Molecular Ecology

WILEY

MOLECULAR ECOLOGY

ORIGINAL ARTICLE OPEN ACCESS

Associations Between Gut Microbiota Diversity and a Host Fitness Proxy in a Naturalistic Experiment Using Threespine Stickleback Fish

Andreas Härer[1] | Ken A. Thompson[2,3,4] | Dolph Schluter[2] | Diana J. Rennison[1]

[1]Department of Ecology, Behavior, & Evolution, School of Biological Sciences, University of California San Diego, La Jolla, California, USA | [2]Department of Zoology, Biodiversity Research Centre, University of British Columbia, Vancouver, British Columbia, Canada | [3]Department of Biology, Stanford University, Stanford, California, USA | [4]Department of Plant Biology, Carnegie Institution for Science, Stanford, California, USA

Correspondence: Andreas Härer (ahaerer@ucsd.edu) | Diana J. Rennison (drennison@ucsd.edu)

Received: 13 February 2024 | **Accepted:** 3 October 2024

Handling Editor: Nick Fountain-Jones

Funding: This study was supported by Natural Sciences and Engineering Research Council of Canada, Human Frontier Science Program, Deutsche Forschungsgemeinschaft (458274593), and University of California, San Diego.

Keywords: 16S rRNA sequencing | Anna Karenina principle | fitness | *Gasterosteus aculeatus* | gut microbiome | host–microbe interaction

ABSTRACT

The vertebrate gut microbiota is a critical determinant of organismal function, yet whether and how gut microbial communities affect host fitness under natural conditions remains largely unclear. We characterised associations between a fitness proxy—individual growth rate—and bacterial gut microbiota diversity and composition in threespine stickleback fish introduced to large semi-natural ponds. We detected a 63% higher richness of bacterial taxa (α-diversity) in the guts of high-fitness fish compared to low-fitness fish, which might be driven by stronger bacterial dispersal among high-fitness fish according to the fit of a neutral community model. Further, microbial communities of high-fitness fish were more similar to one another (i.e., exhibited lower β-diversity) than those of low-fitness fish. The lower β-diversity found to be associated with higher host fitness is consistent with the Anna Karenina principle—that there are fewer ways to have a functional microbiota than a dysfunctional microbiota. Our study links differences in α- and β-diversity to a fitness-related trait in a vertebrate species reared under naturalistic conditions and our findings provide a basis for functional tests of the fitness consequences of host-microbiota interactions.

1 | Introduction

There is increasing recognition of the contribution of the gut microbiota, the microbial community associated with a host's gut, to the evolution of multicellular hosts (Henry et al. 2021; Kolodny and Schulenburg 2020), and host–microbe interactions appear to be a universal feature of eukaryotes (Youngblut et al. 2019; Thompson et al. 2017; Trivedi et al. 2020). Biotic interactions between animals and their gut microbiota can affect host traits crucial for determining host performance and survival via effects on metabolism, regulation of the immune system, disease susceptibility, behaviour and stress tolerance (Turnbaugh et al. 2006; Lathrop et al. 2011; Houwenhuyse et al. 2021; Vuong et al. 2017; Bates et al. 2022). Microbes can also affect niche use, enabling metabolism of novel dietary resources (Henry et al. 2021; Gould et al. 2018). Yet only few studies have explored the precise relationship between fitness-related host traits and the gut microbiota in free-living, non-domesticated, vertebrates (Davidson et al. 2021; Worsley et al. 2021; Risely et al. 2023). In sum, while there is increasing evidence that gut microbiota differences can alter an organism's performance and survival, we still have limited knowledge on the mechanisms underlying this relationship.

There is some evidence linking gut microbiota diversity with host performance and these observed patterns may extend to overall host fitness, which allows making testable predictions. For example, lower bacterial diversity within a host's gut (α-diversity) appears to be commonly associated with disease states (Kriss et al. 2018) and higher α-diversity has been shown to be correlated with higher survival rates (Bestion et al. 2017) and body condition (Stoffel et al. 2020). While the evidence is limited, these results indicate that higher α-diversity may be associated with higher host fitness. Yet, other studies detected no (Worsley et al. 2021; Goertz et al. 2019) or even negative associations between α-diversity and host condition (Davidson et al. 2021). Further, the extent of dissimilarity in gut microbiota composition (β-diversity) might also be associated with host performance. This pattern may be akin to the Anna Karenina principle for animal microbiomes, which states that 'all healthy microbiota are similar; each dysbiotic microbiota is dysbiotic in its own way' (Zaneveld, McMinds, and Vega Thurber 2017). This is based on the prediction that differences in host-associated microbial communities induced by environmental perturbations are stochastic and, thus, lead to greater dissimilarity in gut microbiota composition (i.e., higher β-diversity) among dysbiotic hosts. The Anna Karenina principle was originally developed to explain the observed effects of perturbations and stress on microbiota diversity. However, the framework could be extended to the context of host fitness, leading to the prediction that higher β-diversity is associated with reduced host fitness.

Differences in microbial community dynamics as well as interhost dispersal could also potentially affect host fitness, for example, by sharing beneficial bacteria that allow metabolising a broader range of nutrients or by increasing microbiome diversity which can be beneficial for host organisms (Bestion et al. 2017; Stoffel et al. 2020). Microbial community dissimilarity (β-diversity) can be affected by the extent of bacterial dispersal among hosts. The role of dispersal in shaping structure and function of bacterial communities is increasingly acknowledged (Albright and Martiny 2018), including in animal microbiomes (Burns et al. 2017). For example, studies of wild animal populations revealed that social interactions among hosts, which can facilitate microbial exchange, predict microbiome composition (Tung et al. 2015). These results indicate that the nature and frequency of host–host interactions may affect how commonly microbes are shared. Thus, studying animal microbiomes using metacommunity theory, which describes the movement of species among distinct local communities, can provide crucial insights into microbiome dynamics (Leibold et al. 2004).

