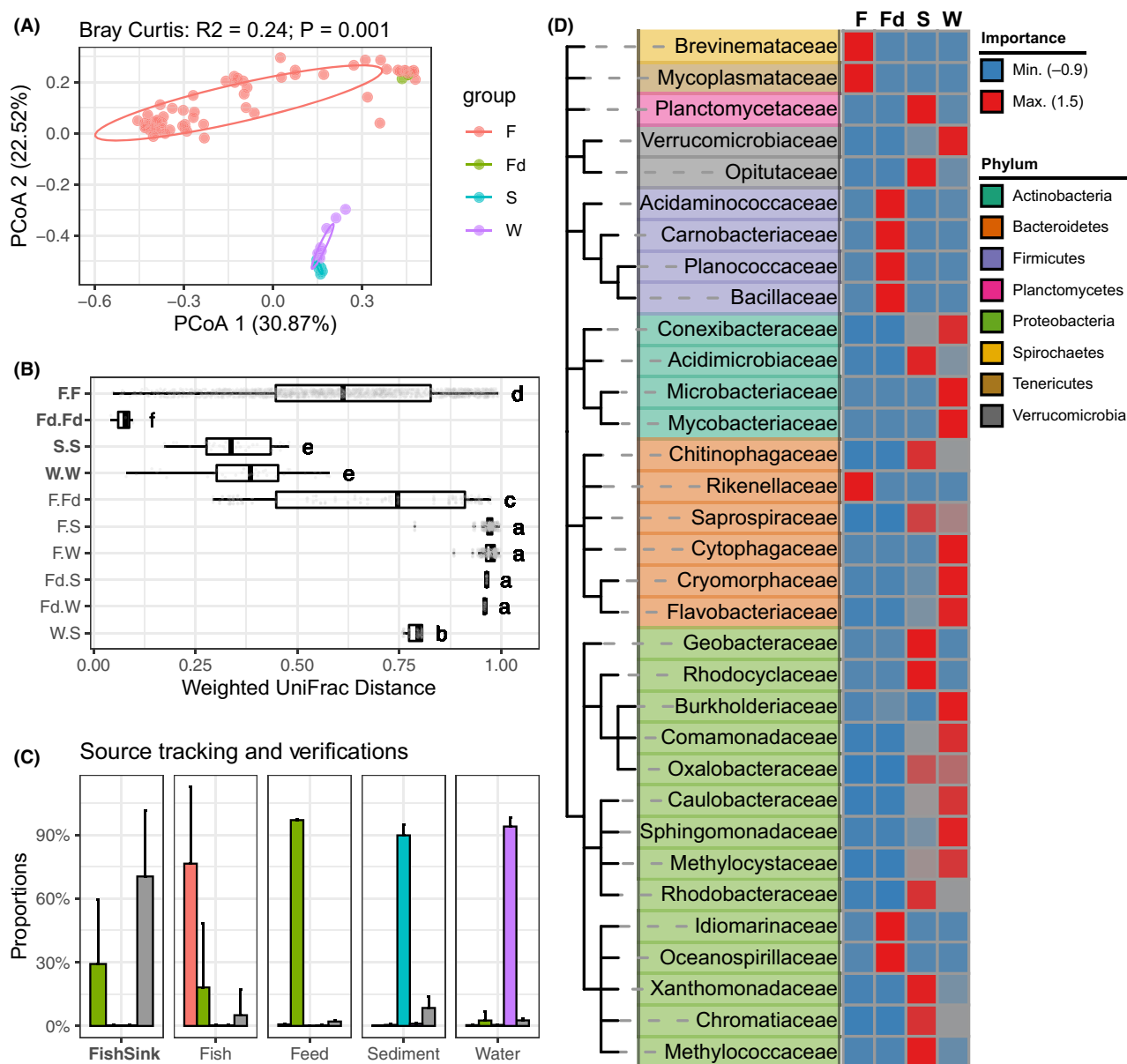
unidentified source (unknown) within each group (four right-side facets in Fig. 1C).

The characteristic taxa were classified as a attribute in distinguishing among different groups of bacterial community using Random-Forest machine-learning approach. Results post 10-fold cross-validation showed that at least 34 bacterial families can be used to classify all four sample groups with 1.22% error rate (Fig. S4, Table S4). Of these, 33 families were taxonomically categorized into eight phylums according to the NCBI Taxonomy (Fig. 1D). Three indicative bacterial families for fugu gut microbiota were identified, namely Brevinemataceae (Spirochaetes), Mycoplasmatataceae (Tenericutes) and Rikenellaceae (Bacteroidetes). However, greater numbers of diverse indicative families for environmental sample groups were determined and brought into correspondence with the observed high bacterial diversity in environmental microbiota.

#### *Community structure in fugu gut is correlated with feed source signal and host body mass*

Fish body mass is of great value in aquaculture for its impact in fish appetite, growth and feed utilization (Breck, 2014). The body weight (unit: gram), fork length (unit: centimetre), liver weight (g) and gonad weight (g) were determined to allow analysis of variance in body mass within the fish population. Individual fugu were clustered based on unsupervised k-means computed from body mass measures. Using gap statistic method, the optimal clustering numbers was determined to be 3, which we defined as 'fast' ( $n = 24$ ), 'medium' ( $n = 20$ ) and 'slow' ( $n = 17$ ) (Fig. 2A and Fig. S5). The quality of representation ( $\cos^2$ ) for body weight, liver weight, fork length and gonad weight on the first principal component (PC1, 88.1% of total variations) were 0.96, 0.93, 0.92 and 0.72 respectively (Table S2). The PC1 value was used as a proxy for individual fish body mass for downstream analysis.

Statistical analyses of beta diversity were performed to infer the potential factors underlying the large inter-individual beta dispersion of gut microbiota. PERMANOVA analysis showed that feed source signals (predicted by SourceTracker) was the best predictor of community composition ( $R^2 = 0.46$ ;  $P = 0.001$ ; MeanSqs = 6.68; strata:Pond; Fig. 2B), followed by fish body mass ( $R^2 = 0.07$ ;  $P < 0.01$ ; MeanSqs = 0.98; strata:'Pond'). Importantly, negative correlation between the fish body mass and the feed source signal was observed (Spearman's Rank  $R = -0.38$ ;  $P < 0.01$ , Fig. 2C). Moreover, the feed source proportions were significantly enriched in slow growth fish (42.2%) comparing to fast one (19.3%) (Tukey HSD,  $P < 0.05$ ). These results suggest the varying degree of effect of feed-associated bacteria

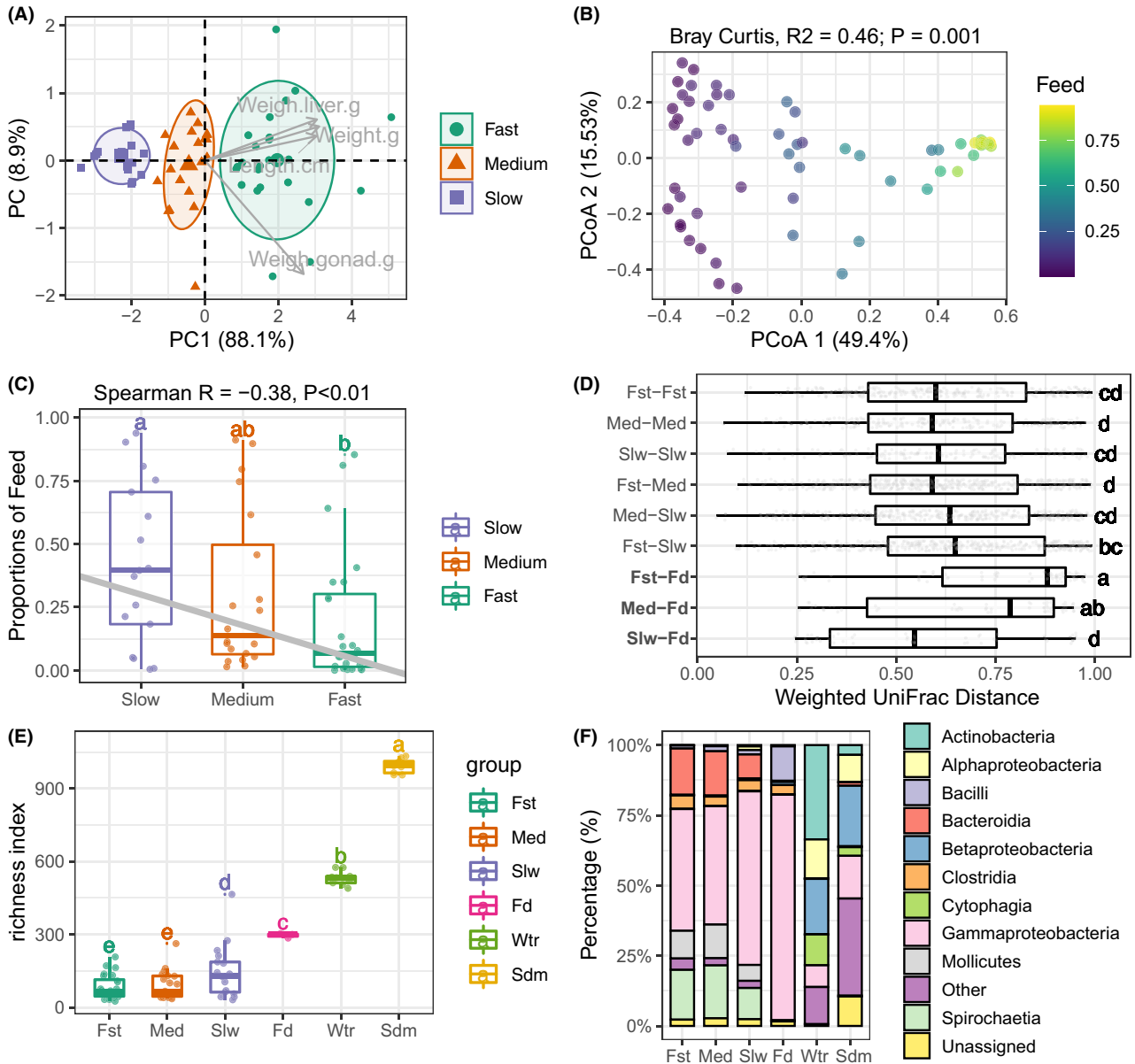


**Fig. 1.** Comparisons of different bacterial communities within the fugu rearing ecosystems. **A.** Two-dimensional scatter plot of all bacterial communities based on principal coordinate analysis (PCoA) on Bray–Curtis dissimilarity. Percentage of explained variance and statistical significance were reported by PERMANOVA. A normal data ellipse for each group was drawn at confidence level of 0.68. Sample groups: ‘F’: Fugu gut; ‘Fd’: feed; ‘S’: sediment; ‘W’: water. **B.** Boxplot shows the intra-/inter-groups pairwise comparisons of weighted UniFrac distances across all sample groups. For each box, the vertical bold bar denotes medians; the width of box denotes the interquartile range (25th percentile–75th percentile); the whiskers mark the values range within 1.5 times interquartile. Lower-cased letters denote statistical significance reported by Mann–Whitney  $U$ -test at confidence level of 0.95. **C.** Source tracking of gut microbiota from rearing environments. The far left facet shows the SourceTracker-estimated proportions of gut microbiota from its surrounding environment; the righter four facets, respectively, show the verifications using one each of the four sample groups as ‘source’. Bars denote average proportions of each source for a indicated community as sink ( $X$ -axis) using trained source samples ( $F = 61$ ,  $Fd = 3$ ,  $S = 9$ ,  $W = 9$ ) rarefied by 1000 ESVs. Error bars represent standard deviations. Bars were coloured according to the indicated sources, of which ‘Unknown’ represent source unidentified from trained source data. Colour code is consistent with (A) excepting the ‘unknown’ source was showed in grey. **D.** Random-Forest identified bacterial families that accurately discriminated bacterial community groups. Featured taxa were hierarchically clustered according to NCBI taxonomy and were colour-coded by phylum ranks. Heatmap showed the centred classic importance of each indicated taxa (row) in discriminating a given bacterial community (column) from others.

on gut microbiota is dependent on host growth rate. Specifically, the gut community dissimilarity distances to feed for slow growth fish were significantly shorter than

that for fast one, as evidenced by the comparisons of abundance-weighted UniFrac distances ( $P < 0.001$ ; Mann–Whitney  $U$ -test, Fig. 2D).





