AMERICAN SOCIETY FOR MICROBIOLOGY | mBio

Check for updates

| Environmental Microbiology | Research Article

Enrichable consortia of microbial symbionts degrade macroalgal polysaccharides in *Kyphosus* fish

Aaron Oliver,[1] Sheila Podell,[1] Linda Wegley Kelly,[2] Wesley J. Sparagon,[3] Alvaro M. Plominsky,[2] Robert S. Nelson,[4] Lieve M. L. Laurens,[4] Simona Augyte,[5] Neil A. Sims,[5] Craig E. Nelson,[3] Eric E. Allen[1,6,7]

AUTHOR AFFILIATIONS See affiliation list on p. 14.

ABSTRACT Coastal herbivorous fishes consume macroalgae, which is then degraded by microbes along their digestive tract. However, there is scarce genomic information about the microbiota that perform this degradation. This study explores the potential of *Kyphosus* gastrointestinal microbial symbionts to collaboratively degrade and ferment polysaccharides from red, green, and brown macroalgae through *in silico* study of carbohydrate-active enzyme and sulfatase sequences. Recovery of metagenome-assembled genomes (MAGs) from previously described *Kyphosus* gut metagenomes and newly sequenced bioreactor enrichments reveals differences in enzymatic capabilities between the major microbial taxa in *Kyphosus* guts. The most versatile of the recovered MAGs were from the *Bacteroidota* phylum, whose MAGs house enzyme collections able to decompose a variety of algal polysaccharides. Unique enzymes and predicted degradative capacities of genomes from the *Bacillota* (genus *Vallitalea*) and *Verrucomicrobiota* (order *Kiritimatiellales*) highlight the importance of metabolic contributions from multiple phyla to broaden polysaccharide degradation capabilities. Few genomes contain the required enzymes to fully degrade any complex sulfated algal polysaccharide alone. The distribution of suitable enzymes between MAGs originating from different taxa, along with the widespread detection of signal peptides in candidate enzymes, is consistent with cooperative extracellular degradation of these carbohydrates. This study leverages genomic evidence to reveal an untapped diversity at the enzyme and strain level among *Kyphosus* symbionts and their contributions to macroalgae decomposition. Bioreactor enrichments provide a genomic foundation for degradative and fermentative processes central to translating the knowledge gained from this system to the aquaculture and bioenergy sectors.

IMPORTANCE Seaweed has long been considered a promising source of sustainable biomass for bioenergy and aquaculture feed, but scalable industrial methods for decomposing terrestrial compounds can struggle to break down seaweed polysaccharides efficiently due to their unique sulfated structures. Fish of the genus *Kyphosus* feed on seaweed by leveraging gastrointestinal bacteria to degrade algal polysaccharides into simple sugars. This study reconstructs metagenome-assembled genomes for these gastrointestinal bacteria to enhance our understanding of herbivorous fish digestion and fermentation of algal sugars. Investigations at the gene level identify *Kyphosus* guts as an untapped source of seaweed-degrading enzymes ripe for further characterization. These discoveries set the stage for future work incorporating marine enzymes and microbial communities in the industrial degradation of algal polysaccharides.

KEYWORDS *Kyphosus*, fish gut microbiome, macroalgal polysaccharides, sulfatase

The *Kyphosus* genus of herbivorous fish, commonly referred to as nenue or rudderfish, graze primarily on macroalgae (1). *Kyphosus* fish serve important ecological roles by

Editor Jennifer B. H. Martiny, University of California, Irvine, Irvine, California, USA

Address correspondence to Aaron Oliver, a3oliver@ucsd.edu, or Eric E. Allen, eallen@ucsd.edu.

The authors declare no conflict of interest.

See the funding table on p. 15.

Received 21 February 2024
Accepted 6 March 2024
Published 27 March 2024

controlling algal cover in Indo-Pacific (2) and Caribbean coral reefs (3), thereby mediating coral-algal competition, overall coral growth, and benthic community composition (4). Their diverse diet includes macroalgae from the three major taxonomic groups: Rhodophyta (red), Chlorophyta (green), and Ochrophyta (brown) (1). Polysaccharides constitute as much as 60% of macroalgal cells by weight (5) and serve roles in both cell structure and energy storage (6). The complex network of linkages in structural polysaccharides resists degradation from chemical and enzymatic stressors and serves as a physical defense mechanism for algal cells (7).

Algal polysaccharides differ from common polysaccharides found in land plants due to the addition of sulfate ester groups (8). Structural polysaccharides from red algae include agar, carrageenan, porphyran, and xylan, which all contain such sulfate groups (9). Brown algae contain the sulfated polysaccharide fucoidan for structure as well as unsulfated alginate as a storage polysaccharide (9). Green algae contain sulfated polysaccharides such as xylan and ulvan but also contain large amounts of unsulfated cellulose common in land plants (9). Algal polysaccharides are depolymerized primarily through the enzymatic activity of bacterial glycoside hydrolases (GHs) and polysaccharide lyases (10), two classes of carbohydrate-active enzymes (CAZymes) (11). Sulfated polysaccharides are particularly recalcitrant to digestion because an additional enzyme class, the sulfatases, is necessary for complete degradation. Full enzyme pathways for the breakdown of various algal polysaccharides have been proposed (9, 12) that include both required CAZyme and sulfatase activities. However, not all algal polysaccharides have well-defined degradation pathways or unique associated CAZymes that enable a high-level connection between gene presence and catabolized substrates. Likewise, sulfatase classes within the SulfAtlas database (13) are primarily classified based on evolutionary history rather than substrate specificity or enzymatic activity, so our ability to evaluate pathway completeness *in silico* is limited.

