nature microbiology

ARTICLES

https://doi.org/10.1038/s41564-019-0560-0

Core gut microbial communities are maintained by beneficial interactions and strain variability in fish

Fotini Kokou[1,2], Goor Sasson[1], Jonathan Friedman[3], Stav Eyal[1], Ofer Ovadia[1], Sheenan Harpaz[2], Avner Cnaani[2] and Itzhak Mizrahi[1*]

The term core microbiome describes microbes that are consistently present in a particular habitat. If the conditions in that habitat are highly variable, core microbes may also be considered to be ecological generalists. However, little is known about whether metabolic competition and microbial interactions influence the ability of some microbes to persist in the core microbiome while others cannot. We investigated microbial communities at three sites in the guts of European seabass under four dietary conditions. We identified generalist core microbial populations in each gut site that are shared across fish, present under multiple diets and persistent over time. We found that core microbes tend to show synergistic growth in co-culture, and low levels of predicted and validated metabolic competition. Within core microbial species, we found high levels of intraspecific variability and strain-specific habitat specialization. Thus, both intraspecific variability and interspecific facilitation may contribute to the ecological stability of the animal core microbiome.

Microbial communities that reside in the vertebrate gastrointestinal tract are tightly connected to many of their host's traits[1–3]. Nevertheless, we are far from understanding the ecological forces that structure these communities, and we therefore need to infer and identify the principles and patterns that shape the microbiome[4]. Stochastic and deterministic processes are thought to shape community composition. Environmental factors, such as pH, host immune system and dietary composition, are considered to be dominant in shaping gut microbial communities, a process also termed environmental selection[5–7]. Furthermore, interspecies interactions—mutualistic or competitive—may further affect the composition of microbial communities and drive their distribution patterns[8,9]. Naturally, these selective forces not only act at the community level but are also bound to affect the individual members of such communities, triggering evolutionary processes such as adaptation and strain variability[10,11]. It is therefore of interest to understand how these ecological forces affect individual species within the community and whether they have equal effects on different community members, such as generalist and specialist species.

From an ecological viewpoint, generalist species can be defined as taxa that inhabit different environments or environmental gradients; within these generalist species are members of the core microbiome—taxa with high occupancy that persist across multiple assemblages associated with a habitat (such as the gut[12]). Thus, the concept of the core microbiome considers persistent and sometimes highly abundant microbes in a microbial community that are also considered to be stable communities[12]. Although research during the past few years has focused on understanding the role of these generalist microbiome members[12–14], their persistence across multiple hosts and habitats is not yet understood. Microbial competition has recently been suggested to drive the stability of microbial communities in the ecosystem[15]; however, mutualism or skewed interspecies interactions (amensalism or commensalism) have been observed in ecology to promote diversity and community stability[13,16–18]. Moreover, alternative ecological theories, suggesting that generalist species are assemblages of more specialized individuals[11,19] through increased variation among individuals, can also explain species persistence[20]. Here, using the fish gut as a model (European seabass), we aimed to identify such generalist core microbial populations that are widespread across multiple habitats and environmental conditions, and to elucidate the mechanisms that enable their patterns of persistence, such as the forces that stabilize them as populations and communities. Our findings suggest that the persistence and coexistence of these core microbes are maintained through low competition and synergistic interactions, as well as intraspecies strain variability.

Results

Habitat filtering by gut part, and not diet, is the major driver shaping fish gut microbial communities. To identify microbial generalists, we first defined habitats/environments and their microbial inhabitants. We analysed the members of the seabass gut microbiome in light of two potential habitat-filtering forces: diet and gut part (study design is described in the Methods and Supplementary Table 1). Diet is known to affect microbiome composition and to act as a habitat filter[21]. The common gut of a bony carnivorous fish, of which European seabass is a good representative, is composed of three main parts: the pyloric caeca, which are finger-like extensions located in the proximal part of the gut; the midgut, which is the main part of the gut; and the hindgut[22] (Fig. 1a). As these parts are connected across the gastrointestinal tube, all members of the microbiome could inhabit them equally on the assumption that only random forces were at play. However, if the parts provide different conditions, instilling habitat filtering, we would expect to see different community compositions in each of the parts[23–26]. On the basis of this premise, we compared the two selective forces and their relative effects on microbiome composition. We sampled the three different gut parts—pyloric caeca,

[1]Department of Life Sciences, Ben-Gurion University of the Negev and the National Institute for Biotechnology in the Negev, Marcus Family Campus, Be'er-Sheva, Israel. [2]Department of Poultry and Aquaculture, Institute of Animal Sciences, Agricultural Research Organization, Rishon LeZion, Israel. [3]Institute of Plant Sciences and Genetics in Agriculture, Faculty of Agriculture, Food and Environment, The Hebrew University of Jerusalem, Rehovot, Israel.
*e-mail: imizrahi@bgu.ac.il

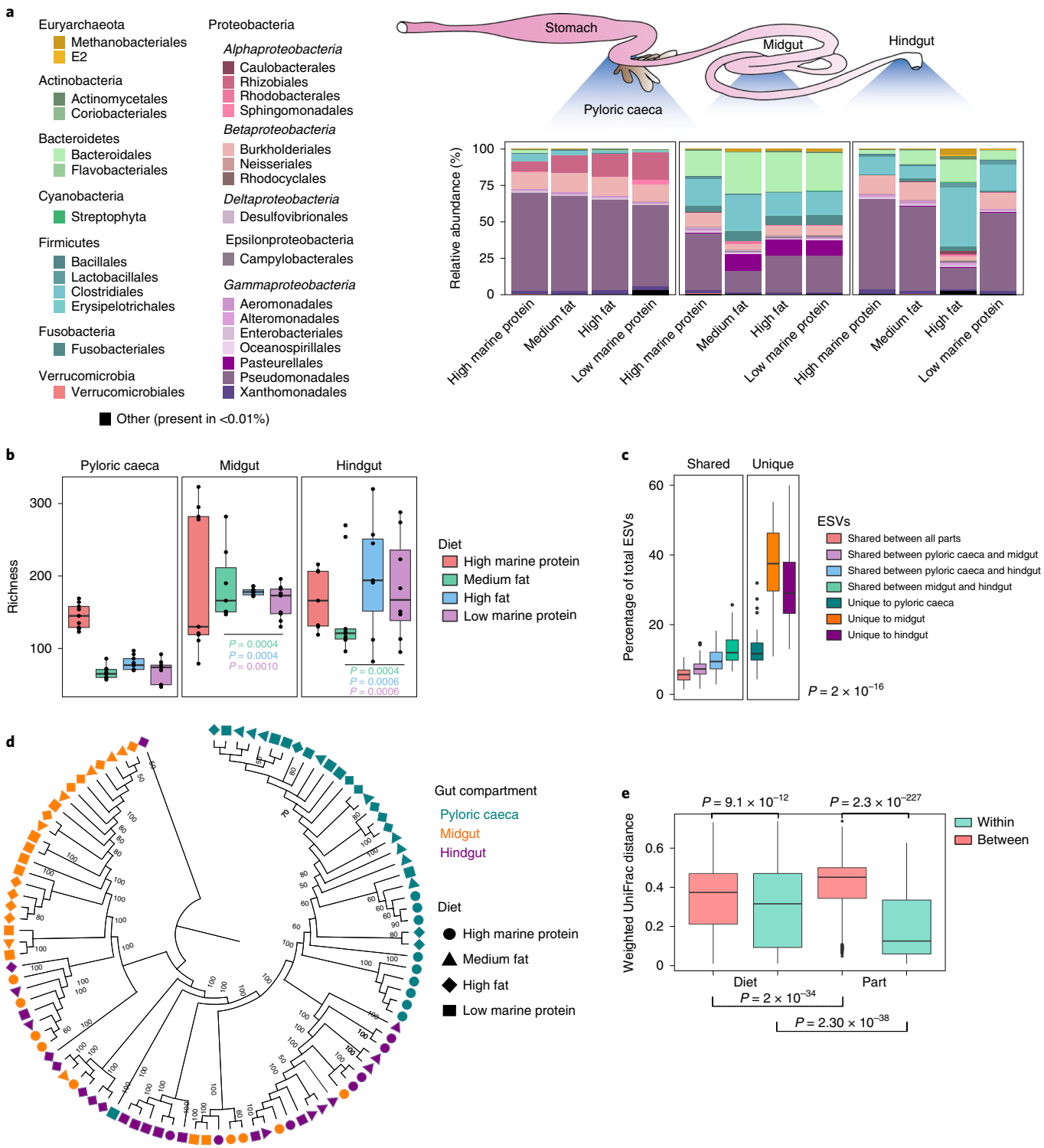

Fig. 1 | Habitat filtering by gut part shapes seabass gut microbial communities. a, The relative abundance of the microbial communities at the order level found in different gut compartments (pyloric caeca, midgut and hindgut) of fish that were fed different diets. **b**, Microbial richness within the different gut compartments of fish fed different diets. *P*-values indicate significant difference from the pyloric caeca (paired two-sided Wilcoxon *t*-test, 95% CI). Data are shown as box plots ($n = 9$ fish individuals); the horizontal line indicates the median and the whiskers indicate the lowest and highest points within 1.5× the interquartile ranges of the lower or upper quartile, respectively. *P* values with different colours denote the values for each diet. **c**, Shared and unique ESVs between the different gut compartments (two-sided Wilcoxon *t*-test, 95% CI). Box plots ($n = 36$ fish individuals) describe the data as in **b**. **d**, Hierarchical clustering dendrogram with jackknife support (numbers on the branches; only values above 50 are shown in the tree) using weighted UniFrac as a metric to compare similarity between different gut compartments. Shapes indicate different diets and colours indicate the different fish gut parts. **e**, Weighted UniFrac distance within and between samples originating from different gut parts or diets (two-sided Wilcoxon *t*-test, 95% CI; for within versus between gut parts, $P < 0.0001$). Box plots ($n = 36$ fish individuals) describe the data as in **b**.

midgut and hindgut—of 36 European seabass individuals that were fed different diets ($n = 9$ per diet).