While some studies have started exploring the fitness effects of host-microbiota interactions in wild vertebrate populations, still insights predominantly come from a few invertebrate model systems reared under laboratory conditions and commonly involve experimental gut microbiota manipulation (e.g., via antibiotic treatment) (Weiland-Brauer et al. 2020; Sison-Mangus, Mushegian, and Ebert 2015). Because many environmental variables in the laboratory are substantially different from natural environments, it is critical to evaluate host fitness-gut microbiota relationships in natural settings to better understand the impact on host biology (Suzuki 2017). Yet, making inferences in wild individuals can be complicated because individual host fitness as well as the gut microbiota are influenced by a range

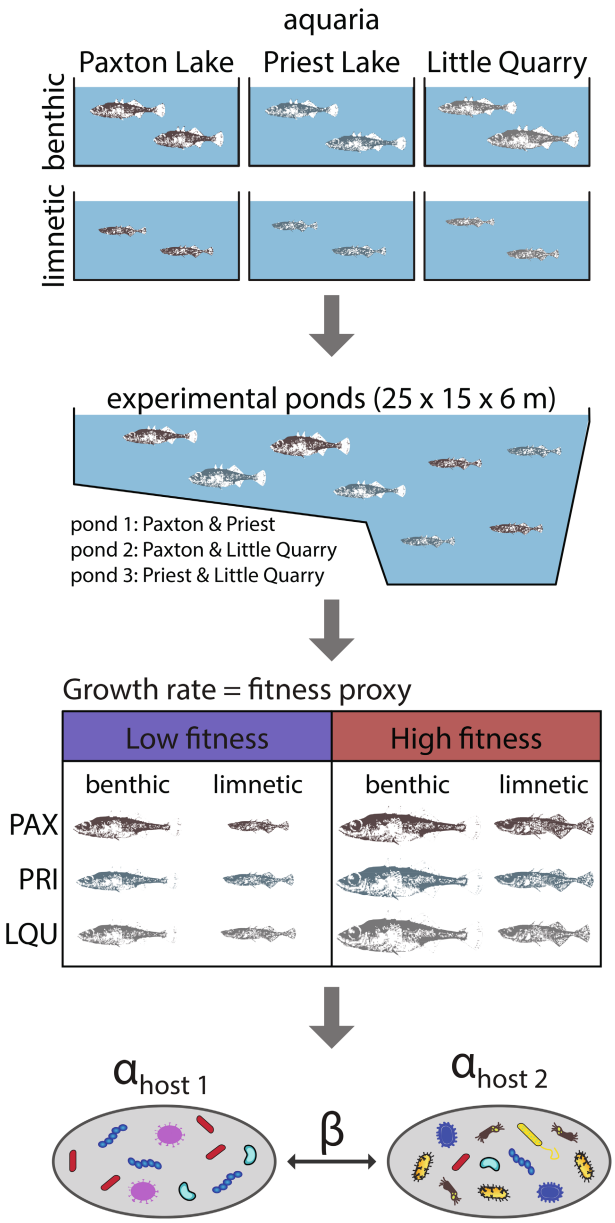

FIGURE 1 | Benthic and limnetic fish from Paxton Lake (PAX), Priest Lake (PRI) and Little Quarry (LQU) were reared in aquaria until an age of 78–119 days post-hatch (dph). Then, they were introduced into three ponds in different combinations and the experiment ran for 72–99 days. Their fitness was assessed by growth rate ($n_{\text{low-fitness}} = 121$, $n_{\text{high-fitness}} = 119$) and gut microbiota α- and β-diversity was quantified based on 16S rRNA gene sequencing data.

of abiotic and biotic factors that are often difficult to account for or measure, such as the rearing environment or age of the organism. By controlling such host and environmental factors, experiments conducted under naturalistic conditions provide a powerful opportunity to investigate associations between the gut microbiota and fitness-related host traits.

Here, we leveraged a unique field-based experimental infrastructure to determine whether bacterial gut microbiota diversity is associated with differences in growth rate—a test of the predictions for α- and β-diversity outlined above (Figure 1). We used threespine stickleback fish (*Gasterosteus aculeatus*;

hereafter 'stickleback'), an emerging model system for eco-evolutionary gut microbiota research (Rennison, Rudman, and Schluter 2019; Bolnick et al. 2014; Smith et al. 2015; Härer and Rennison 2024). Stickleback are well-suited for studying gut microbiota-mediated fitness effects because growth rate can be used as a fitness proxy since body size is strongly positively correlated with fitness via higher fecundity and overwinter survival (Sparkes et al. 2013; Moser, Roesti, and Berner 2012; Thompson and Schluter 2022). To obtain first insights into gut microbiota–host fitness associations, we introduced different combinations of ecologically divergent benthic and limnetic ecotypes originating from three lakes in British Columbia, Canada, into three large experimental ponds and characterised the gut microbiota of individuals with the lowest (low-fitness fish) and highest (high-fitness fish) growth rates from each source population and pond (Figure 1 and Figure S1). In addition to testing associations between host fitness and α- and β-diversity, we further explored the mechanisms shaping gut microbiota dynamics by testing whether bacterial interhost dispersal as well as the contribution of neutral and non-neutral process to gut microbiota community structure differ between fitness groups. Our study provides novel insights by assessing host fitness-gut microbiota associations on the individual host level in free-living animals under naturalistic conditions.

2 | Materials and Methods

2.1 | Experimental Design

All fish used for this study were part of an experiment studying the fitness consequences of hybridisation between 'parallel' ecotypes (no hybrids were used herein), and we refer readers to Thompson and Schluter (Thompson and Schluter 2022) for more detailed information about the experiment. Briefly, parents of the experimental fish were raised from hatching in a common laboratory environment and pure within-population crosses were made between unrelated and lab-hatched fish of benthic and limnetic populations (variously called 'species' or 'ecotypes') from each of Paxton, Priest and Little Quarry Lakes, in British Columbia, Canada (Figure 1). Their offspring—our focal fish—were raised in aquaria until 78–119 days post-hatch (dph) and fed a common diet. Aquaria were either static or connected to a recirculating system, and low and high fitness fish were equally distributed between these two aquaria types ($n_{\text{low-fitness, static}} = 89, n_{\text{high-fitness, static}} = 89, n_{\text{low-fitness, recirculating}} = 32$, $n_{\text{high-fitness, recirculating}} = 30$; see Section 2.2 for information about fitness groups). For the experiment, fish were kept in three unmanipulated semi-natural ponds (25 × 15 m including benthic and limnetic habitats) on the University of British Columbia campus (Arnegard et al. 2014). While we did not measure physiochemical parameters of the ponds (e.g., temperature, salinity, pH or light exposure), we expect these conditions to be similar across ponds for several reasons. All ponds have been filled from the same water source, are not heated or manipulated in any other way, contain the same volume of water and are located just meters from each other. Thus, these ponds are exposed to the same climatic conditions at the research facility including ambient temperature and precipitation and we expect little variation across ponds. More broadly, we treat 'pond' as a variable in our analyses. In each pond, there were combinations of benthic and limnetic populations from two lakes each (e.g., benthic and limnetic fish from Paxton and Priest Lakes in pond 1; Figure 1), and fish were kept in the ponds for 72–99 days. At the end of the experiment, fish were 158–201 dph. Fish were weighed before and after the experiment, and sequential coded wire tags (Northwest Marine Technology, Anacortes, WA, USA) were used for individual identification. There were no differences in initial weight, (two sample *t*-test, $t = 0.422$, $p = 0.673$), age at the start of the experiment ($t = 1.329$, $p = 0.185$) or age at the end of the experiment ($t = 1.016$, $p = 0.311$) between low-fitness and high-fitness fish. Wild progenitors of the fish used in our experiment were collected between 2017 and 2019 under the following permit numbers: SU17-258923, SARA17-PPAC-00002, MRSU18-288855, SARA18-PPAC-00006, MRSU19-454239, SARA19-PPAC-00006. The experiment was conducted in accordance with institutional guidelines under the following animal care permit numbers: A16-0044, A20-0050.