Once complex carbohydrates are broken into subunits by CAZymes and sulfatases, they are utilized by gut microbiota in fermentation reactions to produce short-chain fatty acids (SCFAs) (14). The SCFAs acetate, propanoate, and butyrate have been previously measured in high quantities in *Kyphosus* hindguts (15) and are utilized by the host fish for energy (16). Previous work has suggested correlations between SCFA profiles and bacterial composition (15), but there is no genomic work in algivorous fish pinpointing which microbiota contribute to host nutrition in this way and what pathways are utilized to produce these essential SCFAs.

Our overall understanding of the role of gut microbiota in digestion is still limited in most fishes (17), including *Kyphosus*, in part due to a focus on community composition and diversity rather than function. The genetic study of *Kyphosus* gut symbionts has been limited to 16S rRNA (15, 18) and metabolomic (18) investigations until the incorporation of shotgun metagenomics in a few recent studies (19, 20). What functional profiling has been done in fish guts often relies on extrapolation from amplicon-based taxonomic distributions (21–24), and no study has yet generated a large collection of metagenome-assembled genomes (MAGs) from an algivorous fish gut. A *de novo* genomic investigation of *Kyphosus* symbionts has the potential to reveal degradative capacities that cannot be extrapolated from taxonomic lineage or relatedness to database representatives.

Discoveries from better-studied human gut and terrestrial herbivore systems provide suggestions for how *Kyphosus* symbionts might gain and use such gene pathways. Human gut bacteria have acquired enzymes that degrade sulfated algal polysaccharides through horizontal gene transfer (25, 26). Horizontal gene transfer of antibiotic resistance genes has also been observed in fish gut biofilms (27), but this phenomenon has not yet been reported for carbohydrate-active enzymes in any fish gut symbiont microbe. Once acquired, CAZymes and sulfatases potentially originating from one or multiple organisms may then decompose algal polysaccharides in complex, stepwise pathways. A cooperative division of labor strategy, in which partial breakdown products from one bacterial population serve as a degradative substrate for other bacteria in the community, has

been proposed to occur in human gut microbiota (28) and has been suggested as a way to improve polysaccharide degradation in engineered communities (29). The degree to which collaboration may occur in the herbivorous fish gastrointestinal tract remains unknown.

Exploring functional diversity not only improves our understanding of herbivorous fish digestion but may also enable concrete applications in the fields of aquaculture and bioenergy. Most aquaculture is currently sustained through compound feeds that are composed of fishmeal and fish oils from wild-caught fish (30). Although innovations in aquaculture feed have lowered the trophic levels of captive carnivorous fish and improved overall feed efficiency (31), concerns about sustainability and food security remain. Wan et al. (32) argue that the discovery of efficient methods to degrade complex polysaccharides and enhance nutrient digestibility is a key knowledge gap and barrier limiting macroalgae inclusion into commercial aquafeeds (32). Macroalgal feed additives are also known to counteract methanogenesis in terrestrial ruminants (33) and thus can be applied to reduce methane emissions from livestock husbandry. However, deficiencies in ruminant microbiome digestive capacities may influence the future development and long-term success of seaweed dietary supplementation strategies. Research on *Kyphosus* symbionts and their enzymes can inspire commercializable and scalable methods to break down these barriers in the industry.

Innovations exploiting the experimental propagation of enrichment cultures with *Kyphosus* symbionts can harness these microbial communities for further study and experimentation with commercial outputs in the bioenergy sector as well as the development of macroalgal feed supplements. While a few bacterial isolates have been recovered and sequenced from kyphosid guts (34), no previous study has enriched entire communities from these fishes to investigate their hydrolytic and fermentative capabilities. Hydrolysis of carbohydrates, proteins, and lipids into their monomeric components is a key step in biogas and bioethanol production from macroalgae (35, 36), and the degradation of algal polysaccharides is often the rate-limiting step in anaerobic digestion (37). Milledge et al. (38) call for future studies to look beyond commercially available enzymes to discover candidates that can more efficiently degrade algal polysaccharides (38). The *Kyphosus* gut, with its understudied functional diversity and degradative pathways, offers an untapped source of such enzyme and inoculum candidates.

