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

Xingkun Jin¹, Ziwei Chen¹, Yan Shi¹, Jian-Fang Gui^{1,2}, Zhe Zhao^{1*}.

¹Department of Marine Biology, College of Oceanography, Hohai University, Nanjing 210098, Jiangsu, China; ²State Key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, The Innovation Academy of Seed Design, Chinese Academy of Sciences, Wuhan, 430072, Hubei, China.

Running title: Snapshot of gut microbiota in farmed obscure puffer

Authors' Email address:

Xingkun Jin, jinxingkun@hhu.edu.cn

Ziwei Chen, 191311040017@hhu.edu.cn

Yan Shi, yshi@hhu.edu.cn

Jian-Fang Gui, jfgui@ihb.ac.cn

*To whom correspondence should be addressed: Zhe Zhao, Fax: +86 2583787653;

E-mail: zhezhao@hhu.edu.cn.

Keywords: aquaculture, ecological process, environment, feed-borne bacteria, fish growth, obscure puffer

ABSTRACT

Understanding the ecological processes in controlling the assemblage of gut microbiota becomes an essential prerequisite for a more sustainable aquaculture. Here we used 16S rRNA amplicon sequencing to characterize the hindgut microbiota from cultured obscure puffer *Takifugu obscurus*. The gut microbiota is featured with lower alpha-diversity, greater beta-dispersion and higher average 16S rRNA copy numbers comparing to water and sediment, but far less so to feed. SourceTracker predicted a notable source signature from feed in gut microbiota. Furthermore, effect of varying degrees of feed-associated bacteria on compositional, functional and phylogenetic diversity of gut microbiota were revealed. Coincidentally, considerable increase of species richness and feed source proportions both were observed in slow growth fugu, implying a reduced stability in gut microbiota upon bacterial disturbance from feed. Moreover, quantitative ecological analytic framework was applied and the ecological processes underlying such community shift were determined. In the context of lower degree of feed disturbance, homogeneous selection and dispersal limitation largely contribute to the community stability and partial variations among hosts. Whilst with the degree of feed disturbance increased, variable selection leads to an augmented interaction within gut microbiota, entailing community unstability and shift. Altogether, our findings illustrated a clear diversity-function relationships in fugu gut microbiota, and it has implicated in a strong correlation between feed-borne bacteria and host growth rate. These results provide a new insight into aquaculture of fugu and other economically important fishes, as well as a better understanding of host-microbe interactions in the vertebrate gastrointestinal tract.

IMPORTANCE

Environmental bacteria has 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 environmental bacteria. Through the integrative analysis by community profiling and source tracking, we show that feed-associated bacteria can impose a strong disturbance upon fugu gut microbiota. As a result, marked alterations in the composition and function of gut microbiota in slow growth fugu were observed, which is potentially correlated with the host physiological condition such as gastric evacuation rate. 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.

1. INTRODUCTION

Microbiome can directly or indirectly affect the host's physiological functions such as metabolism, immune defense, growth and development, as evidenced by increasingly large body of microbiome researches (1). 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 (2-4). 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)(1, 5). 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 indepth studies of microbiome-associated ecology, physiology and molecular function (6-8).

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 (9). The development of high-potential microbiome-based innovations offers opportunities for disease control and health promotion in aquaculture practice (10, 11). Therefore, there is an urgent need of establishing an optimized microbial management strategy to achieve a more sustainable aquaculture (12). As a prerequisite, understanding the ecological processes and mechanisms in controlling the assemblage of fish associated microbiome becomes critical (13, 14).

Fish gut can be considered as a dynamic ecosystem in which spatial-temporal disturbance in the microbial nutrient niches shape the bacterial community (15, 16). The potential niche space in fish gut environment can be determined by the host in multiple ways (17), yet it also can be largely determined by the diet as well as it's associated bacteria (18-20). 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 (21-23). Fish-associated microbiota can be potentially acquired from a variety of sources with which organisms come into intimate biological interactions (17).

Takifugu (aka fugu), a genus of pufferfish (*Tetraodontidae*), is better known to be able to accumulate a potent neurotoxin, tetrodotoxin, and possess high toxicity (24). As a vertebrate, fugu displays several peculiarities in its genome, such as shortened intergenic and intronic sequences devoid of repetitive elements (25, 26), thereby makes it distinctive animal model for genome evolution study (27). 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 (28, 29). 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 (24, 28). 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 (30). Meanwhile, feces from farmed fugu can also play a part of bacterial enrichment in rearing water and sediment as seen in other fish species (31, 32). 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, i.e. 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 (19, 33, 34). 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 (13, 35). Despite the

evidences for the broad existence of core gut microbiota, accumulating studies had also shown substantial inter-individual variation in the gut microbial communities for both cultured (9, 36) and wild-captured fish species (37-39). Several factors can contribute to such variability, including environmental filtering either by the host, e.g. variations between regions of the gastrointestinal tract, state of health and nutrition; or by the rearing environments such as diet (40, 41). Such variations should be resolved before the conclusion of comparative analysis of microbiota can be drawn.

Environmental bacteria has a great impact on fish gut microbiota (12, 32, 33), 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. We shown that farmed fugu had a distinct gut microbiota with markedly lower alpha-diversity and higher beta-dispersion in comparison to its rearing environment including feed, water and sediment. Furthermore, effect of varying degrees of feed-associated bacteria on compositional, functional and phylogenetic diversity of gut microbiota were revealed. Furthermore, implications of fish growth-dependent community-level responses to feed-associated bacteria and the underlying relative importance of ecological processes were discussed. 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.

2. 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, datasets 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 (Figure S1).

2.1. 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 (Figure S2). In addition, statistical analysis of beta diversity based on Bray-Curtis distance shown that the pond (n=3) does have certain effect on community dissimilarity ($R^2=0.11$; $P=0.001$; $\text{MeanSqs}=0.97$; Table S2), thereby 'pond' was set as a strata. Further, the community composition was revealed to be significantly

differed among sample groups (PERMANOVA, $R^2=0.24$; $P=0.001$; $\text{MeanSqs}=2.08$; strata='pond'; Figure. 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 (TukeyHSD, $P<0.05$, Table S2), possibly due to the lower samples size of feed ($n=3$) (Figure 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), amongst 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 (Figure 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 Figure 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 (unkown) within each group (four right-side facets in Figure 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 ten-fold cross-validation shown that at least 34 bacterial families can be

used to classify all four sample groups with 1.22% error rate (Figure S4, Table S4). Of these, 33 families were taxonomically categorized into eight phylums according to the NCBI Taxonomy (Figure 1D). Three indicative bacterial families for fugu gut microbiota were identified, namely *Brevinemataceae* (*Spirochaetes*), *Mycoplasmataceae* (*Tenericutes*) and *Rikenellaceae* (*Bacteroidetes*). Whereas, 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.

2.2. Community structure in fugu gut are correlated with feed source signal and host body mass

Individual fugu were clustered based on unsupervised k-means computed from body mass measures including weight (unit: gram) of body, liver and gonad, as well as fork length (unit: centimeter). 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) (Figure 2A and Figure 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 shown feed source signals was the best predictor of community composition ($R^2=0.46$; $P=0.001$; MeanSqs= 6.68; strata:Pond; Figure 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 $R=-0.38$; $P<0.01$, Figure 2C). Moreover, the feed source proportions were significantly enriched in slow growth fish (42.2%) comparing to fast one (19.3%) (TukeyHSD, $P<0.05$). These results suggest the varying degree of effect of feed-associated bacteria 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, Figure 2D).

The estimated indices of alpha-diversity were positively correlated with feed source signals, but negatively correlated with fugu body mass (Figure S6). Further, markedly increased richness of gut microbiota from slow growth fish was observed comparing to fast group ($P<0.05$; Mann–Whitney U, Figure 2E). The taxonomic compositions of bacterial communities were differed among sample groups (Figure S7; Table S5). In the bacterial class level (Figure 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.

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 (Figure 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. Amongst others, *Pseudomonadaceae* was the most pronounced bacterial family ($\log_{10}$ -LDA Scores > 5) that were overrepresented in slow group.

2.3. 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 (42), and can be used as a proxy for phenotypically distinguishing between copiotroph and oligotroph (43, 44). Using the precalculated 16S rRNA counts by PICRUSt, the average 16S rRNA copy number per sample group were estimated. The results shown 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 copiotrophic dominance communities in both gut and feed (Figure 3A).