2.2 | Data Collection

Fish from all ponds were exhaustively collected using minnow traps, dip nets or from the surface of the pond following rotenone application and were shortly thereafter euthanised with an overdose of MS-222. Next, fish had their wire tags recovered, then were weighed, photographed and stored at −20°C in individually labelled 15-mL tubes. Growth rate for each individual fish—our fitness proxy—was calculated based on weight gain during the experiment with linear models using final weight as the response variable and initial weight and the days an individual spent in the pond during the experiment as fixed effects and pond as a random effect. Models for benthic and limnetic ecotypes were fitted separately due to substantial differences in body size and allometry, see Thompson and Schluter for more details (Thompson and Schluter 2022). Sample sizes for each population and pond ranged from 31 to 95 (Figure S1). Out of these, we sampled approximately 20 fish per population within each pond. To maximise our power to detect differences in growth rate between stickleback hosts, we selected the 10–12 slowest growing (i.e., low relative fitness) and the 10–12 fastest growing (i.e., high relative fitness) fish from each source population and pond. Several samples were later excluded due to low sequencing depths, and final sample sizes ranged from 8 to 12 fish per group ($n_{\text{low-fitness}} = 121$, $n_{\text{high-fitness}} = 119$). This sampling design allowed us to test for a difference in a discrete outcome (group assignment) rather than a lower-powered continuous approach. Low- and high-fitness fish differed substantially in growth rate, with high-fitness fish having on average between 30.8% and 73.0% higher growth rates across populations and ponds (Figure S1). Further, our sample sizes were balanced regarding pond ($n_{\text{pond1}} = 79$, $n_{\text{pond2}} = 81$, $n_{\text{pond3}} = 80$), lake of origin ($n_{\text{little quarry}} = 75$, $n_{\text{paxton}} = 81$, $n_{\text{priest}} = 84$), ecotype ($n_{\text{benthic}} = 122$, $n_{\text{limnetic}} = 118$) and host sex ($n_{\text{female}} = 118$, $n_{\text{male}} = 121$, $n_{\text{undetermined}} = 1$).

To collect intestinal tissues, fish were rinsed with 95% EtOH and their whole guts were dissected using sterile equipment. We carefully removed any gut contents and stored samples in sterile tubes at −80°C. DNA was extracted from whole guts with the QIAGEN PowerSoil Pro Kit according to the manufacturer's protocol (Qiagen, Hilden, Germany) under a sterile laminar flow hood.

To characterise microbial communities associated with fish guts, we amplified the V4 region of the 16S rRNA gene with barcoded 515F and 806R primers (https://github.com/SchlossLab/MiSeq_WetLab_SOP/blob/master/MiSeq_WetLab_SOP.md). All PCR reactions were done in triplicate using the Q5 High-Fidelity 2X Master Mix (New England Biolabs, Ipswich, MA) and pooled for each fish after amplification. The PCR had a denaturation step for 60 s at 98°C, 35 amplification cycles with 10 s at 98°C, 20 s at 56°C and 60 s at 72°C, and a final elongation at 72°C for 10 min. To check for successful amplification, we visualised PCR products by gel electrophoresis (2% agarose gel) and DNA concentrations were measured on a Qubit 4 Fluorometer (Thermo Fisher Scientific, Waltham, MA). We included negative controls of sterile water for DNA extraction and PCRs, none of which yielded any detectable DNA amplification. All samples were subsequently pooled in an equimolar manner for the two libraries (samples were either sequenced in 2021 or 2022). At the UC Davis Genome Center, libraries were purified by bead clean-up and DNA quality was checked on a Bioanalzyer (Agilent Technologies, Santa Clara, CA). The final libraries were sequenced on the Illumina MiSeq 600 (PE300) platform.

To obtain information on diet, we collected muscle tissue to determine stable isotope ratios of carbon ($\delta^{13}C$) and nitrogen ($\delta^{15}N$), which allow detecting diet variation associated with benthic and limnetic habitats in stickleback (Arnegard et al. 2014; Bolnick and Ballare 2020). Muscle tissues were dried at 55°C and subsequently homogenised to a powder, 1 mg of each sample was loaded into a tin capsule and combusted in a Elementar vario EL cube elemental analyser interfaced to an Elementar VisION IRMS (Elementar Analysensysteme GmbH, Germany) at the UC Davis Stable Isotope Facility. Laboratory standards indicated measurement errors (SD) of ±0.05‰ for $\delta^{13}C$ and ± 0.07‰ for $\delta^{15}N$. We preserved fin tissue of each fish in 95% EtOH, and after DNA extraction the sex of each fish was determined by PCR following the protocol developed by Peichel et al. (2004).

2.3 | Data Analysis

The data consisted of a total of 9,493,085 raw sequencing reads (mean: 39,554 reads/sample) (Table S1). For some samples, we obtained low sequencing depths, which were further decreased by merging of forward and reverse reads due to filtering during this step. Hence, we chose to use 250 bp of the forward reads for our gut microbiota analyses, since these reads consistently showed higher sequence quality compared to reverse reads yet encompassed 86% of the target locus (250 of 291 bp). All upstream analyses described below were done in QIIME2 (Bolyen et al. 2019). In order to obtain amplicon sequencing variants (ASVs), sequencing reads were checked for quality and corrected, and chimeric sequences were removed using the *dada2* plugin (Callahan et al. 2016). Next, we used FastTree 2.1.3 to assemble a phylogenetic tree of the bacterial lineages (Price, Dehal, and Arkin 2010), and bacterial taxonomy was assigned based on the SILVA 138 ribosomal RNA (rRNA) database at a 99% similarity threshold (Quast et al. 2013). Before downstream analyses, we excluded ASVs with < 10 reads that were detected only in a single sample and ASVs that could not be assigned at the class level. We further filtered out ASVs mapping to chloroplasts, mitochondria, cyanobacteria, or archaea to retain the bacterial gut microbiota only. ASV counts were normalised through scaling with ranked subsampling (SRS) with a C_{min} of 2500 reads (Beule and Karlovsky 2020).

We tested for effects of ecotype, lake-of-origin, carbon and nitrogen stable isotope signature, sex, age, rearing aquarium type (static vs. recirculating), fitness group and pond on α-diversity (i.e., bacterial diversity of individual host fish; ASV richness, Faith's phylogenetic diversity and Shannon diversity) and β-diversity (i.e., gut microbiota dissimilarity between hosts; non-phylogenetic: Bray–Curtis dissimilarity, phylogenetic: unweighted and weighted UniFrac) (Lozupone and Knight 2005; Lozupone et al. 2011). For α-diversity, we used linear mixed effect-models (*lmer* function in the 'lme4' package v1.1-31) (Bates et al. 2015) with pond as random effect and all other variables as fixed effects and produced analysis-of-variance tables to test for statistical significance of the model terms (*anova* function in the 'lmerTest' package v3.1-3) (Kuznetsova, Brockhoff, and Christensen 2017). Adding sequencing read number as an independent variable to the models showed that it had a significant effect on ASV richness and Shannon diversity, but difference in α-diversity between fitness groups were maintained, indicating that the observed effects were not driven by differences in read numbers (Table S2).