**Fig. 2.** Community structure of fugu gut microbiota are correlated with feed source signal and host body mass. **A.** Two-dimensional scatter plot of all fish individual based on principal component analysis (PCA) of four body mass metrics. The grey arrow points showed the correlated body mass measures. Three optimal groups namely 'fast', 'medium' and 'slow' were identified by K-mean clustering followed by Gap statistic. A normal data ellipse for each group was drawn at confidence level of 0.68. **B.** Principal coordinate analysis of gut microbiota based on Bray-Curtis distance. Percentage of explained variance and statistics (PERMANOVA test with 999 permutations) were shown as figure title. Each community point was coloured by the proportion of feed source signal. **C.** SourceTracker-estimated proportions of feed bacteria for each gut community grouped by host body mass. Grey line denotes the linear regression of feed proportions and fugu body mass (Spearman's Rank), statistics was indicated in figure title. **D.** Boxplot of group-paired weighted UniFrac distances showing the dissimilarities between gut and feed were reduced in slow growth fugu (Slw-Fd) comparing to fast growth one (Fst-Fd). **E.** Species richness (Numbers of observed ESVs) among six groups of bacterial community. For each box in (C-E), the bold bar denote medians; the height of box denotes the interquartile range (25th percentile-75th percentile); the whiskers mark the values range within 1.5 times interquartile. Lower-cased letters denote statistical significance reported by Mann-Whitney *U*-test at confidence level of 0.95. **F.** Stacked bars showing the relative abundance of top-10 bacterial classes across six groups of community. Bacterial class which has a lower relative abundance were grouped into 'Other'. The detailed results of replicated samples were shown in Fig. S7.

The estimated indices of alpha diversity were positively correlated with feed source signals, but negatively correlated with fugu body mass (Fig. S6). Further,

markedly increased richness of gut microbiota from slow growth fish was observed comparing to fast group ( $P < 0.05$ ; Mann-Whitney *U*-test, Fig. 2E). The

taxonomic compositions of bacterial communities were differed among sample groups (Fig. S7; Table S5). In the bacterial class level (Fig. 2F), Actinobacteria, Alpha-, Beta-, Gamma-proteobacteria and Cytophagia were commonly found to be dominant in both water (sum 85.9%) and sediment (sum 53.0%), whereas Actinobacteria and Cytophagia were more present in water (33.6% and 10.9%) than to sediment (3.5% and 3.0%). Feed-associated bacterial communities were mainly composed of Gamma-proteobacteria (80.4%), Bacilli (12.2%) and Clostridia (3.4%). The gut bacterial communities of the three fish clusters were mainly composed of Gamma-proteobacteria (avg. 49.1%), Spirochaetia (avg. 16.0%), Bacteroidia (avg. 13.7%), Mollicutes (avg. 9.2%) and Clostridia (avg. 4.0%). Moreover, the fish gut-specific bacterial classes including Bacteroidia, Mollicutes and Spirochaetia were more present in fast and medium groups, whereas Gamma-proteobacteria was more present in the slow group.

Initially, Random-Forest was used to classify the characteristic taxa among these three fish groups. As a result, the estimate of error rate was relatively high (~60%), particularly for the medium growth group (>90%). Alternatively, the LEfSe algorithm was used to identify differentially abundant bacterial taxa that are corresponding with the differences between fugu growth rate. A total of 53 discriminant bacterial clades separating three fish clusters with statistical significance (Kruskal–Wallis sum-rank,  $P < 0.05$ ) were detected (Fig. S8 and Table S6). The overall characteristic bacterial families were more detected in slow group, which is in consistent with its observed high species richness. Among others, Pseudomonadaceae was the most pronounced bacterial family ( $\log_{10}$ -LDA Scores  $> 5$ ) that were overrepresented in slow group.

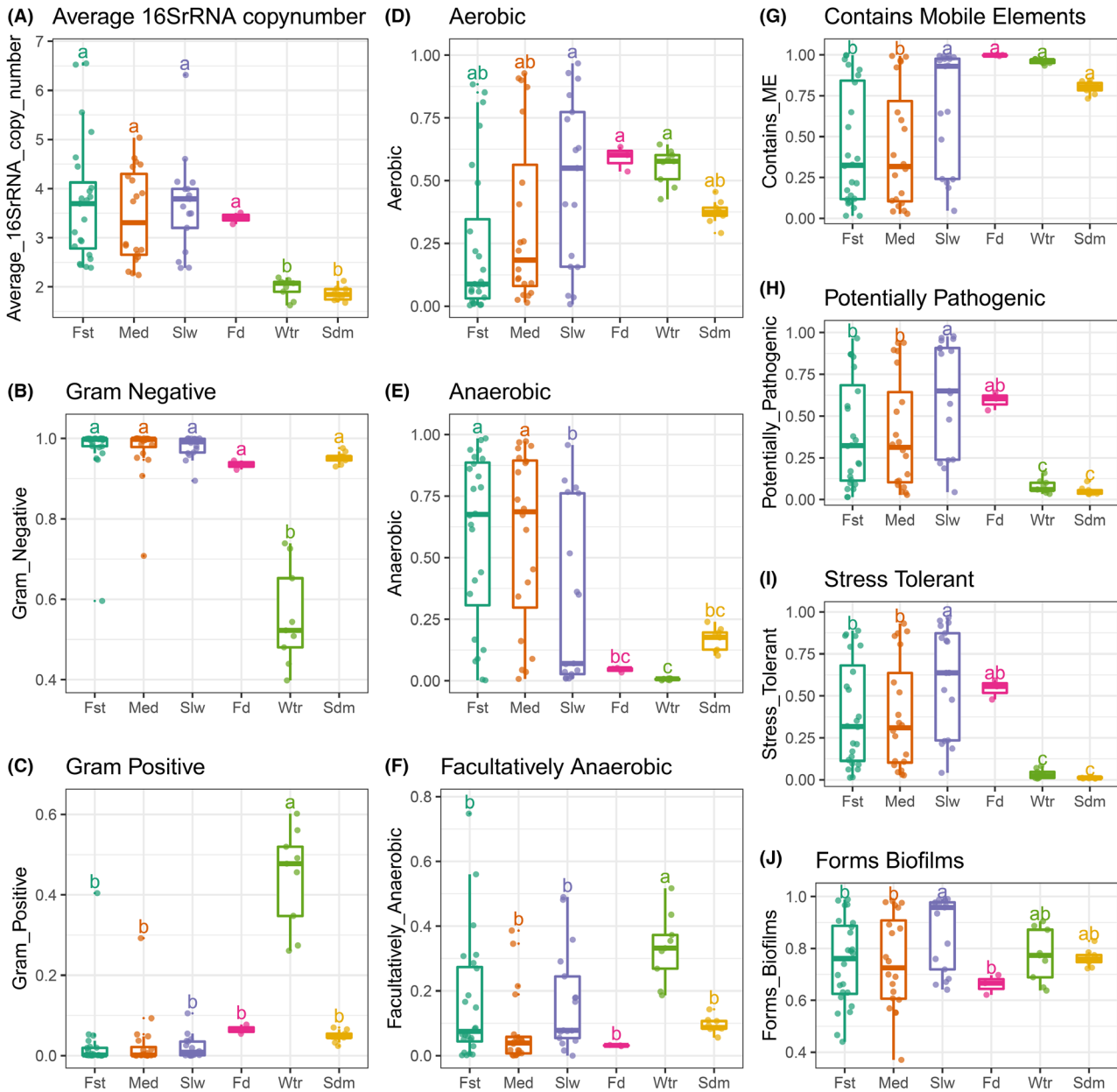
#### *Functional shifts of fugu gut microbiota in response to feed-borne bacteria*

Understanding the phenotypes and functional capacities of the microbes within a community is critical to determine the composition–function relationships. To this end, we used PICRUSt and BugBase, respectively, to infer the functional profiles and organism-level phenotypes of different microbial communities based on 16S rRNA gene. Increased ribosomal content boosts more protein translation and stronger metabolism, thereby the number of rRNA genomic copies is positively correlated with bacterial growth rate (Lankiewicz, *et al.*, 2016), and can be used as a proxy for phenotypically distinguishing between copiotroph and oligotroph (Lauro, *et al.*, 2009; Roller, *et al.*, 2016). Using the precalculated 16S rRNA

counts by PICRUSt, the average 16S rRNA copy number per sample group were estimated. The results showed that bacterial communities from fugu gut (3.6) and feed (3.4) had greater average copy numbers than that from water (2.0) and sediment (1.9), indicating the bacterial communities in both gut and feed are potentially less oligotrophic than those observed in the water and sediment (Fig. 3A).

Further, the phenotypical characteristics including oxygen tolerance, Gram staining and pathogenic potential of bacterial communities were predicted by BugBase algorithm. Compared to the pond water, BugBase predicted the gut microbiota (average in all fish: 97%), feed microbiota (93%) and sediment microbiota (95%) to have a significantly higher proportion of Gram negative (Fig. 3B). However, water microbiota was predicted to be comprised of both Gram negative (56%) and Gram positive (44%) bacteria (Fig. 3B and C, and Fig. S9). Regarding oxygen tolerance traits (Fig. 3D–F), gut was shown to be significantly dominant by anaerobic bacteria (51%) comparing to all the other environments (<20%), where is mostly dominant by aerobic bacteria (59%) as for feed, or by both aerobic (56%) and facultatively anaerobic bacteria (33%) as for water. However, such dominance was significantly decreased as the proportions of aerobic bacteria increased in slow growth fugu (49%) comparing to both fast (24%) and medium (34%) groups. Indeed, the inferred proportions of anaerobic bacteria in gut microbiota are negatively correlated with feed source signals (Spearman's Rank  $R = -0.79$ ;  $P < 0.01$ ), but positively correlated with body mass ( $R = 0.32$ ;  $P < 0.05$ , Fig. S10). The opposite trends were observed for inferred aerobic bacterial proportions. These results suggest the feed source signals is the most pronounced factor in contributing such phenotypic changes in gut microbiota, and it is depended on host body mass.

BugBase also predicted the fugu gut microbiota to have significantly fewer bacteria that contain mobile elements and more bacteria that are potentially pathogenic and oxidative stress tolerant than water and sediment ( $P < 0.05$ , Mann–Whitney  $U$ -test, Fig. 3G–I). Further, bacteria that can form biofilms was not significantly different comparing fugu gut to both water and sediment (Fig. 3J). Within gut microbiota, the bacterial proportions related to these four phenotypes were all positively correlated with feed source signals ( $R > 0.7$ ;  $P < 0.05$ ) but negatively correlated with body mass ( $R < -0.3$ ,  $P < 0.05$ , Fig. S10). Altogether, these results have further emphasized that the functional changes in fugu gut microbiota are driven largely by feed-associated bacteria and depended on host growth rate.



