



OPEN

Gut microbiota analysis of Blenniidae fishes including an algae-eating fish and clear boundary formation among isolated *Vibrio* strains

Masa-aki Yoshida¹✉, Takuma Tanabe², Hideo Akiyoshi² & Makoto Kawamukai²✉

Some marine fishes are algae-feeding, and the microorganisms in their digestive tracts produce carbohydrate hydrolyzing enzymes such as agarase and fucosidase, which are potentially interesting resource for new functional enzymes. The purpose of this study was to establish a method for identifying and utilizing characteristic bacteria from the intestines of two algae-eating fish species: *Andamia tetradactylus*, which exclusively eats algae on the rock surface, and stellar rockskipper *Entomacrodus stellifer*, which feeds on both algae and invertebrates. We tested the species composition of the intestinal bacterial flora and found that Proteobacteria were commonly found both in species as in the common gut communities of marine fish, whereas Spirochaetes and Tenericutes occupied the flora of *A. tetradactylus*. We then performed anaerobic and aerobic cultures and isolated 34 and 44 strains including 48 strains belonged to *Vibrio* species from *A. tetradactylus* and *E. stellifer*. We observed that some *Vibrio* strains formed a clear boundary to avoid contacting other strains of bacteria. Whole-genome sequencing of such two *Vibrio alginolyticus* strains revealed two cyclic chromosomes commonly found in the genome of *Vibrio* species, and some unique genes encoding alginate lyase, chitinases, and type I-F CRISPR-associated endoribonuclease for the first time in *Vibrio alginolyticus*.

Gut flora generally includes bacteria, fungi, viruses (bacteriophages), protozoa, and parasites. Microbes that influence their host organisms are particularly important. The intestinal microbiome has been a subject of an emerging interest, since accumulating evidence indicates that the gut flora of humans significantly affects their health. Gut bacteria provide benefits for host cells by supplying essential nutrients such as vitamin K^{1,2}. Some bacteria may provide a calorific boost to the host by breaking down ingested plant carbohydrates such as porphyrin, a sulfated carbohydrate derived from red algae that human enzymes cannot utilize³. *Bacteroides plebeius*, which breaks down agarose, is predominantly found in human populations that favor algae as a regular diet⁴. The composition of the gut microbiota is governed by a combination of environmental factors, (including diet and drugs) and host genetics.

Although many studies have been focused on human intestinal bacteria, little has been studied in intestinal bacteria of fish. Studies of culturable species do not reflect their actual composition, and non-culturable species are important to consider as a whole. Direct amplification of microbial DNA from gut extracts is a useful method for identifying gut flora. In previous studies, fish gut samples showed that Proteobacteria and Firmicutes were the most dominant phyla⁵⁻⁷. Meta-analysis showed that gut bacterial communities of 25 individual fish species exist depending on trophic level (herbivores, carnivores, or omnivores), habitats (saltwater, freshwater, estuarine, or migratory), and sampling methods⁶. For example, the effect of osmotic pressure on gut microbiota in the environment has been investigated. *Oryzias* is a broad-salinity fish genus that has been used as a model organism for progressive hypotonic transfer experiments⁸. The gut of *Oryzias* fish comprised of *Vibrio* at the genus level, but this was replaced by *Pseudomonas* after fish were transferred to the freshwater. The understanding of

¹Oki Marine Biological Station, Faculty of Life and Environmental Sciences, Shimane University, Oki, Shimane, Japan. ²Faculty of Life and Environmental Sciences, Shimane University, Matsue, Shimane, Japan. ✉email: mayoshida@life.shimane-u.ac.jp; kawamuka@life.shimane-u.ac.jp



Figure 1. (a) *Andamia tetradactylus*, 82 mm. (b) *Entomacrodus stellifer*, 108 mm.

intestinal bacteria in these fish species has been gradually progressing, and is currently being applied to efficient aquaculture.

The digestive system plays an essential role in ingesting, digesting, and absorbing nutrients from food. The digestive systems of teleost fishes, composed of a digestive tract and accessory glands, are generally similar to those of other vertebrates^{9–11}. The digestive tract of most teleosts is composed of the esophagus, stomach, pyloric caeca, and intestine^{12–14}. In teleosts, the stomach is usually well developed, although it is absent in some forms. Stomachless fishes have no gastric glands or pyloric caeca. Unexpectedly, among the representative fish species whose genomes have already been sequenced, four are stomachless (tiger pufferfish, green pufferfish, medaka, and zebrafish) as opposed to gastric species (itoyo, tilapia, and cod). Zebrafish rapidly has become a well-recognized animal model for studying host-microbe-immune interactions because germ-free (GF) individuals can be reared to study host-microbe interactions¹⁵. 16S rRNA amplicon analysis showed abundant bacterial groups represented by 21 operational taxonomic units (OTUs) in normal zebrafish, dominated by members of the Proteobacteria (genera *Aeromonas* and *Shewanella*), followed by Fusobacteria, Firmicutes, Actinobacteria and Bacteroidetes⁵. Several generations of GF zebrafish larvae were found to be mono-associated with *Aeromonas veronii*¹⁶. In addition, reciprocal gut microbiota transplants between zebrafish and mice have been archived as GF recipients¹⁷. After transfer of the mouse microbiota into GF zebrafish, the relative abundance of Proteobacteria increased toward the microbiota composition of zebrafish. It appears that the host gut environment shapes the microbiota.

In the marine environment, a variety of bacteria utilize enzymes to degrade such polysaccharides. Bacteria species, such as *Bacillus*, *Vibrio*, *Pseudomonas*, and *Microbulbifer* are sources of alginate lyases¹⁸ and the industrially essential enzymes used in food, biofuel, and biomedical industries¹⁹. Seaweed is also an effective source of various high-molecular polysaccharides, such as fucoidan and is a useful resource for probiotic products. Degraded microorganisms may be attached to consumed food as they enter the intestine. The long intestine is a special digestive tract that enables a symbiotic environment with these microorganisms, and is an incubator for microorganisms that decompose food, providing sufficient time for decomposition. Blenniidae is unique fish that feed on algae and invertebrates; moreover, fish with algae-feeding habits are exclusively found in the Blenniidae family. The Stellar rockskipper, *Entomacrodus stellifer*, feeds on both algae and invertebrates in seawater. They were observed preying on small crustaceans during aquarium keeping. On the other hand, the rockskipper, *Andamia tetradactylus*, is characteristic as to exclusively eats algae on the rock surface (Fig. 1a). The rockskipper is found in the intertidal to supratidal zones of rocky shore washed by waves. They feed exclusively on algae where there is little competition from other fishes. The gastrointestinal tract also specializes in herbivory. The digestive system of Blenniidae fish, such as rockskippers, lack a stomach and are composed of almost spiral and longer intestines. Electron microscopic observation of the gut of the rockskipper showed that a large number of rod and short rod bacteria, which might contribute to the degradation of algae, were present on the surface of the gut contents (Fig. 2). Since the gut flora of the herbivorous fish have not been studied, we aimed to classify and identify the microorganisms in the intestines of the herbivorous fish species *A. tetradactylus* and *E. stellifer*. Using these two species as a model case, the first step was to identify the characteristic bacteria in their intestines and establish a culture method for them. The composition of the intestinal bacterial flora from the two herbivorous fish species was examined by 16S rDNA amplicon sequencing. We then cultivated bacteria from the two fish species on agar plates, and unexpectedly, found that *Vibrio alginolyticus* strains formed a clear line between species on the agar plate. Finally, whole-genome sequencing of two *V. alginolyticus* strains was conducted, and genes coding alginate lyases were found.