This study leverages metagenome-assembled genomes from *Kyphosus vaigiensis*, *Kyphosus cinerascens*, and *Kyphosus hawaiiensis* gut symbionts and inoculated bioreactor enrichments to connect whole-genome degradative potential of algal polysaccharides to accurate taxonomic lineages and functional roles. The addition of genomes from bioreactor enrichments explores leveraging the metabolic capacities of *Kyphosus* gut consortia in industrial processes. This work extends previous studies of taxonomic-level biogeography (18) and contig-level gene associations (15, 20) in this system using high-quality MAGs, which enables differentiation between processes that can potentially be executed within a single cellular compartment (individual microbial species/population) and those likely to require cooperative action by multiple cells from different species (community impacts). Discoveries in this study provide the foundation for genome-level understanding of microbial contributions to herbivorous fish digestion and beget future investigations to apply these findings toward applications in the aquaculture and bioenergy sectors.

MATERIALS AND METHODS

Sample description and metagenomic assembly

DNA was extracted from liquid samples from 10 anaerobic bioreactors inoculated with gut content from either "Fish 6" (*K. cinerascens*) or "Fish 7" (*K. hawaiiensis*; Table S1) using methods previously described (18) and propagated to enrich degradative properties. Samples were taken 9–10 days after inoculation and incubation at 30°C. A

35 psu Artificial Sea Water (ASW) solution was prepared by dissolving and autoclaving 40 g/L of Instant Ocean sea salts (Instant Ocean, Spectrum Brands, Blacksburg, VA). The Basal Salts Medium solution was then prepared by dissolving and autoclaving 90 mM $MgSO_4$, 6 mM K_2CO_3, 6 mM $CaCO_3$, 20 mM $MgCO_3$, and 1/10 of the final volume of ASW in ultrapure water. Anoxic cultures of 50 mL were processed in a portable anaerobic chamber containing sterile Basal Salts Medium in 150 mL serum bottles, crimp sealed with a rubber septum. These ionic concentrations were selected to simulate the estimated osmolarity of seawater as it passes through the midgut and hindgut (39). Approximately 1 g of fish gut section contents were placed in the bottles along with the indicated substrate (Table S1) and sealed, with no additional feedstock added before sequencing. Substrate selection was focused on polysaccharides and algal species of particular relevance to bioenergy and bioproduct production.

Samples were sequenced using Illumina NovaSeq 6000 technology (Illumina, San Diego, CA). Read trimming was performed using Trimmomatic v. 0.36 (40) with the following parameters: adapter-read alignment settings 2:30:10, LEADING:10, TRAILING:20, HEADCROP:12, SLIDINGWINDOW:4:15, and MINLEN:200. Taxonomic composition of metagenomic reads was determined using Kraken v. 2.0.9 (41), with taxonomic assignment using a protein database based on all amino acid sequences in the NCBI nr database (42) as of April 2022. Cleaned reads were assembled in metaSPAdes v. 3.13 (43) with a minimum contig retention size of 2,000 nucleotides.

Gene calling and functional annotation

Gene boundaries were predicted using prodigal v. 2.6.2 (44) and annotated using prokka v. 1.12 (45). Genes were assigned to CAZy classes from the dbCAN HMMdb v. 10 database (46) based on the CAZy database (11) and to sulfatases classes from the SulfAtlas v. 2.3 database (13), using methods previously described (20). Signal peptides were identified using SignalP v. 6 (47) with default parameters. Additional enzyme classes were annotated with KofamKOALA (48).

Enzyme novelty was evaluated using DIAMOND blastp (49) searches against the NCBI nr database (42) as of April 2022. Some CAZyme classes were grouped into the category of "peptidoglycanases" using the division proposed by López-Mondéjar et al. (50). Distributions of annotated proteins were compared to free-living relatives from the OceanDNA database (51).

Metagenomic binning and biosynthetic gene cluster prediction

Metagenomic binning was performed from both newly assembled bioreactor metagenomes described above and *in vivo* gut metagenomes from *K. vaigiensis*, *K. cinerascens*, and *K. hawaiiensis* lumen contents previously described in Podell et al. (20). Lumen contents were used to maximize microbial biomass while reducing the amount of recovered eukaryotic host DNA. Binning was done through MetaWRAP v. 1.3.2 (52) with a minimum completeness cutoff of 0.7 and a maximum contamination cutoff of 0.05 as determined by CheckM v. 1.0.12 (53). MAG taxonomy was determined using GTDB-Tk v. 1.5.1 (54) with release 202 of the Genome Taxonomy Database (55).

Viral contigs and prophages were identified using DeepVirFinder v. 1.0 (56) using a *q*-score cutoff of 0.94. Viral sequence completeness was determined using CheckV v. 1.5 (57), retaining only regions marked as "high-quality" or "complete." Viral sequences were assigned to host taxonomies using the software VPF-Class (58).

Biosynthetic gene clusters (BGCs) were predicted for each MAG using antiSMASH v. 6.1 (59). Predicted products and BGC classes were annotated using BiG-SLiCE v. 1.1.1 (60). Gene cluster distances were calculated using the BiG-FAM webservice v. 1.0.0 (61), using a novelty distance cutoff of 900 following previous studies (61–63). Short-chain fatty acid gene clusters were annotated using gutSMASH v. 5.0.0 (64).