To compare gut microbiota dissimilarity within and between fitness groups (β-diversity), we used PERMANOVA (*adonis2* function in the 'vegan' package v2.6-2) (Anderson 2001; Oksanen et al. 2019). To test whether the magnitude of gut microbiota dissimilarity measured within each fitness group differed between fitness groups, we determined β-diversity dispersion by calculating the distance of each fish from the centroid of its respective fitness group (PERMDISP; vegan::*betadisper*). We used vegan::*adonis2* to calculate *p*-values for the comparison of β-diversity values between fitness groups. Results were consistent across β-diversity metrics, and we only report unweighted UniFrac statistics in the main text. Statistics for the other two metrics can be found in Table S3.

Higher α-diversity in high-fitness fish could result from differences in the cumulative diversity of microbial communities (γ-diversity) between fitness groups. Thus, we determined γ-diversity based on ASV richness for each fitness group. To test for significant differences in γ-diversity between fitness groups, individual hosts were resampled with replacement 10,000 times for each fitness group, γ-diversity was calculated for each iteration and the γ-diversity ratio between the high-fitness group and the low-fitness group was determined. We then tested for higher γ-diversity in the high-fitness group by calculating the proportion of iterations for which the high-fitness group had a higher γ-diversity and we determined statistical significance from this bootstrap procedure using a two-tailed cut-off of 0.05. Differential abundance of bacterial phyla was assessed by analysis of composition of microbiomes (ANCOM). We further used the Sloan Neutral Community Model, which assumes that community structure is driven solely by chance and dispersal, to test for the contribution of neutral and non-neutral process in shaping gut microbial community composition among stickleback hosts of the same fitness group (Burns et al. 2016; Sloan et al. 2006). Based on this model, we inferred estimated values of bacterial interhost dispersal as well as ASV frequencies and

compared those estimates between fitness groups. Since bacterial dispersal can only occur among fish from the same pond, we ran separate models by pond. All statistical analyses were done in R v4.2.1 (R Core Team 2022).

3 | Results

3.1 | High-Fitness Fish Have a More Diverse Gut Microbiota

In agreement with theoretical predictions, the bacterial diversity associated with single hosts (α-diversity) was higher in high-fitness fish than in low-fitness fish. This pattern was seen for ASV richness (linear mixed-effects model; $F_{1,205.15}=11.25$, $p=0.001$), Faith's phylogenetic diversity ($F_{1,210.91}=8.50$, $p=0.004$) and Shannon diversity ($F_{1,165.87}=6.85$, $p=0.010$) (Figure 2A and Figure S2) after controlling for effects of various covariates (see Section 2; Table S2). Median ASV richness (192 vs. 118) and Faith's phylogenetic diversity (18.78 vs. 13.98) were 63% and 34% higher in high-fitness fish compared to low-fitness fish, but median Shannon diversity was only 4% higher in high-fitness fish (3.85 vs. 3.69). Besides fitness group, we detected an effect of host sex on all three α-diversity metrics (ASV richness: $F_{1,216.98}=7.36$, $p=0.007$, Faith's phylogenetic diversity: $F_{1,216.75}=5.78$, $p=0.017$, Shannon diversity: $F_{1,217.73}=7.95$, $p=0.005$) (Table S2). The higher ASV richness observed in high-fitness fish was not reflected by higher cumulative ASV richness since γ-diversity was similar for the low- and high-fitness groups (4343 and 4387 ASVs; $p_{\text{bootstrapped}}=0.186$); 2374 ASVs were present in both fitness groups whereas 1969 and 2013 ASVs were unique to the low- and high-fitness groups, respectively. While no ASVs were shared among all members of a fitness group, the major bacterial phyla were similar across fitness groups and only the bacterial phylum Planctomycetes was more abundant in high-fitness fish (ANCOM; $W=22$) (Figure 2B).

3.2 | Gut Microbiota Dissimilarity Is Lower Among High-Fitness Fish

Our data support the Anna Karenina principle: gut microbiota dissimilarity (β-diversity) was consistently lower among fish within the high-fitness group than among fish within the low-fitness group across the three β-diversity metrics (PERMDISP; unweighted UniFrac: $F_{1,239}=6.96$, $R^2=0.028$, $p=0.004$) (Figure 2C). We further detected significant gut microbiota dissimilarity between low- and high-fitness fish across the three β-diversity metrics (PERMANOVA; unweighted UniFrac: $F_{1,228}=3.10$, $R^2=0.012$, $p=0.001$), suggesting that microbial community composition differs between fitness groups (Figure 2D and Table S3). Host sex and age, lake-of-origin, and pond also had significant effects on dissimilarity of gut microbiota composition across all three metrics (Table S3). Host ecotype and nitrogen signature showed significant effects based on Bray–Curtis dissimilarity and weighted UniFrac and carbon signature showed a significant effect based on Bray–Curtis dissimilarity (Table S3). Overall, these results indicate that gut microbial communities differ between fitness groups and that they are more similar among high-fitness fish than among low-fitness fish.