**Fig. 3.** Phenotype inference of bacterial communities from the fugu rearing ecosystems. Estimated relative abundances for each indicated bacterial phenotype were compared across different sample groups. A. Averaged 16S rRNA copy numbers were compared among sample sites of bacterial communities. Points denote the mean copy number calculated from a given bacterial community. The following organism-level phenotypes were inferred by BugBase. Relative abundances of bacteria differing in Gram staining were shown in (B) for Gram negative and (C) for Gram positive. Relative abundances of bacteria differing in oxygen tolerance phenotypes were shown in (D) for Aerobic, (E) for Anaerobic and (F) for Facultatively Anaerobic. Relative abundances of bacteria differing in latent pathogenicity phenotypes were shown in (G) for containing mobile elements, (H) for potentially pathogenic, (I) for oxidative stress tolerance and (J) for biofilm formation. See Fig. S9 for the related taxa contributions of the relative abundance of bacteria possessing each phenotype. For each box, the horizontal bold bar denote medians; the height of box denotes the interquartile range (25th percentile–75th percentile); the whiskers mark the values range within 1.5 times interquartile. Lower-case letters denote groups and statistical significance reported by pairwise Mann–Whitney  $U$ -tests with false discovery rate correction at confidence level of 0.95.

#### *Ecological processes underlying bacterial community shift in fugu gut microbiota*

To infer the underlying ecological mechanisms governing the process of bacterial community assemblage in fugu

gut, the phylogenetic signals of gut bacterial communities were estimated using Mantel's correlogram (Oden and Sokal, 1986) and Blomberg's  $K$  statistic (Blomberg, *et al.*, 2003) respectively. We found significant phylogenetic signal ( $P < 0.05$ ) correlating to changes of feed

source signals and fish body mass, but only across very short phylogenetic distances (Fig. S11). On this basis, we proceeded to quantify phylogenetic turnover among close-related bacteria (Stegen, *et al.*, 2012). Considering the ecological influence on bacterial sub-community may varied with its relative abundance (Shade and Gilbert, 2015; Langenheder and Lindström, 2019; Jiao and Lu, 2020), the fugu gut microbiota ('all', 943 ESVs) were further partitioned into 'abundant' (abundance > 0.1%, 63 ESVs) and 'rare' sub-community (abundance < 0.1%, 880 ESVs), respectively. We concentrated on 'all', 'abundant' and 'rare' ESVs to extract phylogenetic distance from a given bacterial community. The results showed that the unweight standardized effect size of mean nearest taxon distances (ses.MNTD) for 'all', 'abundant' and 'rare' sub-communities from gut microbiota were consistently below zero ( $P < 0.01$ , Fig. S12), suggesting the importance of environmental filtering in driving gut bacterial communities to be more phylogenetically clustered. In addition, the average ses.MNTD measures for abundant taxa were markedly higher than that of rare taxa (Fig. S12), suggesting its lower degree of phylogenetic relatedness.

Further, we estimated the phylogenetic beta diversity ( $\beta$ MNTD) and in turn compared it against null model to estimated beta nearest taxon index  $\beta$ NTI (Stegen, *et al.*, 2012; Stegen, *et al.*, 2013). The  $|\beta$ NTI| values > 2 indicate significantly phylogenetic turnover than expected. The Mantel's test showed significant positive correlations between the inter-community  $\beta$ NTI values and the changes of feed source signals, for 'all', 'abundant' and 'rare' sub-communities respectively (Fig. 4A–C). These results suggest that the shift from homogeneous selection to variable selection in contributing to the gut bacterial community assemblage as changes of feed source signals increase.

To quantify the ecological processes in communities that were not governed by selection (i.e.  $|\beta$ NTI| < 2), the number of these pairwise comparisons were fractionated considering  $|\beta$ NTI| < 2 and  $|RC_{bc}| < 0.95$ . The resulting fractions showed that dispersal limitations (40.6%) contribute largely in the process of community assemblage in fugu gut while considering all the taxa, followed by variable selection (28.9%) and homogeneous selection (18.1%). While comparing the sub-communities varied in taxa abundance, the relative importance of deterministic process (homogeneous and variable selection) was larger for 'rare' sub-community, whereas stochastic processes (dispersal limitation and 'undominated') were dominant in 'abundant' sub-community (Fig. 4D). In addition, very few portions of homogenizing dispersal process were observed for all sub-communities. Furthermore, the neutral model explained a larger fraction of variation in the rare sub-community ( $R^2 = 0.639$ ,

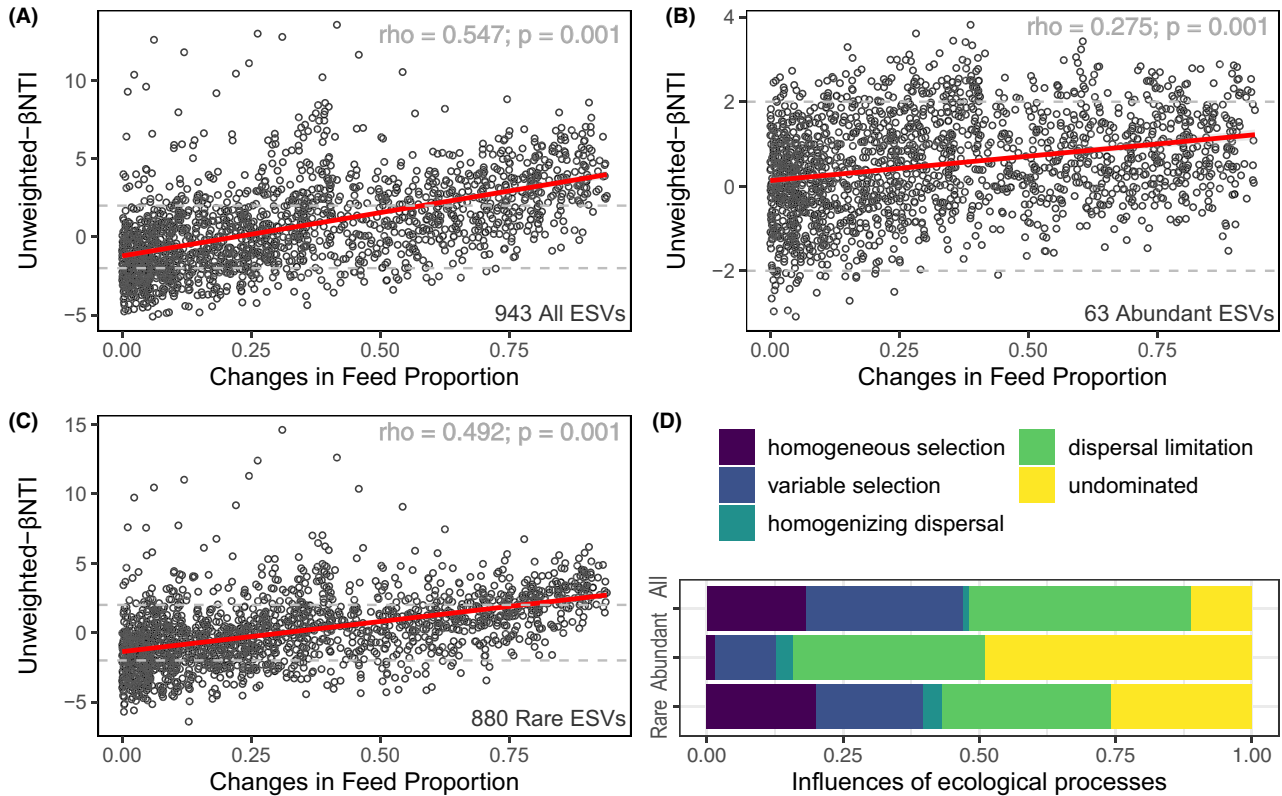
$m = 0.057$ ) than the abundant one ( $R^2 = 0.154$ ,  $m = 0.012$ ), as well as increased migrations rate (Fig. S13).

## Discussion

Here we used 16S rRNA gene amplicon sequencing to explore the relations of bacterial communities between gut of cultured juvenile obscure puffer and its rearing environments. Our results showed that farmed fugu had a distinct gut microbiota with markedly lower richness (alpha diversity) and higher dissimilarity (beta dispersion) in comparison to its rearing environment including feed, water and sediment (Fig. 1). Furthermore, effect of varying degrees of feed-associated bacteria on both composition and function of gut microbiota were revealed, and implications of fish growth-dependent community-level responses and underlying relative importance of ecological processes were found. These results (discussed below) provide a new insight into aquaculture of fugu and other farmed fishes that are of great economic importance, as well as a better understanding of host–microbe interactions in the vertebrate gastrointestinal tract.

### *Distinctions of fugu gut microbiota from its rearing environment*

Herein, the fugu gut microbiota is characterized by lower species richness, higher beta dispersion and rRNA copy numbers in comparison to water and sediment, but far less so to feed (Figs 1 and 3), indicating a strong niche effect on gut microbiota along with considerable diet influence. Indeed, SourceTracker had predicted a notable source signature from feed (~30%) in contributing to fugu gut microbiota (Fig. 1C), although the samples were obtained 1 day post-feeding. This could explain why the community composition differed largely between 'gut & feed' and 'water & sediment', as the metabolic mode of copiotroph and oligotroph differed significantly (Fig. 3A). Moreover, a larger proportion of 'unknown' source signature was also predicted, suggesting an uncharacterized bacterial source possibly (i) derived elsewhere or (ii) acquired before the sampling time. For the first scenario, it is less likely that farmed fugu had much contact with any other environmental source that had not been sampled, as the indoor pond managements and husbandry conditions both were identical within a single closed aquaculture facility. For the second scenario, several studies had pointed that the processes of bacterial colonization in fish gut, particularly during early development, are complex ever since the onset of spawning, and largely determined by the egg adhering microbiota and temporal rearing environment (Giatsis, *et al.*, 2015;



**Fig. 4.** Influence of ecological processes in governing bacterial community assemblage in fugu gut. The correlations between phylogenetic beta diversity  $\beta\text{NTI}$  and differences in feed source signals in fugu gut microbiota for (A) all ESVs, (B) abundant ESVs sub-community and (C) rare ESVs sub-community. Red lines denote fitting of linear regressions and corresponding Spearman's rank correlation coefficients were indicated in the upright corner within each subpanel. Grey horizontal dashed lines denote the cut-off of  $\beta\text{NTI}$  significance between  $-2$  and  $+2$ . D. Depicts the relative importance of different ecological processes in governing community turnover for 'all', 'abundant' and 'rare' sub-communities respectively.

Wong, *et al.*, 2015; Wang, *et al.*, 2018; Abdul Razak and Scribner, 2020). Therefore, it can be hypothesized that the early-life exposure (e.g. from first feeding) to different bacterial communities as seen in other fishes (Romero and Navarrete, 2006; Wang, *et al.*, 2018) might also contribute to the assembly of fugu gut microbiota mentioned here.

Ringø and Birkbeck (Ringø and Birkbeck, 1999) proposed five essential criteria to be considered core (autochthonous) gut microbiota in fishes if they are (i) presented in healthy individuals; (ii) persistent during the whole lifespan of host; (iii) mutually presented in both wild and cultured populations; (iv) anaerobic; and (v) widely distributed in gastrointestinal tract. Our Random-Forest analysis identified three bacterial families, namely Brevinemataceae, Rikenellaceae and Mycoplasmataceae, as the most discriminative taxa for fugu gut (Fig. 1D). In addition, Bugbase inferred that both Brevinemataceae and Rikenellaceae contributed to the most proportions of anaerobic bacteria in fugu gut microbiota (Fig. S9). Although literature concerning the microbiota of wild *T. obscurus* is scarce, recent studies showed that these bacterial families, particularly

Brevinemataceae with greater dominance, can be also observed in many other wild *Takifugu* species, including *T. ocellatus*, *T. bimaculatus*, *T. xanthopterus* (Li, *et al.*, 2020b) and *T. niphobles* (Shiina, *et al.*, 2006) regardless of methodological biases. It is remarkable since unlike these wild species that are native to salt and brackish waters, farmed *T. obscurus* was mainly reared in fresh water after hatching, with the only exceptions of spawning parent fishes and eggs that were transiently acclimated in saline water. Once again, it demonstrates the potential influence of early-life environments in shaping the fugu gut microbiota. Altogether, these bacterial taxa meet most of the criteria to be considered core gut microbiota in fugu, yet further longitudinal experiments concerning multiple host developmental stages are warranted to completely define it.