Further, the phenotypical characteristics including oxygen tolerance, Gram staining and pathogenic potential of bacterial communities were predicted by BugBase algorithm. It inferred the gut most significantly to be the negative bacteria dominance communities (97%), followed by feed (93%) and sediment (95%) (Figure 3B). Whereas water microbiota was predicted to be comprised of both Gram negative (56%) and Gram positive (44%) bacteria (Figure 3B and C, and Figure S9). Regarding oxygen tolerance traits (Figure 3 D~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 as for feed (59%), or by both aerobic and facultatively anaerobic bacteria as for water (56% and 33%). Whereas 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. In deed, the inferred proportions of anaerobic bacteria in gut microbiota is negatively correlated with feed source signals (Spearman's Rank $R=-0.79$; $P<0.01$), but positively correlated with body mass ($R=0.32$; $P<0.05$, Figure 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 aslo 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, Figure 3 G~I). Further, bacteria that can form biofilms was not significantly different comparing fugu gut to both water and sediment (Figure 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 ($R<-0.3$, $P<0.05$, Figure S10). Altogether, these results have further emphasized that the functional changes in fugu gut microbiota is driven largely by feed-associated bacteria, and depended on host growth rate.

2.4. 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 (45) and Blomberg's K statistic (46), 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 (Figure S11). On this basis, we proceeded to quantify phylogenetic turnover among close-related bacteria (47). Considering the ecological influence on bacterial sub-community may varied with its relative abundance (48-50), 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 shown 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 ($P<0.01$, Figure 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 (Figure S12), suggesting its lower degree of phylogenetic relatedness.

Further we estimated the phylogenetic beta diversity (β MNTD) and in turn compared it against null model to estimated beta nearest taxon index β NTI (47, 51). The $|\beta$ NTI| values > 2 indicate significantly phylogenetic turnover than expected. The Mantel's test showed significant positive correlations between the inter-community β NTI values and the changes of feed source signals, for 'all', 'abundant' and 'rare' sub-communities, respectively (Figure 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 shown 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') was dominant in 'abundant' sub-community (Figure 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 (Figure S13).

3. 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 shown 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 (Figure 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.

3.1. 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 (Figure 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 (Figure 1C), although the samples were obtained one 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 (Figure 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 (17, 20, 34, 52). Therefore, it can be hypothesized that the early-life exposure to different bacterial communities as seen in other fishes, might also contribute to the assembly of fugu gut microbiota mentioned here.

Ringø and Birkbeck (53) 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 (Figure 1D). In addition, Bugbase inferred that both *Brevinemataceae* and *Rikenellaceae* contributed to the most proportions of anaerobic bacteria in fugu gut microbiota (Figure S9). Although literature concerning the microbiota of wild *T. obscurus* is scarce, recent studies shown 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* (54) and *T. niphobles* (55) 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.

3.2. 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 (Figure S5), by which they can be rigorously subdivided into three clusters, i.e. fast, medium and slow (Figure 2A). We shown that the feed source signals explained the most variations in the compositional difference in fugu gut microbiota, and its was negatively correlated with fish body mass (Figure 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) (Figure 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 were 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%, Figure S7 and S9). Previous work had shown that *Pseudomonas* is highly persistent during long term storage of fish feed (56). Yet another study found its presence in water, contributed mostly by the fish feces, was positively associated with the dose of administrated commercial fish feed in a recirculating aquaculture system (31). Amongst others, *Pseudomonas* is the most frequently reported gut-associated bacteria across many different freshwater fish species (39). On the other hand, recent studies had also raised concerns that it might represent a common contaminating bacterial taxa in molecular biology reagents (57). Whilst, considering the very rare presence of *Pseudomonas* in all the water and sediment samples (Figure 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, dosens 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 overrepresent feed-borne bacteria in the gut of slow growth fugu were not exclusively due to the influx of DNA from dead or dormant cells (19). Rather, bacteria such as *Pseudomonas*, optimized to grow at high nutrient as copiotroph (with 6~10 average 16S rRNA copy number) (58), 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 (16). Furthermore, the fish size was found to be negatively correlated with gastric emptying time, i.e. larger fish empty their stomachs at a faster rate than smaller ones within a consistent environment (59). And the fish gastric evacuation can be further enhanced by transgenetically overexpressions of host growth hormone gene (60). These findings offer support for our observations that proportions of feed source signals in gut microbiota was negatively correlated with fugu growth rate, 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 (18, 19, 61), yet little is know 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 (Figure 3). Such phenotypical differences are driven mostly by the higher proportion of both *Pseudomonadaceae* and *Enterobacteriaceae* (Figure S9), of which many members are generally considered commensals (61). But like other opportunistic pathogens, certain members from both bacterial families often contain genes for virulence factors and antibiotic resistance (62, 63). 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) (Figure 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 aerobiotic 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.

3.3. 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 ($P < 0.05$, Figure 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 (Figure S12). Whereas, 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 (46, 64). Accordantly, many other studies had shown that robust phylogenetic signal was related to different variables within short distances across ecosystems of diverse types (47, 50, 65-67). On this basis, recent advanced quantitative ecological models (47, 51, 68), 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 (69). 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) (70). 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 (69, 71). Our Mantel's test results shown 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 (Figure 4A). This result further emphasized that niche effect posed a strong selection pressure in determining the community assembly (71), thereby driving bacterial communities into a phylogenetically more closed cluster (64). As a continual inoculum, feeding poses a significant destabilizing force in fish intestinal microbiota and results in a continuous community turnover (18, 19). 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 (72, 73). 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 (Figure 4A). Of note, such pattern was more obviously observed for rare sub-community (Mantel $R=0.492$, Figure 4C) than to abundant one (Mantel $R=0.275$, Figure 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 (74). We observed a large fraction of community shift was governed by dispersal limitation (~40%), whereas very little (<3%) by homogenizing dispersal (Figure 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 (71). The lower degree of homogenizing dispersal (75) 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 (74). 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) (19, 59, 76).

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 unstability 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, e.g. gastric evacuation rate and intestinal transit time of digesta. Nevertheless, we can not 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.

4. MATERIALS AND METHODS

4.1. 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 pellet feeds, 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 thounds 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 post-feeding (1 day) fishes per pond 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 indivisuals. Fishes were anesthetized with 0.05% MS-222 (Sigma-Aldrich, St. Louis, MO, USA), and immediatedly 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-µm pore sized nylon mesh, and the flowthrough was further filtered with 0.22µm Millipore membrane with vacuum pump to collected 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°C before use. The detailed metadata for each sample was recored in Table S3.

4.2 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™ 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., Waltham, MA, USA), 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-µl reaction mix consisted of 40 ng template DNA, 5 µl reaction buffer (5x), 5 µl GC buffer (5x), 2 µl dNTP (2.5mM), 10 pM of barcoded forward and reverse primers, and 0.75 U Q5® High-Fidelity DNA Polymerase (NEB, Ipswich, MA, USA). The thermocycling program consisted of one

hold at 98 °C for 2 min, followed by three-step 25 cycles of 15 s at 95 °C, 30 s at 55 °C and 30 s at 72 °C, and one final fold at 72°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).

4.3 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 (77). 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 (78, 79) using USEARCH v.10.0 (80). In brief, the paired-end reads were merged and relabeled 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, i.e. 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 were assigned by the syntax

algorithm using the Ribosomal Database Project (RDP) classifier (RDP training set v16) (81, 82), and chloroplast ESVs were removed after taxonomy-based filtering. Sequencing depth was normalized by subsample using 'otutab_norm', yielding 7,000 sequences per sample to keep all the samples for downstream analysis. Community diversity was analyzed using '-alpha_div', '-alpha_div_rare', '-cluster_agg' and '-beta_div' respectively.

4.4 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)(83) 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 (84).

To further identify the potential origins for the compositional structure for each indicated sample group, SourceTracker (85) 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 analyzed 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 (86).

4.5 Phenotypic inferring of bacterial communities

For metagenomics inference, the ESVs table normalized by 16S rRNA copy numbers were annotated by PICRUST (87) with closed reference picking (Greengenes v.13_5)(88). 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 (89, 90). 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 (91). Taxonomic level was set as bacterial family.

4.6 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 (46) and Mantel's Correlogram (45) were used to

measure phylogenetic signal using *phyloSignal()* and *phyloCorrelogram()* implemented in R package 'PhyloSignal' (92). 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 (93), 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 (64). Ecological null model was used to quantify the relative importance of ecological processes in driving microbial community assemblage (51). Briefly, both metrics of β -nearest taxon index (β 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 (β MNTD) using the *comdistnt()*. Further, the β NTI, i.e. the degree to which observed β MNTD deviates from the mean of the null distribution (999 randomization) was computed according to Stegen et al. (51). The influence of deterministic community turnover including homogeneous selection and variable selection (70) is estimated by inferring the fraction of pairwise community comparisons with significance thresholds of β NTI < -2 and β NTI > 2, respectively (51). 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 β -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).