Phylogenomics and enzyme phylogenetics

A phylogenetic tree of MAGs was generated using PhyloPhlAn v. 3.0.2 (65) using a concatenated universal set of 400 marker genes (66). MAGs containing at least 100 marker genes underwent concatenated alignment using MAFFT v. 7.505 (67). The phylogenetic tree was built using RaxML v. 8.2.12 (68) and visualized using R v. 4.2.0 (69) packages treeio v. 1.20.0 (70), ggtree v. 3.4.0 (71), and ggtreeExtra v. 1.6.0 (72).

Multiple sequence alignments for genes belonging to CAZy class GH86 were made using MUSCLE v. 3.8.31 (73) and visualized using the R package ggmsa v. 1.2.0 (74). Gene trees were created using FastTree v. 2.1.10 (75). Additional reference genes were included in the tree based on DIAMOND blastp matches to the NCBI nr database as of April 2022. Protein domains were analyzed with the CDD webservice (76). Three-dimensional protein structures for CAZymes were predicted using ColabFold v. 1.3.0 (77) and visualized using ChimeraX v. 1.3 (78). Residue conservation was visualized using the WebLogo (79) webservice.

RESULTS

A (meta)genome catalog of enrichable symbionts in the *Kyphosus* gut

New data derived from *K. cinerascens* and *K. hawaiiensis* enrichment cultures expand the diversity of previous *K. cinerascens*, *K. hawaiiensis*, and *K. vaigiensis* gut metagenomes (20). This more complete catalog of *Kyphosus* gut microbiota provides additional details on the metabolic potential of taxa that were rare in the *in vivo* gut metagenome samples and highlights potential challenges in harnessing gastrointestinal microbiota for industrial processes. The fish inoculum species, gut location, and feedstock that were combined to establish each enrichment sample are described in Table S1. The taxonomic classification of unassembled metagenomic reads revealed high-level consistency at the phylum level between the *in vivo* gut microbiomes (20) and enrichment samples (Fig. 1). *Bacillota*, *Bacteroidota*, and *Gammaproteobacteria* constitute the dominant bacterial lineages in most samples, although the *Desulfovibrionales* order (phylum *Thermodesulfobacteriota*) was highly abundant in two enrichment samples.

Seventy-four medium- and high-quality MAG bins were obtained from newly assembled enrichment metagenomes, along with 137 new bins from previously described wild fish gut metagenomes (Fig. S1). These MAGs all met the minimum of 70% completion and a maximum of 5% redundancy standards (80). Assembly statistics for enrichment metagenomes are shown in Table S2, and MAG summary metrics outlined by the Genomic Standards Consortium (80) are provided in Table S3. Consistent with the unassembled read-based taxonomic profiles of the metagenomes, most MAGs were assigned to the phyla *Bacillota* (78 MAGs), *Bacteroidota* (72 MAGs), the class *Gammaproteobacteria* (31 MAGs), or the order *Desulfovibrionales* (13 MAGs), along with phylum *Verrucomicrobiota* (6 MAGs). The enrichments provide information on microbial members that were not as abundant in the fish gut metagenomes and vice versa. In one example, bins containing the *Verrucomicrobiota* order *Kiritimatiellales* were recovered in *K. cinerascens* gut samples but not in enrichment metagenomes. These dissimilarities were also reflected in nucleotide similarities, as only 9 of the 74 (12%) enrichment MAGs match MAGs generated from *in vivo* fish gut metagenomes at the species level. Enrichment samples averaged approximately 6% eukaryotic reads, while adult fish gut samples averaged 13%, and juvenile fish gut samples averaged 46%, possibly due to the technical limitations of collecting ample microbial biomass from smaller fish.

Viral and archaeal sequences comprised less than 0.5% of all unassembled metagenomic reads, with 69 viral contigs and 3 prophages identified as either high quality or complete. Within these viral elements, 30 auxiliary metabolic genes found on potential prophage regions were annotated as CAZymes and 13 as sulfatases, suggesting a potential role for viral dissemination of these genes across the bacterial community. The taxa *Bacillota*, *Bacteroidota*, and *Gammaproteobacteria* were the most frequently predicted viral hosts (Table S4), which is consistent with the taxonomic abundances of