#### *Effects of feed-borne bacteria on gut microbiota differed among fugu at inconsistent growth rate*

We next asked which factors (either intrinsic or extrinsic) potentially correlate with the large inter-individual beta dispersion of gut bacterial community. Here, the gut

contents were individually sampled from the hindgut of farmed fugu that were feed with identical commercial fish feed. In addition, the fish used here were offsprings from a full-sib family, thereby the potential genetic variations were kept to a minimum. Moreover, the environmental bacterial communities for both water and sediment were found to be consistent across different ponds (data not shown), thereby the likelihood of such extrinsic variations could also be ruled out. In fact, we observed a large dispersion of body mass among the sampled fish population (Fig. S5), by which they can be rigorously subdivided into three clusters, that is fast, medium and slow (Fig. 2A). We showed that the feed source signals explained the most variations in the compositional difference in fugu gut microbiota, and it was negatively correlated with fish body mass (Fig. 2B and C). Across the fishes clustered by body mass, both the composition and function of gut bacterial communities were found to be varied significantly, as evidenced by the comparative analysis of richness (alpha diversity), dissimilarity (beta diversity), discriminated abundant bacterial taxa (LEfSe) and organism-level phenotypes (BugBase) (Figs 2 and 3) respectively. These results illustrated a clear diversity–function relationships in gut bacterial communities, and more importantly, it implicated a strong correlation between gut bacterial communities and host growth rate. Specifically, markedly increased species richness driven by feed source proportions were observed in slow growth fugu, where a greater level of correlation between feed-associated bacteria and gut microbiota was implied.

Such coincidence of altered microbiota composition and function in the gut of slow growth fish was largely driven by a gamma-proteobacterial genus *Pseudomonas* (LEfSe,  $\log_{10}$ -LDA Score > 5,  $P < 0.05$ ), notably a group of widespread aerobic bacteria found to be predominant in fish feed herein (>50%, Fig. S7 and S9). Previous work had showed that *Pseudomonas* is highly persistent during long term storage of fish feed (Zmyslowska and Lewandowska, 1999). Yet another study found its presence in water, contributed mostly by the fish faeces, was positively associated with the dose of administrated commercial fish feed in a recirculating aquaculture system (Gołaś, *et al.*, 2019). Among others, *Pseudomonas* is the most frequently reported gut-associated bacteria across many different freshwater fish species (Egerton, *et al.*, 2018). On the other hand, recent studies had also raised concerns that it might represent a common contaminating bacterial taxa in molecular biology reagents (Li, *et al.*, 2020a). While considering the very rare presence of *Pseudomonas* in all the water and sediment samples (Fig. S7E), along with the negative amplification signal in our blank (no template) control experiment, it is highly unlikely that this bacteria taxa was introduced during libraries preparation as a contamination, rather than

a biological clue reflecting the heterogeneous sample sources. Nevertheless, we could not explicitly estimate the extent to which the alteration of gut microbiota was solely due to the introduced feed-borne bacteria, since the DNA-based amplicon sequencing, as was used here, is unable to ascertain (transcriptionally) active bacteria. Of note, dozens of ESVs filtered out as chloroplastid were mostly derived from the feed (~13.7% of total reads), but barely detected in gut (~0.1%). Such disproportion indicates that the over-represent feed-borne bacteria in the gut of slow growth fugu were not exclusively due to the influx of DNA from dead or dormant cells (Parris, *et al.*, 2019). Rather, bacteria such as *Pseudomonas*, optimized to grow at high nutrient as copiotroph (with 6–10 average 16S rRNA copy number) (Lladó and Baldrian, 2017), are known to be more competitive than oligotrophy in rapid growing or efficiently utilizing resource in the intestinal environment. Importantly, we found the response of individual gut microbiota to feed-borne bacteria varied with the fish growth rate and became more pronounced in slow growth fugu.

Previous studies suggested that the transit rate and residence time of diet in the gastrointestinal tract may influence the gut microbiota and host–microbial interactions (Ringø, *et al.*, 2016). Furthermore, gastric evacuation times varied substantially across fish species and developmental stages, and many factors like temperature, fish size and diet characteristics may contribute to such variation (Flowerdew and Grove, 1979; Kristiansen, 1998; Nikolopoulou, *et al.*, 2011; Ortiz-Monís, *et al.*, 2018; Gilannejad, *et al.*, 2019). For instance, the gastric evacuation half-time in the gastrointestinal tract of gilt-head seabream *Sparus aurata* was markedly prolonged (~3 h) in juvenile (Nikolopoulou, *et al.*, 2011; Gilannejad, *et al.*, 2019) than that (~30 min) in larvae (Ortiz-Monís, *et al.*, 2018). In the other study, two age-equal (half yearlings) cohorts of rainbow trout (*Oncorhynchus mykiss*) and brown trout (*Salmo trutta*), respectively, were found to have prolonged residence time of feed in the gastrointestinal tract of the smaller juvenile fishes and then decreased with body weight (Kristiansen, 1998). Moreover, the fish gastric evacuation can be further enhanced by transgenetically over-expressions of host growth hormone gene (Kim, *et al.*, 2018). These findings offer support for our observations that proportions of feed source signals in gut microbiota was negatively correlated with fugu growth rate within an age-equal cohorts, which could be largely due to the inconsistent gastric evacuation rate among different sized hosts. Several studies had suggested that the feed-associated microbiota could play a crucial role in host health and nutrition (Ruzauskas, *et al.*, 2018; Parris, *et al.*, 2019; Wilkes Walburn, *et al.*, 2019), yet little is known about the functional impact and the extent to which it exerted.



Herein, BugBase predicted the gut of slow growth fugu (with higher feed source signal) to have significantly more bacteria that are potentially pathogenic and biofilms forming than other fish clusters (Fig. 3). Such phenotypical differences are driven mostly by the higher proportion of both Pseudomonadaceae and Enterobacteriaceae (Fig. S9), of which many members are generally considered commensals (Ruzauskas, *et al.*, 2018). But like other opportunistic pathogens, certain members from both bacterial families often contain genes for virulence factors and antibiotic resistance (Cazares, *et al.*, 2020; Hamza, *et al.*, 2020). It's interesting that, the Pseudomonadaceae that was considered as presumptive feed-borne bacteria had very high relative abundance in feed (65.9%), whereas Enterobacteriaceae is relatively rare in feed (0.3%) in comparison to gut (3.8% in slow growth fugu) (Fig. S7). Therefore, we hypothesized that the observed alteration of fugu gut microbiota was determined by a combinational effects of (i) intestinal residence time allowing feed-borne bacteria to proliferate and (ii) response outcomes from interactions between allochthonous and autochthonous bacteria. The intuitive expectation is that, in slow growth fugu, further reduced gastric evacuation rate and more inadequate anaerobic niche in intestinal lumen allowed the faster grow of aerobic feed-borne bacteria (mainly *Pseudomonas*) and simultaneously influenced the abundance of autochthonous bacteria by either enhancing or diminishing their growth. As an further extreme assumption, the increased feed-borne bacteria in gut microbiota may lead to an accumulation of negative effects (e.g. pathogenesis) for the host, thereby stimulated intestinal stress response and hampered metabolic and digestion in host can be expected, which consequently could lead to a further reduce scope for the fish growth (lower fitness). Future studies concerning the ingested feed-borne bacteria and its interactions with both gut bacterial symbiosis and its host are critical for a better understanding of fish health and metabolism.

#### *Quantitative estimation of ecological processes underlying the gut microbiota response to feed-borne bacteria*

We found significant phylogenetic signals related to traits of feed source and body mass ( $P < 0.05$ , Fig. S11). In addition, the mean within-community phylogenetic distances (ses.MNTD) in gut microbiota were all found to be negative (t-test,  $P < 0.001$ ), further confirming all the ESVs were phylogenetically correlated more closely than expected by chance (Fig. S12). However, all the K values were estimated to be positive but closer to zero (Table S2), indicating a random or convergent pattern of evolution for the phylogenetic niche conservatism

relating to aforementioned traits (Webb, *et al.*, 2002; Blomberg, *et al.*, 2003). Accordantly, many other studies had showed that robust phylogenetic signal was related to different variables within short distances across ecosystems of diverse types (Stegen, *et al.*, 2012; Yan, *et al.*, 2016; Tripathi, *et al.*, 2018; Jiao and Lu, 2020; Liu, *et al.*, 2020b). On this basis, recent advanced quantitative ecological models (Stegen, *et al.*, 2012; Stegen, *et al.*, 2013; Dini-Andreote, *et al.*, 2015) can be applied to compare observed between-community diversity of both phylogeny and taxonomy against a null model, thereby quantify the relative importance of ecological processes underlying the gut microbiota in response to feed-borne bacteria.

The turnover of an ecological community (fugu gut microbiota in our case) is driven by two forces, determinacy and stochasticity (Hubbell, 2001). Determinacy indicates that the presence/absence and relative abundances of species is governed by abiotic and biotic factors, which are associated with ecological selection (niche effect) (Vellend, 2010). In contrast, stochasticity indicates that the relative abundances (ecological drift) of species are determined by probabilistic dispersal and randomness (neutral effect), which are not caused by environmentally determined fitness (Hubbell, 2001; Chase and Myers, 2011). Our Mantel's test results showed that severe differences in feed source signals act as a strong habitat filter and lead to significant phylogenetic clustering ( $|\beta_{NTI}| > 2$ ), whereas the degree to which community composition influenced by such filtering was weakened under moderate feed source signal (Fig. 4A). This result further emphasized that niche effect posed a strong selection pressure in determining the community assembly (Chase and Myers, 2011), thereby driving bacterial communities into a phylogenetically more closed cluster (Webb, *et al.*, 2002). As a continual inoculum, feeding poses a significant destabilizing force in fish intestinal microbiota and results in a continuous community turnover (Parris, *et al.*, 2019; Wilkes Walburn, *et al.*, 2019). On this basis, community invasibility, the successful colonization of allochthonous organisms such as feed-borne bacteria in a given community like fish gut, can be indicative of community stability (Litchman, 2010; Shade, *et al.*, 2012). In supporting this notion, we found that with the changes in feed-signals increases, the assembly processes of gut bacterial community were gradually shifted from homogeneous selection, to stochasticity, then to variable selection (Fig. 4A). Of note, such pattern was more obviously observed for rare sub-community (Mantel  $R = 0.492$ , Fig. 4C) than to abundant one (Mantel  $R = 0.275$ , Fig. 4B), indicating the abundant bacterial taxa (mostly gamma-proteobacterial copiotroph) might be more competitively dominant and persistent in gut niche. Specifically, the variable selection

process could be possibly explained by the increased habitat complexities of tripartite interactions among host, bacteria and nutrient, considering accumulated ingested feed-borne bacteria can cause more disturbance, thereby imposed a larger impact on rare sub-community as was evident by the increased alpha diversity in slow growth fugu (more prone to feed disturbance). As for the homogeneous selection, we inferred that the deterministic factors, particularly by the host and autochthonous bacteria, became more dominant and stable (e.g. lower alpha diversity in fast growth fugu) as the changes of feed source signal reduced, thereby jointly exerted more convergent force of selection.