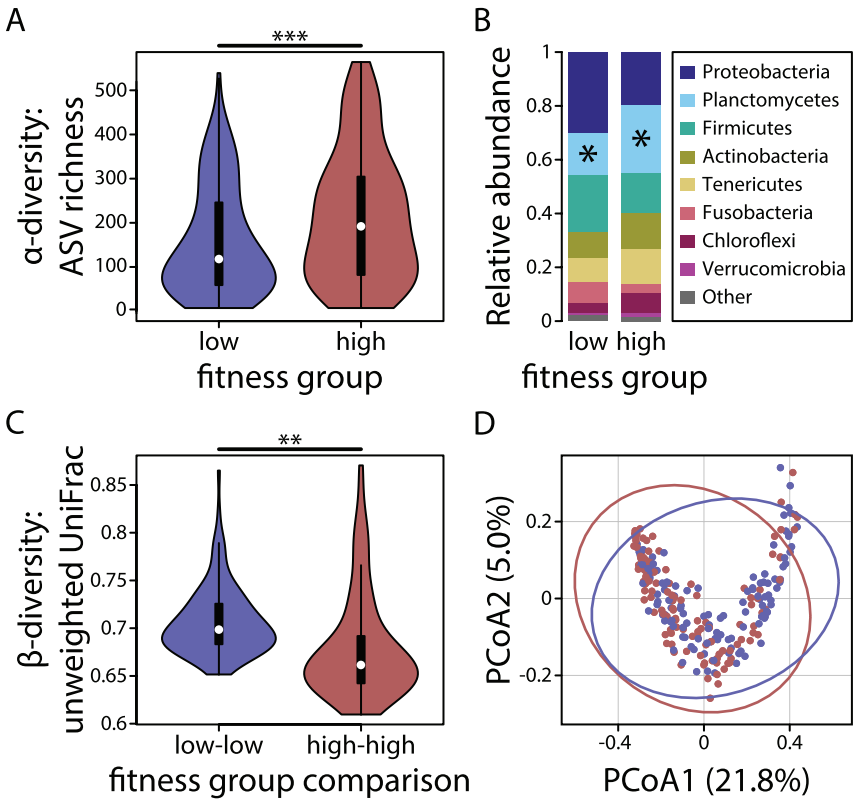

FIGURE 2 | High-fitness fish had higher α-diversity (A), higher abundance of Planctomycetes as indicated by the asterisk (*) (B), and lower β-diversity (distance from fitness group centroid) (C) compared to low-fitness fish. Low- and high-fitness fish further differed significantly in their gut microbiota composition (D). Significance thresholds as shown in (A) and (C): $^{**}p<0.01$, $^{***}p<0.001$.

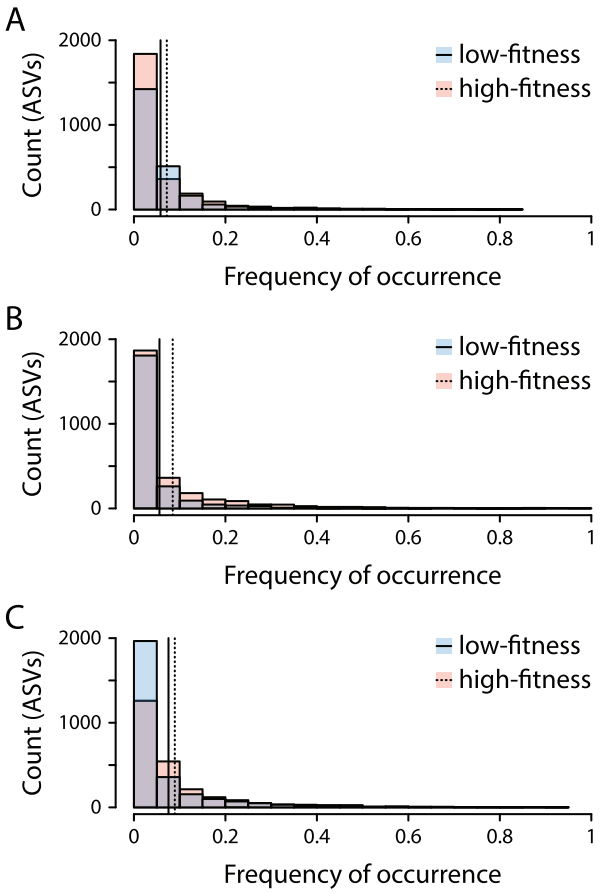

FIGURE 3 | Frequency of occurrence of ASVs in low- and high-fitness fish separated by pond. On average, ASVs were shared among more hosts in high-fitness fish (dashed vertical line) compared to low-fitness fish (solid vertical line), indicating higher dispersal among high-fitness fish.

The higher α-diversity and lower β-diversity observed in high-fitness fish might be explained by differences in bacterial interhost dispersal within each fitness group. Using a neutral community model to estimate bacterial dispersal rates among members of each fitness group (Burns et al. 2016; Sloan et al. 2006), we detected consistent evidence for higher estimated rates of bacterial dispersal among high-fitness fish, suggesting stronger dispersal limitation among low-fitness fish (Figure 3 and Table S4). At the same time, we found inconsistent patterns regarding the fit of the neutral community model for explaining the distribution of ASVs within each fitness group. Neutral processes better explained structure of gut microbial communities for low-fitness fish in one pond, for high-fitness fish in another pond, whereas model fit was very similar in the third pond (Table S4).

4 | Discussion

In agreement with theoretical predictions and expectations from prior empirical studies, we found high-fitness fish, that is, the fastest growing fish, to have higher α-diversity and lower β-diversity compared with low-fitness fish which might be explained by differential bacterial interhost dispersal. Thus, our study represents an important step in establishing a connection between individual host fitness and gut microbiota composition in free-living vertebrates. We chose individual growth rate as a fitness proxy because of previous knowledge that this trait is positively correlated with survival and fecundity (Sparkes et al. 2013; Moser, Roesti, and Berner 2012; Thompson and Schluter 2022). We found strong evidence that variation in gut microbial communities is clearly associated with a fitness-related host trait in our system, suggesting that host–gut microbiota interactions could therefore alter evolutionary trajectories of their hosts (Henry et al. 2021).

Our results imply that higher α-diversity might be beneficial for the host, and similar results have been found in a small number of previous studies (Bestion et al. 2017; Stoffel et al. 2020). But how exactly higher α-diversity could promote host growth requires further investigation. No single bacterial ASV was detected among all members of the high-fitness group, and we found little evidence for differentially abundant bacterial phyla, which in combination indicates that faster growth is unlikely to be driven by a few beneficial taxa. Further, both fitness groups harbour a similarly diverse—and partially overlapping—pool of bacteria (γ-diversity) with equal proportions of unique taxa, illustrating that the less diverse bacterial communities of low-fitness fish do not merely represent a subset of the more diverse bacterial communities of high-fitness fish. However, high-fitness fish, on average, capture a larger proportion of the γ-diversity within each host individual. The higher α-diversity and larger proportion of γ-diversity captured within high-fitness fish could be driven by the higher bacterial dispersal that we detected among high-fitness fish (Figure 3 and Table S4). This is in line with theoretical predictions that α-diversity increases with dispersal rate (Mouquet and Loreau 2003).

Despite the observed higher bacterial dispersal rates among high-fitness fish, our analysis using a neutral community model (Burns et al. 2016; Sloan et al. 2006) did not detect consistent evidence that gut microbiota structure within each fitness group is differentially shaped by neutral (e.g., chance and dispersal) and non-neutral processes (e.g., microbe–microbe interactions and host selection) (Table S4). This neutral model predicts the distribution of microbial taxa across local communities (i.e., individual host organisms) and the metacommunity (i.e., a population of host organisms) by solely incorporating effects of random dispersal. While we consistently detected higher bacterial dispersal rates among high fitness fish (i.e., higher frequencies of occurrence of bacterial ASVs), the relative contributions of neutral and non-neutral processes to microbiome structure do not seem to fundamentally differ between fitness groups. The exact mechanisms by which higher interhost dispersal among high-fitness fish is achieved remain to be determined, but differences in host–host interactions and social behaviour might play a role (Burns et al. 2017; Martinez et al. 2015). Our results imply that a more diverse gut microbiota per se could be beneficial, for example, by allowing hosts to metabolise a wider range of nutrients and therefore exploit novel or broader trophic niches (Moeller and Sanders 2020).