4.7 Statistics

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 datasets were tested using *shapiro.test()* and *bartlett.test()*, respectively. In turn, parametric or non-parametric tests were applied accordingly. For parametric datasets, group means were compared by one-way ANOVA using *aov()*, followed by *post-hoc* test using *TukeyHSD()*. For non-parametric datasets, 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 dataset 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 (Principle 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 (94). 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 (95), followed by visualization according to principal components analysis by *prcomp()*. Spearman-correlations

between the phylogenetic beta diversity (β NTI matrix) and indicated traits difference (euclidean-based trait distance matrix) we assessed using *mantel()* implemented in 'ecodist' package (96) with 9,999 permutations. The *neutral.fit()* in the 'MicEco' package (97) was used to evaluate the fitness of neutral model (98) on the gut bacterial community, and to determine the contribution of neutral processes in community assembly.

AVAILABILITY OF DATA

The data set generated and analyzed for this study is available in the NCBI sequencing reads archive (SRA), under BioProject accession number PRJNA658467.

AUTHOR CONTRIBUTIONS

XJ conceived and designed the experiments. ZZ, YS provided the initial framework and acquired fund for the study. XJ conducted the sampling and experimental work. XJ and ZC performed data analysis and visualization. XJ, ZC, YS, JFG and ZZ interpreted the results. XJ wrote the manuscript with input of all authors. All authors read and approved the final manuscript.

ACKNOWLEDGEMENTS

This work was supported 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 [2019]477). 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 declare that we have no competing interests.

REFERENCES

1. Berg G, Rybakova D, Fischer D, Cernava T, Vergès M-CC, Charles T, Chen X, Coccolin L, Eversole K, Corral GH, Kazou M, Kinkel L, Lange L, Lima N, Loy A, Macklin JA, Maguin E, Mauchline T, McClure R, Mitter B, Ryan M, Sarand I, Smidt H, Schelkle B, Roume H, Kiran GS, Selvin J, Souza RSCd, van Overbeek L, Singh BK, Wagner M, Walsh A, Sessitsch A, Schloter M. 2020. Microbiome definition re-visited: old concepts and new challenges. *Microbiome* 8:103.
2. Bordenstein SR, Theis KR. 2015. Host Biology in Light of the Microbiome: Ten Principles of Holobionts and Hologenomes. *PLOS Biology* 13:e1002226.
3. Theis KR, Dheilly NM, Klassen JL, Brucker RM, Baines JF, Bosch TCG, Cryan JF, Gilbert SF, Goodnight CJ, Lloyd EA, Sapp J, Vandenkoornhuyse P, Zilber-Rosenberg I, Rosenberg E, Bordenstein SR. 2016. Getting the Hologenome Concept Right: an Eco-Evolutionary Framework for Hosts and Their Microbiomes. *mSystems* 1:e00028-16.
4. Zilber-Rosenberg I, Rosenberg E. 2008. Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. *FEMS Microbiology Reviews* 32:723-735.
5. Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R. 2012. Diversity, stability and resilience of the human gut microbiota. *Nature* 489:220-230.
6. Lloyd-Price J, Mahurkar A, Rahnavard G, Crabtree J, Orvis J, Hall AB, Brady A, Creasy HH, McCracken C, Giglio MG, McDonald D, Franzosa EA, Knight R, White O, Huttenhower C. 2017. Strains, functions and dynamics in the expanded Human Microbiome Project. *Nature* 550:61-66.
7. Sunagawa S, Coelho LP, Chaffron S, Kultima JR, Labadie K, Salazar G, Djahanschiri B, Zeller G, Mende DR, Alberti A, Cornejo-Castillo FM, Costea PI, Cruaud C, Ovidio F, Engelen S, Ferrera I, Gasol JM, Guidi L, Hildebrand F, Kokoszka F, Lepoivre C, Lima-Mendez G, Poulain J, Poulos BT, Royo-Llonch M, Sarmiento H, Vieira-Silva S, Dimier C, Picheral M, Searson S, Kandels-Lewis S, Bowler C, de Vargas C, Gorsky G, Grimsley N, Hingamp P, Iudicone D, Jaillon O, Not F, Ogata H, Pesant S, Speich S, Stemmann L, Sullivan MB, Weissenbach J, Wincker P, Karsenti E, Raes J, Acinas SG, Bork P. 2015. Structure and function of the global ocean microbiome. *Science* 348:1261359.
8. Hatzenpichler R, Krukenberg V, Spietz RL, Jay ZJ. 2020. Next-generation physiology approaches to study microbiome function at single cell level. *Nature Reviews Microbiology* 18:241-256.
9. Perry WB, Lindsay E, Payne CJ, Brodie C, Kazlauskaitė R. 2020. The role of the gut microbiome in sustainable teleost aquaculture. *Proceedings of the Royal Society B: Biological Sciences* 287:20200184.
10. de Bruijn I, Liu Y, Wiegertjes GF, Raaijmakers JM. 2017. Exploring fish microbial communities to mitigate emerging diseases in aquaculture. *FEMS Microbiology Ecology* 94.
11. Chao R, Yu L, Xufa M, Yadong X, Mingxu X, Yuting Z, Wei Z, Yalin Y, Zhen Z, Li Z, Kaijian W, Zhigang Z. Interactions between commensal bacteria and viral infection: insights for viral disease control in farmed animals. *SCIENCE CHINA Life Sciences* doi:<https://doi.org/10.1007/s11427-020-1721-5>.
12. Lieke T, Meinelt T, Hoseinifar SH, Pan B, Straus DL, Steinberg CEW. 2020. Sustainable aquaculture requires environmental-friendly treatment strategies for fish diseases. *Reviews in Aquaculture* 12:943-965.
13. De Schryver P, Vadstein O. 2014. Ecological theory as a foundation to control pathogenic invasion in aquaculture. *The ISME Journal* 8:2360-2368.
14. Zhang Q-Y, Gui J-F. 2018. Diversity, evolutionary contribution and ecological roles of aquatic viruses. *Science China Life Sciences* 61:1486-1502.
15. Pereira FC, Berry D. 2017. Microbial nutrient niches in the gut. *Environmental Microbiology* 19:1366-1378.
16. Ringø E, Zhou Z, Vecino JLG, Wadsworth S, Romero J, Kroghdal Å, Olsen RE, Dimitroglou A, Foey A, Davies S, Owen M, Lauzon HL, Martinsen LL, De Schryver P, Bossier P, Sperstad S, Merrifield DL. 2016. Effect of dietary components on the gut microbiota of aquatic animals. A never-ending story? *Aquaculture Nutrition* 22:219-282.
17. Wang AR, Ran C, Ringø E, Zhou ZG. 2018. Progress in fish gastrointestinal microbiota research. *Reviews in Aquaculture* 10:626-640.
18. Wilkes Walburn J, Wemheuer B, Thomas T, Copeland E, O'Connor W, Booth M, Fielder S, Egan S. 2019. Diet and diet-associated bacteria shape early microbiome development in Yellowtail Kingfish (*Seriola lalandi*). *Microbial Biotechnology* 12:275-288.
19. Parris DJ, Morgan MM, Stewart FJ. 2019. Feeding Rapidly Alters Microbiome Composition and Gene Transcription in the Clownfish Gut. *Applied and Environmental Microbiology* 85:e02479-18.
20. Abdul Razak S, Scribner KT. 2020. Ecological and Ontogenetic Components of Larval Lake