Results

Intestinal flora of herbivorous fish. High-throughput sequencing yielded 89,615–128,023-reads for each sample. The QIIME 2-based microbial community analysis yielded 749 representative OTUs. Based on phylogenetic classification using the Greengenes 13_8 16S rRNA gene database, most reads were classified as Proteobacteria, Spirochaetes, and Tenericutes (class 2, Fig. 3). A rarefaction analysis was carried out at species level. When numbers of observed OTUs plotted against numbers of reads, all samples except for the ethanol-

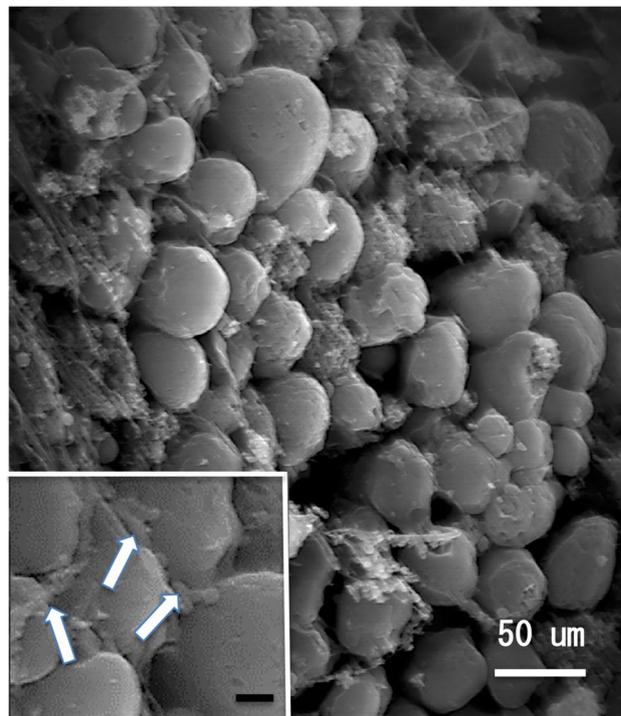


Figure 2. Scanning electron microscopic image of diatoms in the intestine of *Andamia tetradactylum*. Inset: The club-shaped bacteria (arrows) were seen on the surfaces of diatoms. Bar = 20 μm .

fixed sample of *E. stellifer* reached a plateau, and the difference between samples was small (Fig. S1). This indicates that the sequencing depth was sufficient to carry out a thorough description of each sample. To statistically analyze the diversity of the herbivorous gut microbial communities, we examined alpha and beta diversity metrics implemented in QIIME 2. Community dissimilarity between individuals was statistically examined between the following two categories: species (*A. tetradactylum* or *E. stellifer*) and fixation method (ethanol or frozen). According to the diversity analysis of QIIME2, there was no significant difference in diversity between species or fixation (Fig. S2). While the number of observed OTUs was higher in the ethanol-fixed sample ($p = 0.046$, Kruskal–Wallis test; Fig. S2b), Shannon diversity for *A. tetradactylum* and *E. stellifer* was not significantly different ($p = 0.83$, Fig. S2d). Thus, *A. tetradactylum* and *E. stellifer* have gut microbiomes that contain similar numbers of OTUs.

In the *A. tetradactylum* microbial community, Proteobacteria, Spirochaetes, and Tenericutes (class 2) were dominant, accounting for more than 80% of the bacterial diversity in all samples. In contrast, *E. stellifer* was particularly rich in Proteobacteria (Fig. 3a). Among Proteobacteria, Vibrionales was found to be the most abundant in both hosts, Vibrionales alone accounted for more than 20% of the total bacterial flora in both hosts, regardless of the presence or absence of fixation (Fig. S3). Among the Proteobacteria, Rhodospirillales was the second most abundant (Fig. S4). Campylobacteriales were also abundant, but they were characterized by a high variability among the samples.

We attempted to determine how similar are the samples obtained from the different sites in terms of their identified microbial communities. The fecal microbiota for the *A. tetradactylum* and *E. stellifer* groups were clearly separated by Principal Coordinate Analysis (PCoA) plots of beta diversity of unweighted UniFrac distance (Fig. 3b). *Andamia* samples were generally similar regardless of the individual or experimental technique, whereas *Entomacrodus* showed a different microflora profile. We also used linear discriminant analysis (LDA) of effect size (LEfSe) to determine the taxa that most likely explained the differences between the *A. tetradactylum* and *E. stellifer*. When performing the LEfSe analysis, we found that the two genera *Propionigenium* and *Ferriomonas* were significantly enriched in the samples of *E. stellifer* (Fig. 3d, LDA scores = 4.33 and 3.76, respectively, and $p < 0.05$). The family Pirellulaceae and associated unknown genus also showed enrichment with a high LDA score in the *E. stellifer* samples (Fig. 3d, LDA score = 3.34 and $p < 0.05$). Although undetectable in LEfSe analysis, we note that the following two families showed significant variations in beta-diversity metrics: Brachyspiraceae, and Spirochaetaceae (class 5, $P < 0.05$, ANCOM results). The two families of Spirochaetes were found only in *A. tetradactylum* (Figs. S5, S6), consistent with the prominence of Spirochaetes in *A. tetradactylum*. The high frequency of Tenericutes in *A. tetradactylum* was due to the abundance of Mycoplasma species (Fig. S7); however, this difference was not statistically significant.

To evaluate differences in community functional attributes, we used PICRUST2. We found 118 EC gene families with statistically significant differences between the two groups of species (FDR $q < 0.05$) (Fig. S8). EC:2.7.7.71 (biosynthesis of GDP-D-glycero-alpha-D-manno-heptose) was most significantly increased in the *Entomacrodus* samples. We also found 132 COG with statistically significant differences between the two groups of species

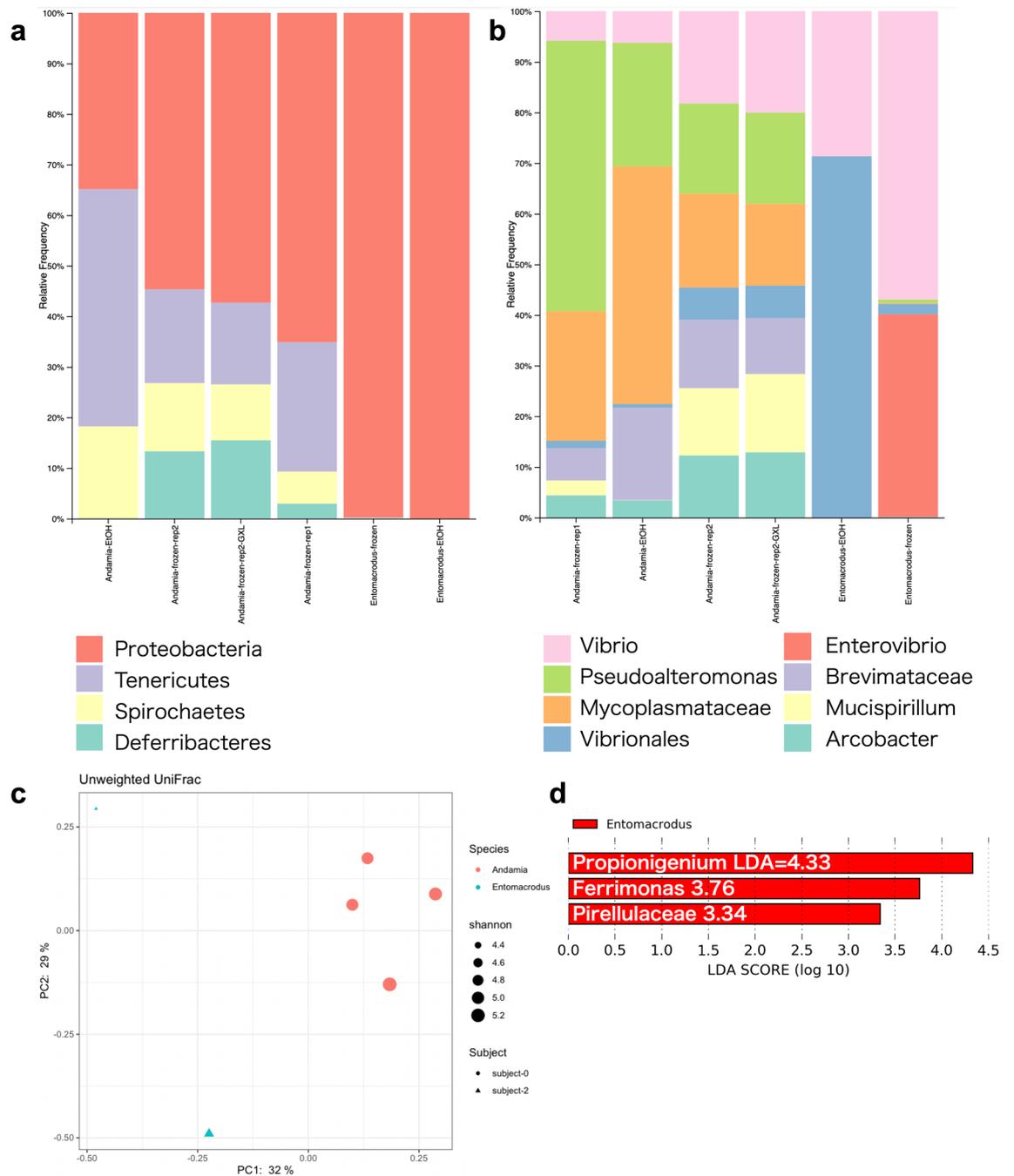


Figure 3. Major groups of bacteria detected by 16S rDNA amplification from *Andamia tetradactylus* and *Entomacrodus stellifer*. **(a)** Relative abundance of phylum (class 2) category was shown in individual gut microbiome samples. Categories with less than 10,000 reads are not shown to represent a major proportion of the bacterial population. **(b)** Relative abundance of genus (class 6) category was shown in individual gut microbiome samples. For the three groups, Mycoplasmataceae, Vibrionales, and Brevimataceae, their family or order names are given since they have not yet been identified as genera. **(c)** Principal Coordinate Analysis (PCoA) plots using distance matrices for beta-diversity (unweighted UniFrac distance). **(d)** A forest plot showing differentially abundant bacterial taxa that were significantly different between species groups as determined using the Kruskal–Wallis test. LDA score (effect size) indicating significant differences in bacterial taxa (LDA score > 3.0; alpha value $p < 0.05$).