Amplicon sequencing of the microbial communities across the gut of the seabass revealed significant changes in composition between gut parts (Supplementary Table 2a–d). The gut part had a stronger effect on shaping microbial communities than the diet, as we observed more similarity between the microbial composition within gut parts, regardless of diet (Fig. 1a). Moreover, we found only a small proportion of shared species between the gut parts, although these parts are connected, and a high degree of shared species might be expected (Fig. 1c; Wilcoxon t-test, $P = 2 \times 10^{-16}$). Among the three gut parts, the pyloric caeca showed significantly lower species richness (paired Wilcoxon t-test, $P < 0.05$; Fig. 1b, Supplementary Fig. 1) and diversity (Shannon index and phylogenetic diversity; Supplementary Fig. 2a,b). Furthermore, there was a gradual increase in the individual variation across the gut, with pyloric caecal communities exhibiting the lowest variability in richness within treatments (Fig. 1b; linear trend for the coefficient of variance, $r^2 = 0.37$, $P = 0.04$). This may suggest that physiological conditions along the gut, such as a lower pH measured in the pyloric caeca compared with the other gut parts, can act as a selection force or habitat filter (Supplementary Fig. 2c) that potentially constrains microbial communities and decreases the inter-individual variation. Indeed, when we analysed our data using clustering analysis of beta diversity with the jackknife approach, we found strong support for microbiome clustering by gut part (Fig. 1d, Supplementary Table 2). Specifically, we observed dominance of gut part over diet in determining microbiome composition, which was shown by clustering of the communities coming from different diets, primarily by gut part (weighted UniFrac, Fig. 1d, Supplementary Fig. 3; permutational analysis of variance (PERMANOVA) analysis, Supplementary Table 2). Moreover, we found a stronger impact of the gut part compared with the diet on the beta-diversity variation of samples (Fig. 1e). These findings further strengthen the idea that the conditions along the gut serve as strong environmental filters, enabling the establishment of distinct microbial communities in the different parts.

Facilitation and positive interactions are prevalent among identified generalist core microbes. After identifying the effects of different environmental conditions—diet and gut part—on the microbial communities, we characterized the niche width of individual microbes to classify them as generalists or specialists. We defined generalists as microbes that inhabit a wide range of environments along the different samples, such as different diets and gut parts, and specialists as microbes that have a narrower range of occupancy across these different environments. We therefore measured the niche width as previously described[11,27] for each microbe using three different methods: the Shannon diversity index, the occupancy across habitats (richness) and the Levin's measure of niche breadth[28] (see the 'Niche width' section in the Methods). Taken together, these methods enabled us to assess the niche-width distribution across the microbiomes, looking at each dietary regime and gut part as a potential habitat. In this analysis, microbes with the highest niche width were defined as generalists that occupy multiple habitats, whereas microbes with the lowest niche width were defined as specialists that occupy a specific and limited number of habitats. We therefore define generalist microbes as those that are shared between the upper 10th percentile of the distribution of all of the calculated niche-width methods (Fig. 2a, Supplementary Fig. 4). We found 8 such generalist species that fell into this definition (8 clusters of 97% similarity originating from 12 exact sequence variants (ESVs) identified as core; Fig. 2b). In contrast to most microbes, these generalist microbes were found to be present in more than 90% of the different gut parts and diets, therefore falling into the core microbe definition. Although these generalist microbes contributed almost 60% of the total relative abundance (Fig. 2c), they comprised less than 0.5% of the overall richness. The distribution of these microbes clearly varied in our system (Supplementary Fig. 5), and was affected by gut part, dietary treatment or both (Supplementary Fig. 6).

We then examined the persistence of these seabass core gut microbes over time. We conducted a time-series experiment in which we sampled the three gut parts across different diets. The rationale for this experiment stemmed from the possibility that some gut microbes are transient residents of gut ecosystems, as previously reported[29]. Such transient taxa, which are usually rare members of the microbiome, can occasionally bloom and, therefore, could be misleadingly counted as stable residents of the microbiome[29]. The distribution of transient microbiome residents over time is expected to follow a bimodal distribution, whereas stable residents will show a uniform distribution. Taking this into consideration, we implemented a previously described method for detecting such patterns of transient or resident taxa in the microbiome time-series data[29] (see the 'Transient taxa' section in the Methods). This analysis, which was performed across time, diets and gut parts, revealed that the identified core microbes are stable residents of the seabass gut microbiome, as were most of the gut microbes in this study (Supplementary Table 3). Interestingly, when we examined other datasets from different studies (Supplementary Table 15), these microbes, at the genus level annotation, could be found in many gut systems of other fish species (Fig. 2d). Thus, although we did not select these core generalist microbes using such criteria, they show temporal stability and presence in other fish species, suggesting strong association with the fish gut, and therefore rendering them good candidates to study the core microbiome.

Notably, when we examined the correlations and co-occurrence patterns between these core microbes (Fig. 3a), we found mainly positive interactions. These results contrasted with the mutual-exclusion patterns observed for other microbiome members in this analysis and suggested that the coexistence of these microbes is driven by a non-competitive relationship. Indeed, evaluation of the metabolic competitive potential of these species, by their metabolically overlapping pathways using NetCmpt[30], showed low levels of competition (Fig. 3b; see the 'Shared and unique microbiomes' section in the Methods). These findings suggested that niche partitioning contributes to the persistence of these core microbes by enabling low levels of competition and potentially positive interactions between them. To test this hypothesis, we isolated and obtained all of the core microbes and determined their interactions using two approaches: minimal media containing various carbon sources and media containing in vitro digested feed to simulate gut conditions (Supplementary Fig. 7; see the 'In vitro interactions among the core microbes' section in the Methods). This provided a further understanding of the potential interactions between core microbes, considering interference competition and other types of interaction that can occur between microbes[31] but are overlooked by the co-occurrence networks and the NetCmpt tool. Following these analyses, we could not detect any interference interactions on the minimal media containing five carbon sources (Supplementary Tables 4–9). When we compared the utilization of various substrates by the core microbes (see the 'Substrate utilization' section in the Methods), we observed that core microbes differ in their preferences for the examined substrates (Supplementary Fig. 9). This further suggested that these core microbes partition their niches, thereby reducing interspecies competition. We then investigated their interactions in simulated gut conditions using the fish's actual in vitro stomach-digested feed to better understand coexistence under natural conditions (see the 'In vitro interactions among the core microbes' section in the Methods; Supplementary Fig. 7). In these experiments, we found either positive or low-level competitive interactions among the core microbes, with a significantly higher degree of positive interactions (Fig. 3c–e). We also noted