- Sturgeon Gut Microbiota Assembly, Successional Dynamics, and Ecological Evaluation of Neutral Community Processes. *Applied and Environmental Microbiology* 86:e02662-19.
21. Roeselers G, Mittge EK, Stephens WZ, Parichy DM, Cavanaugh CM, Guillemin K, Rawls JF. 2011. Evidence for a core gut microbiota in the zebrafish. *The ISME Journal* 5:1595-1608.
22. Wong S, Waldrop T, Summerfelt S, Davidson J, Barrows F, Kenney PB, Welch T, Wiens GD, Snekvik K, Rawls JF, Good C. 2013. Aquacultured Rainbow Trout (*Oncorhynchus mykiss*) Possess a Large Core Intestinal Microbiota That Is Resistant to Variation in Diet and Rearing Density. *Applied and Environmental Microbiology* 79:4974.
23. Riiser ES, Haverkamp THA, Borgan Ø, Jakobsen KS, Jentoft S, Star B. 2018. A Single Vibrionales 16S rRNA Oligotype Dominates the Intestinal Microbiome in Two Geographically Separated Atlantic cod Populations. *Frontiers in Microbiology* 9.
24. Zhang X, Zong J, Chen S, Li M, Lu Y, Wang R, Xu H. 2020. Accumulation and Elimination of Tetrodotoxin in the Pufferfish *Takifugu obscurus* by Dietary Administration of the Wild Toxic Gastropod *Nassarius semiplicata*. *Toxins* 12.
25. Brenner S, Elgar G, Sanford R, Macrae A, Venkatesh B, Aparicio S. 1993. Characterization of the pufferfish (*Fugu*) genome as a compact model vertebrate genome. *Nature* 366:265-268.
26. Aparicio S, Chapman J, Stupka E, Putnam N, Chia J-m, Dehal P, Christoffels A, Rash S, Hoon S, Smit A, Gelpke MDS, Roach J, Oh T, Ho IY, Wong M, Detter C, Verhoef F, Predki P, Tay A, Lucas S, Richardson P, Smith SF, Clark MS, Edwards YJK, Doggett N, Zharkikh A, Tavtigian SV, Pruss D, Barnstead M, Evans C, Baden H, Powell J, Glusman G, Rowen L, Hood L, Tan YH, Elgar G, Hawkins T, Venkatesh B, Rokhsar D, Brenner S. 2002. Whole-Genome Shotgun Assembly and Analysis of the Genome of *Fugu rubripes*. *Science* 297:1301.
27. Van de Peer Y. 2004. Tetraodon genome confirms *Takifugu* findings: most fish are ancient polyploids. *Genome Biology* 5:250.
28. Wang T, Wei X, Chen T, Wang W, Xia X, Miao J, Yin S. 2019. Studies of the mechanism of fatty liver formation in *Takifugu fasciatus* following copper exposure. *Ecotoxicology and Environmental Safety* 181:353-361.
29. Gao F-X, Shi Y, Duan W, Lu W-J, Huang W, Zhang X-J, Zhao Z, Zhou L, Gui J-F. 2020. A rapid and reliable method for identifying genetic sex in obscure pufferfish (*Takifugu obscurus*). *Aquaculture* 519:734749.
30. Okabe T, Oyama H, Kashitani M, Ishimaru Y, Suo R, Sugita H, Itoi S. 2019. Toxic Flatworm Egg Plates Serve as a Possible Source of Tetrodotoxin for Pufferfish. *Toxins* 11.
31. Gołaś I, Szmyt M, Potorski J, Łopata M, Gotkowska-Płachta A, 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.
32. Muziasari WI, Pitkänen LK, Sørum H, Stedtfeld RD, Tiedje JM, Virta M. 2017. The Resistome of Farmed Fish Feces Contributes to the Enrichment of Antibiotic Resistance Genes in Sediments below Baltic Sea Fish Farms. *Frontiers in Microbiology* 7.
33. Eichmiller JJ, Hamilton MJ, Staley C, Sadowsky MJ, Sorensen PW. 2016. Environment shapes the fecal microbiome of invasive carp species. *Microbiome* 4:44.
34. Giatsis C, Sipkema D, Smidt H, Heilig H, Benvenuti G, Verreth J, Verdegem M. 2015. The impact of rearing environment on the development of gut microbiota in tilapia larvae. *Scientific Reports* 5:18206.
35. Troussellier M, Escalas A, Bouvier T, Mouillot D. 2017. Sustaining Rare Marine Microorganisms: Macroorganisms As Repositories and Dispersal Agents of Microbial Diversity. *Frontiers in Microbiology* 8.
36. Li X, Zhou L, Yu Y, Ni J, Xu W, Yan Q. 2017. Composition of Gut Microbiota in the Gibel Carp (*Carassius auratus gibelio*) Varies with Host Development. *Microbial Ecology* 74:239-249.
37. Star B, Haverkamp THA, Jentoft S, Jakobsen KS. 2013. Next generation sequencing shows high variation of the intestinal microbial species composition in Atlantic cod caught at a single location. *BMC Microbiology* 13:248.
38. Riiser ES, Haverkamp THA, Varadharajan S, Borgan Ø, Jakobsen KS, Jentoft S, Star B. 2020. Metagenomic Shotgun Analyses Reveal Complex Patterns of Intra- and Interspecific Variation in the Intestinal Microbiomes of Codfishes. *Applied and Environmental Microbiology* 86:e02788-19.
39. Egerton S, Culloty S, Whooley J, Stanton C, Ross RP. 2018. The Gut Microbiota of Marine Fish. *Frontiers in Microbiology* 9.
40. Clements KD, Angert ER, Montgomery WL, Choat JH. 2014. Intestinal microbiota in fishes: what's known and what's not. *Molecular Ecology* 23:1891-1898.
41. Wong S, Waldrop T, Summerfelt S, Davidson J, Barrows F, Kenney PB, Welch T, Wiens GD,

- Snekvik K, Rawls JF, Good C. 2013. Aquacultured Rainbow Trout (*Oncorhynchus mykiss*) Possess a Large Core Intestinal Microbiota That Is Resistant to Variation in Diet and Rearing Density. *Applied and Environmental Microbiology* 79:4974.
42. Lankiewicz TS, Cottrell MT, Kirchman DL. 2016. Growth rates and rRNA content of four marine bacteria in pure cultures and in the Delaware estuary. *The ISME Journal* 10:823-832.
43. Roller BRK, Stoddard SF, Schmidt TM. 2016. Exploiting rRNA operon copy number to investigate bacterial reproductive strategies. *Nature Microbiology* 1:16160.
44. Lauro FM, McDougald D, Thomas T, Williams TJ, Egan S, Rice S, DeMaere MZ, Ting L, Ertan H, Johnson J, Ferriera S, Lapidus A, Anderson I, Kyrpides N, Munk AC, Detter C, Han CS, Brown MV, Robb FT, Kjelleberg S, Cavicchioli R. 2009. The genomic basis of trophic strategy in marine bacteria. *Proceedings of the National Academy of Sciences* 106:15527.
45. Oden NL, Sokal RR. 1986. Directional Autocorrelation: An Extension of Spatial Correlograms to Two Dimensions. *Systematic Zoology* 35:608-617.
46. Blomberg SP, Garland JR. T, Ives AR. 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution* 57:717-745.
47. Stegen JC, Lin X, Konopka AE, Fredrickson JK. 2012. Stochastic and deterministic assembly processes in subsurface microbial communities. *The ISME Journal* 6:1653-1664.
48. Shade A, Gilbert JA. 2015. Temporal patterns of rarity provide a more complete view of microbial diversity. *Trends in Microbiology* 23:335-340.
49. Langenheder S, Lindström ES. 2019. Factors influencing aquatic and terrestrial bacterial community assembly. *Environmental Microbiology Reports* 11:306-315.
50. Jiao S, Lu Y. 2020. Soil pH and temperature regulate assembly processes of abundant and rare bacterial communities in agricultural ecosystems. *Environmental Microbiology* 22:1052-1065.
51. Stegen JC, Lin X, Fredrickson JK, Chen X, Kennedy DW, Murray CJ, Rockhold ML, Konopka A. 2013. Quantifying community assembly processes and identifying features that impose them. *The ISME Journal* 7:2069-2079.
52. Wong S, Stephens WZ, Burns AR, Stagaman K, David LA, Bohannan BJM, Guillemin K, Rawls JF. 2015. Ontogenetic Differences in Dietary Fat Influence Microbiota Assembly in the Zebrafish Gut. *mBio* 6:e00687-15.
53. Ringø E, Birkbeck TH. 1999. Intestinal microflora of fish larvae and fry. *Aquaculture Research* 30:73-93.
54. Li Z, Tian J, Lai Y, Lee C-H, Cai Z, Yu C-F. 2020. Puffer Fish Gut Microbiota Studies Revealed Unique Bacterial Co-Occurrence Patterns and New Insights on Tetrodotoxin Producers. *Marine Drugs* 18.
55. Shiina A, Itoi S, Washio S, Sugita H. 2006. Molecular identification of intestinal microflora in Takifugu niphobles. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics* 1:128-132.
56. Zmyslowska I, Lewandowska D. 1999. Survival of Bacterial Strains in Fish Feeds Stored in Different Temperatures. *Polish Journal of Environmental Studies* 8:447-451.
57. Li Y, Bruni L, Jaramillo-Torres A, Gajardo K, Kortner TM, Krogdahl Å. 2020. 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* doi:10.1101/2020.05.08.083899:2020.05.08.083899.
58. Lladó S, Baldrian P. 2017. Community-level physiological profiling analyses show potential to identify the copiotrophic bacteria present in soil environments. *PLOS ONE* 12:e0171638.
59. Flowerdew MW, Grove DJ. 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. *Journal of Fish Biology* 14:229-238.
60. Kim J-H, Chatchaiphon S, Crown MT, White SL, Devlin RH. 2018. Effect of growth hormone overexpression on gastric evacuation rate in coho salmon. *Fish Physiology and Biochemistry* 44:119-135.
61. Ruzauskas M, Klimiene I, Armalyte J, Bartkiene E, Siugzdiniene R, Skerniskyte J, Krasauskas R, Suziedeliene E. 2018. Composition and antimicrobial resistance profile of Gram-negative microbiota prevalent in aquacultured fish. *Journal of Food Safety* 38:e12447.
62. Hamza D, Dorgham S, Ismael E, El-Moez SIA, Elhariri M, Elhelw R, Hamza E. 2020. Emergence of β -lactamase- and carbapenemase- producing Enterobacteriaceae at integrated fish farms. *Antimicrobial Resistance & Infection Control* 9:67.
63. Cazares A, Moore MP, Hall JPJ, Wright LL, Grimes M, Emond-Rhéault J-G, Pongchaikul P, Santanirand P, Levesque RC, Fothergill JL, Winstanley C. 2020. A megaplasmid family driving dissemination of multidrug resistance in *Pseudomonas*. *Nature Communications* 11:1370.