By inferring the turnover in ESV composition ( $RC_{bc}$ ), the relative importance of stochastic processes can be further fractionated into homogenizing dispersal, dispersal limitation and the undominated fraction (Stegen, *et al.*, 2015). We observed a large fraction of community shift was governed by dispersal limitation (~40%), whereas very little (<3%) by homogenizing dispersal (Fig. 4D). Higher level of dispersal limitation can be viewed as a result of host effect, by which the degree to which bacteria move among individual gut communities was limited, even though the influence from the feed is strong. In line with these, both the fitness of neutral model ( $R^2 = 0.154$ ) and migration rate ( $m = 0.012$ ) for abundant sub-community were lower than that of rare sub-community ( $R^2 = 0.639$ ;  $m = 0.057$ ), further emphasizing the limited levels of dispersal leads to dissipation of diversity into local gut communities, thus decreased alpha diversity (Chase and Myers, 2011). The lower degree of homogenizing dispersal (Cadotte, 2006) is consistent with our observed higher degree of beta dispersion in fugu gut microbiota. In addition, an undominated fraction was also observed for all sub-communities, particularly to a larger extent in abundant one (49%). The inflated undominated fraction could be viewed as the results of diminished selection and/or dispersal rates, that both are opposite to that of variable selection (Stegen, *et al.*, 2015). As such, it is tempting to speculate that the observed undominated fraction here is likely reflecting the inconsistent selection from either feed-borne bacteria and host filtering (or both), which possibly due to the varied host physiology (e.g. gastric evacuation and growth) and/or gut nutrition (e.g. feeding fluctuation and frequency) (Flowerdew and Grove, 1979; Bolnick, *et al.*, 2014; Parris, *et al.*, 2019).

In conclusion, this snapshot study shown here characterized the compositional, functional and phylogenetical stability of gut microbiota from farmed fugu within a local pond environment. This stability is accompanied largely by homogeneous selection and dispersal limitation, thereby, a reduced intraspecies competition, enabling community stability and partial variations among hosts.

However, presumptive attaching bacteria introduced by feed can pose a strong restructuring force upon gut bacterial communities. Such disturbance involved with variable selection, therefore, an augmented interaction between autochthonous and allochthonous species, entailing community instability and shift. Moreover, we observed marked alterations in the composition and function of gut microbiota in slow growth host, potentially correlated to the different host physiological conditions, for example gastric evacuation rate and intestinal transit time of digesta. Nevertheless, we cannot completely rule out the possibility that such community shift in gut microbiota can be due to the efflux DNA of feed-borne bacteria, considering the lack of measurement of active bacteria for both gut microbiota and feed in our experiment. Future longitudinal studies concerning the ingested feed-borne bacteria and its interactions with both gut symbiosis and its host are needed to determine whether the observed correlations can be maintained or any causality can be confirmed. Altogether, our findings emphasized the intricate linkage between feed and gut microbiota, and highlighted the essential prerequisite to resolve the signal from feed-associated bacteria before the conclusions of comparative analysis of microbiota can be drawn.

## Experimental procedures

### Sample collection

We collected samples from 7-month-old juvenile obscure puffer (*T. obscurus*) in December 2017 from a single commercial supplier and raised in a fishfarm located in Yangzhong, East China. After acclimating to commercial extruded compound feed pellet (composite of fish meal, soybean meal, corn gluten meal, wheat flour, fish oil, soybean phospholipid oil, calcium dihydrogen phosphate, vitamin premix, mineral premix produced by Marine Fish Feed 891 Series), the fish offsprings from the full-sib family were introduced into indoor greenhouse ponds (80m x 25m x 1.8m) with a stocking density of 25 thousands per pond since April. Each pond had a separate water inlet and outlet. Three adjacent ponds were selected as the objects of study. The practical pond management is identical in terms of daily feeding rate (twice per day, i.e. 10am and 2pm), and pond water exchange rate (twice per month, 10–20%). Approximately 20 individual juvenile fugu (1 day post-feeding) per pond with comparable hepatosomatic index (10–15%) with other studies on *T. obscurus* growth performance (Cheng, *et al.*, 2018; Liu, *et al.*, 2020a) were randomly captured using a fishing net along the different sites of each pond. No obvious fin damage or skin injury were observed across the body among individuals. Fishes were anesthetized with 0.05% MS-222 (Sigma-

Aldrich, St. Louis, MO, USA) and immediately weighted and fork length measured. After being surface sterilized with 75% ethanal, the 2 cm long part of distal intestine was dissected out to collect the gut content using sterilized spatula. Subsequently, other tissues including the whole liver and gonad were also dissected and rapidly weighted. All samples were placed in individually labelled tubes and kept well in dry ice before shipping. All animal experiments were performed following the protocol approved by the Ethics Committee of Experimental Animals at Hohai University, and were in direct accordance with the Animal Care Guidelines issued by the Ministry of Science and Technology, the People's Republic of China.

For the environmental samples including pond water and sediment, three longitudinal distributed sites (similar locations among the ponds) per pond were chose respectively. Approximately 2.0 L of pond water was collected from 10 to 20 cm below surface per sites within each pond respectively. Water sample was prefiltered by 100- $\mu$ m pore sized nylon mesh, and the flowthrough was further filtered with 0.22 $\mu$ m Millipore membrane with vacuum pump to collect the water planktonic microbes. The sediment was collected from the same site as water sample and was packaged in airtight sterile plastic bags. All the microbial samples were preserved with dry ice during the shipment and stored at  $-80^{\circ}\text{C}$  before use. The detailed metadata for each sample was recorded in Table S3.

#### DNA isolation and bacterial 16S rRNA amplicon sequencing

All the fugu gut contents ( $n = 61$ ), feed pellet ( $n = 3$ ), water ( $n = 9$ ) and sediment ( $n = 9$ ) samples were subjected to bacterial 16S rRNA amplicon profiling by Illumina sequencing. DNA was extracted using FastDNA<sup>TM</sup> SPIN Kit for Soil (MP Biomedicals, Irvine, CA, USA) according to manufacturer's instructions and was verified by 0.8% agarose gel electrophoresis. The integrity and concentration of DNA were determined by NanoDrop ND 2000 spectrophotometer (Thermo Fisher Scientific Inc., Waltham, MA, USA) and Quant-iT PicoGreen dsDNA assay kit (Thermo Fisher Scientific Inc.) respectively. The V4-V5 hypervariable region of bacterial 16S rRNA gene were amplified using the primer pair 515F (5'-CTGCCAGCMGCCGCGGTAA-3') and 926R (5'-CCGTCAATTCMTTTRAGTTT-3'), and the unique eight-nucleotide barcode sequences were incorporated into each sample. Each sample was tested in triplicate (together with No template controls) with the 25- $\mu$ l reaction mix consisted of 40 ng template DNA, 5  $\mu$ l reaction buffer (5x), 5  $\mu$ l GC buffer (5x), 2  $\mu$ l dNTP (2.5mM), 10 pM of barcoded forward and reverse primers, and 0.75 U Q<sub>5</sub><sup>®</sup> High-Fidelity DNA Polymerase (NEB, Ipswich, MA,

USA). The thermocycling programme consisted of one hold at 98  $^{\circ}\text{C}$  for 2 min, followed by three-step cycles of 15 s at 95  $^{\circ}\text{C}$ , 30 s at 55  $^{\circ}\text{C}$  and 30 s at 72  $^{\circ}\text{C}$ , and one final hold at 72 $^{\circ}\text{C}$  for 5 min.

All PCR amplicons were subjected to 2% agarose gel electrophoresis and purified with the AMPure XP Kit (Beckman Coulter GmbH, USA), quality checked using Agilent High Sensitivity DNA Kit (Agilent, Santa Carla, CA, USA) and Quant-iT PicoGreen dsDNA assay kit. DNA libraries for purified PCR amplicons were constructed using an Illumina TruSeq Nano DNA LT Library Prep Kit (Illumina, San Diego, CA, USA) according to the manufacturer's instruction, and 300-bp pair-ended insertion were sequenced using MiSeq Reagent Kit V3 chemistry on Illumina MiSeq platform owned by Personal Biotechnology Co., Ltd. (Shanghai, China).

#### Reads processing and bacterial 16S rRNA gene profiling

The adapters (indexers) in the paired-end raw reads were trimmed out by the quality control tool Trim-Galore (a wrapper tool based on Cutadapt v.1.4.2 and FastQC v.0.10.1) for high-throughput sequence data, as set-up by default quality threshold of Q20 (Martin, 2011). The FastQC reports from both before- and after-trimming were checked. Further, the 16S rRNA amplicon raw reads were processed according to the previous literatures (Zhang, *et al.*, 2018; Zhang, *et al.*, 2019) using USEARCH v.10.0 (Edgar, 2010). In brief, the paired-end reads were merged and relabelled using '-fastq\_mergepairs'; both barcodes and primers were removed using '-fastx\_truncate'; reads with low-quality (error rates > 0.01) and redundancy were filtered and dereplicated using '-fastq\_filter' and '-fastx\_uniques' respectively. The *de novo* biological sequences, that is ESVs (exact sequence variants), were clustered, and chimeras were filtered using '-unoise3'. Subsequently, the operational taxonomic units (OTUs) table was created by '-otutab'. The taxonomy for each representative sequences was assigned by the SYNTAX algorithm using the Ribosomal Database Project (RDP) classifier (RDP training set v16) (Wang, *et al.*, 2007; Cole, *et al.*, 2013), and chloroplast ESVs were removed after taxonomy-based filtering. Sequencing depth was normalized by subsample using 'otutab\_norm', yielding 7000 sequences per sample to keep all the samples for downstream analysis. Community diversity was analysed using '-alpha\_div', '-alpha\_div\_rare', '-cluster\_agg' and '-beta\_div' respectively.

#### Featured taxa identifying and source tracking across different bacterial communities

To discriminate the compositional difference and identify the featured taxa across different bacterial communities,

R package RandomForest (v.4.6-14) (Liaw and Wiener, 2002) was used to classify the relative abundances of bacterial taxa in the family level across different sample source using parameters of 'ntree = 1000, importance = TRUE, proximity = TRUE', followed by cross-validation using *rfcv()* function for feature selection. Important features were clustered according to the NCBI taxonomy and visualized by Interactive Tree of Life, iTOL (Letunic and Bork, 2019).

To further identify the potential origins for the compositional structure for each indicated sample group, SourceTracker (Knights, *et al.*, 2011) was used to estimate the proportions of source bacteria from environments in each community. All the samples ( $n = 82$ ) were first rarefied at 1000 sequences as training set, and SourceTracker object was trained by *sourcetracker()* function. Each indicated sample group was used as test set and of which the source proportions (from SourceTracker object) were estimate by *predict()* function with the alpha values tuned to 0.001.

The ESV table was further analysed to identify featured bacterial taxa that are specific to indicated group. ESVs were filtered with abundance greater than 0.0001% across gut samples. Significant changes in relative ESV abundance were identified with threshold of logarithmic LDA (linear discriminant analysis) score greater than 2 combining with effect size (LEfSe) algorithm (Segata, *et al.*, 2011).

#### *Phenotypic inferring of bacterial communities*

For metagenomics inference, the ESVs table normalized by 16S rRNA copy numbers were annotated by PICRUSt (Langille, *et al.*, 2013) with closed reference picking (Greengenes v.13\_5) (DeSantis, *et al.*, 2006). Abundance of bacterial communities with two trophic modes of bacterial life namely copiotroph and oligotroph were calculated by computing the mean number of ribosomal operon copies in the genome across all bacterial ESVs presented in each sample as previously described (Nemergut, *et al.*, 2016; Prest, *et al.*, 2018). Firstly, rRNA copy number per ESVs (predicted by PICRUSt) within a given sample was multiplied by the corresponding relative abundance and summed for the total copy number per sample. Secondly, mean community copy number was averaged per indicated sample group (fish gut and environments) and were subjected to group-wised multiple comparisons using indicated statistical approach as will be described below. Biologically interpretable phenotypes such as oxygen tolerance, Gram staining and pathogenic potential, within each indicated community were predicted based on the Greengenes OTU table using BugBase (Ward, *et al.*, 2017). Taxonomic level was set as bacterial family.