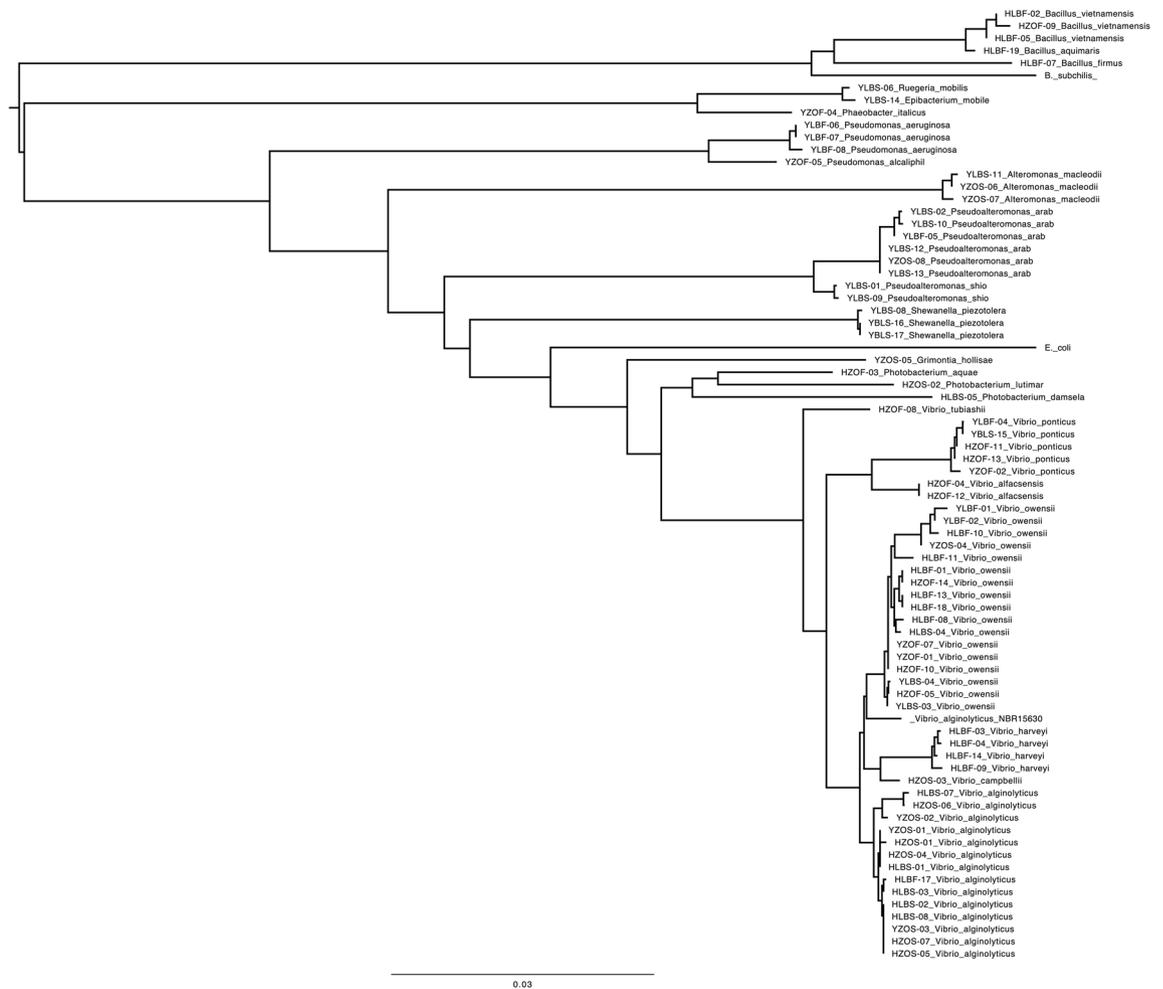


Figure 4. Phylogeny of isolated bacteria from *Andamia tetradactylus* and *Entomacrodus stellifer*. The 35 and 40 strains were isolated from *A. tetradactylus* and *E. stellifer*, respectively. 16S rDNA of those strains were amplified by PCR and determined by sequencing. Based on those sequences, the phylogenetic tree was constructed by the ClustalW program (DDBJ, Japan).

(FDR $q < 0.05$) (Fig. S9). COG4637 (Predicted ATPase) was most significantly increased in the *Entomacrodus* samples.

Isolation of bacteria by culturing on agar plates. We cultured bacteria from the intestines of *A. tetradactylus* and *E. stellifer* on Zobell and LB plates. Bacterial colonies were picked and isolated as single colonies. Single colonies were then examined for 16S rDNA after amplification of the 16S rDNA region by PCR and sequencing. A total of 35 and 40 strains were isolated from *A. tetradactylus* (Table S1) and *E. stellifer* (Table S2), respectively. Some strains were indistinguishable from other strains only by 16S rDNA sequences and indicated, as in Tables S1 and S2. The phylogenetic tree of the isolated bacteria is shown in Fig. 4. Most of the *A. tetradactylus* guts were concentrated in four genera: *Vibrio* spp., *Pseudoalteromonas* spp., *Pseudomonas* spp., and *Shewanella* spp., whereas *E. stellifer* isolates were concentrated in three genera: *Vibrio* spp., *Photobacterium* spp., and *Bacillus* spp. There were also single species isolates, such as *Alteromonas macleodii*, *Paeobacter italicus*, *Ruegeria mobilis*, *Epibacterium mobile*, and *Grimontia* sp. from *A. tetradactylus*. The number of isolated species were larger in *A. tetradactylus* than in *E. stellifer*, although the numbers of isolates were opposite. A wide range of *Vibrio* spp. including *V. alginolyticus*, *V. harveyi*, *V. owensii*, *V. alfacensis*, and *V. ponticus*, were found in the guts of both fish species. *Vibrio* spp. have been frequently identified in other marine organisms. We previously isolated an agar-degrading bacterium from the gastrointestinal tracts of *A. tetradactylus*, but to the date, no agar-degrading bacteria have been isolated from the intestine.

Among the postulated *Vibrio* spp. isolated in this study, 14 strains are migratory. The isolated *Vibrio* spp. were spotted on the agar plate and they started moving on the agar plate by flagella until they met with other strains (Fig. 5a). We found that these *Vibrio* strains exhibited characteristic phenotypes of their boundary formation. HLBS-07 and YZOS-03 formed a very clear boundary line (Fig. 5a). They did not invade the territories of the other strains not like HLBS-07 itself or YZOS-3 itself (Fig. 5a). We then tested the boundary formation of YZOS-03 and HLBS-07 with HLBS-01, HLBS-02, HLBS-03, HLBS-08, HLB-16, HLB-17, YZOS-01, YZOS-02, HZOS-04, HZOS-05 and HZOS-07 strains, which all belong to *Vibrio* spp. All strains formed boundary lines