64. Webb CO, Ackerly DD, McPeck MA, Donoghue MJ. 2002. Phylogenies and Community Ecology. *Annual Review of Ecology and Systematics* 33:475-505.
65. Yan Q, Li J, Yu Y, Wang J, He Z, Van Nostrand JD, Kempner ML, Wu L, Wang Y, Liao L, Li X, Wu S, Ni J, Wang C, Zhou J. 2016. Environmental filtering decreases with fish development for the assembly of gut microbiota. *Environmental Microbiology* 18:4739-4754.
66. Tripathi BM, Stegen JC, Kim M, Dong K, Adams JM, Lee YK. 2018. Soil pH mediates the balance between stochastic and deterministic assembly of bacteria. *The ISME Journal* 12:1072-1083.
67. Liu J, Zhu S, Liu X, Yao P, Ge T, Zhang X-H. 2020. Spatiotemporal dynamics of the archaeal community in coastal sediments: assembly process and co-occurrence relationship. *The ISME Journal* 14:1463-1478.
68. Dini-Andreote F, Stegen JC, van Elsas JD, Salles JF. 2015. Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proceedings of the National Academy of Sciences* 112:E1326.
69. Hubbell SP. 2001. *The Unified Neutral Theory of Biodiversity and Biogeography* doi:10.2307/j.ctt7rj8w. Princeton University Press.
70. Vellend M. 2010. Conceptual Synthesis in Community Ecology. *The Quarterly Review of Biology* 85:183-206.
71. Chase JM, Myers JA. 2011. Disentangling the importance of ecological niches from stochastic processes across scales. *Philosophical Transactions of the Royal Society B: Biological Sciences* 366:2351-2363.
72. Litchman E. 2010. Invisible invaders: non-pathogenic invasive microbes in aquatic and terrestrial ecosystems. *Ecology Letters* 13:1560-1572.
73. Shade A, Peter H, Allison S, Baho D, Berga M, Buergermann H, Huber D, Langenheder S, Lennon J, Martiny J, Matulich K, Schmidt T, Handelsman J. 2012. Fundamentals of Microbial Community Resistance and Resilience. *Frontiers in Microbiology* 3.
74. Stegen JC, Lin X, Fredrickson JK, Konopka AE. 2015. Estimating and mapping ecological processes influencing microbial community assembly. *Frontiers in Microbiology* 6.
75. Cadotte Marc W. 2006. Dispersal and Species Diversity: A Meta-Analysis. *The American Naturalist* 167:913-924.
76. Bolnick DI, Snowberg LK, Hirsch PE, Lauber CL, Knight R, Caporaso JG, Svanbäck R. 2014. Individuals' diet diversity influences gut microbial diversity in two freshwater fish (threespine stickleback and Eurasian perch). *Ecology Letters* 17:979-987.
77. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal*; Vol 17, No 1: Next Generation Sequencing Data Analysis doi:10.14806/ej.17.1.200.
78. Zhang J, Liu Y-X, Zhang N, Hu B, Jin T, Xu H, Qin Y, Yan P, Zhang X, Guo X, Hui J, Cao S, Wang X, Wang C, Wang H, Qu B, Fan G, Yuan L, Garrido-Oter R, Chu C, Bai Y. 2019. NRT1.1B is associated with root microbiota composition and nitrogen use in field-grown rice. *Nature Biotechnology* 37:676-684.
79. Zhang J, Zhang N, Liu Y-X, Zhang X, Hu B, Qin Y, Xu H, Wang H, Guo X, Qian J, Wang W, Zhang P, Jin T, Chu C, Bai Y. 2018. Root microbiota shift in rice correlates with resident time in the field and developmental stage. *Science China Life Sciences* 61:613-621.
80. Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26:2460-2461.
81. Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, Brown CT, Porras-Alfaro A, Kuske CR, Tiedje JM. 2013. Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Research* 42:D633-D642.
82. Wang Q, Garrity GM, Tiedje JM, Cole JR. 2007. Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Applied and Environmental Microbiology* 73:5261.
83. Liaw A, Wiener M. 2002. Classification and Regression by randomForest. *R News* 2:5.
84. Letunic I, Bork P. 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research* 47:W256-W259.
85. Knights D, Kuczynski J, Charlson ES, Zaneveld J, Mozer MC, Collman RG, Bushman FD, Knight R, Kelley ST. 2011. Bayesian community-wide culture-independent microbial source tracking. *Nature Methods* 8:761-763.
86. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C. 2011. Metagenomic biomarker discovery and explanation. *Genome Biology* 12:R60.
87. Langille MGI, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, Clemente JC, Burkpile DE, Vega Thurber RL, Knight R, Beiko RG, Huttenhower C. 2013. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature Biotechnology*

- 31:814-821.
88. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL. 2006. Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. *Applied and Environmental Microbiology* 72:5069.
89. Nemergut DR, Knelman JE, Ferrenberg S, Bilinski T, Melbourne B, Jiang L, Violle C, Darcy JL, Prest T, Schmidt SK, Townsend AR. 2016. Decreases in average bacterial community rRNA operon copy number during succession. *The ISME Journal* 10:1147-1156.
90. Prest TL, Kimball AK, Kueneman JG, McKenzie VJ. 2018. Host-associated bacterial community succession during amphibian development. *Molecular Ecology* 27:1992-2006.
91. Ward T, Larson J, Meulemans J, Hillmann B, Lynch J, Sidiropoulos D, Spear JR, Caporaso G, Blekhman R, Knight R, Fink R, Knights D. 2017. BugBase predicts organism-level microbiome phenotypes. *bioRxiv* doi:10.1101/133462:133462.
92. Keck F, Rimet F, Bouchez A, Franc A. 2016. phylosignal: an R package to measure, test, and explore the phylogenetic signal. *Ecology and Evolution* 6:2774-2780.
93. Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, Blomberg SP, Webb CO. 2010. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26:1463-1464.
94. Oksanen J, Blanchet F, Guillaume, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin R, Peter, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H. 2019. vegan: Community Ecology Package. CRAN-The Comprehensive R Archive Network.
95. Kassambara A, Mundt F. 2020. factoextra: Extract and Visualize the Results of Multivariate Data Analyses. CRAN-The Comprehensive R Archive Network.
96. Goslee CS, Urban LD. 2007. The ecodist Package for Dissimilarity-based Analysis of Ecological Data. *Journal of Statistical Software* 22:19.
97. Russel J. 2020. Russel88/MicEco: v0.9.11. Zenodo doi:<http://doi.org/10.5281/zenodo.3945943>.
98. Sloan WT, Lunn M, Woodcock S, Head IM, Nee S, Curtis TP. 2006. Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environmental Microbiology* 8:732-740.

FIGURE LEGENDS

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 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. (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) rarified 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. Color code is consistent with Figure 1A excepting the 'unknown' source was shown 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 color-coded by phylum ranks. Heatmap shown the centered classific importance of each indicated taxa (row) in discriminating a given bacterial community (column) from others.

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 Components Analysis (PCA) of four body mass metrics. The grey arrow points shown 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 Figure S7.

FIG 3. Phenotype inference of bacterial communities from the fugu rearing ecosystems. Estimated relative abundances for each indicated bacterial phenotype was 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 Figure 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-cased letters denote groups and statistical significance reported by pair-wise Mann-Whitney U tests with false discovery rate correction at confidence level of 0.95.

FIG 4. Influence of ecological processes in governing bacterial community assemblage in fugu gut. The correlations between phylogenetic beta diversity β 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 cutoff of β 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.

SUPPLEMENTAL MATERIAL

TABLE LEGENDS

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

ID	Size	ID	Size	ID	Size	ID	Size
F1	36736	F9	26043	Fb9	25070	FD1	36815
F10	22333	Fb1	25583	Fc1	21975	FD2	37651
F11	27447	Fb10	24496	Fc10	62291	FD3	38229
F12	28627	Fb11	23487	Fc11	24664	S1In	10336
F13	35318	Fb12	22294	Fc12	22261	S1Mid	9619
F14	33645	Fb13	27105	Fc13	55646	S1Out	9122
F15	31020	Fb14	20734	Fc14	28143	S2In	10294
F16	32521	Fb15	23462	Fc15	24754	S2Mid	8410
F17	31784	Fb16	25183	Fc16	24317	S2Out	12520
F18	30501	Fb17	27676	Fc17	26723	S3In	8888
F19	33761	Fb18	26288	Fc18	23858	S3Mid	9550
F2	30753	Fb19	24460	Fc19	27784	S3Out	7753
F20	25821	Fb20	24798	Fc2	23387	W1In	47616
F21	26256	Fb21	26507	Fc20	59896	W1Mid	25924
F3	31846	Fb22	21946	Fc3	21910	W1Out	22123
F4	28077	Fb3	23116	Fc4	28096	W2In	16548
F5	29201	Fb5	29513	Fc5	27971	W2Mid	23136
F6	30302	Fb6	26927	Fc6	20612	W2Out	30361
F7	28165	Fb7	31320	Fc7	23442	W3In	33120
F8	36612	Fb8	22443	Fc8	22271	W3Mid	28363
				Fc9	66205	W3Out	29713

TABLE.S2 Statistical results for each indicated figure.