#### *Quantifying relative importance of ecological processes in gut bacterial community assemblage*

The existence and degree of phylogenetic relatedness in species traits were assessed through the calculation of the phylogenetic signal of a given trait. To this end, both Blomberg's K statistic (Blomberg, *et al.*, 2003) and Mantel's Correlogram (Oden and Sokal, 1986) were used to measure phylogenetic signal using *phyloSignal()* and *phyloCorrelogram()* implemented in R package 'PhyloSignal' (Keck, *et al.*, 2016). To assess the phylogenetic structure of gut bacterial communities, the standardized effect size of mean nearest taxon distances (ses.MNTD) per community was calculated using *ses.mntd()* function implemented in 'Picante' package (Kembel, *et al.*, 2010), supplied with options of 'taxa.labels' as null model, unweighted and 999 randomization. Ses.MNTD values that were significant less or greater than 0 indicate phylogenetically close-related or unrelated than expected by chance respectively (Webb, *et al.*, 2002). Ecological null model was used to quantify the relative importance of ecological processes in driving microbial community assemblage (Stegen, *et al.*, 2013). Briefly, both metrics of  $\beta$ -nearest taxon index ( $\beta$ NTI) and Bray–Curtis-based Raup–Crick ( $RC_{BC}$ ) were used to quantify the contributions from deterministic and stochastic processes respectively. First, the phylogenetic beta diversity were measured by computing unweighted inter-community MNTD metric ( $\beta$ MNTD) using the *comdistnt()*. Further, the  $\beta$ NTI, that is the degree to which observed  $\beta$ MNTD deviates from the mean of the null distribution (999 randomization) was computed according to Stegen *et al.* (Stegen, *et al.*, 2013). The influence of deterministic community turnover including homogeneous selection and variable selection (Vellend, 2010) is estimated by inferring the fraction of pairwise community comparisons with significance thresholds of  $\beta$ NTI < -2 and  $\beta$ NTI > 2 respectively (Stegen, *et al.*, 2013). In turn, the non-significant fractions with  $|\beta$ NTI| < 2 that indicate community turnover governed by stochastic ecological processes were further subdivided according to the taxonomic  $\beta$ -diversity metrics of  $RC_{BC}$ . Within this subset ( $|\beta$ NTI| < 2),  $RC_{BC}$  metric normalizes the deviation between observed Bray–Curtis and the null distribution (9,999 randomization) to vary between -1 and + 1. The turnover of microbial community with significant  $RC_{BC}$  values that are less than -0.95 or greater than + 0.95 were interpreted as governed by homogenizing dispersal and dispersal limitation respectively. The ones with non-significant values (i.e.  $|RC_{BC}| < 0.95$ ) was regarded as community turnover governed by undominated processes without a profound effect by any given ecological pressure (i.e. weak dispersal and weak selection).

### Statistical analysis

All statistical analyses were performed under R environment (v3.6.0) through RStudio (v1.1.463). Before analysis, the normality and homogeneity of variances of data sets were tested using *shapiro.test()* and *bartlett.test()* respectively. In turn, parametric or non-parametric tests were applied accordingly. For parametric data sets, group means were compared by one-way ANOVA using *aov()*, followed by *post hoc* test using *TukeyHSD()*. For non-parametric data sets, rank sum were compared by Kruskal–Wallis test using *kruskal.test()*, followed by Wilcoxon rank sum test using *wilcox.test()*. Significance were considered while FDR-adjusted *P* values were less than 0.05, and the detailed statistical test applied in each data set is elaborated in the figure legend.

The homogeneity of multivariate dispersion of groups was evaluated by (PERMDISP) test with *betadisper()* followed by permutation test with *permutest()*. Ordination (principal coordinate analyses, PCoA) were performed based on Bray–Curtis dissimilarity metric using *cmdscale()* followed by permutational multivariate analyses of variance (PERMANOVA) test by *adonis()* implemented in ‘vegan’ package (Oksanen, *et al.*, 2019). The clustering analysis based on individual fish body mass were determined by *eclust ()* using K-means distance and *clusGap()* conducting gap statistic algorithm implemented in ‘factoextra’ package (Kassambara and Mundt, 2020), followed by visualization according to principal components analysis by *prcomp ()*. Spearman correlations between the phylogenetic beta diversity ( $\beta$ NTI matrix) and indicated traits difference (euclidean-based trait distance matrix) we assessed using *mantel()* implemented in ‘ecodist’ package (Goslee and Urban, 2007) with 9999 permutations. The *neutral.fit()* in the ‘MicEco’ package (Russel, 2020) was used to evaluate the fitness of neutral model (Sloan, *et al.*, 2006) on the gut bacterial community and to determine the contribution of neutral processes in community assembly.

### Acknowledgements

This work was supported by the National Key Research and Development Program of China (2018YFD0900200); the National Natural Science Foundation of China (31872597); the Fundamental Research Funds for the Central Universities from China (B200202141); and the Earmarked Fund for Jiangsu Agricultural Industry Technology System JATS[2020]465). We thank Mr Zhu Jikun for his deep knowledge of pufferfish husbandry, and Song Jing, Gao Tianheng, Gao Fanxiang, Wu Chao, Liu Yupeng and Wen Duan for their help in sampling. We would like to thank the editor and the reviewer(s), who

provided comments and suggestions that helped to substantially improve the manuscript.

### Conflict of interest

The authors declare no conflict of interest.

### References

- Abdul Razak, S., and Scribner, K.T. (2020) Ecological and ontogenetic components of larval lake sturgeon gut microbiota assembly, successional dynamics, and ecological evaluation of neutral community processes. *Appl Environ Microbiol* **86**: e02662–02619.
- Aparicio, S., Chapman, J., Stupka, E., Putnam, N., Chia, J.-M., Dehal, P., *et al.* (2002) whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*. *Science* **297**: 1301.
- Berg, G., Rybakova, D., Fischer, D., Cernava, T., Vergès, M.-C.C., Charles, T., *et al.* (2020) Microbiome definition re-visited: old concepts and new challenges. *Microbiome* **8**: 103.
- Blomberg, S.P., Garland, T., and Ives, A.R. (2003) Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution* **57**: 717–745.
- Bolnick, D.I., Snowberg, L.K., Hirsch, P.E., Lauber, C.L., Knight, R., Caporaso, J.G., and Svanbäck, R. (2014) Individuals’ diet diversity influences gut microbial diversity in two freshwater fish (threespine stickleback and Eurasian perch). *Ecol Lett* **17**: 979–987.
- Bordenstein, S.R., and Theis, K.R. (2015) Host biology in light of the microbiome: ten principles of holobionts and hologenomes. *PLoS Biol* **13**: e1002226.
- Breck, J.E. (2014) Body composition in fishes: body size matters. *Aquaculture* **433**: 40–49.
- Brenner, S., Elgar, G., Sanford, R., Macrae, A., Venkatesh, B., and Aparicio, S. (1993) Characterization of the pufferfish (*Fugu*) genome as a compact model vertebrate genome. *Nature* **366**: 265–268.
- de Bruijn, I., Liu, Y., Wiegertjes, G.F., and Raaijmakers, J.M. Exploring fish microbial communities to mitigate emerging diseases in aquaculture. *FEMS Microbiol Ecol* **94**(1): fix161. <https://doi.org/10.1093/femsec/fix161>
- Cadotte, M.W. (2006) Dispersal and species diversity: a meta-analysis. *Am Nat* **167**: 913–924.
- Cazares, A., Moore, M.P., Hall, J.P.J., Wright, L.L., Grimes, M., Emond-Rhéault, J.-G., *et al.* (2020) A megaplasmid family driving dissemination of multidrug resistance in *Pseudomonas*. *Nat Commun* **11**: 1370.
- Chao, R., Yu, L., Xufa, M., Yadong, X., Mingxu, X., Yuting, Z., *et al.* Interactions between commensal bacteria and viral infection: insights for viral disease control in farmed animals. *Science China Life Sciences*.
- Chase, J.M., and Myers, J.A. (2011) Disentangling the importance of ecological niches from stochastic processes across scales. *Phil Transac Roy Soc B: Biol Sci* **366**: 2351–2363.
- Cheng, C.-H., Guo, Z.-X., Ye, C.-X., and Wang, A.-L. (2018) Effect of dietary astaxanthin on the growth performance, non-specific immunity, and antioxidant capacity of

- pufferfish (*Takifugu obscurus*) under high temperature stress. *Fish Physiol Biochem* **44**: 209–218.
- Clements, K.D., Angert, E.R., Montgomery, W.L., and Choat, J.H. (2014) Intestinal microbiota in fishes: what's known and what's not. *Mol Ecol* **23**: 1891–1898.
- Cole, J.R., Wang, Q., Fish, J.A., Chai, B., McGarrell, D.M., Sun, Y., *et al.* (2013) Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res* **42**: D633–D642.
- De Schryver, P., and Vadstein, O. (2014) Ecological theory as a foundation to control pathogenic invasion in aquaculture. *ISME J* **8**: 2360–2368.
- DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., *et al.* (2006) Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. *Appl Environ Microbiol* **72**: 5069.
- Dini-Andreote, F., Stegen, J.C., van Elsas, J. D., and Salles, J.F. (2015) Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proc Natl Acad Sci USA* **112**: E1326.
- Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460–2461.
- Egerton, S., Culloty, S., Whooley, J., Stanton, C., and Ross, R.P. (2018) The gut microbiota of marine fish. *Front Microbiol* **9**: 873.
- Eichmiller, J.J., Hamilton, M.J., Staley, C., Sadowsky, M.J., and Sorensen, P.W. (2016) Environment shapes the fecal microbiome of invasive carp species. *Microbiome* **4**: 44.
- Flowerdew, M.W., and Grove, D.J. (1979) Some observations of the effects of body weight, temperature, meal size and quality on gastric emptying time in the turbot, *Scophthalmus maximus* (L.) using radiography. *J Fish Biol* **14**: 229–238.
- Gao, F.-X., Shi, Y., Duan, W., Lu, W.-J., Huang, W., Zhang, X.-J., *et al.* (2020) A rapid and reliable method for identifying genetic sex in obscure pufferfish (*Takifugu obscurus*). *Aquaculture* **519**: 734749.
- Giatsis, C., Sipkema, D., Smidt, H., Heilig, H., Benvenuti, G., Verreth, J., and Verdegem, M. (2015) The impact of rearing environment on the development of gut microbiota in tilapia larvae. *Sci Rep* **5**: 18206.
- Gilannejad, N., Silva, T., Martínez-Rodríguez, G., and Yúfera, M. (2019) Effect of feeding time and frequency on gut transit and feed digestibility in two fish species with different feeding behaviours, gilthead seabream and Senegalese sole. *Aquaculture* **513**: 734438.
- Golaś, I., Szmyt, M., Potorski, J., Łopata, M., Gotkowska-Płachta, A., and Glińska-Lewczuk, K. (2019) Distribution of pseudomonas fluorescens and aeromonas hydrophila bacteria in a recirculating aquaculture system during farming of european grayling (*Thymallus thymallus* L.) Broodstock. *Water* **11**: 376.
- Goslee, C.S., and Urban, L.D. (2007) The ecodist package for dissimilarity-based analysis of ecological data. *J Stat Softw* **22**: 19.
- Hamza, D., Dorgham, S., Ismael, E., El-Moez, S. I. A., Elhariri, M., Elhelw, R., and Hamza, E. (2020) Emergence of  $\beta$ -lactamase- and carbapenemase- producing Enterobacteriaceae at integrated fish farms. *Antimicrob Resist Infect Control* **9**: 67.
- Hatzenpichler, R., Krukenberg, V., Spietz, R.L., and Jay, Z. J. (2020) Next-generation physiology approaches to study microbiome function at single cell level. *Nat Rev Microbiol* **18**: 241–256.
- Hubbell, S.P. (2001) *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton, NJ: Princeton University Press.
- Jiao, S., and Lu, Y. (2020) Soil pH and temperature regulate assembly processes of abundant and rare bacterial communities in agricultural ecosystems. *Environ Microbiol* **22**: 1052–1065.
- Kassambara, A., and Mundt, F. (2020) *factoextra: Extract and Visualize the Results of Multivariate Data Analyses*, CRAN-The Comprehensive R Archive Network.
- Keck, F., Rimet, F., Bouchez, A., and Franc, A. (2016) phyloSignal: an R package to measure, test, and explore the phylogenetic signal. *Ecol Evol* **6**: 2774–2780.
- Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D., *et al.* (2010) Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**: 1463–1464.
- Kim, J.-H., Chatchaiphan, S., Crown, M.T., White, S.L., and Devlin, R.H. (2018) Effect of growth hormone overexpression on gastric evacuation rate in coho salmon. *Fish Physiol Biochem* **44**: 119–135.
- Knights, D., Kuczynski, J., Charlson, E.S., Zaneveld, J., Mozer, M.C., Collman, R.G., *et al.* (2011) Bayesian community-wide culture-independent microbial source tracking. *Nat Methods* **8**: 761–763.
- Kristiansen, H.R. (1998) Effects of handling, discrete meals and body weight on the individual variation of gastric emptying parameters. *Aquac Res* **29**: 717–729.
- Langenheder, S., and Lindström, E.S. (2019) Factors influencing aquatic and terrestrial bacterial community assembly. *Environ Microbiol Rep* **11**: 306–315.
- Langille, M.G.I., Zaneveld, J., Caporaso, J.G., McDonald, D., Knights, D., Reyes, J.A., *et al.* (2013) Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol* **31**: 814–821.
- Lankiewicz, T.S., Cottrell, M.T., and Kirchman, D.L. (2016) Growth rates and rRNA content of four marine bacteria in pure cultures and in the Delaware estuary. *The ISME Journal* **10**: 823–832.
- Lauro, F.M., McDougald, D., Thomas, T., Williams, T.J., Egan, S., Rice, S., *et al.* (2009) The genomic basis of trophic strategy in marine bacteria. *Proc Natl Acad Sci USA* **106**: 15527.
- Letunic, I., and Bork, P. (2019) Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* **47**: W256–W259.
- Li, Y., Bruni, L., Jaramillo-Torres, A., Gajardo, K., Kortner, T.M., and Krogdahl, Å. (2020a) Differential response of digesta- and mucosa-associated intestinal microbiota to dietary black soldier fly (*Hermetia illucens*) larvae meal in seawater phase atlantic salmon (*Salmo salar*), *bioRxiv*. <https://doi.org/10.1101/2020.05.08.083899>
- Li, Z., Tian, J., Lai, Y., Lee, C.-H., Cai, Z., and Yu, C.-F. (2020b) puffer fish gut microbiota studies revealed unique bacterial co-occurrence patterns and new insights on tetrodotoxin producers. *Marine Drugs* **18**: 278.