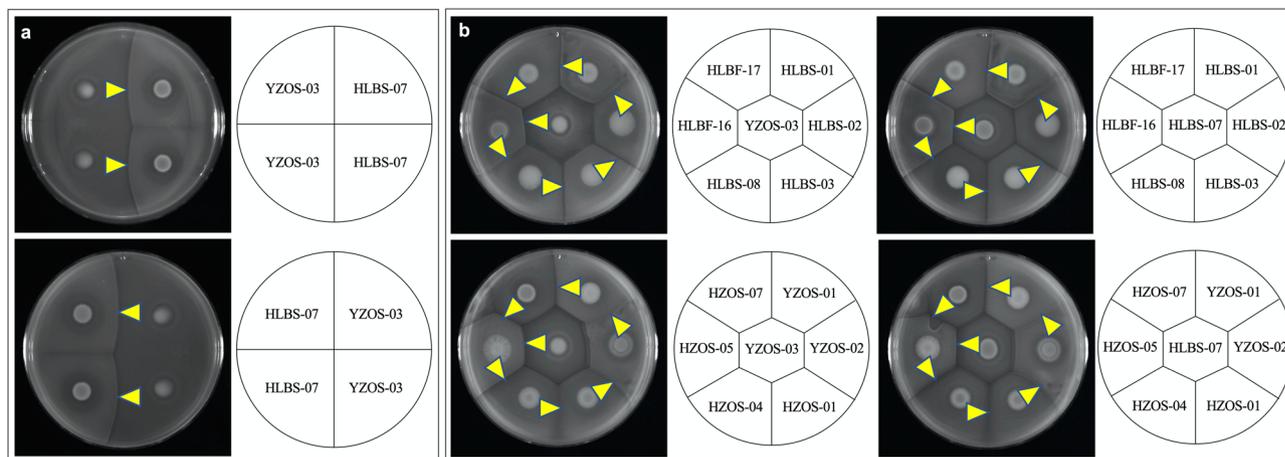


Figure 5. The formation of territory among the bacteria isolated from *Andamia tetradactylus* and *Entomacrodus stellifer*. (a) Two individual colonies of HLBS-07 and YZOS-03 strains belonging to *V. alginolyticus* were grown on Zobell medium at 30 °C for 2 days. (b) The indicated strains belonging to *V. alginolyticus* were grown on Zobell medium at 30 °C for 2 days. Arrowheads point to the boundary between strains.

Assembler + Polishers	YZOS-03					HLBS-07		
	Canu	Flye	Canu + Racon	Canu + MarginPolish	Canu+MarginPolish+Medaka	Canu	Canu + MarginPolish	Canu+MarginPolish+Medaka
# contigs	3	2	3	3	3	2	2	2
Total length	5,195,096	5,145,162	5,195,786	5,196,640	5,196,188	5,336,637	5,338,931	5,338,402
BUSCO score*	94.7	97.4	87.6	95.5	98.9	91.5	95.3	98.5
# Complete	1368	1407	1265	1379	1429	1322	1377	1423
#Fragmented	49	21	107	43	9	83	43	13
#Missing	28	17	73	23	7	40	25	9

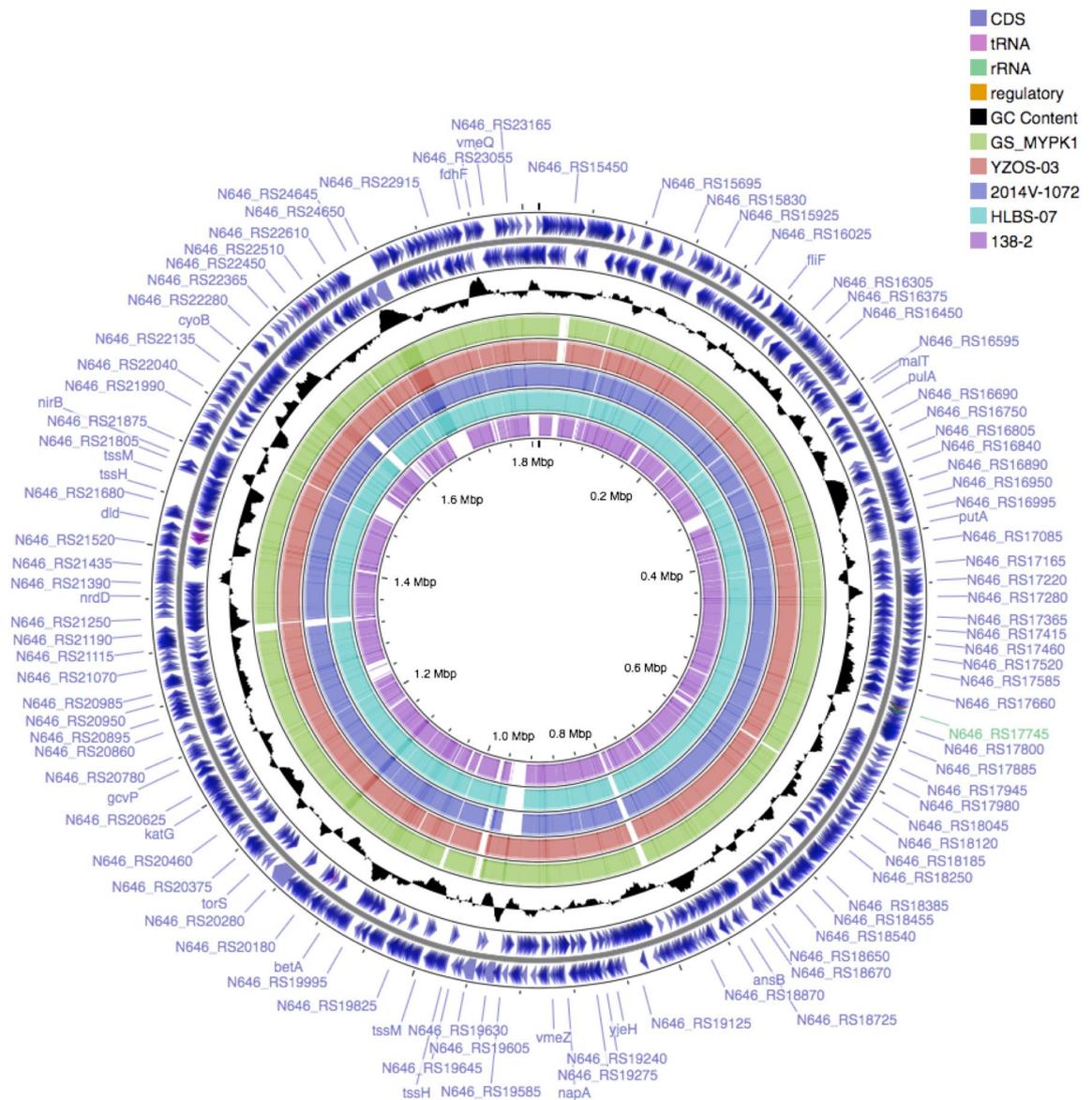
Table 1. Comparison of BUSCO completeness of genome assemblies. *Results from dataset vibrionales_odb.

with HLBS-07 and YZOS-03 on the Zobell medium (Fig. 5b). These boundary lines were less clear between HLBS-01 and HLBS-02 strains, or between YZOS-01, and YZOS-02 strains. Clear territory formation occurred only between different isolates and not within the same *Vibrio* isolates.

Complete genome sequencing of *Vibrio* strains cloned from the herbivorous fish. For YZOS-03, 1,427,401 nanopore reads (11,170,207,911 bp) were obtained with an N50 length of 13.9 kbp. For HLBS-07, 1,150,246 reads (9,836,928,300 bp) were obtained with an N50 length of 15.0 kbp. The subsampled reads of YZOS-03 were assembled with Canu v2.0 into three circular contigs (3,351,300, 1,901,164 and 95,862 bp). We discarded the shorter contig because this contig was 100% identical to one part of the largest contig. Therefore, we selected Canu assembly for subsequent steps. Consensus assembly was called using the MarginPolish with 1,427,401 reads, followed by the Medaka polishing. Racon polishing worsened the BUSCO score and was not used (Table 1). We compared the BUSCO score with that of the Flye assembly and selected the Canu + MarginPolish + Medaka polished assembly for the subsequent analysis (Table 1). After the polishing process, the total nucleotide genome of YZOS-03 was 5,145,821 bp, with a G + C content of 44.7% and BUSCO score of 98.9%. The subsampled reads of HLBS-07 were also assembled with Canu v2.0 into two contigs (3,794,970 and 11,408 bp). The longer contig was circularized by searching overlapping fragments from the raw reads, although the shorter contigs was circular. After the polishing process, the total nucleotide genome of HLBS-07 was 5,338,402 bp with a G + C content of 44.5%.

DFAST identified 4631 predicted protein-coding sequences, 37 rRNA genes, and 129 tRNA genes for YZOS-03. In addition, 4856 predicted protein-coding sequences, 40 rRNA genes, and 130 tRNA genes were identified for HLBS-07. Circular visualization of the genomes was carried out on the CGview server v1.0 (Figs. 6, 7). Comparison of genomic sequences of YZOS-03 and HLBS-07 with the type strain of *V. alginolyticus* NBRC 15630 (= ATCC 17749, accession numbers NC_022349.1 and NC_022359.1) revealed that HLBS-07 had a more similar overall nucleotide identity (99.14% for Chr1, 98.23% for Chr2) than YZOS-03 (98.66% for Chr1, 98.46% for Chr2). The closest strain to HLBS-07 based on chromosome identity was 2014 V-1072 (99.62% for Chr1, 98.88% for Chr2, Fig. 6) and that of YZOS-03 was G3_MYPK1 (99.08% for Chr1, 98.62% for Chr2, Fig. 7).