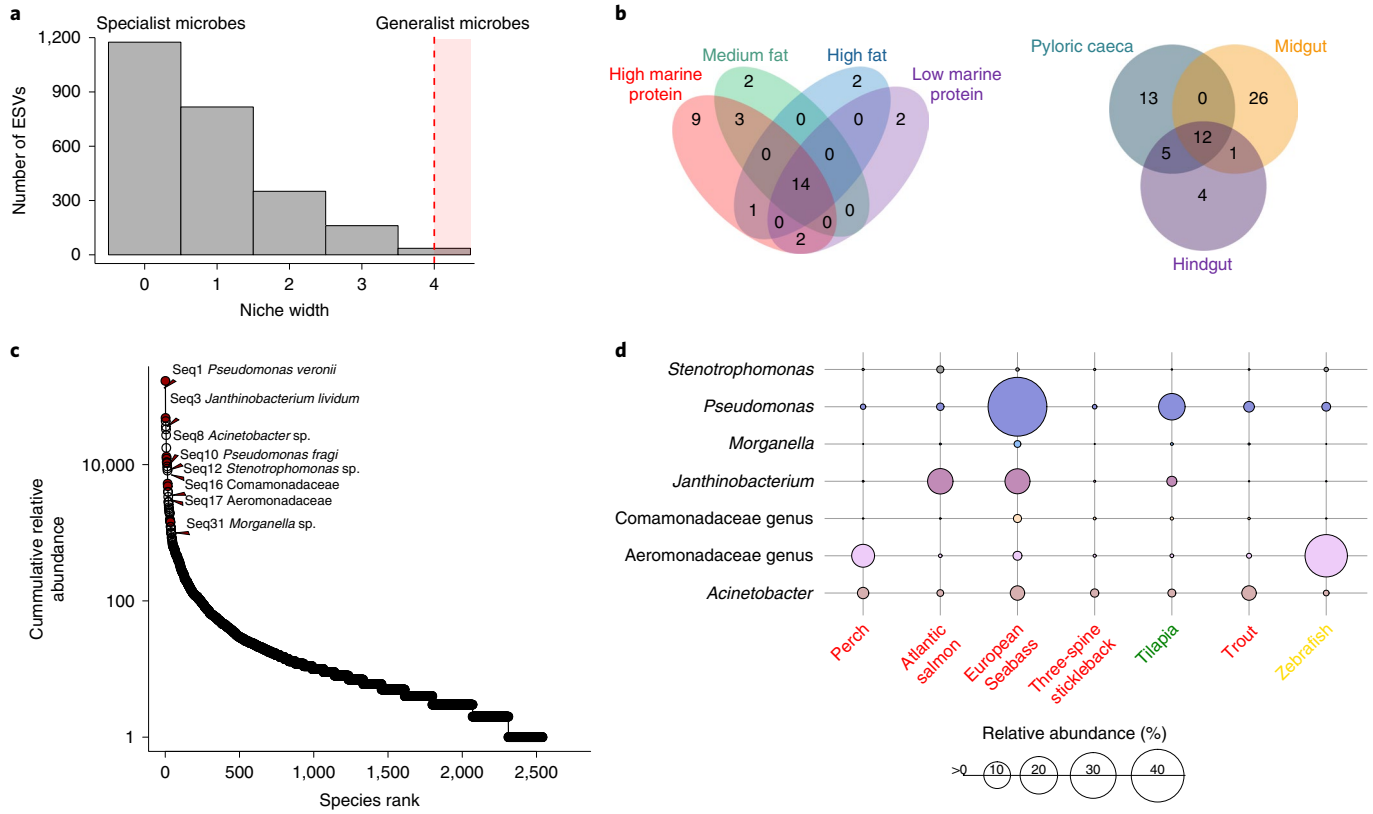

Fig. 2 | A core microbial community composed of 8 abundant generalist species persists across different habitats of diets and gut parts, and across different species of fish. a, Niche width of the overall microbial communities calculated using Shannon index *H*′. Generalists are defined as microbes with the largest niche width (top 10th percentile), and specialists as microbes with the smallest niche width (bottom 10th percentile). **b**, Venn diagram showing shared and unique ESVs between dietary treatments and gut parts (habitats), present in more than 90% of the samples for each specific habitat (see the 'Shared and unique microbiomes' section in the Methods). At the species level, 8 microbes that were shared among all diets and gut parts and fell into the generalist niche width were defined as core microbes (4 out of 12 ESVs belonged to the same species at 97% sequence identity). **c**, Rank-relative abundance plot of the overall microbial communities on a log scale. Red circles indicate the core microbes. **d**, Balloon plot showing the abundance of the core generalist communities across different fish species at the genus level. Red, green and yellow text indicates carnivorous fish, herbivores and omnivores, respectively. Data are from the studies cited in Supplementary Table 15.

that, in most cases, these pairwise interactions were synergistic; the total cell number in the co-cultures exceeded the expected cell number in the corresponding individual cultures. This was concluded by analysing the total 16S rRNA gene copy number in the co-cultures compared with the sum of 16S rRNA gene copy numbers in each individual culture (Fig. 3e, Supplementary Fig. 8), further supporting facilitation of the coexistence of these core species by positive interactions (Fig. 3c–e). Taken together, these findings from both comparative metabolic potential and co-culturing experiments were in agreement and indicated that positive and low-level competitive interactions support stabilization and coexistence of these core microbes, potentially through resource partitioning.

Generalist core species show higher strain variability congruent with their habitat preferences. According to ecological theories, such as the niche variation hypothesis proposed by Van Valen, we would expect generalists to tend towards more variability to reduce intraspecific competition[11,20], especially in depauperate environments. Inspired by this, we examined the intraspecies variation of the clusters that these species belong to and compared the variability of the core microbes to that of other members of the microbiome. To this end, we clustered all of our ESVs into species-level clusters (97% sequence similarity) and measured the strain variability within each species cluster across our dataset. We found that core species have higher strain variability compared with non-core species (Fig. 4a).

We next examined whether the different strains of the core species show a tendency to inhabit one habitat over another, which would indicate their preference for a specific habitat. We analysed the presence–absence patterns of these strains across the different gut parts using the Jaccard similarity metric. Our analysis revealed significant clustering of strains of all generalist species according to the gut part (PERMANOVA R-statistic = 0.10339, P = 0.001, 95% confidence interval (CI); Fig. 4b, Supplementary Fig. 13). Moreover, when examining the richness of core strains across parts, we observed an opposite pattern compared with the overall richness of the microbiome (Figs. 1b and 4c). This phenomenon occurred for most of the core and some of the highly prevalent microbes (Fig. 4c, Supplementary Fig. 10, Supplementary Table 16, Supplementary Data for Supplementary Table 16). These findings show that generalist microbes have higher strain variability with habitat specification.

We then examined whether the habitat tendency patterns are the result of radiation of these strains from an ancestor that carried this trait or whether this trait developed in parallel in different strains, regardless of their phylogenetic similarity. We created a phylogenetic tree for all of the core species strain clusters. Our results did not show agreement between habitat tendency and phylogeny (Supplementary Fig. 14). Moreover, we found that the strains were less similar within a fish than between fish (Fig. 4d), suggesting that this preference is an outcome of multiple evolutionary trajectories that converged into this phenotype; nevertheless, many other