Figure 1A

(1) adonis setting pond as factor: adonis (dis ~ Pond, data = design, by = NULL)						
	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
Pond	2	2.9104	0.97014	3.2346	0.11064	0.001
Residuals	79	23.3941	0.29992	0.88936		
Total	81	26.3045	1			
(2) adonis setting sample groups (F, Fd, W, S) as factor: adonis (dis ~ group, strata = design\$Pond, data = design, by=NULL)						
	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
group	3	6.2479074	2.0826358	8.09934898	0.2375221	0.001
Residuals	78	20.0566234	0.2571362	NA	0.7624779	NA
Total	81	26.3045308	NA	NA	1	NA

Figure 1B

Pairwise comparisons using Wilcoxon rank sum test, P value adjustment method: BH									
	F.F	F.Fd	F.S	F.W	Fd.Fd	Fd.S	Fd.W	S.S	W.S
F.Fd	0.00010227	NA	NA	NA	NA	NA	NA	NA	NA
F.S	1.32E-245	7.18E-81	NA	NA	NA	NA	NA	NA	NA
F.W	1.72E-249	7.25E-82	1.16E-05	NA	NA	NA	NA	NA	NA
Fd.Fd	0.00330435	0.00342397	0.00329105	0.00329105	NA	NA	NA	NA	NA
Fd.S	2.25E-16	1.34E-14	3.66E-06	9.11E-05	0.00186231	NA	NA	NA	NA
Fd.W	1.12E-15	1.34E-14	5.41E-09	1.44E-06	0.00421957	1.54E-10	NA	NA	NA
S.S	2.90E-13	2.18E-12	3.40E-23	3.40E-23	0.00502153	1.70E-11	2.28E-11	NA	NA
W.S	2.08E-11	0.28852238	4.18E-45	4.75E-47	0.00356081	1.53E-14	1.61E-14	2.17E-17	NA
W.W	1.41E-10	1.08E-09	3.40E-23	3.40E-23	0.00056274	1.70E-11	2.28E-11	0.16972718	2.17E-17

Figure2A

The quality of representation (cos2) for each indicated body mass indice on the principal components (PC1~4)				
	PC1	PC2	PC3	PC4
Weight.g	0.962038403	0.025239425	0.003224548	0.009497624
Length.cm	0.917584928	0.012917832	0.068747184	0.000750056
Weigh.gonad.g	0.718414987	0.280904081	0.000676607	4.33E-06
Weigh.liver.g	0.925801286	0.036780732	0.032483293	0.004934688

Figure 2B

(1) adonis analysis on feed proportions predicted by SourceTracker as continuous variable, and pond as strata: adonis (dis ~ feed, strata = design\$Pond, data = design, by = NULL)						
	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
feed	1	6.67712586	6.67712586	50.4357901	0.46087107	0.001
Residuals	59	7.81093	0.13238864	NA	0.53912893	NA
Total	60	14.4880559	NA	NA	1	NA
(2) adonis analysis on body mass index (PC1) as continuous variable, and pond as strata: adonis (dis ~ Comp.1, strata = design\$Pond, data = design, by = NULL)						
	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
Comp.1	1	0.97960187	0.97960187	4.27854366	0.06761445	0.003
Residuals	59	13.508454	0.22895685	NA	0.93238555	NA
Total	60	14.4880559	NA	NA	1	NA

Figure 2C

Multiple comparisons of proportions of feed (predicted by SourceTracker) using TukeyHSD.				
	diff	lwr	upr	p adj
Medium-Fast	0.106935	-0.1083305	0.32220053	0.461004
Slow-Fast	0.22879706	0.00340915	0.45418497	0.04588741
Slow-Medium	0.12186206	-0.1126851	0.35640922	0.4292129

Figure 2D

Pairwise comparisons using Wilcoxon rank sum test (P value adjustment method: BH)								
	Fst-Fd	Fst-Fst	Fst-Med	Fst-Slw	Med-Fd	Med-Med	Med-Slw	Slw-Fd
Fst-Fst	1.05E-05	NA	NA	NA	NA	NA	NA	NA
Fst-Med	7.01E-07	0.74256729	NA	NA	NA	NA	NA	NA
Fst-Slw	0.00036736	0.07494946	0.01044206	NA	NA	NA	NA	NA
Med-Fd	0.11633715	0.09130052	0.03107713	0.35913902	NA	NA	NA	NA
Med-Med	1.05E-06	0.69882918	0.86343745	0.03107713	0.03107713	NA	NA	NA
Med-Slw	3.08E-06	0.69882918	0.36686535	0.14483626	0.0693612	0.36686535	NA	NA
Slw-Fd	0.00036736	0.22086864	0.36686535	0.06239833	0.09560009	0.37318151	0.28345795	NA
Slw-Slw	1.33E-05	0.92290083	0.86343745	0.12770101	0.07494946	0.77457806	0.69882918	0.36686535

Figure 2E

Pairwise comparisons using Wilcoxon rank sum test (P value adjustment method: BH)					
	Fd	Fst	Med	Sdm	Slw
Fst	0.011500301	NA	NA	NA	NA
Med	0.011779595	0.804463546	NA	NA	NA
Sdm	0.022131884	9.18E-05	9.18E-05	NA	NA
Slw	0.02484192	0.04352426	0.072248752	0.000105016	NA
Wtr	0.022131884	9.18E-05	9.18E-05	0.000877415	0.000105016

Figure 4 Blomberg's K statistical output.

#ALL_fugu	K	PIC.variance.obs	PIC.variance.rnd.mean	PIC.variance.P	PIC.variance.Z
PC1	0.05398214	61.668551	205.5198	0.001	-8.668789
feed	0.05049687	2.065222	6.30681	0.001	-8.897068
#Fast_fugu	K	PIC.variance.obs	PIC.variance.rnd.mean	PIC.variance.P	PIC.variance.Z
PC1	0.04644161	16.148637	47.220684	0.001	-5.047588
feed	0.08689312	1.788734	9.770652	0.001	-8.297465
#medium_fugu	K	PIC.variance.obs	PIC.variance.rnd.mean	PIC.variance.P	PIC.variance.Z
PC1	0.02844361	9.942098	15.908873	0.001	-3.913017
feed	0.05284201	3.271343	9.773969	0.001	-7.303169
#slow_fugu	K	PIC.variance.obs	PIC.variance.rnd.mean	PIC.variance.P	PIC.variance.Z
PC1	0.03317783	5.482231	11.00622	0.001	-3.705141
feed	0.05582301	1.262326	4.26296	0.001	-8.018897

TABLE.S3 Metadata.