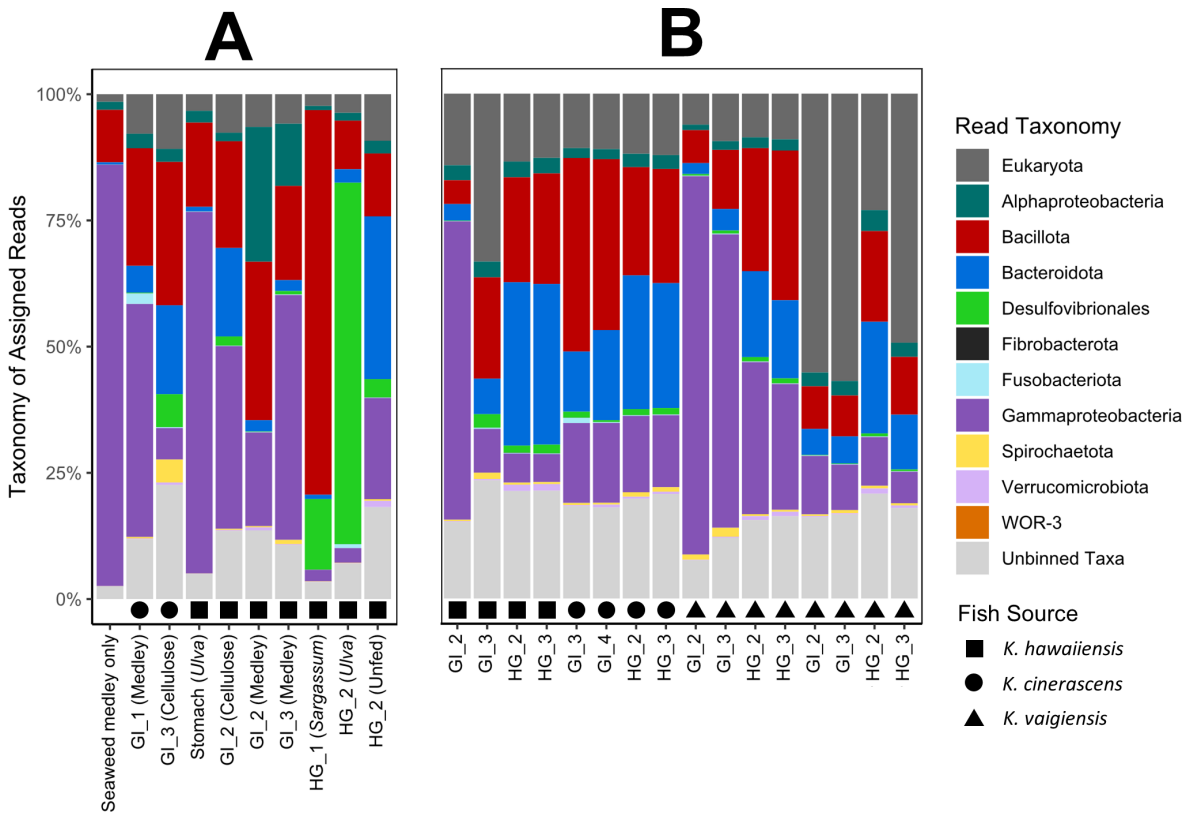

FIG 1 Taxonomic distribution of enrichment and fish gut samples. Unassembled metagenomic reads were classified using Kraken2. (A) Enrichment samples are labeled with inoculant fish taxa, gut region, and bioreactor feed. (B) Wild fish gut metagenomic samples previously assembled by Podell et al. (20). Shapes along the x-axis denote the species of *Kyphosus* whose gut was either (A) used as the inoculant or (B) directly sequenced. Abbreviations: GI, midgut; HG, hindgut; medley, a combination of *Ulva*, *Sargassum*, and *Agardhiella* seaweed.

classified unassembled metagenomic reads and recovered MAGs. Despite the presence of numerous auxiliary metabolic genes annotated as mediating more general polysaccharide degradation, none of the viral sequences we detected appeared to specifically target large, complex sulfated macroalgal polysaccharides.

Genome capacities reveal metabolic specialization among gut symbionts of *Kyphosus* fish

The distribution of CAZymes and sulfatases was correlated with the phylogeny of fish gut and enrichment MAGs, as shown using a concatenated marker gene tree (Fig. 2A). This assessment revealed that among the MAGs generated in this study, the *Bacteroidota* genomes contained the majority of CAZymes and sulfatases (Fig. 2B). Algal degradation-specific CAZyme-rich genomes among the MAGs from other phyla were restricted either to a single order, *Kiritimatiellales* (*Verrucomicrobiota*), or a single genus, *Vallitalea* (*Bacillota*). Recovered *Gammaproteobacteria* and *Desulfovibrionales* genomes lacked enzymes required for digesting sulfated algal polysaccharides despite the relatively high abundance of these taxonomic groups in classified unassembled reads and the recovered MAGs. However, the *Gammaproteobacteria* MAGs contained more peptidoglycanases than other taxa, suggesting a potential niche in digesting alternative dietary components. This analysis also showed that CAZymes targeting ulvan, a green algal polysaccharide, were less prevalent among the symbiotic MAGs associated with wild fish than CAZymes targeting red and brown algae-associated polysaccharides (Fig. 2B), consistent with previous results quantifying relative amounts of these algae types consumed by the *Kyphosus* fish included in this study (20). The most abundant phyla yielded binned MAGs from both *in vivo* and enrichment samples (Fig. 2C).