- Li, X., Zhou, L., Yu, Y., Ni, J., Xu, W., and Yan, Q. (2017) Composition of Gut Microbiota in the Gibel Carp (*Carassius auratus gibelio*) Varies with Host Development. *Microb Ecol* **74**: 239–249.
- Liaw, A., and Wiener, M. (2002) Classification and Regression by randomForest. *R News* **2**: 5.
- Lieke, T., Meinelt, T., Hoseinifar, S.H., Pan, B., Straus, D.L., and Steinberg, C.E.W. (2020) Sustainable aquaculture requires environmental-friendly treatment strategies for fish diseases. *Rev Aquacult* **12**: 943–965.
- Litchman, E. (2010) Invisible invaders: non-pathogenic invasive microbes in aquatic and terrestrial ecosystems. *Ecol Lett* **13**: 1560–1572.
- Liu, Q., Ye, H., Xu, M., Mai, K., Sun, Z., Tan, X., et al. (2020a) The effect and mechanism of dietary monocalcium phosphate restriction on hepatic lipid deposition and immune status in obscure puffer, *Takifugu obscurus*. *Aquaculture* **524**: 735261.
- Liu, J., Zhu, S., Liu, X., Yao, P., Ge, T., and Zhang, X.-H. (2020b) Spatiotemporal dynamics of the archaeal community in coastal sediments: assembly process and co-occurrence relationship. *ISME J* **14**: 1463–1478.
- Lladó, S., and Baldrian, P. (2017) Community-level physiological profiling analyses show potential to identify the copiotrophic bacteria present in soil environments. *PLoS One* **12**: e0171638.
- Lloyd-Price, J., Mahurkar, A., Rahnavard, G., Crabtree, J., Orvis, J., Hall, A.B., et al. (2017) Strains, functions and dynamics in the expanded Human Microbiome Project. *Nature* **550**: 61–66.
- Lozupone, C.A., Stombaugh, J.I., Gordon, J.I., Jansson, J.K., and Knight, R. (2012) Diversity, stability and resilience of the human gut microbiota. *Nature* **489**: 220–230.
- Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.J* **17**: 10.
- Muziasari, W.I., Pitkänen, L.K., Sørum, H., Stedtfeld, R.D., Tiedje, J.M., and Virta, M. (2017) The resistome of farmed fish feces contributes to the enrichment of antibiotic resistance genes in sediments below baltic sea fish farms. *Front Microbiol* **7**: 2137.
- Nemergut, D.R., Knelman, J.E., Ferrenberg, S., Bilinski, T., Melbourne, B., Jiang, L., et al. (2016) Decreases in average bacterial community rRNA operon copy number during succession. *ISME J* **10**: 1147–1156.
- Nikolopoulou, D., Moutou, K.A., Fountoulaki, E., Venou, B., Adamidou, S., and Alexis, M.N. (2011) Patterns of gastric evacuation, digesta characteristics and pH changes along the gastrointestinal tract of gilthead sea bream (*Sparus aurata* L.) and European sea bass (*Dicentrarchus labrax* L.). *Comp Biochem Physiol A: Mol Integr Physiol* **158**: 406–414.
- Oden, N.L., and Sokal, R.R. (1986) Directional autocorrelation: an extension of spatial correlograms to two dimensions. *Syst Zool* **35**: 608–617.
- Okabe, T., Oyama, H., Kashitani, M., Ishimaru, Y., Suo, R., Sugita, H., and Itoi, S. (2019) Toxic flatworm egg plates serve as a possible source of tetrodotoxin for pufferfish. *Toxins* **11**: 402.
- Oksanen, J., Blanchet, F., Guillaume, F.M., Kindt, R., Legendre, P., McGlenn, D., et al. (2019) *vegan: Community Ecology Package, CRAN-The Comprehensive R Archive Network*.
- Ortiz-Monís, M.A., Mancera, J.M., and Yúfera, M. (2018) Determining gut transit rates in gilthead seabream larvae fed microdiets. *Aquaculture* **495**: 523–527.
- Parris, D.J., Morgan, M.M., and Stewart, F.J. (2019) Feeding rapidly alters microbiome composition and gene transcription in the clownfish gut. *Appl Environ Microbiol* **85**: e02479–02418.
- Pereira, F.C., and Berry, D. (2017) Microbial nutrient niches in the gut. *Environ Microbiol* **19**: 1366–1378.
- Perry, W.B., Lindsay, E., Payne, C.J., Brodie, C., and Kazlauskaitė, R. (2020) The role of the gut microbiome in sustainable teleost aquaculture. *Proc R Soc B: Biol Sci* **287**: 20200184.
- Prest, T.L., Kimball, A.K., Kueneman, J.G., and McKenzie, V.J. (2018) Host-associated bacterial community succession during amphibian development. *Mol Ecol* **27**: 1992–2006.
- Riiser, E.S., Haverkamp, T.H.A., Borgon, Ø., Jakobsen, K.S., Jentoft, S., and Star, B. (2018) A single vibriomera 16S rRNA oligotype dominates the intestinal microbiome in two geographically separated atlantic cod populations. *Front Microbiol* **9**: 873.
- Riiser, E.S., Haverkamp, T.H.A., Varadharajan, S., Borgon, Ø., Jakobsen, K.S., Jentoft, S., and Star, B. (2020) Metagenomic shotgun analyses reveal complex patterns of intra- and interspecific variation in the intestinal microbiomes of codfishes. *Appl Environ Microbiol* **86**: e02788–02719.
- Ringø, E., and Birkbeck, T.H. (1999) Intestinal microflora of fish larvae and fry. *Aquac Res* **30**: 73–93.
- Ringø, E., Zhou, Z., Vecino, Jlg, Wadsworth, S., Romero, J., Krogdahl, Å., et al. (2016) Effect of dietary components on the gut microbiota of aquatic animals. A never-ending story? *Aquac Nutr* **22**: 219–282.
- Roeselers, G., Mittge, E.K., Stephens, W.Z., Parichy, D.M., Cavanaugh, C.M., Guillemin, K., and Rawls, J.F. (2011) Evidence for a core gut microbiota in the zebrafish. *ISME J* **5**: 1595–1608.
- Roller, B.R.K., Stoddard, S.F., and Schmidt, T.M. (2016) Exploiting rRNA operon copy number to investigate bacterial reproductive strategies. *Nat Microbiol* **1**: 16160.
- Romero, J., and Navarrete, P. (2006) 16S rDNA-based analysis of dominant bacterial populations associated with early life stages of Coho salmon (*Oncorhynchus kisutch*). *Microb Ecol* **51**: 422–430.
- Russel, J. (2020) *Russel88/MicEco: v0.9.11, Zenodo*.
- Ruzauskas, M., Klimiene, I., Armalyte, J., Bartkiene, E., Siugzdiniene, R., Skerniskyte, J., et al. (2018) Composition and antimicrobial resistance profile of Gram-negative microbiota prevalent in aquacultured fish. *J Food Saf* **38**: e12447.
- Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W.S., and Huttenhower, C. (2011) Metagenomic biomarker discovery and explanation. *Genome Biol* **12**: R60.
- Shade, A., and Gilbert, J.A. (2015) Temporal patterns of rarity provide a more complete view of microbial diversity. *Trends Microbiol* **23**: 335–340.