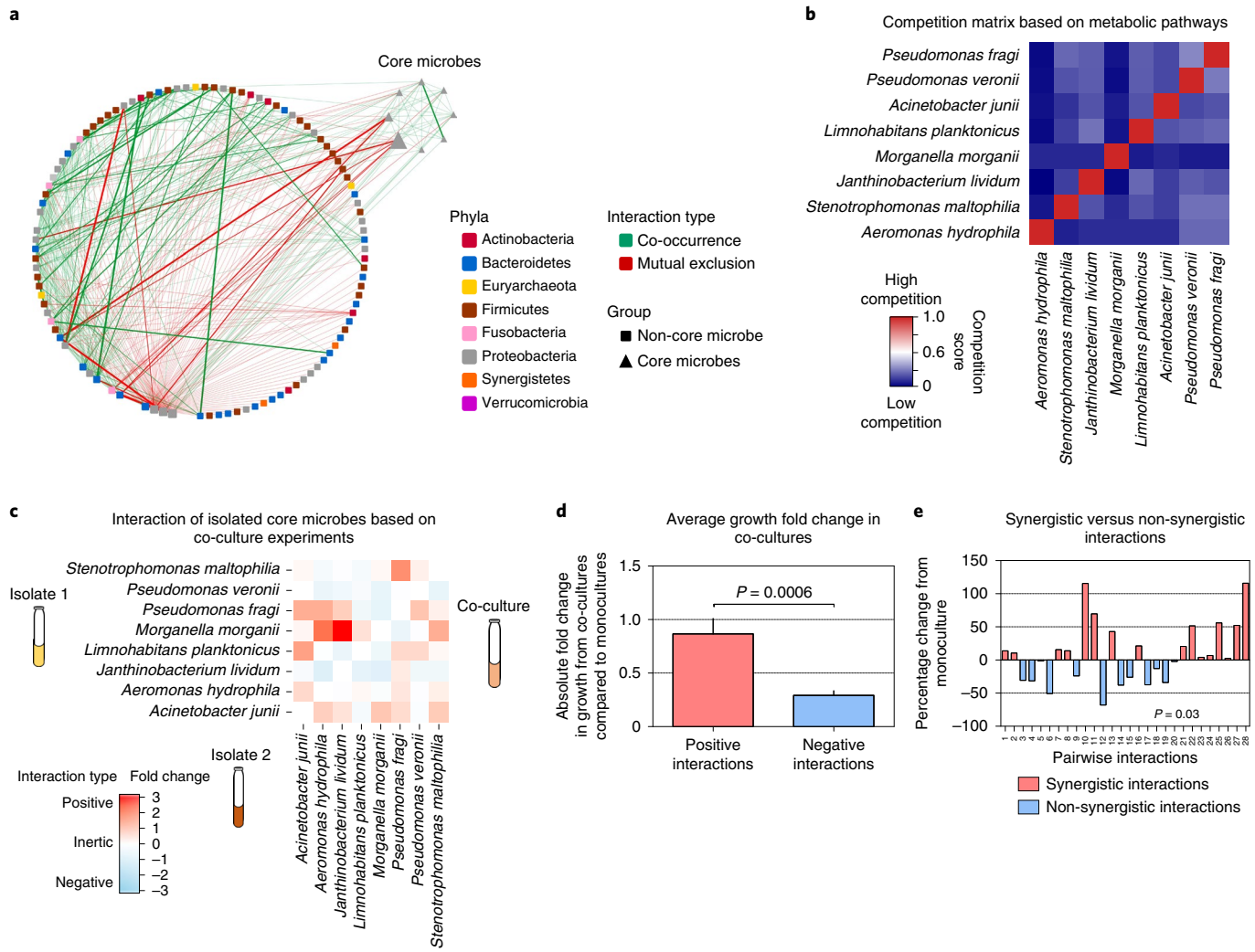

Fig. 3 | Facilitation and positive interactions are prevalent among identified generalist core microbes. a, Network of co-occurring ESVs based on correlation analysis. Core microbes have strong positive correlations among them. A connection stands for a strong (1,000 most extreme values—negative or positive—based on of five methods of ensemble reference: Spearman, Pearson, mutual information, Bray–Curtis and Kullback–Leibler dissimilarity; default settings) and significant correlation ($q < 0.05$, after Benjamini–Hochberg correction; bootstrap, $n = 100$; permutations, $n = 100$). The size of each node is proportional to the relative abundance of the ESV. The edge thickness corresponds to the statistical significance (the *P* value) of the correlation—the thicker the edge, the lower the *P* value—whereas the edge transparency denotes the strength of the correlation (*R*). Green edges represent co-occurrence between two ESVs, red edges represent mutually excluded ESVs. Nodes are coloured by phylum taxonomy. Nodes of core microbes are indicated as triangles. **b**, Competition matrix based on the analysis using NetCmpt (0, no competition; 1, competitive interactions). The matrix is not symmetrical as the pairwise interactions between the ESVs may differentially affect each of the microbes. **c**, Pairwise co-cultivation of the core microbes in digested in vitro feed extract for 24 h (experimental design is provided in Supplementary Fig. 7). Colour intensity indicates fold-change increase (positive interaction; red) or decrease (negative interaction; blue) relative to single-microbe growth (monoculture). **d**, Box plots showing the absolute fold change in 16S rRNA gene copies from co-cultures in comparison to the single cultures. Positive interactions originating from the co-cultures are shown in red and negative interactions are shown in blue. Statistical analysis was performed using two-sided Wilcoxon rank-sum tests, 95% CI. For the box plots ($n = 8$ core microbe in pairwise interactions), the horizontal line indicates the median and the whiskers indicate the lowest and highest points within 1.5× the interquartile ranges of the lower or upper quartile, respectively. **e**, Synergistic interactions are calculated by analysing the total 16S rRNA gene copy number in the co-cultures compared with the sum of 16S rRNA gene copy numbers in each individual culture (see Methods; Supplementary Fig. 8). Bar plots represent the percentage change in growth (either positive or negative) of each co-culture compared with the monocultures, showing that synergistic interactions (pink, synergistic; blue, non-synergistic interactions) are prevalent in the core microbes. We found that the percentage change originating from positive and negative interactions was significant; the statistical analysis was performed using a Wilcoxon *t*-test.

factors can contribute to variation among hosts, including host genotype and stochastic processes.

Discussion

In recent years, the term core microbiome was coined to describe taxa with high occupancy that persist in most of an animal cohort, across multiple habitats or environments[32,33]. These core microorganisms can be viewed as ecologically generalist species that are found across multiple changing environmental conditions, such as diets and individual hosts, or through time[34]. We found eight such generalist core microbes by defining their niche width using three different methods (Figs. 2a and 3b, Supplementary Fig. 4). These

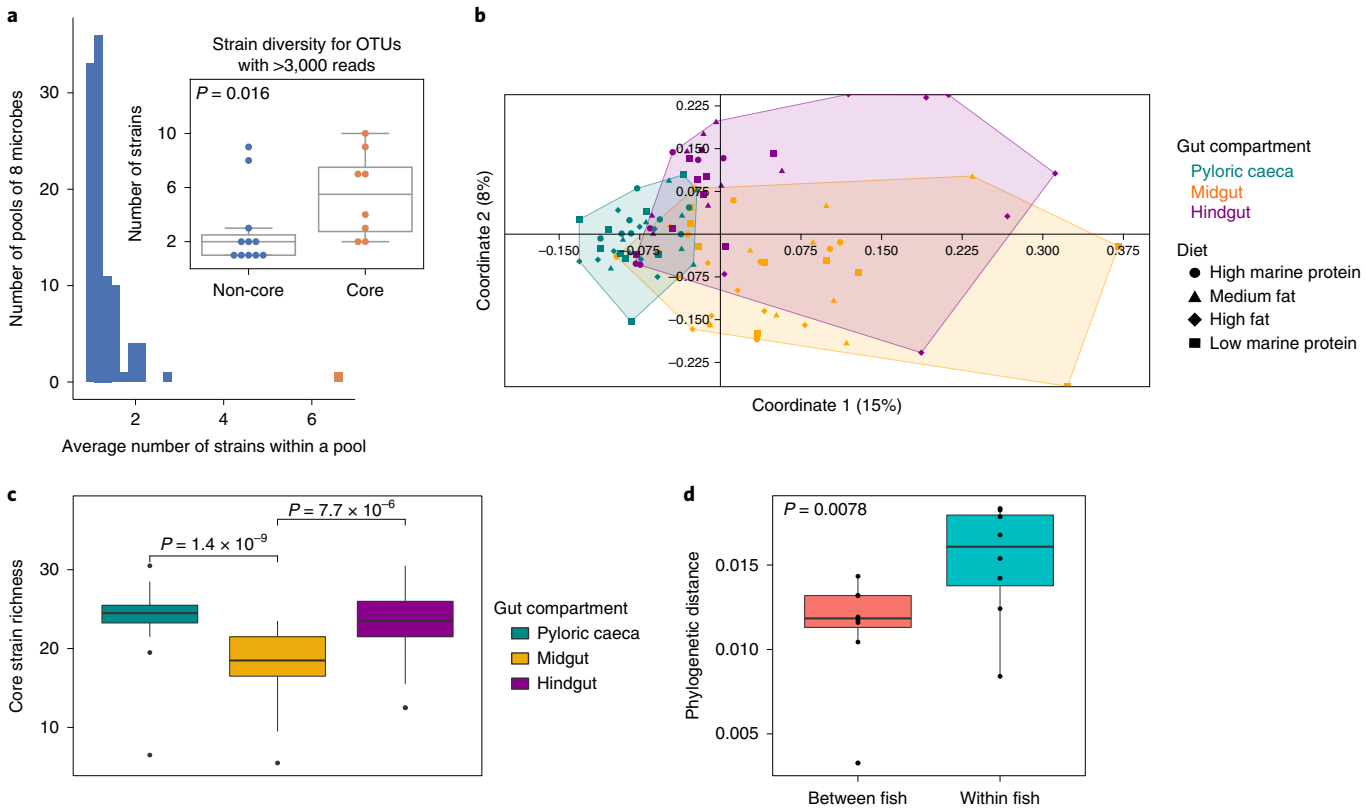

Fig. 4 | Generalist core species show higher strain variability, which is congruent with their habitat preferences. a, A histogram showing the strain richness within a species (97% sequence similarity clusters) comparing core versus randomly sampled non-core species (see the 'Strain variation' section in the Methods). Data are average strain richness in a pool of 8 microbes. The red bar shows the average number of strains of the core microbes; the blue bars denote the average number of strains of randomly selected pools of 8 microbes, iterated 100 times. Significance ($P=0.0099$) was assessed using one-sided Mann–Whitney *U*-tests (95% CI). Inset: the box plots show the strain richness within a species (97% sequence similarity clusters) comparing core versus non-core species with more than 3,000 sequence read depth (see the 'Strain variation' section in the Methods). Significance ($P=0.016$) was assessed using two-sided Mann–Whitney *U*-tests (95% CI). Data are shown as box plots; the horizontal line indicates the median of the sum of strains for all of the core species and the whiskers indicate the lowest and highest point within 1.5× the interquartile ranges of the lower or upper quartile, respectively. **b**, Principal coordinates analysis of samples (PERMANOVA, $P=0.001$, 95% CI; $n=9$ fish per diet and 3 parts per fish; see Methods) based on the presence–absence patterns of strains (Jaccard metric) originating from the core microbe cluster by gut part (colour) and diet (shape). **c**, Strain richness originating from all of the core microbes across the gut parts ($n=51$ total strains from 8 core microbes). Statistical analyses were performed using paired two-sided Wilcoxon rank-sum tests with Benjamini–Hochberg correction ($P<0.05$, 95% CI). Box plots describe the data as the inset of **a**. **d**, The average phylogenetic distances of strains shared between fish and strains ($n=8$ core microbes) within a fish. Statistical analyses were performed using paired two-sided Wilcoxon rank-sum tests ($P=0.0078$, 95% CI). Box plots describe the data as the inset of **a**.

core species were persistent through time, as determined by our time-series experiments (Supplementary Table 3). Moreover, at the genus level annotation, the identified core microbes were found to be present across different fish hosts that inhabit different environments, such as sea water and fresh water, and with various feeding habits, such as herbivores, carnivores and omnivores (Fig. 2d). Interestingly, these core microbes shared only positive co-occurrence patterns, in contrast to other members of the microbiome, in which co-exclusion was mostly observed (Fig. 3b). The strong co-occurrence patterns between these core microbes might be an indication of low levels of competition, potentially due to niche partitioning, which would enable these species to coexist. Indeed, genomic analysis of metabolic potential and resource utilization of the isolates that correspond to these core microbes supported this assumption, which is probably the driver of their high prevalence and consistency across the individual guts (Fig. 3c–e, Supplementary Fig. 8). Moreover, the outcome of our co-culturing experiments between these core microbes showed a tendency towards facilitation and low-level competition (Fig. 3d,e). It has been suggested that competition promotes stability in the gut ecosystem and that cooperation creates dependencies that foster instability in microbial communities[15]. Indeed, we observed a high number of negative interactions among many members of the gut microbial community as well as between the core microbes and other community members in this study (Fig. 3b). However, it should be noted that despite the positive interactions among them, these core microbes did not reveal any interdependency for growth as they were also growing as isolates; thus, this type of cooperation among the core microbes does not seem to be obligatory and is therefore less likely to destabilize the community. Our findings concerning the metabolic potential and co-culturing assays therefore highlight niche partitioning and microbe facilitation as important determinants of core-microbe interactions that potentially contribute to their persistence.