Next, we analyzed genes that were common to YZOS-03 and HLBS-07, but not in the genome of the type strain *V. alginolyticus* NBRC15630 (GCF_000354175.2). Among the genes in HLBS-07, 234 genes (e-value threshold $1e^{-10}$) were not detected in BLASTP in either the YZOS-03 or NBRC15630 strains. Similarly, 181 genes were unique to YZOS-03. We found that both strains encoded two alginate lyases, which hydrolyze the major



Vibrio alginolyticus YZOS-03 & HLBS-07 (chromosome 2)

Figure 7. Structure of chromosome2 of *Vibrio alginolyticus*. Chromosomal sequences of *V. alginolyticus* G3_MYPK1, 2014V-1072, 138-2 HLBS-07 and YZOS-03 strains were mapped based on NBRC15630. CDSs were estimated with Prokka on the CGview.

may be due to the effect of suspension of the intestinal tract during extraction. Fusobacteria were observed in the EtOH-fixed samples but not in frozen samples. This may stem from difference in the sampling methods. The presence or absence of ethanol-fixation and type of PCR enzymes were not significant. In particular, in the case of *E. stellifer*, the highest diversity was detected in ethanol-fixed samples, indicating that the similar detection is possible with ethanol-fixed fish intestinal contents.

Previous studies have indicated that Proteobacteria, in addition to Bacteroidetes and Firmicutes, comprise 90% of the fish intestinal microbiota²³. Actinobacteria are also found at high frequencies²⁴. The gut of *A. tetractylus*, which is rich in Spirochaetes and Tenericutes, is very different from these previous findings. In marine fish, the genera *Vibrio*, *Photobacterium*, and *Clostridium* were often reported to be dominant. Proteobacteria are commonly found in herbivorous and carnivorous fish⁶. A meta-analysis of the gut communities of marine fish revealed that Vibrionales bacteria, which include the genera *Vibrio* and *Photobacterium*, accounted for 70% of sequence reads⁶. It is reasonable to consider that freshwater fish do not harbor *Vibrio* spp. in their guts, because *Vibrio* are generally found in seawater. In the present study, Firmicutes containing *Clostridium* were only found in very small numbers in half of the six samples, and were rarely found in the algae-eating fish in this study. The other genera identified included *Aeromonas*, *Photobacterium*, *Pseudomonas*, which have all been previously identified in fish gut microbiota that might aid digestion²⁵. *Aeromonas* and *Pseudomonas* are notable in the fish gut

as being both pathogenic and probiotic bacteria, but were only found in very small numbers in the two species in our study. In summary, the absence of Firmicutes and Actinobacteria was what distinguished the algae-eating fish from other species. However, it is known that *Clostridia* also dominate the gut microbial flora in different marine herbivorous fish species²⁶. Among all fish species tested, marine herbivorous fishes have been shown to have a high diversity of gut bacteria⁶. This is in good agreement with the diversity of gut bacteria in the absolute algae-eating *Andamia*, where even the most abundant species of the bacterial phyla are as low as 30%. Marine herbivorous fishes also harbor few of the typical environmental bacteria, yet many close relatives of bacteria from mammalian guts⁶. The presence of short-chain fatty acids in the gut of marine herbivores suggests that herbivorous fish and mammals are similar in the process of enteric fermentation²⁷.

Le and Wang²⁸ found that Brevinemataceae (phylum Spirochaetes) and Mycoplasmataceae (phylum Tenericutes) are the most abundant species in the gut microbiota of the mullet. Diverse Spirochaetes are regularly found deep in marine sediments and soils, and in the digestive tracts of arthropods and several species of mammals. Spirochaetes are commonly detected at low concentrations in the gut microbiota of fish; however, Givens et al.²⁹ reported that one of the three barracuda individuals contained a community comprised of 99% Spirochaetes and the three *Mahi mahi* (common dolphinfish) individuals contained 64–98% Spirochaetes. This feature has also been reported in old gilthead snappers³⁰ and puffer fish³¹. The common diet of these fish species is difficult to clarify, but may be related to the predation of the whole sediment.

The composition of Fusobacteria and Proteobacteria in *E. stellifer* is similar to that found in other marine fish, such as *Silurus* and *Carassium*²⁴, and to results reported for freshwater garpike³². A few studies have shown Fusobacteria to be the dominant members of the gut microbiota of freshwater fish^{33,34}. Although *E. stellifer* is stomachless, this species is considered to be an omnivore, and the above tendency is similar to that of zooplankton-eating fish species. In addition, Fusobacteria are anaerobic gram-negative rods that produce butyric acid³⁵, a short-chain fatty acid that is the end product of fermentation of carbohydrates, including those found in mucins^{36,37}. This fatty acid has been found in the guts of herbivorous and omnivorous fish^{27,38}. Nuez-Ortín et al.³⁹ demonstrated the ability of butyric acid to inhibit potential pathogens in freshwater fish, and butyrate sodium is now marketed as a food additive to promote fish health and growth. This is likely to be the case for similar probiotics in marine fish, and further validation of their efficacy is expected.

Interestingly, the novel Mycoplasma phylotype was found to be predominant in wild Atlantic salmon in Scotland and in fish raised in pens, while *Acinetobacter junii* was predominant in farmed fish in Norway. The farmed fish in these two locations were fed different diets. In another study that examined changes in the gut microbiota of salmon over their life cycle, it was observed that Proteobacteria were predominant at all stages, with an enrichment of Tenericutes (especially *Mycoplasma* spp.⁴⁰). Tenericutes predominant gut microbiota is also found in some fish species, but is difficult to culture and has not yet been studied.

In this study, many *Vibrio* spp. were isolated, in addition to some strains of *Shewanella*, *Photobacterium*, and *Bacillus*, which were obtained only in *E. stellifer*. In comparison with NGS, minor strains were isolated, but strains that matched the culture conditions may have been prioritized. Among the strains, we found a very clear boundary that resembled an artificial line drawn that did not merge among the isolated *Vibrio* spp. The coordinated behavior of swarming was known to enable *Vibrio* spp. to colonize surfaces, coordinate behavior, and form multicellular communities⁴¹. The peritrichous lateral flagella is essential for swarming and the expression of flagella forming genes are regulated by the quorum sensing regulators *aphA* and *opaR*⁴². We found *opaR* in the genomes of *V. alginolyticus* YZOS-03 and HLBS-07, annotated as *luxR*, and the sequence similarity between these strains is 100% among *Vibrio* spp. On the other hand, *aphA* is apparently absent both in the YZOS-03 and HLBS-07 strains. Swarming ability of *Vibrio* spp. including YZOS-03 and HLBS-07 strains appears to be necessary for boundary formation, but its mechanistic analysis in *Vibrio* spp. has not been conducted. The molecular mechanism of boundary formation as a self-recognition mechanism has been studied mostly in *Proteus mirabilis*. A six-gene operon, *idsA-idsF*, of *P. mirabilis* has been identified as important for self versus non-self discrimination by forming a boundary between two strains⁴³. We did not find apparent orthologs of *idsA-F* genes in the genomes of YZOS-03 and HLBS-07. Curiously, we found type I-F CRISPR-associated endoribonuclease Cas6/Csy4 and CRISPR-associated protein Cys3 in HLBS-07, but not in YZOS-03 or any other *V. alginolyticus* strains. This is a unique feature of the HLBS-07 strain.

V. alginolyticus is a halophilic anaerobic gram-negative bacterium frequently found in marine environments, and some species cause epidemic vibriosis. Whole genome sequencing of *V. alginolyticus* has been previously reported^{44–48}. The genome size of 71 reported *V. alginolyticus* strains varied between 6.17 and 3.94 Mb with predicted ORF between 5619 and 3908. Among the 71 published *V. alginolyticus* genomes, YZOS-03 (5.14 Mb) is relatively small and HLBS-07 (5.33 Mb) is a medium-sized strain. Among these *Vibrio* genomes, the YZOS-03 and HLBS-07 strains were close to G3_MYPK1 and 2014V-1072, respectively. YZOS-03 and HLBS-07 were 98.7% identical. Both strains encoded two alginate lyases, which are commonly found in many *Vibrio* spp.¹⁸. The genes encoding alginate lyases, amylases, chitinases, and mannosidase are present in both strains, suggesting that the bacteria assist in the digestion of polysaccharides.

In summary, we investigated the bacteria of two stomachless fish species, *A. tetractylus* and *E. stellifer*, and found a diversity in their habitats based on the direct sequencing of 16S rDNA. The isolated bacteria had some specific features including boundary formation in the two species. Combined with genomic DNA analysis of representative two *Vibrio* spp., the current analysis is useful for understanding of such mechanisms and for future application in the biotechnology field.