- Shade, A., Peter, H., Allison, S., Baho, D., Berga, M., Buergmann, H., *et al.* (2012) Fundamentals of microbial community resistance and resilience. *Front Microbiol* **3**: 417.
- Shiina, A., Itoi, S., Washio, S., and Sugita, H. (2006) Molecular identification of intestinal microflora in *Takifugu niphobles*. *Comp Biochem Physiol D: Genomics Proteomics* **1**: 128–132.
- Sloan, W.T., Lunn, M., Woodcock, S., Head, I.M., Nee, S., and Curtis, T.P. (2006) Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environ Microbiol* **8**: 732–740.
- Star, B., Haverkamp, T.H.A., Jentoft, S., and Jakobsen, K.S. (2013) Next generation sequencing shows high variation of the intestinal microbial species composition in Atlantic cod caught at a single location. *BMC Microbiol* **13**: 248.
- Stegen, J.C., Lin, X., Fredrickson, J.K., Chen, X., Kennedy, D.W., Murray, C.J., *et al.* (2013) Quantifying community assembly processes and identifying features that impose them. *ISME J* **7**: 2069–2079.
- Stegen, J.C., Lin, X., Fredrickson, J.K., and Konopka, A.E. (2015) Estimating and mapping ecological processes influencing microbial community assembly. *Front Microbiol* **6**: Article 370.
- Stegen, J.C., Lin, X., Konopka, A.E., and Fredrickson, J.K. (2012) Stochastic and deterministic assembly processes in subsurface microbial communities. *ISME J* **6**: 1653–1664.
- Sunagawa, S., Coelho, L.P., Chaffron, S., Kultima, J.R., Labadie, K., Salazar, G., *et al.* (2015) Structure and function of the global ocean microbiome. *Science* **348**: 1261359.
- Theis, K.R., Dheilly, N.M., Klassen, J.L., Brucker, R.M., Baines, J.F., Bosch, T.C.G., *et al.* (2016) Getting the hologenome concept right: an eco-evolutionary framework for hosts and their microbiomes. *mSystems* **1**: e00028–00016.
- Tripathi, B.M., Stegen, J.C., Kim, M., Dong, K., Adams, J.M., and Lee, Y.K. (2018) Soil pH mediates the balance between stochastic and deterministic assembly of bacteria. *ISME J* **12**: 1072–1083.
- Troussellier, M., Escalas, A., Bouvier, T., and Mouillot, D. (2017) Sustaining rare marine microorganisms: macroorganisms as repositories and dispersal agents of microbial diversity. *Front Microbiol* **8**: 947.
- Van de Peer, Y. (2004) Tetraodon genome confirms Takifugu findings: most fish are ancient polyploids. *Genome Biol* **5**: 250.
- Vellend, M. (2010) Conceptual synthesis in community ecology. *Q Rev Biol* **85**: 183–206.
- Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. (2007) Naïve bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* **73**: 5261.
- Wang, A.R., Ran, C., Ringø, E., and Zhou, Z.G. (2018) Progress in fish gastrointestinal microbiota research. *Rev Aquacult* **10**: 626–640.
- Wang, T., Wei, X., Chen, T., Wang, W., Xia, X., Miao, J., and Yin, S. (2019) Studies of the mechanism of fatty liver formation in *Takifugu fasciatus* following copper exposure. *Ecotoxicol Environ Saf* **181**: 353–361.
- Ward, T., Larson, J., Meulemans, J., Hillmann, B., Lynch, J., Sidiropoulos, D., *et al.* (2017) BugBase predicts organism-level microbiome phenotypes. *bioRxiv*. <https://doi.org/10.1101/133462>
- Webb, C.O., Ackerly, D.D., McPeck, M.A., and Donoghue, M.J. (2002) Phylogenies and community ecology. *Annu Rev Ecol Syst* **33**: 475–505.
- Wilkes Walburn, J., Wemheuer, B., Thomas, T., Copeland, E., O'Connor, W., Booth, M., *et al.* (2019) Diet and diet-associated bacteria shape early microbiome development in Yellowtail Kingfish (*Seriola lalandi*). *Microb Biotechnol* **12**: 275–288.
- Wong, S., Stephens, W.Z., Burns, A.R., Stagaman, K., David, L.A., Bohannan, B.J.M., *et al.* (2015) Ontogenetic differences in dietary fat influence microbiota assembly in the zebrafish gut. *MBio* **6**: e00687–00615.
- Wong, S., Waldrop, T., Summerfelt, S., Davidson, J., Barrows, F., Kenney, P. B., *et al.* (2013) Aquacultured rainbow trout (*Oncorhynchus mykiss*) possess a large core intestinal microbiota that is resistant to variation in diet and rearing density. *Appl Environ Microbiol* **79**: 4974.
- Xiong, J., Zhu, J., Dai, W., Dong, C., Qiu, Q., and Li, C. (2017) Integrating gut microbiota immaturity and disease-discriminatory taxa to diagnose the initiation and severity of shrimp disease. *Environ Microbiol* **19**: 1490–1501.
- Yan, Q., Li, J., Yu, Y., Wang, J., He, Z., Van Nostrand, J.D., *et al.* (2016) Environmental filtering decreases with fish development for the assembly of gut microbiota. *Environ Microbiol* **18**: 4739–4754.
- Zhang, Q.-Y., and Gui, J.-F. (2018) Diversity, evolutionary contribution and ecological roles of aquatic viruses. *Sci China Life Sci* **61**: 1486–1502.
- Zhang, J., Liu, Y.-X., Zhang, N., Hu, B., Jin, T., Xu, H., *et al.* (2019) NRT1.1B is associated with root microbiota composition and nitrogen use in field-grown rice. *Nat Biotechnol* **37**: 676–684.
- Zhang, J., Zhang, N., Liu, Y.-X., Zhang, X., Hu, B., Qin, Y., *et al.* (2018) Root microbiota shift in rice correlates with resident time in the field and developmental stage. *Sci China Life Sci* **61**: 613–621.
- Zhang, K., Zheng, X., He, Z., Yang, T., Shu, L., Xiao, F., *et al.* (2020a) Fish growth enhances microbial sulfur cycling in aquaculture pond sediments. *Microb Biotechnol* **13**: 1597–1610.
- Zhang, X., Zong, J., Chen, S., Li, M., Lu, Y., Wang, R., and Xu, H. (2020b) Accumulation and elimination of tetrodotoxin in the pufferfish *takifugu obscurus* by dietary administration of the wild toxic gastropod *nassarius semiplicata*. *Toxins* **12**: 278.
- Zilber-Rosenberg, I., and Rosenberg, E. (2008) Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. *FEMS Microbiol Rev* **32**: 723–735.
- Zmyslowska, I., and Lewandowska, D. (1999) Survival of bacterial strains in fish feeds stored in different temperatures. *Pol J Environ Stud* **8**: 447–451.

## Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Fig. S1.** Evaluation of sequencing depth of bacterial communities across different sample groups. (A) Rarefaction curves of detected bacterial species (numbers of observed ESVs) within each indicated microbiota sample tend to be ‘flat’ as the numbers of sampling increase, suggesting the sampling covers the most bacterial species (saturated). (B) Rarefaction curves of detected bacterial species (numbers of observed ESVs) in each indicated sample groups. Vertical bars denote standard error. Samples size for each indicated groups are as follows: fish ( $n = 61$ ), including fast ‘Fst’ ( $n = 24$ ), Medium ‘Med’ ( $n = 20$ ) and Slow ‘Slw’ ( $n = 17$ ); feed ‘Fd’ ( $n = 3$ ); water ‘Wtr’ ( $n = 9$ ); sediment ‘Sdm’ ( $n = 9$ ).

**Fig. S2.** Calculated alpha-diversity based on indices of (A) Chao1; (B) Shannon\_e (logarithm base  $e$ ); (C) Simpson among six groups of bacterial community. For each box, the horizontal bold bar denote medians; the height of box denotes the interquartile range (25th percentile – 75th percentile); the whiskers mark the values range within 1.5 times interquartile. Lower-cased letters denote statistical significance reported by Mann–Whitney  $U$  Test at confidence level of 0.95.

**Fig. S3.** Differences in group homogeneities was examined by betadisper() implemented in R package ‘vegan’. (A) Principal coordinates plot showing the Bray–Curtis dissimilarity distances between each sample and its group centroids. Samples groups: F: Fugu gut; FD: feed pellet; S: sediment; W: water. (B) Boxplot of distances to the group centroid for the four groups of bacterial community. For each box, the horizontal bold bar denote medians; the height of box denotes the interquartile range (25th percentile – 75th percentile); the whiskers mark the values range within 1.5 times interquartile.

**Fig. S4.** Bars of variable importance as measured by a Random Forest. The y-axis represents the top-34 bacterial families that accurately discriminated different groups of bacterial community. Bacterial families are ranked in the order of importance for group classification, from top to bottom.

**Fig. S5.** K-mean clustering based on body mass metrics. Boxplots show the differences of (A) body weight; (B) fork length; (C) liver weight and (D) gonad weight across all fish groups namely “fast”, “medium” and “slow”. For each box, the vertical bold bar denote medians; the width of box denotes the interquartile range (25th percentile – 75th percentile); the whiskers mark the values range within 1.5 times interquartile. Lower-cased letters denote statistical significance reported by Mann–Whitney  $U$  test at confidence level of 0.95. (E) Line chart depicts the optimal numbers of clusters suggested by Gap statistic method. (F) Bars depict the silhouette width for each sample, and red dash line shows the average silhouette width for k-means clustering. Bars are coloured according to the fish groups as the same in (A–D). (G) Box plot shows the stratification of growth performance (body weight) within each pond. (H) The same two-dimensional scatter plot of all fish individual based on

Principal Components Analysis (PCA) of four body mass metrics as Fig. 2A but was color-coded by ‘Pond’ (P: Pond; Digitals: Pond numbers).

**Fig. S6.** Heatmaps showing the correlations between each indicated alpha-diversity indice of fugu gut microbiota and (A) Feed source proportions (‘feed’) predicted by SourceTracker; or (B) body mass indexes (‘Comp.1’) computed by PCA analysis. Color scale denotes the degree of correlatedness, only statistical significant ( $P < 0.05$ ) Spearman’s Rho were shown in each heatmap pixel.

**Fig. S7.** Stackbars showing the relative abundance of top-10 bacterial (A) phylums, (B) classes, (C) orders, (D) families and (E) genus across all biologically replicated samples. The bacterial taxonomic ranks which have lower relative abundance were grouped into “Other”.

**Fig. S8.** Bars show the LEfSe-estimated discriminative bacterial families that are differential among three fugu cluters with statistical significance reported by Kruskal–Wallis sum-rank test ( $P < 0.05$ ). Bar width denotes ranks of the log10-transformed LDA effect size.

**Fig. S9.** Stackbars showing the relative abundance of bacterial families in contributing to the proportions of bacterial phenotypes that are inferred by BugBase. Different sample groups of bacterial communities were shown in x-axis.

**Fig. S10.** Heatmaps showing the correlations between the relative abundance of each indicated bacterial phenotype inferred by BugBase and (A) Feed source proportions (‘feed’) predicted by SourceTracker; or (B) body mass indexes (‘Comp.1’) computed by PCA analysis. Color scale denotes the degree of correlatedness, only statistical significant ( $P < 0.05$ ) Spearman’s Rho were shown in each heatmap pixel.

**Fig. S11.** Calculation of the phylogenetic signal of traits including feed source proportions (‘feed’) predicted by SourceTracker and body mass indexes (‘PC1’) computed by PCA analysis. The solid bold black line denotes Mantel’s statistic of autocorrelation, and the dashed black lines denote the lower and upper bounds of the confidence envelop (0.95). The horizontal black line denotes the expected value of Mantel’s statistic under the null hypothesis of no phylogenetic autocorrelation. The colored bar show whether the autocorrelation is significant (based on the confidence interval): red for significant positive autocorrelation, black for nonsignificant autocorrelation, and blue for significant negative autocorrelation.

**Fig. S12.** Scatterplots show the LOESS smooth curve fitting (Local Polynomial Regression) between unweighted ses.MNTD metrics of indicated sub-community with a given trait (feed source or body mass). Regressions of (A) feed source and (B) body mass with ses.MNTD for all 943 ESVs. Regressions of (C) feed source and (D) body mass with ses.MNTD for 63 abundant ESVs. Regressions of (E) feed source and (F) body mass with ses.MNTD for 880 rare ESVs.

**Fig. S13.** Fitness of the neutral model for (A) all 943 ESVs; (B) 63 abundant ESVs and (C) 880 rare ESVs. Neutral (fit), overrepresented and underrepresented ESVs are coloured grey, red and blue respectively. The solid black line denotes model prediction, and the dashed black lines denote the lower and upper bounds of the confidence

envelop (0.95). The fitness of the neutral model ( $R^2$ ) and migration rate ( $m$ ) are shown as plot title of each subpanel.

**Table S1.** Numbers of reads that had passed quality control for each sample.

**Table S2.** Statistical results for each indicated figure.

**Table S3.** Metadata.

**Table S4.** Importance value of Random-Forest predicted bacterial families in classifying sample groups.

**Table S5.** Mean relative abundance of bacterial classes from each indicated sample group.

**Table S6ss** Discriminative bacterial taxa identified by LEfSe.