When we looked at the species level, we found that these core microbes have higher strain variability than other microbiome members, and this variability was in agreement with habitat conditions across the gut. Looking at the strain level of these core microbes, we observed a differential distribution along the gut, visualized as clear strain clusters according to the gut parts, especially the pyloric caeca, indicating habitat specialization (Fig. 4b). Furthermore, we

show that these core species have higher strain (intraspecies) variability than other species (Fig. 4a), and that this increase is exhibited when the overall microbiome richness decreases (Figs. 1b and 4c). Together, these two patterns suggest that the higher variability of these core microbes reduces intraspecific competition, potentially enabling them to persist across habitats and multiple hosts. According to the niche variation and expansion hypothesis—which has only been examined for macroorganisms such as birds[35–37], carnivorous mammals[38] and lizards[39]—generalist species should have higher individual variability than species with lower occupancy[11,20]. Here we provide support for this theory in the microbial world, as these core microbiome species, which are considered to be generalists, were more variable than their specialist counterparts. According to this theory, we would also expect the variability of the core species to increase even further when competition with other species is reduced, for example, when environmental selection decreases the number of species in a given habitat (Fig. 1b, Supplementary Fig. 2).

We found that the strains between fish are more phylogenetically related than those within fish. Moreover, the specification process was not associated with closer phylogenetic distances between the strains of the different species (Supplementary Fig. 14), suggesting that the evolutionary trajectories to achieve this adaptation are diverse and, therefore, different strains can develop it in parallel and through independent evolutionary adaptation processes. Although there is evidence from both microbial communities in general, and gut-associated communities in particular, that greater 16S rRNA gene amplicon variation in this range of similarity implies greater differences in gene content[40–42], exploring the genomes of these strains for their specific functional adaptations would shed light on these evolutionary trajectories.

In this study, we show that core gut microbes tend to be more ecologically and genetically variable. This variability is accompanied by habitat specialization and, therefore, a reduction in intraspecies competition, enabling ecological stability and dispersal across hosts. Moreover, we show facilitation and low levels of interspecies competition between these core microbes, potentially due to resource partitioning. These findings could explain the high occupancy of certain microbial species in gut environments, species that are often referred to as core microbes.

Methods

Experimental design and sampling. The experimental procedures used in this study were approved by the Animal Policy and Welfare Committee of the Agricultural Research Organization (approval number IL-241/10) and were in accordance with the guidelines of the Israel Council on Animal Care. In this study, European seabass (*Dicentrarchus labrax*), an agriculturally important carnivorous fish with a typical gastrointestinal tract, was obtained from a commercial hatchery (Maagan Michael, Israel) and was housed in 250 l experimental indoor tanks equipped with recirculating systems. Seabass individuals were randomly and evenly distributed in triplicate groups (randomization) and, after adaptation to experimental conditions, the groups were fed 4 experimental diets for 6 weeks: a high-marine-protein diet, a medium-fat diet, a high-fat diet and a low-marine-protein diet (Supplementary Table 1). As the focus of our broad analysis around different microbial species distribution and their potential relationship to physiological functions, we aimed for a representative sample size that would cover the different microbial taxa in the different fish gut parts. For this purpose, we performed a rarefaction curve, taking into account different sample sizes and read depths, and we reached a visible sample rarefaction plateau (Supplementary Fig. 1). At the end of the experiment, three fish from each tank were randomly selected and their guts were dissected using sterile instruments and separated into pyloric caeca, foregut and hindgut (Fig. 1a). After dissection, each sample was ground, frozen and stored at −80 °C for further analysis. Sampling was performed after a 1 d fasting period. Each gut part of an individual fish represents a research sample. These samples were taken from seabass intestinal tract that originated from the same population (same parents). The age of the animals was approximately 6–7 months (approximately 10 g in weight). At this stage, all of the animals are sexually immature. Throughout the trial, all of the treatments were handled in a similar manner by the same trained research technician of the institute. Samples for each time point were collected during the same day by trained scientific personnel. After anaesthesia using clove oil, fish from each tank were randomly selected and their guts were dissected using sterile instruments. The researchers analysed the data using the same computational tools and procedures and therefore blinding was irrelevant.

DNA extraction. Bacterial DNA was isolated from gut samples using the protocol described previously by Roeselers et al.[43] with some modifications[44]. Excised intestines were combined in 2.0 ml screw-cap tubes with 0.5 mm and 1 mm silica beads (Biospec), 400 ml 50 mM sodium phosphate buffer (pH 8.0) and 200 ml lysis solution containing 5% (w/v) sodium dodecyl sulfate, 0.5 M Tris-HCl (pH 8.0) and 0.1 M NaCl. Samples were homogenized in a bead beater for 5 min on high speed. The supernatant was transferred to new tubes and lysozyme (Sigma) was added up to a final concentration of 2 mg ml^{-1}, followed by incubation at 42 °C for 1 h and then 37 °C for 1 h. The solution was then sequentially extracted using TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA), saturated phenol, phenol–chloroform (1:1 v/v) and chloroform–isoamyl alcohol (24:1 v/v). Finally, DNA in the aqueous phase was precipitated with 0.1 volume 3 M sodium acetate (pH 5.2) and 0.7 volume isopropanol. The concentration of DNA in the solution was measured using a Nanodrop 2000 UV-Vis spectrophotometer (Thermo Scientific) and the DNA was stored at −20 °C for further analysis. Only samples that resulted in a high yield of high-quality DNA were used for subsequent analyses.

Sequencing of the gut microbiome. Sequencing of the PCR-amplified V4 region of 16S rRNA was performed using a MiSeq 2000 Next Generation system (Illumina). First, amplification of the V4 region using the primers 515F (5′-GTGCCAGCMGCCGCGGTAA-3′) and 806R (each reverse primer contained a different 12 bp index) and the enzyme HotStar Taq (5 U µl^{-1}; Qiagen), was performed under the following conditions: 94 °C for 15 min, followed by 35 cycles of 94 °C for 45 s, 50 °C for 60 s and 72 °C for 90 s, and a final elongation step at 72 °C for 10 min. The PCR product (380 bp) was cleaned using a PCR clean-up kit (DNA Clean & Concentrator, Zymo Research) and quantified for fragments containing the Illumina adapters. Amplification involved initial denaturation at 95 °C for 15 min and then 40 cycles at 95 °C for 10 s followed by annealing at 60 °C for 20 s and extension at 72 °C for 30 s. The product was quantified using a standard curve with serial DNA concentrations (0.1–10 nM). Finally, the samples were equimolarly diluted to a concentration of 0.4 nM and prepared for sequencing according to the manufacturer's instructions. Data quality control and analyses were performed using the DADA2 and QIIME 2 pipeline (v.11.2018; https://qiime2.org). DADA2[45] was applied to model and correct Illumina-sequenced amplicon errors, using the option of pooling all of the samples in the paired-end merged reads (https://github.com/benjjneb/dada2). DADA2 resolves differences at the single-nucleotide level, and the end product is an amplicon sequence variant table, which is a higher-resolution analogue of the traditional operational taxonomic unit (OTU) table, recording the number of times each ESV was observed in each sample (100% sequence identity). Taxonomy was assigned using the Ribosomal Database Project Classifier[46] against the 16S rRNA gene reference Greengenes database (v.13.8)[47]. Owing to variation of sequence depths among samples, all of the samples were normalized to the lowest depth by subsampling (6,000 reads per sample); samples below 4,000 reads were discarded.

Quantification of sequencing noise. To quantify the variation of sequencing noise in our dataset, we used two approaches as follows.