ID	Proportions of source predicted by SourceTracker				Group	PCA coordination based on body mass metrics				Fish Body Mass				Pond
	feed	sediment	water	Unknown		PC1	PC2	PC3	PC4	Weight.g	Length.cm	Weight.gonad.g	Weight.liver.g	
F1	0.812	0.000	0.003	0.185	Fast	2.462	1.034	0.353	0.149	94.740	14.700	0.060	12.920	P1
F10	0.002	0.000	0.000	0.997	Fast	1.939	1.941	0.047	0.237	94.330	15.000	0.020	12.150	P1
F11	0.407	0.000	0.005	0.588	Fast	1.386	0.207	0.464	-0.427	62.260	13.200	0.070	11.250	P1
F12	0.404	0.004	0.003	0.589	Slow	-2.028	0.515	-0.108	0.101	25.320	10.200	0.000	2.370	P1
F13	0.706	0.001	0.003	0.290	Slow	-2.137	0.489	0.214	0.060	24.250	9.500	0.000	2.910	P1
F14	0.939	0.002	0.004	0.055	Slow	-2.013	0.537	0.035	0.033	24.920	10.000	0.000	2.890	P1
F15	0.610	0.002	0.002	0.387	Slow	-2.117	-0.091	0.323	0.014	19.960	9.100	0.020	2.560	P1
F16	0.372	0.001	0.002	0.626	Slow	-3.010	0.068	0.476	0.149	12.200	7.800	0.000	1.000	P1
F17	0.754	0.007	0.005	0.234	Slow	-2.268	0.127	0.268	0.051	19.910	9.100	0.010	2.330	P1
F18	0.796	0.000	0.001	0.202	Medium	-0.261	-1.869	0.144	-0.010	31.850	10.800	0.110	3.710	P1
F19	0.747	0.001	0.003	0.249	Medium	-1.289	-0.307	-0.078	-0.022	27.320	10.700	0.040	3.160	P1
F2	0.349	0.000	0.001	0.650	Fast	1.759	0.342	-0.050	0.007	75.160	14.400	0.070	10.080	P1
F20	0.017	0.000	0.001	0.981	Medium	-0.295	-0.690	0.049	-0.137	36.620	11.500	0.070	5.200	P1
F21	0.912	0.016	0.016	0.057	Medium	-0.966	-0.735	0.022	-0.012	29.450	10.700	0.060	3.460	P1
F3	0.095	0.001	0.000	0.905	Fast	2.264	-0.324	-0.278	0.025	76.780	15.100	0.100	9.720	P1
F4	0.021	0.000	0.000	0.979	Fast	1.750	0.308	-0.361	0.089	75.070	14.900	0.070	9.160	P1
F5	0.616	0.001	0.002	0.381	Medium	-1.077	0.378	-0.316	0.111	36.840	11.600	0.020	3.630	P1
F6	0.053	0.000	0.001	0.946	Fast	1.746	0.006	-0.318	0.267	76.810	14.600	0.080	8.470	P1
F7	0.000	0.000	0.000	1.000	Fast	1.001	-0.362	-0.207	0.242	63.810	13.400	0.080	5.810	P1
F8	0.457	0.000	0.003	0.539	Medium	0.004	0.323	-0.359	0.096	50.990	12.900	0.040	5.670	P1
F9	0.350	0.001	0.001	0.648	Fast	1.429	-0.447	-0.252	0.113	66.250	14.000	0.090	7.750	P1
Fb1	0.105	0.000	0.001	0.894	Medium	0.051	0.312	-0.776	-0.092	45.660	13.800	0.040	5.260	P2
Fb10	0.054	0.000	0.001	0.945	Medium	-0.385	0.713	-0.535	-0.023	45.020	13.000	0.020	5.260	P2
Fb11	0.396	0.032	0.025	0.548	Slow	-1.992	0.214	-0.316	0.079	22.240	10.500	0.010	1.660	P2
Fb12	0.904	0.003	0.007	0.086	Slow	-3.380	-0.104	0.652	0.116	5.940	7.000	0.000	0.500	P2
Fb13	0.106	0.000	0.002	0.891	Medium	-0.431	-0.461	0.044	-0.032	38.540	11.400	0.060	4.970	P2
Fb14	0.237	0.000	0.002	0.761	Medium	-0.840	0.210	-0.220	-0.036	36.290	11.700	0.030	4.430	P2
Fb15	0.808	0.001	0.004	0.187	Slow	-2.950	0.095	0.426	0.089	11.720	8.000	0.000	1.180	P2
Fb16	0.164	0.001	0.002	0.834	Medium	0.060	0.045	-0.455	-0.065	45.950	13.100	0.050	5.580	P2
Fb17	0.281	0.001	0.002	0.717	Medium	0.058	0.390	0.029	-0.131	49.250	12.400	0.040	7.210	P2
Fb18	0.084	0.001	0.001	0.914	Medium	-0.995	-0.749	0.016	-0.069	27.730	10.700	0.060	3.500	P2
Fb19	0.516	0.001	0.002	0.482	Slow	-1.724	-0.214	0.067	-0.064	21.710	10.000	0.030	2.810	P2
Fb20	0.898	0.002	0.017	0.083	Medium	-0.674	-0.292	-0.056	-0.081	34.860	11.400	0.050	4.590	P2
Fb21	0.098	0.000	0.001	0.900	Fast	5.075	0.641	1.214	-0.130	126.500	16.100	0.120	20.160	P2
Fb22	0.854	0.001	0.003	0.142	Fast	2.829	-0.615	0.041	0.313	90.160	14.900	0.120	10.540	P2
Fb3	0.016	0.000	0.001	0.983	Fast	4.246	-0.447	0.534	0.109	109.090	15.800	0.140	15.120	P2
Fb5	0.014	0.000	0.000	0.986	Medium	-0.358	0.427	-0.525	-0.042	43.020	12.900	0.030	5.020	P2
Fb6	0.047	0.000	0.001	0.952	Slow	-2.918	0.128	0.617	0.098	13.490	7.700	0.000	1.680	P2
Fb7	0.642	0.001	0.003	0.355	Fast	1.016	-0.030	-0.017	-0.130	59.430	13.400	0.070	8.500	P2
Fb8	0.007	0.000	0.000	0.992	Fast	1.183	0.374	0.187	-0.140	65.360	13.400	0.060	9.810	P2
Fb9	0.255	0.000	0.001	0.744	Medium	-0.179	0.262	-0.048	-0.105	45.220	12.200	0.040	6.350	P2
Fc1	0.003	0.000	0.001	0.996	Fast	2.207	0.889	0.123	-0.126	83.280	14.900	0.060	12.340	P3
Fc10	0.283	0.000	0.001	0.716	Fast	1.455	-0.133	-0.259	0.065	67.800	14.200	0.080	8.290	P3
Fc11	0.055	0.000	0.001	0.944	Fast	2.282	0.016	0.048	-0.028	79.860	14.700	0.090	11.010	P3
Fc12	0.079	0.000	0.001	0.920	Fast	2.865	-1.501	-0.056	-0.013	77.030	14.900	0.150	10.010	P3
Fc13	0.258	0.000	0.002	0.740	Slow	-2.090	-0.134	-0.317	-0.041	15.880	10.300	0.020	1.290	P3
Fc14	0.051	0.000	0.002	0.948	Slow	-1.978	-0.034	0.158	-0.021	20.770	9.600	0.020	2.630	P3
Fc15	0.133	0.000	0.002	0.864	Fast	0.856	-0.123	-0.144	-0.113	56.300	13.400	0.070	7.750	P3
Fc16	0.113	0.001	0.001	0.886	Medium	-0.928	-0.130	-0.173	-0.099	31.540	11.400	0.040	4.070	P3
Fc17	0.004	0.000	0.000	0.996	Fast	1.501	-0.082	0.053	0.082	70.610	13.700	0.080	9.090	P3
Fc18	0.027	0.000	0.002	0.971	Fast	1.846	-1.715	-0.003	0.079	63.480	13.500	0.140	7.610	P3
Fc19	0.035	0.000	0.001	0.963	Medium	-0.192	0.555	-0.072	-0.195	45.120	12.400	0.030	6.840	P3
Fc2	0.041	0.000	0.001	0.958	Medium	-0.089	-0.023	-0.405	-0.247	39.790	12.900	0.050	5.710	P3
Fc20	0.287	0.002	0.001	0.711	Fast	0.884	-0.094	0.053	0.062	61.500	13.000	0.070	7.880	P3
Fc3	0.067	0.000	0.001	0.932	Medium	0.055	0.378	-0.059	-0.028	50.890	12.500	0.040	6.770	P3
Fc4	0.008	0.000	0.000	0.992	Slow	-1.976	-0.320	0.313	-0.025	19.430	9.200	0.030	2.620	P3
Fc5	0.004	0.000	0.000	0.995	Fast	1.508	-0.399	-0.197	-0.066	64.160	14.100	0.090	8.490	P3
Fc6	0.054	0.000	0.002	0.944	Fast	1.768	0.643	-0.118	-0.065	75.570	14.700	0.060	10.490	P3
Fc7	0.183	0.000	0.001	0.816	Slow	-1.993	-0.345	0.129	-0.018	18.290	9.500	0.030	2.140	P3
Fc8	0.004	0.000	0.000	0.996	Slow	-1.667	0.102	-0.066	0.018	25.860	10.400	0.020	2.850	P3
Fc9	0.213	0.001	0.001	0.785	Slow	-2.287	0.099	0.049	-0.021	16.920	9.500	0.010	1.940	P3