Materials and methods

Sample collection and rearing. We collected twenty Rockskipper, *A. tetradactyla*, from the coast of the Iriomote Island, Okinawa Prefecture, twelve Stellar Rockskipper, *E. stellifer*, from in the coast of Shimane Peninsula, Shimane Pref., in Japan.

To control for the influence of seasonal differences and growth stages, only adult-stage specimens were used. Because they are wild animals, it was difficult to randomize the sex and age of the animals. All specimens were caught using hand nets between the months of June and November from 2007 to 2019. The Stellar Rockskipper were maintained in 60 cm glass tanks (60 L) with running water under optimal oxygenation conditions, and the temperature was maintained at 22.0 ± 1.0 °C throughout the 3-day experimental period. The study protocol was approved by the animal ethics committee of Shimane University, and was conducted in strict adherence with the guidelines for the care and use of research animals set out by the committee and in compliance with the ARRIVE guidelines v1.0 since our experiments have done prior to ARRIVE guidelines v2.0 published on 2020⁴⁹. Our experiments have done are planned with the following points in mind. To avoid overexploitation of wild populations, we used a minimal experimental scale including preliminary experiments, 2–3 individuals each for morphological observations, gut microbe isolation and metagenomic analysis, and totally 10 individuals for both species. The main purpose of this study was to describe the gut bacteria, and no study was conducted that required a control group. The age of the fish was 1 year old adult fish, and egg-carrying individuals were not used in the experiment, so that the individuals used in the experiment were males. After ice-cold anesthesia, the animals were quickly decapitated and subjected to humane killing.

Tissue preparation. Excised samples of digestive tracts fixed with 1.5% glutaraldehyde were rinsed, dehydrated, conductive-stained by 2.0% tannic acid and 1.0% osmium tetroxide, and freeze-dried with t-butyl alcohol. Samples were subsequently coated with platinum and observed under a digital scanning electron microscope (S-4800; Hitachi High-Technologies Corp.).

Flora analysis. Whole flesh frozen bodies of *A. tetradactylus* and *E. stellifer* were utilized for feces collection. For the ethanol-fixed samples, frozen fish was directly immersed in the 70% ethanol/milliQ water. Before dissection, tweezers were sterilized with 70% ethanol and UV light irradiation for 10 min. Fish bodies were dissected under a binocular microscope. The feces were roughly homogenized using tweezers in the C1 buffer included in the DNA Powersoil kit (QIAGEN). For the small *Andamia* fish, the whole gut was resuspended in the Powerbeads tube and utilized for the following procedure. DNA extraction was performed according to the manufacturer's instruction.

The V4 region of the bacterial 16S rRNA gene was amplified using a 2 ng of DNA template with primers 515 F (5'-GTGCCAGCMGCCGCGTAA-3') and 806R (5'-GGACTACHVGGGTWTCTAAT-3'). PCR reactions were performed in 20 µL with ExTaq HS (Takara Bio Inc.) or PrimeSTAR GXL DNA Polymerase (Takara Bio Inc.) at an annealing temperature of 50 °C for 20 cycles. The PCR products were purified with the FastGene PCR Extraction kit and sent to Fasmac Inc, Japan. The amplicon library was sequenced by Illumina MiSeq 2 × 300 bp paired-end platform according to the manufacturer's instruction. The MiSeq fastq reads are available in the DNA Databank of Japan (DDBJ) Sequence Read Archive (DRA) under the accession number DRA011079.

Read data was analyzed using QIIME 2⁵⁰. Paired-end sequences imported into QIIME 2 were quality-controlled and combined using DADA2 (–p-trunc-len-f 240 –p-trunc-len-r 200 –p-trim-left-f 19 –p-trim-left-r 20)⁵¹. The settings for quality control were based on the reads' quality distribution along the length of the sequence. Alpha rarefaction analysis, OTU, alpha diversity (number of observed OTUs, Shannon diversity, Faith phylogenetic diversity), and beta diversity (Jaccard distance, Bray–Curtis distance, unweighted and weighted UniFrac distance) were analyzed using QIIME 2. To evaluate the impact of the difference in the number of reads between samples, the alpha rarefaction curve was plotted with 5000 sampling depths. For taxonomic classification, the “Greengenes 13_8 99% OTUs” dataset⁵² was utilized as 16S rRNA gene databases. In order to make a comparison excluding bacteria with low frequency of appearance, we analyzed the readings with low frequency of appearance less than 1,500 times (–p-min-frequency 1500). The beta diversity metric is an estimation of the between-sample diversity of the microbial profile and it was calculated by the QIIME 2 “diversity beta-group-significance” script. The PCoA analysis of has been done with the qiime2R scripts (qiime2R: Importing QIIME2 artifacts and associated data into R sessions. Jordan E Bisanz (2018) <https://github.com/jbisanz/qiime2R>.) Statistical analyses for diversity metrics and ANCOM were also done through QIIME 2. Box plots in a Qurro visualization are generated by Altair-4.1.0⁵³ with vega-datasets-0.9.0⁵⁴ in Qurro's Python code.

Linear discriminant analysis Effect Size (LefSe) was used to identify OTUs that explain differences between treatments using Kruskal–Wallis tests and a Linear Discriminant Analysis (LDA) threshold of 3.0⁵⁵. p-values below 0.05 were considered significant. For functional inferences of the microbial community, we conducted Phylogenetic Investigation of Communities by Reconstruction of Unobserved States 2 (PICRUSt2) (v2.2.0-b)⁵⁶ with ASVs according to the instructions published at <https://github.com/picrust/picrust2/wiki>. PICRUSt2 predictions were supported by Enzyme Classification numbers (EC numbers) and COGs—Clusters of Orthologous Groups (as of 13 Dec 2021). We generated PICRUSt2 EC and COG gene family predictions. The results were visualized in statistical analysis of taxonomic and functional profiles (STAMP) version 2.1.3⁵⁷ and tested using Welch's *t* test for two groups of species. All predictions were corrected for multiple testing (Benjamini–Hochberg method, FDR $q < 0.05$).

Isolation of intestinal bacteria from the gastrointestinal tract of *Andamia tetradactylus* and *Entomacrodus stellifer*. Isolation of bacteria from *A. tetradactylum* was carried out on 15th December in 2017 and 3rd April in 2018, and it was on 20th July and 6th August in 2018 for *E. stellifer*. The gastrointestinal

tracts were processed as follows. The gastrointestinal tract was finely chopped with scissors and suspended in 10 mL of seawater diluent and vortexed. 100 µL of them was spread in LB and Zobell solid media. They were incubated at 25 °C for about 2–7 days. Several colonies of distinct color and colony shape were isolated. The isolated strains were stocked in LB 15% Glycerol medium or Zobell 15% Glycerol medium at – 80 °C.

As a result of bacterial isolation, a total of 78 bacterial strains were isolated including 20 and 14 strains in the LB and Zobell media, respectively, from *A. tetradactylus*, and 24 and 20 strains in the LB and Zobell media, respectively, from *E. stellifer*. From these strains, genome DNA was extracted and a region of 16S rDNA was amplified to determine the species using the following primer set (5'-AGAGTTTGATCCTGGCTCAG-3' and 5'-GGTTACCTGTTACGACTT-3'). A phylogenetic tree was constructed by the ClustalW program (DDBJ, Japan) and visualized by FigTree⁵⁸.

Genome sequencing. Genomic DNAs of YZOS-03 isolated from *A. tetradactylus* and HLBS-07 strain isolated from *E. stellifer* were purified by zirconia bead crushing and subsequent ethanol precipitation. Sequencing libraries were prepared using the Rapid Sequencing Kit (SQK-RAD004, Oxford Nanopore Technologies (ONT)) from unfragmented genomic DNA (approximately 400 ng input). Long-read sequencing was performed using the minION platform with an R9.4.1 flow cell (FLO-MIN106). Sequencing quality was monitored with the MinKNOW interface 19.12.5. After 24 h sequencing output fast5 files were basecalled using the Guppy software v3.6.0 (via docker GenomicParisCentre/guppy-gpu⁵⁹).