We detected higher similarity in gut microbiota composition (i.e., lower β-diversity) among high-fitness fish (Figure 2C). This pattern might be produced by higher bacterial dispersal among high-fitness fish (Figure 3 and Table S4), thereby homogenising

microbial community composition as expected based on empirical (Burns et al. 2017) as well as theoretical studies (Mouquet and Loreau 2003). In accordance with the Anna Karenina principle, our results suggest that 'all microbiota conferring high host fitness are similar; each microbiota conferring low host fitness does so in its own way'. Evidence for the Anna Karenina principle has been mixed and appears to be context-dependent (Ahmed et al. 2019; Li et al. 2023; Ma 2020). For example, Anna Karenina principle effects were detected in approximately 50% of human microbiome associated diseases and the strength of effects appeared to differ across bacterial lineages with varying abundance (Ma 2020). Further, a meta-analysis across terrestrial and aquatic hosts found that Anna Karenina principle effects were more likely to be observed in hosts with low initial microbiome diversity (Li et al. 2023). Clearly, additional research is needed to determine the conditions under which it might be expected. Here, we found a pattern consistent with the Anna Karenina principle when investigating gut microbiota diversity in relation to a fitness-related host trait. To our knowledge, this is the first study to provide evidence for the Anna Karenina principle with regard to gut microbiota β-diversity effects on within-generation host fitness. Albeit speculative at this point, our results suggest that patterns consistent with the Anna Karenina principle may emerge in a broader range of contexts than previously appreciated (Zaneveld, McMinds, and Vega Thurber 2017), but more studies are needed to confirm whether this assumption is valid.

The high- and low-fitness groups harbour significantly dissimilar gut microbial communities. This pattern is consistent with observed gut microbiota differences related to host performance and fitness found between treatment groups in laboratory settings (e.g., Gould et al. 2018; Fontaine, Mineo, and Kohl 2022). For example, a study in tadpoles showed that experimental manipulation of the microbiota led to differences in thermal tolerance, strongly indicating that variation in microbial communities affects their hosts' survival (Fontaine, Mineo, and Kohl 2022). Similarly, composition and complexity of microbial communities affects a range of fitness-related traits (development, reproduction and lifespan) in fruit flies (Gould et al. 2018). While the results of these previous studies clearly show that microbiota variation can affect various aspects of host fitness, our work is so far only correlational. Thus, it remains to be tested whether gut microbiota variation can indeed cause variation in host fitness in stickleback fish. Further, the exact gut microbiota configurations that promote higher host fitness remain to be explored, but our study is in line with results from a small number of studies that demonstrated such associations in wild vertebrate populations (e.g., Worsley et al. 2021). Yet, results from wild populations can be confounded by a range of environmental variables that are often difficult to account for. To control for such unmeasured variation, we performed this experiment under naturalistic conditions. Crucially, the setting of our study standardised both the early-life diet and our analysis accounted for the age of fish while allowing fish to occupy and forage in different microhabitats; dietary patterns in our experimental ponds (inferred from variation in carbon and nitrogen isotope values; Figure S4 and Table S5) correspond to those observed in wild stickleback populations, suggesting biologically relevant niche use (Matthews et al. 2010). This is important because the effects of gut microbiota variation on host fitness can depend on environmental factors (especially diet) (Cooper, Vavra, and Cressler 2021), emphasising that the ecological context may strongly affect host-microbiota interactions. Lastly, since gut microbes have been shown to shape a range of host phenotypes and life history traits (Henry et al. 2021; Suzuki 2017; Sharon et al. 2013; Smith et al. 2017), future studies should explore whether variation in gut microbiota composition impacts other fitness-related host traits. Such work would be of particular importance since our analyses incorporated only one fitness proxy (growth rate), which limits our ability to draw solid conclusions about whether and how gut microbial communities may directly affect host fitness. Thus, we want to highlight that future studies should strive to test for effects of gut microbiota diversity on other potential fitness-related traits (e.g., metabolism and immune system function) or, ideally, direct fitness measures (survival and reproduction) to obtain a more comprehensive understanding of gut microbiota-mediated fitness effects. Finding consistent associations between the gut microbiota and various potential fitness-related host traits will be necessary to strengthen the case that microbial communities play a crucial role in determining the fitness of their animal hosts. We demonstrate that variation in gut bacterial communities is associated with a proxy for stickleback fitness, but it remains to be determined whether the observed variation is a cause or consequence of differences in host fitness. One potential opportunity to determine casualty would be through gut microbiota manipulation experiments, a method that has been established in several model systems and yielded groundbreaking results (Turnbaugh et al. 2006; Smith et al. 2017). Specifically, transferring gut microbial communities of low- and high-fitness fish to germ-free recipients could be a powerful approach to link gut microbiota variation directly to host fitness. Stickleback can be instrumental for this type of research since protocols for rearing germ-free stickleback larvae and producing gnotobiotic fish have previously been developed (Milligan-Myhre et al. 2016). Such work would further allow studying the mechanisms by which gut microbes affect host fitness in more detail, for example, by investigating host phenotypes such as immune system function or metabolism that could affect growth rate, body condition, and survival. Our work provides a foundation for future work aiming to establish causal relationships through gut microbiota manipulation (Smith et al. 2017) and we advocate that such studies be conducted in a diverse range of host organisms to determine the generality of host fitness-gut microbiota interactions.

Author Contributions

D.S. and K.A.T. designed and conducted the experiment. A.H., K.A.T. and D.J.R. developed the project idea. A.H. prepared the samples for sequencing, analysed the data, and wrote the manuscript with substantial input from all co-authors.

Acknowledgements

This work was supported by funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—project number 458274593 to A.H. and from the University of California San Diego to D.J.R. K.A.T. was supported by a Human Frontier Science Program Long-Term Fellowship. D.S. was supported by the Natural Sciences and Engineering Research Council of Canada. We would like to thank D. Rubenstein for alerting us to the Anna Karenina principle.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The raw sequencing reads (https://doi.org/10.6084/m9.figshare.22802993.v1) and all data files and R scripts (https://doi.org/10.6084/m9.figshare.23713482.v1) have been deposited on figshare.

References

Ahmed, H. I., M. Herrera, Y. J. Liew, and M. Aranda. 2019. "Long-Term Temperature Stress in the Coral Model Aiptasia Supports the "Anna Karenina Principle" for Bacterial Microbiomes." *Frontiers in Microbiology* 10: 975.

Albright, M. B. N., and J. B. H. Martiny. 2018. "Dispersal Alters Bacterial Diversity and Composition in a Natural Community." *ISME Journal* 12, no. 1: 296–299.