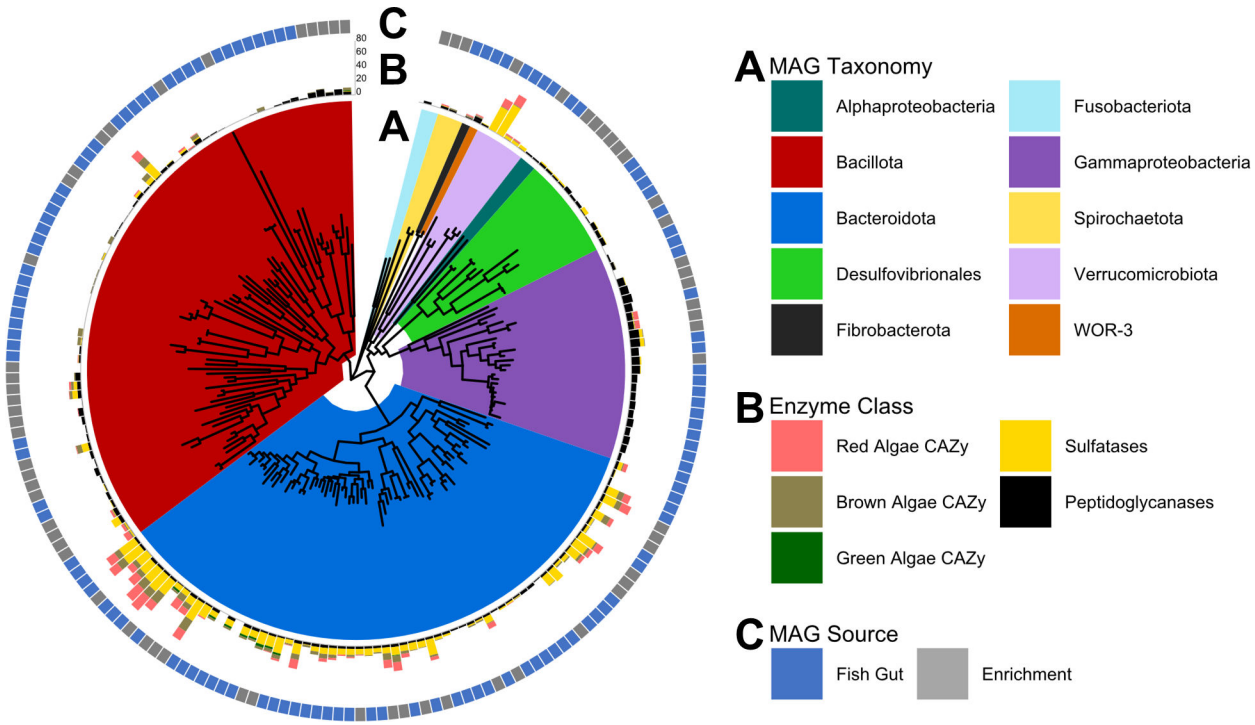

FIG 2 Genomic CAZyme distributions reveal connections between metabolic strategies and taxonomic lineage. (A) The gene tree shows a concatenated alignment of 400 PhyloPhlAn universal marker genes for each recovered MAG, with branches colored by assigned MAG taxonomy. (B) The inner ring displays genomic gene counts for sulfatases and carbohydrate-active enzymes that specifically target algal polysaccharides or peptidoglycans. (C) Environmental source of each MAG.

A search was performed for genes involved in mannitol metabolism to determine whether this sugar alcohol, known to be abundant in brown algae, might be used for fermentation. In support of this hypothesis, genes predicted to encode mannitol 2-dehydrogenases, mannitol-1-phosphate 5-dehydrogenases, mannitol-specific phosphotransferase system (PTS) enzymes, and mannitol operon repressors were detected in both MAGs and metagenomes from natural fish gut samples as well as enrichment cultures (Table S5). This metabolic potential was not lineage-specific, as MAG representatives from *Bacteroidota*, *Bacillota*, *Gammaproteobacteria*, and *Verrucomicrobiota* all contained these genes, and 22% of our recovered MAGs contained at least one of the two major enzyme classes thought to contribute to mannitol to fructose conversion in *Kyphosus* guts (19). Even though not all genes were present in all samples, it was not possible to conclude whether differences between samples might be significant due to unavoidable variability in overall community complexity, assembly efficiency, MAG completeness, and uneven representation of less abundant taxa.

An assessment of SCFA production gene pathways of recovered MAGs using gutSMASH (64) revealed that most of the *Kyphosus* gut symbiotic taxa (67% of fish gut MAGs and 77% of enrichment MAGs) can potentially contribute to host nutrition through the production of SCFAs (Fig. 3). One hundred thirty-nine genomes from analyzed kyphosid fish gut microbial communities contained pathways for producing acetate, but only six genomes contained pathways for butyrate production. The pyruvate formate lyase and pyruvate:ferredoxin oxidoreductase pathways were the most abundant overall, present in 126 MAGs, while *Bacteroidota* contained the most gene clusters (39) related to propanoate production.