In the first approach, we sequenced five biological samples with four technical replicates (different barcodes) as we would expect that technical replicates should give rise to similar frequencies of sequencing errors. We followed the DADA2 approach, as described in the 'Sequencing of gut microbiome' and 'Strain variation' sections. We then calculated the coefficient of variation for the number of strains between the technical and biological replicates to measure the overall sequencing variation at the strain level (ESV; Supplementary Fig. 11a). Next, we aimed to quantify similarities between biological and technical replicates using a phylogenetic distance matrix. As biological replicates, we randomly selected samples originating from different parts and different dietary treatments (that is, midgut from fish originating from high-marine-protein and low-marine-protein diets); detailed information is provided in Supplementary Table 19. We used the weighted UniFrac distance metric to calculate the similarities of these samples on the basis of the phylogenetic distances of their strains. Our results showed a higher similarity in the phylogenetic distances of strains coming from the technical versus biological replicates, whereas there was a higher variation in phylogenetic distances within the biological replicates (Supplementary Fig. 11b).

In the second approach, we followed the method described by Jing et al.[48]. We tested whether random sequencing errors can account for the abundance of strains. DADA2[45] has been reported to have an error rate between 10^{-6} to 10^{-8}. We therefore took a conservative and stringent approach and accounted for an error rate of both 10^{-5} and 10^{-6} to evaluate whether the strain variation was originating from random sequencing errors, which was our null hypothesis. Our results showed that the null hypothesis was rejected for around 87% and 93% of the overall species (97% OTU clusters) for these two error rates. The results for each species are presented in Supplementary Table 17. Note that in this file, 'true' means that the *P* value was significant, and that the null hypothesis was rejected and there are no probable sequencing errors within this species. Specifically, all of the core species were not suspected to have sequencing errors.

Comparison of gut communities. Alpha diversity was calculated using richness (number of observed species), Shannon and phylogenetic diversity. Cluster analyses exploring the similarities among gut-community compositions of different samples were performed using phylogeny-based approaches (UniFrac[49]). A neighbour-joining tree of all of the processed reads was created using FastTree[50] (v.2.1, default settings), which was subsequently used to calculate the weighted UniFrac matrix. The resulting distance matrix was visualized using a UPGMA dendrogram in MEGA 6.0[51]. To directly measure the robustness of individual UPGMA clusters, we performed jackknifing by repeatedly resampling a subset of 4,000 reads from each sample.

Transient taxa. To identify potential transient taxa in our dataset, we included a time-series experiment. It consisted of a 6 week feeding trial with different diets for European seabass. During these 6 weeks, samples were taken from the different diets (4 diets; Supplementary Table 1) and different gut parts (pyloric caeca, midgut and hindgut) at several time points. We sequenced the 395 samples originating from three time points—1, 2 and 6 weeks from the start of the trial. Data were submitted to the Sequence Read Archive (SRA) database under submission number SUB3969450. For the analysis, we applied a method described by Shade et al.[29] to detect transient taxa, which follow a bimodal distribution over time (rare taxa that, depending on the conditions, become temporarily highly abundant). Specifically, the statistical method for detecting a bimodal distribution computes the coefficient of bimodality, *b*. We calculated the *b* coefficient for all of the ESVs in our dataset across time (in all of the different parts and different diets).

From the distribution of a taxon's levels of abundance with time, the coefficient of bimodality, *b*, is calculated as follows (where *x* is the abundance of the taxon, and *i* gives the time point:

$$b = \frac{\left(1 + \text{skewness}^2\right)}{(\text{kurtosis} + 3)}$$

where skewness is defined as:

$$\frac{\sum_{i=1}^{n}(xi - \bar{x})^3/n}{\left[\sum_{i=1}^{n}(xi - \bar{x})^2/n\right]^{3/2}}$$

and kurtosis is defined as:

$$\frac{\sum_{i=1}^{n}(xi - \bar{x})^4/n}{\left[\sum_{i=1}^{n}(xi - \bar{x})^2/n\right]^2}$$

A taxon is considered to be transient or conditionally rare[29] when its *b* is more than 0.9 and its abundance is more than 0.01%.

Niche width. We used three methods to analyse niche width. The first method was Shannon diversity, as previously used in studies in macroecology[11]; this method reflects both the number of different habitats that each species occupies and the evenness with which they occur and is also suitable for a high number of samples. The second method was the unweighted richness of habitats, and the third was Levins' niche breadth[28], which defines habitat specialization as a function of uniformity of the distribution of species abundance among habitats[28,52]. Within each habitat (samples originating from different diets and gut parts), we used the frequencies of microbial taxa across the seabass gut to calculate the different indices as measures of population niche width[11,53]. Taxa with higher niche width values are those that use a broader range of habitats (that is, species that are more equally distributed across samples and are found in more of them). Thus, taxa with higher and lower values of niche width can be considered to be generalists and specialists, respectively. We then calculated the niche width and found that 8 microbes, defined as core microbes (12 ESVs—shared between the parts and diets—clustered into 8 clusters at 97% similarity), fall within the upper tail distribution of all three indices, which we defined as generalist microbes (Supplementary Fig. 4, Supplementary Table 10).

Shared and unique microbiomes. Analysis of shared and unique microbes (ESVs) was conducted on the basis of the table generated by DADA2. The shared microbes were defined as those that were present in at least 90% of the samples for either each fish gut part or each diet. The unique microbes were arbitrarily defined as those that were present in more than 90% of only one of the fish gut parts or diet samples and were not found in the other types of sample. The closest bacterial genomes of the core microbes, as annotated by BLAST[54], were obtained from NCBI and uploaded to the RAST server[55] (default settings; Supplementary Table 11), in which their predicted genes were functionally annotated by KEGG[56]. These predicted functional profiles in the form of enzyme category (EC) numbers were supplied to NetCmpt[30] to calculate the competitive potential between each species pair. The NetCmpt tool enables calculation of the competitive potential between bacterial pairs by simply providing the species-specific content of EC reactions. In brief, given the EC list input, the NetCmpt tool constructs species-specific networks and metabolic environments. Then, for each species pair (considering all of the possible pairwise combinations), it constructs pair-specific environments comprising all of the non-overlapping metabolites of a given pairwise combination.

In vitro interactions among the core microbes. To isolate the core microbes from seabass gut, different microbiological media were used (nutrient agar, Lysogeny Broth, King B; Supplementary Table 12), following serial dilutions (0–10^3) of homogenized gut tissue in 0.9% saline solution. The identity of each microbe was verified by Sanger sequencing of the V3–V4 region of the 16S rDNA (Hylabs) using the primers 27F (5′-AGAGTTTGATCMTGGCTCAG-3′) and 1525R (5′-AAGGAGGTGWTCCARCC-3′) and annotation with BLAST[54]. The bacteria that we did not succeed in isolating were obtained from DSMZ (Supplementary Table 13). To explore competitive/inhibitory interactions among the core microbes, in vitro cross-streak pairwise assays were performed in different media—nutrient agar and Davis minimal media—by adding different carbon sources: glucose, cellulose, phosphatidylcholine, cysteine and casein. Triplicates of bacterial strain pairs, freshly prepared and adjusted to the same optical density at 600 nm (OD_{600}) after washing in sterile PBS (phosphate buffer and saline solution) (pH 7.0), were streaked (10 µl) across different media plates and the plates were then incubated for 72 h at 28 °C. Macroscopic growth was observed for potential inhibitory activity.

A co-culture assay between pairs of core microbes was performed using in vitro digested seabass feed. The digested feed was prepared by adding 5 g of a commercial European seabass feed to sterile 50 ml tubes containing 0.1 N HCl and 0.2% (w/v) pepsin (~pH 2–3). The feed was digested overnight in a shaker incubator at room temperature. The next day, the tubes were centrifuged at 10,500*g* for 20 min; the supernatant pH was adjusted to 7 by titrating with sodium bicarbonate and filtered to sterile using 0.22-µm-pore filters. At the start of the co-culture experiment, single colonies were grown for 24 h in 3 ml nutrient agar medium. Cells were centrifuged at 6,000*g* for 15 min and washed three times with 0.9% sterile saline to remove any excess medium. The cells were then adjusted to an OD_{600} of 0.2 using 0.9% sterile saline and added to 96-well sterile flat-bottom plates containing 100 µl of feed extract per well. Each well contained different pairs of core microbes (50 µl of each microbe) in triplicate (36 pairwise interactions, and 8 controls—single microbes). The co-cultures were grown for 24 h at 28 °C and mixed on a shaker rotating at 250 r.p.m. To measure relative abundance, DNA was extracted from a 10 µl sample from each well using the Prepman Ultra Kit (Applied Biosystems) following the manufacturer's instructions. The relative abundance of each microbe was then measured by quantitative PCR, using species-specific primers (Supplementary Table 14). Cross-reactivity of the designed primers was validated with 1% agarose gels for each specific primer pair using the extracted DNA from each of the 8 core microbes (Supplementary Fig. 12). The fold change in growth of each microbe was calculated in comparison to the growth of single microbes as the control.

Substrate utilization. After isolating or obtaining the core microbes (see the 'In vitro interactions among the core microbes' section; Supplementary Fig. 7, Supplementary Tables 12 and 13), we evaluated their potential for utilization of various substrates, using Ecoplate (BIOLOG). Ecoplate is a fingerprinting method that elaborates different substrates to reveal the physiological profile of different microbes[57]. The core bacteria were grown in liquid cultures containing nutrient agar overnight at 28 °C, their growth was evaluated by OD_{600}. After being washed with sterile PBS and centrifuged to remove any nutrient residue, they were inoculated in 96-microwell plates in triplicates per microbe and were grown for 72 h, according to BIOLOG instructions for Ecoplate.