For the assembly process, a total of 1.0 Gbps reads were quality-filtered and subsampled by Filtlong v0.2.0 to filter by lengths of 5000 bp (-min_length 5000) and quality of 1 (-min_mean_q 90), resulting in around 200× expected coverage. The subsampled reads were assembled with Canu v2.0⁶⁰ or Flye v2.7.1⁶¹. Consensus assembly was called using MarginPolish v1.3.dev-5492204⁶² or Racon (four times repetition⁶³) with a total of raw reads, followed by Medaka v0.11.5, sequence correction software provided by ONT Research (via docker nanozoo/medaka, <https://github.com/nanoporetech/medaka>). Completeness of the assemblies of each step was tested with BUSCO v4.6.0 with lineage option “vibrionales_odb10”. Annotation was performed using DFAST v. 1.2.4 (2019.11.7) from the DDBJ⁶⁴. Circular visualization of the genomes was carried out in the CGview server v1.0⁶⁵.

Data availability

The MiSeq fastq reads for Flora analysis are available in the DRA under the accession number DRA011079. The chromosome sequences were deposited in DDBJ under the accession numbers AP023185 and AP023186 for YZOS-03, and AP023187 and AP023188 for HLBS-07. Raw ONT reads are also available in the DRA under the accession number DRA010239 and the BioProject accession number PRJDB9917 for YZOS-03, and DRA010240 and the BioProject accession number PRJDB9917 for HLBS-07.

Received: 21 June 2021; Accepted: 7 March 2022

Published online: 17 March 2022

References

- Ramotar, K., Conly, J. M., Chubb, H. & Louie, T. J. Production of menaquinones by intestinal anaerobes. *J. Infect. Dis.* **150**, 213–218 (1984).
- Conly, J. M. & Stein, K. The production of menaquinones (vitamin K2) by intestinal bacteria and their role in maintaining coagulation homeostasis. *Prog. Food Nutr. Sci.* **16**, 307–343 (1992).
- Duttaroy, A. K. Chapter 9—Polysaccharides on the gut microbiome and epigenome. In *Evidence-Based Nutrition and Clinical Evidence of Bioactive Foods in Human Health and Disease* (ed. Duttaroy, A. K.) 129–137 (Academic Press, 2021).
- Hehemann, J. H. *et al.* Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. *Nature* **464**(7290), 908–912. <https://doi.org/10.1038/nature08937> (2010).
- Roeselers, G. *et al.* Evidence for a core gut microbiota in the zebrafish. *ISME J.* **5**, 1595–1608 (2011).
- Sullam, K. E. *et al.* Environmental and ecological factors that shape the gut bacterial communities of fish: A meta-analysis. *Mol. Ecol.* **21**(13), 3363–3378. <https://doi.org/10.1111/j.1365-294X.2012.05552.x> (2012).
- Wu, S. *et al.* Composition, diversity, and origin of the bacterial community in grass carp intestine. *PLoS ONE* **7**(2), e30440. <https://doi.org/10.1371/journal.pone.0030440> (2012).
- Lai, K. P. *et al.* Osmotic stress induces gut microbiota community shift in fish. *Environ. Microbiol.* **22**(9), 3784–3802. <https://doi.org/10.1111/1462-2920.15150> (2020).
- Akiyoshi, H. & Inoue, A. Comparative histological study of teleost livers in relation to phylogeny. *Zool. Sci.* **21**(8), 841–850. <https://doi.org/10.2108/zsj.21.841> (2004).
- Kimura, S. *New Atlas of Fish Anatomy* 1st edn. (Midori Shobo Publishing, 2010).
- Wilson, J. M. & Castro, L. F. C. Morphological diversity of the gastrointestinal tract in fishes. In *Fish Physiology: The Multifunctional Gut of Fish* (eds Grosell, M. *et al.*) 1–55 (Academic Press, 2010).
- Hossain, A. M. & Dutta, H. M. Phylogeny, ontogeny, structure and function of digestive tract appendages (caeca) in teleost fish. In *Fish Morphology Horizon of New Research* (eds Munshi, J. S. D. & Dutta, H. M.) 59–76 (A.A Balkema Publishers, 1996).
- Kent, G. C. & Carr, R. K. Digestive system. In *Comparative Anatomy of the Vertebrates* 4th edn (eds Kent, G. C. & Carr, R. K.) 235–255 (McGraw Hill, 2000).
- Romer, A. S. & Parsons, T. S. Digestive system. In *The Vertebrate Body* 5th edn (eds Romer, A. S. & Parsons, T. S.) 306–325 (Holt Rinehart & Winston, 2007).
- López Nadal, A. *et al.* Feed, microbiota, and Gut immunity: Using the zebrafish model to understand fish health. *Front. Immunol.* **11**, 114. <https://doi.org/10.3389/fimmu.2020.00114> (2020).
- Robinson, C. D. *et al.* Experimental bacterial adaptation to the zebrafish gut reveals a primary role for immigration. *PLoS Biol.* **16**(12), e2006893. <https://doi.org/10.1371/journal.pbio.2006893> (2018).
- Rawls, J. F., Mahowald, M. A., Ley, R. E. & Gordon, J. I. Reciprocal gut microbiota transplants from zebrafish and mice to germ-free recipients reveal host habitat selection. *Cell* **127**(2), 423–433. <https://doi.org/10.1016/j.cell.2006.08.043> (2006).
- Kawamoto, H. *et al.* Cloning and sequencing analysis of alginate lyase genes from the marine bacterium *Vibrio* sp. O2. *Mar. Biotechnol. (New York)* **8**(5), 481–490. <https://doi.org/10.1007/s10126-005-6157-z> (2006).