The overall prevalence of acetate production pathways was lower than that previously reported in human gut microbiota (81). The total absence of some alternate fermentation pathways from our MAGs, such as choline utilization, suggests that those processes are not core to dominant members of the *Kyphosus* gut microbiome. Only one

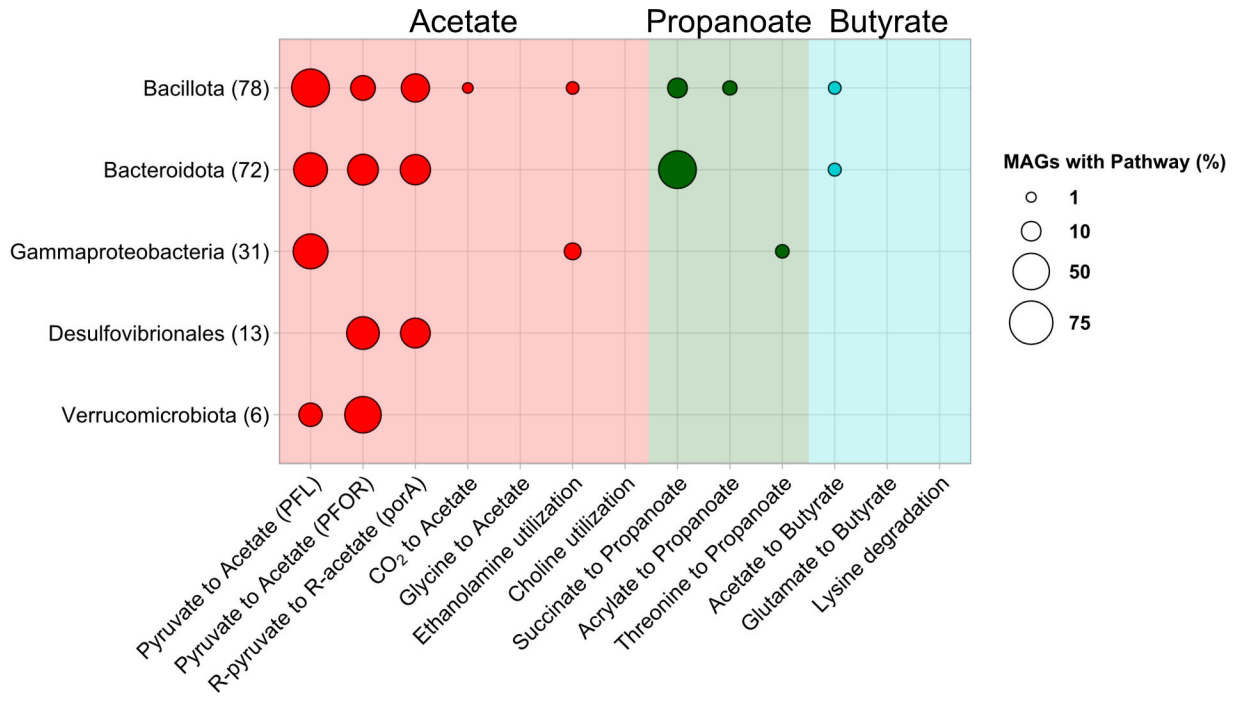

FIG 3 *Kyphosus* gut symbionts use diverse metabolic strategies to produce SCFAs. The bubble plot displays the presence of pathways in MAGs from different taxonomic groups as determined by gutSMASH. Circles increase in size based on the proportion of binned genomes that are annotated with each pathway. Bubbles and chart backgrounds are colored based on their association with the production of the fatty acids acetate (red), propanoate (green), or butyrate (cyan). Counts next to taxonomic names denote the number of MAGs analyzed from that taxa.

genome from this study contained fermentation pathways involving the degradation of amino acids such as glycine, threonine, and lysine, suggesting that *Kyphosus* gut microbiota do not rely directly on dietary proteins for energy. Such lessened reliance on nitrogen-based substrates for fermentation is consistent with a low-protein, algae-based diet rich in available polysaccharides and limited in available nitrogen.

Functional adaptations to life in the *Kyphosus* gut

Adaptations to environmental conditions in herbivorous fish gut microbes are reflected in the high abundance of CAZyme classes specifically targeting algal polysaccharides (20). Figure 4A shows that the amino acid sequences of CAZyme classes abundant in the MAGs of this study are well conserved across *Kyphosus* gut symbiont genomes. However, such enzymes are poorly represented in both specialty and general databases of previously described sequences, with closest enzyme homologs averaging less than 60% sequence identity for most of the highlighted CAZyme classes. Similar trends are observed for the sulfatase subclasses in *Kyphosus* gut symbiont genomes (Fig. 4B). Both cases demonstrate the extent that this study expands known sequence diversity within these enzyme classes, underscoring unusual domains that may not be captured by current databases.