Anderson, M. J. 2001. "A New Method for Non-Parametric Multivariate Analysis of Variance." *Australian Journal of Ecology* 26, no. 1: 32–46.

Arnegard, M. E., M. D. McGee, B. Matthews, et al. 2014. "Genetics of Ecological Divergence During Speciation." *Nature* 511, no. 7509: 307–311.

Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software* 67, no. 1: 1–48.

Bates, K. A., U. Sommer, K. P. Hopkins, et al. 2022. "Microbiome Function Predicts Amphibian Chytridiomycosis Disease Dynamics." *Microbiome* 10, no. 1: 44.

Bestion, E., S. Jacob, L. Zinger, et al. 2017. "Climate Warming Reduces Gut Microbiota Diversity in a Vertebrate Ectotherm." *Nature Ecology & Evolution* 1: 0161.

Beule, L., and P. Karlovsky. 2020. "Improved Normalization of Species Count Data in Ecology by Scaling With Ranked Subsampling (SRS): Application to Microbial Communities." *PeerJ* 8: e9593.

Bolnick, D. I., and K. M. Ballare. 2020. "Resource Diversity Promotes Among-Individual Diet Variation, but Not Genomic Diversity, in Lake Stickleback." *Ecology Letters* 23, no. 3: 495–505.

Bolnick, D. I., L. K. Snowberg, P. E. Hirsch, et al. 2014. "Individuals' Diet Diversity Influences Gut Microbial Diversity in Two Freshwater Fish (Threespine Stickleback and Eurasian Perch)." *Ecology Letters* 17, no. 8: 979–987.

Bolyen, E., J. R. Rideout, M. R. Dillon, et al. 2019. "Reproducible, Interactive, Scalable and Extensible Microbiome Data Science Using QIIME 2." *Nature Biotechnology* 37, no. 8: 852–857.

Burns, A. R., E. Miller, M. Agarwal, et al. 2017. "Interhost Dispersal Alters Microbiome Assembly and Can Overwhelm Host Innate Immunity in an Experimental Zebrafish Model." *Proceedings of the National Academy of Sciences of the United States of America* 114, no. 42: 11181–11186.

Burns, A. R., W. Z. Stephens, K. Stagaman, et al. 2016. "Contribution of Neutral Processes to the Assembly of Gut Microbial Communities in the Zebrafish Over Host Development." *ISME Journal* 10, no. 3: 655–664.

Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, and S. P. Holmes. 2016. "DADA2: High-Resolution Sample Inference From Illumina Amplicon Data." *Nature Methods* 13, no. 7: 581–583.

Cooper, R. O., J. M. Vavra, and C. E. Cressler. 2021. "Targeted Manipulation of Abundant and Rare Taxa in the Daphnia Magna Microbiota With Antibiotics Impacts Host Fitness Differentially." *Msystems* 6, no. 2: e0091620.

Davidson, G. L., S. E. Somers, N. Wiley, et al. 2021. "A Time-Lagged Association Between the Gut Microbiome, Nestling Weight and Nestling Survival in Wild Great Tits." *Journal of Animal Ecology* 90, no. 4: 989–1003.

Fontaine, S. S., P. M. Mineo, and K. D. Kohl. 2022. "Experimental Manipulation of Microbiota Reduces Host Thermal Tolerance and Fitness Under Heat Stress in a Vertebrate Ectotherm." *Nature Ecology & Evolution* 6, no. 4: 405–417.

Goertz, S., A. B. de Menezes, R. J. Birtles, et al. 2019. "Geographical Location Influences the Composition of the Gut Microbiota in Wild House Mice (Mus Musculus Domesticus) at a Fine Spatial Scale." *PLoS One* 14, no. 9: e0222501.

Gould, A. L., V. Zhang, L. Lamberti, et al. 2018. "Microbiome Interactions Shape Host Fitness." *Proceedings of the National Academy of Sciences of the United States of America* 115, no. 51: E11951–E11960.

Härer, A., and D. J. Rennison. 2024. "Gut Microbiota Uniqueness Is Associated With Lake Size – A Proxy for Diet Diversity – In Stickleback Fish." *American Naturalist* 203, no. 2: 284–291.

Henry, L. P., M. Bruijning, S. K. G. Forsberg, and J. F. Ayroles. 2021. "The Microbiome Extends Host Evolutionary Potential." *Nature Communications* 12, no. 1: 5141.

Houwenhuyse, S., R. Stoks, S. Mukherjee, and E. Decaestecker. 2021. "Locally Adapted Gut Microbiomes Mediate Host Stress Tolerance." *ISME Journal* 15, no. 8: 2401–2414.

Kolodny, O., and H. Schulenburg. 2020. "Microbiome-Mediated Plasticity Directs Host Evolution Along Several Distinct Time Scales." *Philosophical Transactions of the Royal Society B* 375: 20190589.

Kriss, M., K. Z. Hazleton, N. M. Nusbacher, C. G. Martin, and C. A. Lozupone. 2018. "Low Diversity Gut Microbiota Dysbiosis: Drivers, Functional Implications and Recovery." *Current Opinion in Microbiology* 44: 34–40.

Kuznetsova, A., P. B. Brockhoff, and H. B. Christensen. 2017. "lmerTest Package: Tests in Linear Mixed Effects Models." *Journal of Statistical Software* 82, no. 13: 1–26.

Lathrop, S. K., S. M. Bloom, S. M. Rao, et al. 2011. "Peripheral Education of the Immune System by Colonic Commensal Microbiota." *Nature* 478, no. 7368: 250–255.

Leibold, M. A., N. Holyoak, N. Mouquet, et al. 2004. "The Metacommunity Concept: A Framework for Multi-Scale Community Ecology." *Ecology Letters* 7, no. 7: 601–613.

Li, J., K. A. Bates, K. L. Hoang, T. E. Hector, S. C. L. Knowles, and K. C. King. 2023. "Experimental Temperatures Shape Host Microbiome Diversity and Composition." *Global Change Biology* 29, no. 1: 41–56.

Lozupone, C., and R. Knight. 2005. "UniFrac: A New Phylogenetic Method for Comparing Microbial Communities." *Applied and Environmental Microbiology* 71, no. 12: 8228–8235.

Lozupone, C., M. E. Lladser, D. Knights, J. Stombaugh, and R. Knight. 2011. "UniFrac: An Effective Distance Metric for Microbial Community Comparison." *ISME Journal* 5, no. 2: 169–172.

Ma, Z. S. 2020. "Testing the Anna Karenina Principle in Human Microbiome-Associated Diseases." *iScience* 23, no. 4: 101007.

Martinez, I., J. C. Stegen, M. X. Maldonado-Gomez, et al. 2015. "The Gut Microbiota of Rural Papua New Guineans: Composition, Diversity Patterns, and Ecological Processes." *Cell Reports* 11, no. 4: 527–538.