19. Dharani, S. R., Srinivasan, R., Sarath, R. & Ramya, M. Recent progress on engineering microbial alginate lyases towards their versatile role in biotechnological applications. *Folia Microbiol.* **65**(6), 937–954. <https://doi.org/10.1007/s12223-020-00802-8> (2020).
20. Park, J.-K. *et al.* Molecular cloning, nucleotide sequencing, and regulation of the *chiA* gene encoding one of chitinases from *Enterobacter* sp. G-1. *J Ferment Bioeng.* **84**, 493–501 (1997).
21. Amakata, D. *et al.* *Mitsuaria chitosanitabida* gen. nov., sp. Nov., an aerobic, chitosanase-producing member of the “Betaproteobacteria”. *Int. J. Syst. Evol. Microbiol.* **55**(Pt 5), 1927–1932. <https://doi.org/10.1099/ij.s.0.63629-0> (2005).
22. Kawamukai, M. Biosynthesis and applications of prenylquinones. *Biosci. Biotechnol. Biochem.* **82**(6), 963–977. <https://doi.org/10.1080/09168451.2018.1433020> (2018).
23. Ghanbari, M., Kneifel, W. & Domig, K. J. A new view of the fish gut microbiome: Advances from next-generation sequencing. *Aquaculture* **448**, 464–475 (2015).
24. Talwar, C., Nagar, S., Lal, R. & Negi, R. K. Fish gut microbiome: Current approaches and future perspectives. *Indian J. Microbiol.* **58**(4), 397–414. <https://doi.org/10.1007/s12088-018-0760-y> (2018).
25. Ray, A. K., Ghosh, K. & Ringø, E. Enzyme-producing bacteria isolated from fish gut: A review. *Aquac. Nutr.* **18**, 465–492 (2012).
26. Clements, K. D., Pasch, I. B. Y., Moran, D. & Turner, S. J. Clostridia dominate 16S rRNA gene libraries prepared from the hindgut of temperate marine herbivorous fishes. *Mar. Biol.* **150**, 1431–1440 (2007).
27. Clements, K. D. & Choat, J. H. Fermentation in tropical marine herbivorous fishes. *Physiol. Zool.* **68**(3), 355–378 (1995).
28. Le, M. H. & Wang, D. Structure and membership of gut microbial communities in multiple fish cryptic species under potential migratory effects. *Sci. Rep.* **10**(1), 7547. <https://doi.org/10.1038/s41598-020-64570-8> (2020).
29. Givens, C. E., Ransom, B., Bano, N. & Hollibaugh, J. T. Comparison of the gut microbiomes of 12 bony fish and 3 shark species. *Mar. Ecol. Prog. Ser.* **518**, 209–223 (2015).
30. Piazzon, M. C. *et al.* Sex, age, and bacteria: How the intestinal microbiota is modulated in a protandrous hermaphrodite fish. *Front. Microbiol.* **10**, 2512. <https://doi.org/10.3389/fmicb.2019.02512> (2019).
31. Li, Z. *et al.* Puffer fish gut microbiota studies revealed unique bacterial co-occurrence patterns and new insights on tetrodotoxin producers. *Mar. Drugs* **18**(5), 278. <https://doi.org/10.3390/md18050278> (2020).
32. Méndez-Pérez, R. *et al.* High-throughput sequencing of the 16S rRNA gene to analyze the gut microbiome in juvenile and adult tropical gar (*Atractosteus tropicus*). *Lat. Am. J. Aquat. Res.* **48**(3), 456–479 (2020).
33. van Kessel, M. A. *et al.* Pyrosequencing of 16S rRNA gene amplicons to study the microbiota in the gastrointestinal tract of carp (*Cyprinus carpio* L.). *AMB Express* **1**, 41. <https://doi.org/10.1186/2191-0855-1-41> (2011).
34. Di Maiuta, N., Schwarzenruber, P., Schenker, M. & Schoelkopf, J. Microbial population dynamics in the faeces of wood-eating loricariid catfishes. *Lett. Appl. Microbiol.* **56**(6), 401–407. <https://doi.org/10.1111/lam.12061> (2013).
35. Bennett, K. W. & Eley, A. Fusobacteria: New taxonomy and related diseases. *J. Med. Microbiol.* **39**(4), 246–254. <https://doi.org/10.1099/00222615-39-4-246> (1993).
36. Titus, E. & Ahearn, G. A. Short-chain fatty acid transport in the intestine of a herbivorous teleost. *J. Exp. Biol.* **135**, 77–94 (1988).
37. von Engelhardt, W., Bartels, J., Kirschberger, S., Meyer-zu-Düttlingdorf, H. D. & Busche, R. Role of short-chain fatty acids in the hind gut. *Vet. Q.* **20**, S52–S59 (1998).
38. Clements, K. D., Gleeson, V. P. & Slaytor, M. Short-chain fatty acid metabolism in temperate marine herbivorous fish. *J. Comp. Physiol. B* **164**, 372–377 (1994).
39. Nuez-Ortín, W. G., Prado, S. & Toranzo, A. E. Antimicrobial properties of butyric acid and other organic acids against pathogenic bacteria affecting the main aquatic species. In *Conference proceedings aqua conference 2012, Prague, Czech Republic* (2012).
40. Llewellyn, M. S. *et al.* The biogeography of the atlantic salmon (*Salmo salar*) gut microbiome. *ISME J.* **10**(5), 1280–1284. <https://doi.org/10.1038/ismej.2015.189> (2016).
41. McCarter, L. L. Polar flagellar motility of the Vibrionaceae. *Microbiol. Mol. Biol. Rev.* **65**(3), 445–462. <https://doi.org/10.1128/MMBR.65.3.445-462.2001> (2001).
42. Lu, R. *et al.* The quorum sensing regulator OpaR is a repressor of polar flagellum genes in *Vibrio parahaemolyticus*. *J. Microbiol. (Seoul)*. <https://doi.org/10.1007/s12275-021-0629-3> (2021).
43. Gibbs, K. A., Urbanowski, M. L. & Greenberg, E. P. Genetic determinants of self identity and social recognition in bacteria. *Science (New York)* **321**(5886), 256–259. <https://doi.org/10.1126/science.1160033> (2008).
44. Liu, X. F., Cao, Y., Zhang, H. L., Chen, Y. J. & Hu, C. J. Complete genome sequence of *Vibrio alginolyticus* ATCC 17749T. *Genome Announc.* **3**(1), e01500-e1514. <https://doi.org/10.1128/genomeA.01500-14> (2015).
45. Bhotra, T. & Singh, D. V. Whole-genome sequence of *Vibrio alginolyticus* isolated from the Mucus of the Coral Fungia danai in the Andaman Sea, India. *Genome Announc.* **4**(3), e00339. <https://doi.org/10.1128/genomeA.00339-16> (2016).
46. Deng, Y. *et al.* Complete genome sequence of *Vibrio alginolyticus* ZJ-T. *Genome Announc.* **4**(5), e00912-e916. <https://doi.org/10.1128/genomeA.00912-16> (2016).
47. Wang, P., Wen, Z., Li, B., Zeng, Z. & Wang, X. Complete genome sequence of *Vibrio alginolyticus* ATCC 33787(T) isolated from seawater with three native megaplasmids. *Mar. Genomics* **28**, 45–47. <https://doi.org/10.1016/j.margen.2016.05.003> (2016).
48. Gao, M. *et al.* Characteristics and complete genome sequence of the virulent *Vibrio alginolyticus* phage VAP7, isolated in Hainan, China. *Arch. Virol.* **165**(4), 947–953. <https://doi.org/10.1007/s00705-020-04535-4> (2020).
49. Kilkenny, C., Browne, W. J., Cuthill, I. C., Emerson, M. & Altman, D. G. Improving bioscience research reporting: The ARRIVE guidelines for reporting animal research. *PLoS Biol.* **8**(6), e1000412. <https://doi.org/10.1371/journal.pbio.1000412> (2010).
50. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**(8), 852–857. <https://doi.org/10.1038/s41587-019-0209-9> (2019).
51. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**(7), 581–583. <https://doi.org/10.1038/nmeth.3869> (2016).
52. DeSantis, T. Z. *et al.* Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **72**(7), 5069–5072. <https://doi.org/10.1128/AEM.03006-05> (2006).
53. VanderPlas, J. *et al.* Altair: Interactive statistical visualizations for python. *J. Open Source Softw.* **3**, 1057 (2018).
54. Fedarko, M. W. *et al.* Visualizing ‘omic feature rankings and log-ratios using Qurro. *NAR Genomics Bioinform.* **2**(2), 023. <https://doi.org/10.1093/nargab/lqaa023> (2020).
55. Segata, N. *et al.* Metagenomic biomarker discovery and explanation. *Genome Biol.* **12**, R60. <https://doi.org/10.1186/gb-2011-12-6-r60> (2011).
56. Douglas, G. M. *et al.* PICRUSt2: An improved and customizable approach for metagenome inference. *BioRxiv*. <https://doi.org/10.1101/672295> (2020).
57. Parks, D. H., Tyson, G. W., Hugenholtz, P. & Beiko, R. G. STAMP: Statistical analysis of taxonomic and functional profiles. *Bioinformatics* **30**, 3123–3124 (2014).
58. Rambaut, A. *FigTree v1.4.2. Computer Program* (2009). Retrieved 4 Oct 2016, from <http://tree.bio.ed.ac.uk/software/figtree/>.
59. Wick, R. R., Judd, L. M. & Holt, K. E. Performance of neural network basecalling tools for Oxford Nanopore sequencing. *Genome Biol.* **20**(1), 129. <https://doi.org/10.1186/s13059-019-1727-y> (2019).
60. Koren, S. *et al.* Canu: Scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* **27**(5), 722–736. <https://doi.org/10.1101/gr.215087.116> (2017).
61. Kolmogorov, M., Yuan, J., Lin, Y. & Pevzner, P. A. Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* **37**(5), 540–546. <https://doi.org/10.1038/s41587-019-0072-8> (2019).

62. Shafin, K. *et al.* Nanopore sequencing and the Shasta toolkit enable efficient de novo assembly of eleven human genomes. *Nat. Biotechnol.* **38**(9), 1044–1053. <https://doi.org/10.1038/s41587-020-0503-6> (2020).
63. Vaser, R., Sović, I., Nagarajan, N. & Šikić, M. Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res.* **27**(5), 737–746. <https://doi.org/10.1101/gr.214270.116> (2017).
64. Tanizawa, Y., Fujisawa, T. & Nakamura, Y. DFAST: A flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics (Oxford)* **34**(6), 1037–1039. <https://doi.org/10.1093/bioinformatics/btx713> (2018).
65. Stothard, P. & Wishart, D. S. Circular genome visualization and exploration using CGView. *Bioinformatics (Oxford)* **21**(4), 537–539. <https://doi.org/10.1093/bioinformatics/bti054> (2005).

Acknowledgements

The work by H.A., M.K. and M.A.Y. was supported by Grant-in-Aid for Scientific Research (C) [17KT0148] from Japan Society for the Promotion of Science (JSPS), Nissui fund for M.K. and Mishima Kaiun Memorial Foundation for M.A.Y. Computations were partially performed on the NIG supercomputer at ROIS National Institute of Genetics. We thank Yawei Lu, Miyako Toyama, Hiromu Kashima, Hayato Hironaka, Hikaru Morita and Masashi Kimura for technical assistance. We thank Dr. Eiji Ishii (Osaka Univ.) for his opinion on boundary formation. We also thank the Faculty of Life and Environmental Sciences at Shimane University for the financial support for publishing this report.

Author contributions

H.A. and M.K. conceived and planned the experiments. M.A.Y., H.A. and M.K. obtained research funding. M.A.Y. and T.T. carried out the experiments for informatics and cultures, respectively. H.A. and T.T. contributed to sample preparation. M.A.Y. took the lead in writing the manuscript. All authors contributed to the interpretation of the results and provided critical feedback and helped developing the research and the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-08511-7>.

Correspondence and requests for materials should be addressed to M.A.Y. or M.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022