Fish-gut metagenomes. To compare our findings with data from previously published fish-gut metagenome studies, we reanalysed those published sequencing datasets using the QIIME[58] closed reference protocol against the 97% similarity to the Greengenes database[47] (most recent). The studies included in the analysis are listed in Supplementary Table 15.

Strain variation. The 16S rRNA gene is commonly used as a molecular marker for microbial community composition and structure analysis. Such analysis generally relies on classification-based approaches that make taxonomic assignments by comparing each DNA sequence to reference databases, or clustering-based methods that group together multiple sequences as taxon-independent units using a sequence-similarity threshold; both approaches include intrinsic critical limitations. Analyses that classify sequence reads by similarity to taxonomic database entries may provide poorly resolved descriptions of diversity, especially for samples that are collected from high-diversity environments. Although the 16S rRNA gene has limited specificity (for example, two distant organisms may have almost identical genes), it is very sensitive—a single nucleotide difference at the gene level can reflect vast genomic differences[59]. Unravelling the relationships between bacteria and their environment often requires information about microbial diversity on a finer scale. Sequence similarity of 97% at the 16S rRNA gene is accepted to be a measure that clusters strains of the same species. Thus, very similar, but slightly distinct, gene sequences represent different microbes in a community. We therefore quantified the number of strains—that is, ESVs—within each 97% species–OTU cluster. More specifically, the total subsampled ESVs were clustered into species at 97% sequence identity using the 'vsearch cluster-features-de-novo' command in QIIME 2 (v.2018-11). Richness was calculated for each species measuring the number of ESVs that exist within each. Phylogenetic distances between ESVs were calculated using FastTree[60] (v.2.1, default settings).

Statistical analysis. Kruskal–Wallis one-way analysis of variance and the pairwise comparison Wilcoxon paired *t*-test were used to test whether the means and standard deviations of alpha-diversity values between different diet categories were significantly different ($P < 0.05$) and to test strain richness (using the wilcox.test R package). Wilcoxon rank-sum test was used to test for shared and unique taxa within the different gut parts. Corrections for *P* values were applied whenever stated using R p.adjust. Clustering significance for the Jaccard metric was evaluated with ANOSIM and two-way PERMANOVA for the factors gut part and diet (R package vegan). Heat maps and graphs were created using R (packages vegan and ggplot2). The CoNet tool was used[61] to identify and visualize significant co-occurrence patterns within the gut microbial communities. All of the experiments were performed in triplicates.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Sequencing data are provided at the NCBI (SRA) database under the study accession code SRP118834. The non-sequencing related data are provided in Supplementary Data 1.

Received: 1 October 2017; Accepted: 12 August 2019;
Published online: 23 September 2019

References

1. Shabat, S. K. B. et al. Specific microbiome-dependent mechanisms underlie the energy harvest efficiency of ruminants. *ISME J.* **10**, 2958–2972 (2016).
2. Friedman, N. et al. Diet-induced changes of redox potential underlie compositional shifts in the rumen archaeal community. *Environ. Microbiol.* **19**, 174–184 (2017).
3. Zeevi, D. et al. Structural variation in the gut microbiome associates with host health. *Nature* **568**, 43–48 (2019).
4. Kayser, J., Schreck, C. F., Yu, Q., Gralka, M. & Hallatschek, O. Emergence of evolutionary driving forces in pattern-forming microbial populations. *Proc. R. Soc. B* **373**, 20170106 (2018).
5. Donaldson, G. P., Lee, S. M. & Mazmanian, S. K. Gut biogeography of the bacterial microbiota. *Nat. Rev. Microbiol.* **14**, 20–32 (2016).
6. Jeraldo, P. et al. Quantification of the relative roles of niche and neutral processes in structuring gastrointestinal microbiomes. *Proc. Natl Acad. Sci. USA* **109**, 9692–9698 (2012).
7. Shaani, Y., Zehavi, T., Eyal, S., Miron, J. & Mizrahi, I. Microbiome niche modification drives diurnal rumen community assembly, overpowering individual variability and diet effects. *ISME J.* **12**, 2446–2457 (2018).
8. Ghoul, M. & Mitri, S. The ecology and evolution of microbial competition. *Trends Microbiol.* **24**, 833–845 (2016).
9. Pacheco, A. R., Moel, M. & Segrè, D. Costless metabolic secretions as drivers of interspecies interactions in microbial ecosystems. *Nat. Commun.* **10**, 103 (2019).
10. Travisano, M. & Rainey, P. B. Studies of adaptive radiation using model microbial systems. *Am. Nat.* **156**, S35–S44 (2000).
11. Bolnick, D. I., Svanbäck, R., Araújo, M. S. & Persson, L. Comparative support for the niche variation hypothesis that more generalized populations also are more heterogeneous. *Proc. Natl Acad. Sci. USA* **104**, 10075–10079 (2007).
12. Shade, A. & Handelsman, J. Beyond the Venn diagram: the hunt for a core microbiome. *Environ. Microbiol.* **14**, 4–12 (2011).
13. Thomas, T. et al. Diversity, structure and convergent evolution of the global sponge microbiome. *Nat. Commun.* **7**, 11870 (2016).
14. Astudillo-García, C. et al. Evaluating the core microbiota in complex communities: a systematic investigation. *Environ. Microbiol.* **19**, 1450–1462 (2017).
15. Coyte, K. Z., Schluter, J. & Foster, K. R. The ecology of the microbiome: networks, competition, and stability. *Science* **350**, 663–666 (2015).
16. Bastolla, U. et al. The architecture of mutualistic networks minimizes competition and increases biodiversity. *Nature* **458**, 1018–1020 (2009).
17. Thébault, E. & Fontaine, C. Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science* **329**, 853–856 (2010).
18. Lurgi, M., Montoya, D. & Montoya, J. M. The effects of space and diversity of interaction types on the stability of complex ecological networks. *Theor. Ecol.* **9**, 3–13 (2016).
19. Bolnick, D. I. & Doebeli, M. Sexual dimorphism and adaptive speciation: two sides of the same ecological coin. *Evolution* **57**, 2433–2449 (2003).
20. Van Valen, L. Morphological variation and width of ecological niche. *Am. Nat.* **99**, 377–390 (1965).
21. Hacquard, S. et al. Microbiota and host nutrition across plant and animal kingdoms. *Cell Host Microbe* **17**, 603–616 (2015).
22. Wilson, J. & Castro, L. Morphological diversity of the gastrointestinal tract in fishes. *Fish. Physiol.* **30**, 1–55 (2010).
23. Lanan, M. C., Rodrigues, P. A. P., Agellon, A., Jansma, P. & Wheeler, D. E. A bacterial filter protects and structures the gut microbiome of an insect. *ISME J.* **10**, 1866–1876 (2016).
24. Godoy-Vitorino, F. et al. Comparative analyses of foregut and hindgut bacterial communities in hoatzins and cows. *ISME J.* **6**, 531–541 (2012).
25. Zhang, Z. et al. Spatial heterogeneity and co-occurrence patterns of human mucosal-associated intestinal microbiota. *ISME J.* **8**, 881–893 (2014).
26. Looft, T. et al. Bacteria, phages and pigs: the effects of in-feed antibiotics on the microbiome at different gut locations. *ISME J.* **8**, 1566–1576 (2014).
27. Sriswasdi, S., Yang, C.-c & Iwasaki, W. Generalist species drive microbial dispersion and evolution. *Nat. Commun.* **8**, 1162 (2017).
28. Levins, R. *Evolution in Changing Environments: Some Theoretical Explorations* (Princeton Univ. Press, 1968).
29. Shade, A. et al. Conditionally rare taxa disproportionately contribute to temporal changes in microbial diversity. *mBio* **5**, e01371-14 (2014).
30. Kreimer, A., Doron-Faigenboim, A., Borenstein, E. & Freilich, S. NetCmpt: a network-based tool for calculating the metabolic competition between bacterial species. *Bioinformatics* **28**, 2195–2197 (2012).
31. Cordero, O. X. et al. Ecological populations of bacteria act as socially cohesive units of antibiotic production and resistance. *Science* **337**, 1228–1231 (2012).
32. Bolnick, D. I. Intraspecific competition favours niche width expansion in *Drosophila melanogaster*. *Nature* **410**, 463–466 (2001).
33. Costa, G. C., Mesquita, D. O., Colli, G. R. & Vitt, L. J. Niche expansion and the niche variation hypothesis: does the degree of individual variation increase in depauperate assemblages? *Am. Nat.* **172**, 868–877 (2008).
34. Muller, E. E. L. Determining microbial niche breadth in the environment for better ecosystem fate predictions. *mSystems* **4**, e00080-19 (2019).
35. Soule, M. & Stewart, B. R. The "niche-variation" hypothesis: a test and alternatives. *Am. Nat.* **104**, 85–97 (1970).
36. Dennison, M. D. & Baker, A. J. Morphometric variability in continental and Atlantic island populations of chaffinches (*Fringilla coelebs*). *Evolution* **45**, 29–39 (1991).
37. Feinsinger, P. & Swarm, L. A. "Ecological release," seasonal variation in food supply, and the hummingbird *Amazilia tobaci* on Trinidad and Tobago. *Ecology* **63**, 1574–1587 (1982).
38. Meiri, S., Dayan, T. & Simberloff, D. Variability and sexual size dimorphism in carnivores: testing the niche variation hypothesis. *Ecology* **86**, 1432–1440 (2005).
39. Lister, B. C. & McMurtrie, R. E. On size variation in anoline lizards. *Am. Nat.* **110**, 311–314 (1976).
40. Konstantinidis, K. T. & Tiedje, J. M. Genomic insights that advance the species definition for prokaryotes. *Proc. Natl Acad. Sci. USA* **102**, 2567–2572 (2005).
41. Zaneveld, J. R., Lozupone, C., Gordon, J. I. & Knight, R. Ribosomal RNA diversity predicts genome diversity in gut bacteria and their relatives. *Nucleic Acids Res.* **38**, 3869–3879 (2010).
42. Chaffron, S., Rehrauer, H., Pernthaler, J. & von Mering, C. A global network of coexisting microbes from environmental and whole-genome sequence data. *Genome Res.* **20**, 947–959 (2010).
43. Roeselers, G. et al. Evidence for a core gut microbiota in the zebrafish. *ISME J.* **5**, 1595–1608 (2011).
44. Sun, H., Jami, E., Harpaz, S. & Mizrahi, I. Involvement of dietary salt in shaping bacterial communities in European sea bass (*Dicentrarchus labrax*). *Sci. Rep.* **3**, 1558 (2013).
45. Callahan, B. J. et al. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581 (2016).
46. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73**, 5261–5267 (2007).
47. McDonald, D. et al. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J.* **6**, 610–618 (2012).
48. Jing, X. et al. The bacterial communities in plant phloem-sap-feeding insects. *Mol. Ecol.* **23**, 1433–1444 (2014).
49. Lozupone, C. & Knight, R. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl. Environ. Microbiol.* **71**, 8228–8235 (2005).
50. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5**, e9490 (2010).
51. Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725–2729 (2013).
52. Colwell, R. K. & Futuyma, D. J. On the measurement of niche breadth and overlap. *Ecology* **52**, 567–576 (1971).
53. Roughgarden, J. *Theory of Population Genetics and Evolutionary Ecology: An Introduction* (1979).
54. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).