Matthews, B., K. B. Marchinko, D. I. Bolnick, and A. Mazumder. 2010. "Specialization of Trophic Position and Habitat Use by Sticklebacks in an Adaptive Radiation." *Ecology* 91, no. 4: 1025–1034.

Milligan-Myhre, K., C. M. Small, E. K. Mittge, et al. 2016. "Innate Immune Responses to Gut Microbiota Differ Between Oceanic and Freshwater Threespine Stickleback Populations." *Disease Models & Mechanisms* 9, no. 2: 187–198.

Moeller, A. H., and J. G. Sanders. 2020. "Roles of the Gut Microbiota in the Adaptive Evolution of Mammalian Species." *Philosophical Transactions of the Royal Society B* no. 375: 20190597.

Moser, D., M. Roesti, and D. Berner. 2012. "Repeated Lake-Stream Divergence in Stickleback Life History Within a Central European Lake Basin." *PLoS One* 7, no. 12: e50620.

Mouquet, N., and M. Loreau. 2003. "Community Patterns in Source-Sink Metacommunities." *American Naturalist* 162, no. 5: 544–557.

Oksanen, J., G. F. Blanchet, M. Friendly, et al. 2019. "Vegan: Community Ecology Package." R Package Version 2.5-6.

Peichel, C. L., J. A. Ross, C. K. Matson, et al. 2004. "The Master Sex-Determination Locus in Threespine Sticklebacks Is on a Nascent Y Chromosome." *Current Biology* 14, no. 16: 1416–1424.

Price, M. N., P. S. Dehal, and A. P. Arkin. 2010. "FastTree 2-Approximately Maximum-Likelihood Trees for Large Alignments." *PLoS One* 5, no. 3: e9490.

Quast, C., E. Pruesse, P. Yilmaz, et al. 2013. "The SILVA Ribosomal RNA Gene Database Project: Improved Data Processing and Web-Based Tools." *Nucleic Acids Research* 41, no. D1: 590–596.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rennison, D. J., S. M. Rudman, and D. Schluter. 2019. "Parallel Changes in Gut Microbiome Composition and Function During Colonization, Local Adaptation and Ecological Speciation." *Proceedings of the Royal Society B: Biological Sciences* no. 286: 20191911.

Risely, A., N. Muller-Klein, D. W. Schmid, et al. 2023. "Climate Change Drives Loss of Bacterial Gut Mutualists at the Expense of Host Survival in Wild Meerkats." *Global Change Biology* 29, no. 20: 5816–5828.

Sharon, G., D. Segal, J. M. Ringo, A. Hefetz, I. Zilber-Rosenberg, and E. Rosenberg. 2013. "Commensal Bacteria Play a Role in Mating Preference of *Drosophila melanogaster*." *Proceedings of the National Academy of Sciences of the United States of America* 107, no. 46: 20051–20056.

Sison-Mangus, M. P., A. A. Mushegian, and D. Ebert. 2015. "Water Fleas Require Microbiota for Survival, Growth and Reproduction." *ISME Journal* 9, no. 1: 59–67.

Sloan, W. T., M. Lunn, S. Woodcock, I. M. Head, S. Nee, and T. P. Curtis. 2006. "Quantifying the Roles of Immigration and Chance in Shaping Prokaryote Community Structure." *Environmental Microbiology* 8, no. 4: 732–740.

Smith, C. C. R., L. K. Snowberg, J. G. Caporaso, R. Knight, and D. I. Bolnick. 2015. "Dietary Input of Microbes and Host Genetic Variation Shape Among-Population Differences in Stickleback Gut Microbiota." *ISME Journal* 9, no. 11: 2515–2526.

Smith, P., D. Willemsen, M. Popkes, et al. 2017. "Regulation of Life Span by the Gut Microbiota in the Short-Lived African Turquoise Killifish." *eLife* 6: 6.

Sparkes, T. C., V. Rush, D. A. Kopp, and S. A. Foster. 2013. "Reproductive Success in a Natural Population of Male Three-Spined Stickleback Gasterosteus Aculeatus: Effects of Nuptial Colour, Parasites and Body Size." *Journal of Fish Biology* 82, no. 5: 1720–1727.

Stoffel, M. A., K. Acevedo-Whitehouse, N. Morales-Duran, et al. 2020. "Early Sexual Dimorphism in the Developing Gut Microbiome of Northern Elephant Seals." *Molecular Ecology* 29, no. 11: 2109–2122.

Suzuki, T. A. 2017. "Links Between Natural Variation in the Microbiome and Host Fitness in Wild Mammals." *Integrative and Comparative Biology* 57, no. 4: 756–769.

Thompson, K. A., and D. Schluter. 2022. "Heterosis Counteracts Hybrid Breakdown to Forestall Speciation by Parallel Natural Selection." *Proceedings of the Royal Society B: Biological Sciences* no. 289: 20220422.

Thompson, L. R., J. G. Sanders, D. McDonald, et al. 2017. "A Communal Catalogue Reveals Earth's Multiscale Microbial Diversity." *Nature* 551, no. 7681: 457–463.

Trivedi, P., J. E. Leach, S. G. Tringe, T. Sa, and B. K. Singh. 2020. "Plant-Microbiome Interactions: From Community Assembly to Plant Health." *Nature Reviews Microbiology* 18, no. 11: 607–621.

Tung, J., L. B. Barreiro, M. B. Burns, et al. 2015. "Social Networks Predict Gut Microbiome Composition in Wild Baboons." *eLife* 4: e27014.

Turnbaugh, P. J., R. E. Ley, M. A. Mahowald, V. Magrini, E. R. Mardis, and J. I. Gordon. 2006. "An Obesity-Associated Gut Microbiome With Increased Capacity for Energy Harvest." *Nature* 444, no. 7122: 1027–1031.

Vuong, H. E., J. M. Yano, T. C. Fung, and E. Y. Hsiao. 2017. "The Microbiome and Host Behavior." *Annual Review of Neuroscience* 40: 21–49.

Weiland-Brauer, N., N. Pinnow, D. Langfeldt, et al. 2020. "The Native Microbiome Is Crucial for Offspring Generation and Fitness of Aurelia Aurita." *MBio* 11, no. 6: e0233620.

Worsley, S. F., C. S. Davies, M. E. Mannarelli, et al. 2021. "Gut Microbiome Composition, Not Alpha Diversity, Is Associated With Survival in a Natural Vertebrate Population." *Animal Microbiome* 3, no. 1: 84.

Youngblut, N. D., G. H. Reischer, W. Walters, et al. 2019. "Host Diet and Evolutionary History Explain Different Aspects of Gut Microbiome Diversity Among Vertebrate Clades." *Nature Communications* 10: 2200.

Zaneveld, J. R., R. McMinds, and R. Vega Thurber. 2017. "Stress and Stability: Applying the Anna Karenina Principle to Animal Microbiomes." *Nature Microbiology* 2: 17121.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.