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

Family	F	FD	S	W
Cytophagaceae	-0.512	-0.512	-0.475	1.500
Chitinophagaceae	-0.792	-0.782	1.298	0.275
Conexibacteraceae	-0.768	-0.772	0.209	1.331
Caulobacteraceae	-0.814	-0.794	0.349	1.259
Mycoplasmataceae	1.500	-0.499	-0.500	-0.500
Oxalobacteraceae	-0.859	-0.866	0.978	0.747
Comamonadaceae	-0.777	-0.781	0.243	1.314
Cryomorphaceae	-0.609	-0.610	-0.260	1.479
Brevinemataceae	1.500	-0.500	-0.500	-0.500
Methylocystaceae	-0.812	-0.818	0.401	1.229
Sphingomonadaceae	-0.681	-0.680	-0.077	1.438
Mycobacteriaceae	-0.521	-0.521	-0.458	1.499
Planctomycetaceae	-0.542	-0.555	1.496	-0.399
Rhodobacteraceae	-0.740	-0.843	1.286	0.297
Rhodocyclaceae	-0.507	-0.508	1.500	-0.485
Saprospiraceae	-0.841	-0.838	1.139	0.540
Acidimicrobiaceae	-0.683	-0.739	1.410	0.012
Chromatiaceae	-0.793	-0.793	1.284	0.302
Opitutaceae	-0.622	-0.622	1.474	-0.231
(Unassigned)	-0.879	-0.714	1.270	0.323
Flavobacteriaceae	-0.691	-0.705	-0.027	1.423
Verrucomicrobiaceae	-0.662	-0.659	-0.132	1.453
Rikenellaceae	1.500	-0.500	-0.500	-0.500
Burkholderiaceae	-0.580	-0.260	-0.640	1.479
Geobacteraceae	-0.518	-0.463	1.499	-0.518
Microbacteriaceae	-0.509	-0.517	-0.474	1.500
Xanthomonadaceae	-0.668	-0.681	1.443	-0.094
Methylococcaceae	-0.596	-0.600	1.484	-0.287
Planococcaceae	-0.490	1.500	-0.504	-0.506
Carnobacteriaceae	-0.500	1.500	-0.503	-0.497
Idiomarinaceae	-0.500	1.500	-0.500	-0.500
Bacillaceae_2	-0.500	1.500	-0.500	-0.500
Oceanospirillaceae	-0.500	1.500	-0.500	-0.500
Acidaminococcaceae	-0.529	1.498	-0.547	-0.423

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

Order	Fst	Med	Slw	Fd	Wtr	Sdm
Gammaproteobacteria	43.4	42.2	61.8	80.4	7.8	15.3
Spirochaetia	17.8	18.9	11.2	0.0	0.0	0.4
Bacteroidia	16.5	15.8	8.6	0.5	0.0	1.3
Mollicutes	9.9	12.0	5.7	0.0	0.0	0.0
Betaproteobacteria	0.3	0.4	0.6	1.1	19.8	21.5
Actinobacteria	0.1	0.2	0.5	0.2	33.6	3.5
Clostridia	4.8	3.4	3.8	3.4	0.0	0.4
Unassigned	2.2	2.7	2.4	1.6	0.7	10.5
Alphaproteobacteria	0.1	0.2	1.3	0.3	13.8	9.7
Cytophagia	0.0	0.0	0.1	0.0	10.9	3.0
Bacilli	1.0	1.7	1.5	12.2	0.2	0.0
Other	4.0	2.5	2.5	0.3	13.2	34.5

1195 **Table.S6 Discriminative bacterial taxa identified by LEfSe.**

Taxa	Log10LDA	Group	Pval
Proteobacteria.Gammaproteobacteria.Pseudomonadales	5.652	Slw	0.047
Proteobacteria.Gammaproteobacteria.Pseudomonadales.Pseudomonadaceae	5.634	Slw	0.043
Proteobacteria.Gammaproteobacteria.Pseudomonadales.Pseudomonadaceae.Pseudomonas	5.634	Slw	0.043
Proteobacteria.Gammaproteobacteria.Enterobacteriales	4.580	Slw	0.050
Proteobacteria.Gammaproteobacteria.Enterobacteriales.Enterobacteriaceae	4.580	Slw	0.050
Proteobacteria.Gammaproteobacteria.Enterobacteriales.Enterobacteriaceae_	4.435	Slw	0.034
Proteobacteria.Alphaproteobacteria	4.115	Slw	0.050
Proteobacteria.Alphaproteobacteria.Rhodobacterales	3.888	Slw	0.046
Proteobacteria.Alphaproteobacteria.Rhodobacterales.Rhodobacteraceae	3.888	Slw	0.046
Proteobacteria.Alphaproteobacteria.Rhizobiales.Brucellaceae.Pseudochrobactrum	1.529	Slw	0.018
Firmicutes.Bacilli.Lactobacillales.Enterococcaceae.Vagococcus	1.772	Slw	0.005
Proteobacteria.Betaproteobacteria.Burkholderiales.Comamonadaceae.Comamonas	2.041	Slw	0.045
Proteobacteria.Deltaproteobacteria.Desulfobacterales.Desulfobulbaceae.Desulfoprunum	1.706	Slw	0.019
Proteobacteria.Alphaproteobacteria.Rhodobacterales.Rhodobacteraceae.Tabrizicola	3.305	Slw	0.010
Firmicutes.Clostridia.Clostridiales.Ruminococcaceae_	2.205	Slw	0.001
Firmicutes.Bacilli.Lactobacillales.Streptococcaceae.Streptococcus	3.285	Slw	0.048
Firmicutes.Bacilli.Bacillales.Staphylococcaceae_	2.662	Fst	0.031
Actinobacteria.Actinobacteria.Actinomycetales_	3.223	Slw	0.033
Actinobacteria.Actinobacteria.Actinomycetales_	3.223	Slw	0.033
Proteobacteria.Deltaproteobacteria_	1.927	Slw	0.045
Actinobacteria.Actinobacteria.Actinomycetales.Microbacteriaceae.Microbacterium	2.302	Med	0.008
Proteobacteria.Gammaproteobacteria.Pseudomonadales.Moraxellaceae.Psychrobacter	2.580	Slw	0.013
Firmicutes.Bacilli.Bacillales.Bacillaceae_1.Bacillus	3.120	Slw	0.043
Proteobacteria.Deltaproteobacteria.Desulfobacterales.Desulfobulbaceae	2.007	Slw	0.021
Proteobacteria.Deltaproteobacteria_	1.927	Slw	0.045
Proteobacteria.Deltaproteobacteria_	1.927	Slw	0.045
Proteobacteria.Deltaproteobacteria.Desulfobacterales	2.007	Slw	0.021
Bacteroidetes.Bacteroidia.Bacteroidales.Porphyromonadaceae	3.270	Med	0.015
Acidobacteria.Acidobacteria_Gp10_	2.157	Slw	0.018
Proteobacteria.Betaproteobacteria.Burkholderiales.Comamonadaceae	3.238	Slw	0.040
Bacteroidetes.Bacteroidia.Bacteroidales.Porphyromonadaceae_	3.131	Med	0.033
Bacteroidetes.Bacteroidia.Bacteroidales.Porphyromonadaceae.Barnesiella	2.720	Fst	0.018
Proteobacteria.Betaproteobacteria.Burkholderiales.Comamonadaceae_	3.172	Slw	0.037
Proteobacteria.Alphaproteobacteria_	2.598	Slw	0.010
Actinobacteria.Actinobacteria.Acidimicrobiales	2.783	Slw	0.025
Proteobacteria.Alphaproteobacteria_	2.598	Slw	0.010
Planctomycetes.Planctomycetia.Planctomycetales.Planctomycetaceae	2.801	Slw	0.023
Planctomycetes	2.801	Slw	0.023
Planctomycetes.Planctomycetia.Planctomycetales	2.801	Slw	0.023
Planctomycetes.Planctomycetia.Planctomycetales.Planctomycetaceae_	2.666	Slw	0.022
Acidobacteria.Acidobacteria_Gp10_..Gp10	2.102	Slw	0.018
Planctomycetes.Planctomycetia	2.801	Slw	0.023
Proteobacteria.Alphaproteobacteria_	2.598	Slw	0.010
Proteobacteria.Gammaproteobacteria.Enterobacteriales.Enterobacteriaceae.Cronobacter	2.340	Slw	0.033
Chloroflexi_	2.374	Slw	0.041
Proteobacteria.Gammaproteobacteria.Xanthomonadales.Xanthomonadaceae_	2.006	Slw	0.018
Acidobacteria.Acidobacteria_Gp10	2.157	Slw	0.018
Acidobacteria.Acidobacteria_Gp10_	2.157	Slw	0.018
Chloroflexi_	2.374	Slw	0.041
Chloroflexi_	2.374	Slw	0.041
Chloroflexi_	2.374	Slw	0.041
Chloroflexi.Caldilineae.Caldilineales.Caldilineaceae.Litorilinea	2.659	Slw	0.043
Proteobacteria.Deltaproteobacteria	2.529	Slw	0.014

1196

1197

1198

1199

1200

1201

1202

FIGURE LEGENDS

Figure 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).

Figure 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.

Figure 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.

Figure 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.

Figure 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).

Figure 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.

Figure 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”.

Figure 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.

Figure 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.

Figure 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.

Mantel correlogram PC1+Feed

Figure 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.

Figure 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.

Figure 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 (R²) and migration rate (*m*) are shown as plot title of each subpanel.