55. Aziz, R. K. et al. The RAST Server: rapid annotations using subsystems technology. *BMC Genom.* **9**, 75 (2008).
56. Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
57. Gryta, A., Frąc, M. & Oszust, K. The application of the Biolog EcoPlate approach in ecotoxicological evaluation of dairy sewage sludge. *Appl. Biochem. Biotechnol.* **174**, 1434–1443 (2014).
58. Caporaso, J. G. et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**, 335–336 (2010).
59. Ward, D. M. et al. Genomics, environmental genomics and the issue of microbial species. *Heredity* **100**, 207–219 (2008).
60. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* **26**, 1641–1650 (2009).
61. Faust, K. & Raes, J. CoNet app: inference of biological association networks using Cytoscape. *F1000res* **5**, 1519 (2016).

Acknowledgements

The research described here was supported by grants from the European Research Council under the European Union's Horizon 2020 research and innovation program (grant no. 640384), the Israel Science Foundation (grant no. 1947/19) and the Chief Scientist of the Ministry of Agriculture and Rural Development (grant no. 356-0665-14).

Author contributions

F.K. conducted the experiments, molecular work and sequencing, analysed and interpreted the sequencing and molecular data, and wrote the manuscript. G.S. analysed some of the sequencing data. J.F. analysed some of the sequencing data and provided advice on analysis. S.E. conducted some of the microbiology experiments and isolation. O.O. consulted on the interpretation of the data. S.H. and A.C. helped with animal experiments and with the supervision and financial support of the project. I.M. obtained the funding, supervised the study, analysed and interpreted the data, financially managed the project and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41564-019-0560-0.

Correspondence and requests for materials should be addressed to I.M.

nature research

Corresponding author(s): Itzhak Mizrahi

Last updated by author(s): Aug 6, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Metagenomes were collected from MG-RAST: http://www.mcs.anl.gov/project/mg-rast-metagenomics-rast-server, and the SRA database, and information on their source (published work) is referred in Supplementary Table 15.
Data analysis	All the software used for data analysis is available online: R software - version 3.5.1 - packages: vegan, ggplot2, heatmap3, bipartite, picante, MASS. NetCmpt - Kreimer, A., Doron-Faigenboim, A., Borenstein, E. & Freilich, S. NetCmpt: a network-based tool for calculating the metabolic competition between bacterial species. Bioinformatics 28, 2195-2197, doi:10.1093/bioinformatics/bts323 (2012). CoNet - version 1.1.1 beta and Cytoscape - version 3.7.1. QIIME2 - https://qiime2.org/ DADA2 - package in R FastTree, 2.1 version MEGA, 6.0 version

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

A data availability statement is provided in the main manuscript, Line 452-456:

Data availability. Sequencing data can be found at the NCBI (SRA) database under the study accession code SRP118834 and SUB396945. Non-sequencing data are available in Supplementary Data Tables 1, 16 and 17.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	The animal trial included a 6 week experiment with 3 sampling points (1st, 2nd and last week). Four different diets (in triplicate groups) consisting of different nutrient levels (protein and fat), as described in the Methods section, were administered in juvenile European sea bass. At the end of the trial, 9 randomly selected animals per group were sampled and their intestine was dissected into 3 parts (pyloric caeca, midgut, hindgut; See Figure 1). Thus, 3 intestinal samples were collected per each individual fish - a total of 36 fish - and subjected to 16S rRNA gene amplicon sequencing analysis (after DNA extraction).
Research sample	Each gut part of an individual fish represents a research sample. These samples were taken from European sea bass (Dicentrarchus labrax) intestinal tract that were brought to the laboratory from a commercial hatchery (Maagan Michael, Israel), originating from the same population (same parents). The age of the animals was approximately 6-7 months (approximately 10 g in weight). At this stage, the animals are all sexually immature.
Sampling strategy	As the focus of our broad analysis around different microbial species distribution and their potential relationship to physiological functions, we aimed at a representative sample size that would cover the different microbial taxa in the different fish gut parts. For this purpose, we did a rarefaction curve taking different sample sizes and read depths and we reached a visible sample rarefaction plateau (Supplementary Figure S1).
Data collection	Throughout the trial all treatments were handled in a similar manner by the same trained research technician of the institute. Samples for each time point were collected during the same day, by trained scientific personnel. After anesthesia using clove oil, fish from each tank were randomly selected and their guts were dissected using sterile instruments and separated into pyloric caeca, foregut and hindgut. Samples were taken both for DNA extraction and bacteria isolation kept in glycerol.
Timing and spatial scale	The animal experimental trial lasted in total 6 weeks. Samples were taken after week 1, 2, and at the end point (6th week). The experimental setup consisted of three systems, each containing six 250-l tanks and a central main biofilter of 350 l, were the fish were kept throughout both the whole experimental trial and the acclimatization period (1 month before).
Data exclusions	Four sample were discarded due to insufficient sequencing depths. After quality filtering, reads that were much below a 4000 sequence depth were excluded. This sequence depth was selected based on the rarefaction curves, reaching the minimum depth for a plateau.
Reproducibility	The animal experiment was performed successfully in triplicate groups for each treatment. Concerning the microbial interaction trials (liquid and agar media), the trials were performed successfully in triplicates (96-well plates and agar plates, respectively).
Randomization	Concerning the animal experiment, animals (fish) were randomly allocated and were in equal numbers in each experimental group (diet). The size of the animals was measured and animals deviating from the mean were removed, in order not to have significantly different size between the groups. The samples used for the analysis originated from randomly selected fish from each experimental group. The experimental setup consisted of three systems, each containing six 250-l tanks and a central main biofilter of 350 l. The experimental feed was given in triplicate for each diet with random assignment. Each of the diets was represented in each of the three systems used to assess the influence of biofilter on intestinal bacterial community. Moreover, ANOSIM did not reveal any clustering of samples according to biofilter (Supplementary Table 18).
Blinding	The experimental procedure was applied by research technicians who were unaware to the meaning of the different treatments and was therefore unaffected by it. Additionally, the investigators analyzed the data using the same computational tools and procedures and therefore make blinding irrelevant.

Did the study involve field work? ☐ Yes ☒ No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	For the study, European sea bass (Dicentrarchus labrax) were involved, brought to the laboratory from a commercial hatchery (Maagan Michael, Israel), originating from the same population. The age of the animals was approximately 6-7 months (approximately 10 g in weight). At this stage, the animals are all sexually immature.
Wild animals	The study did not involve wild animals.
Field-collected samples	The study did not involve samples collected from the field.
Ethics oversight	The experimental procedures used in the present study were approved by the Animal Policy and Welfare Committee of the Agricultural Research Organization (approval number IL-241/10) and were in accordance with the guidelines of the Israel Council on Animal Care.

Note that full information on the approval of the study protocol must also be provided in the manuscript.