The discovery of novel enzyme sequences in these enzyme classes presents numerous opportunities to expand our understanding of marine polysaccharide degradation. One example using the phylogeny of CAZy class GH86, consisting of β-agarases and β-porphyranases, illustrates previously unappreciated cryptic variability within this enzyme family. A gene tree of class GH86 CAZyme examples from this study plus closest GenBank homologs (Fig. 5A) shows that many of these genes are associated with *Bacteroidota*, consistent with the high abundance of CAZymes and sulfatases found among MAGs from this phylum in *Kyphosus* guts (Fig. 2). Binned MAG genes annotated as β-porphyranases all originate from hindgut or enrichment samples, consistent with

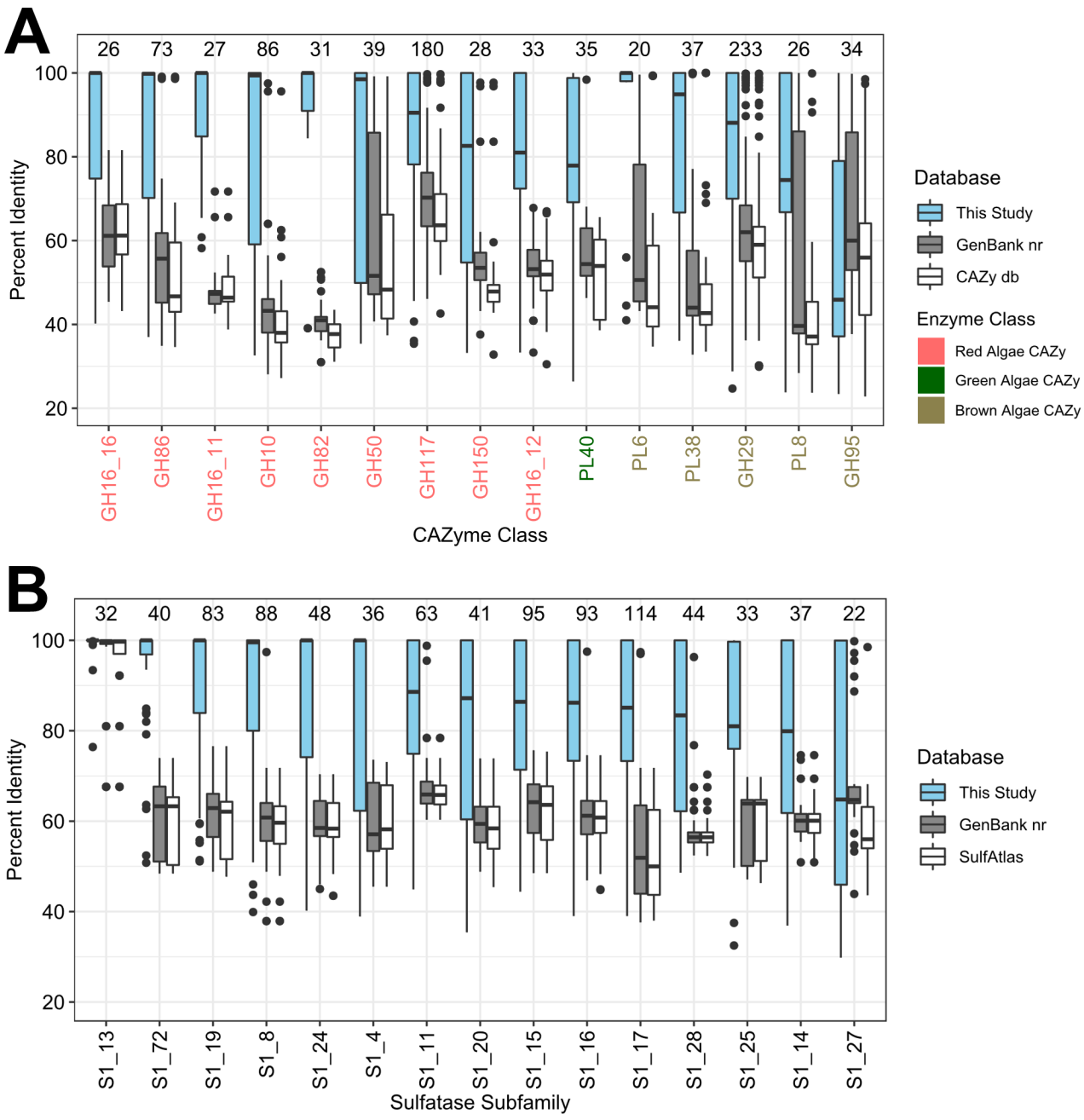

FIG 4 *Kyphosus* gut symbionts encode CAZymes and sulfatases divergent from other data sets and environments. Percent identity of binned (A) CAZymes and (B) sulfatases to best blast matches found in the following databases: all genes from MAGs in this study (blue), the GenBank nr database (gray), and either (A) the CAZy database or (B) the SulfAtlas database (white). CAZyme classes are colored based on the degradation of red, green, or brown algal polysaccharides. Each group is labeled by the number of genes with that enzyme annotation found in our MAGs.

previously reported physiological localization of polysaccharide degradation capabilities (20). Surprisingly, two GH86 genes recovered in *Bacillota* MAGs from bioreactor enrichments and two *Bacillota* homologs from the NCBI nr database nested within a clade of genes from phylum *Verrucomicrobiota*. This unexpected pattern of association between genes from very distant microbial taxa has not been described in prior literature and may be indicative of horizontal gene transfer.

Amino acid insertions in this unique clade might either extend the signal peptide or contribute additional catalytic functionality (82). Among NCBI nr homologs, only

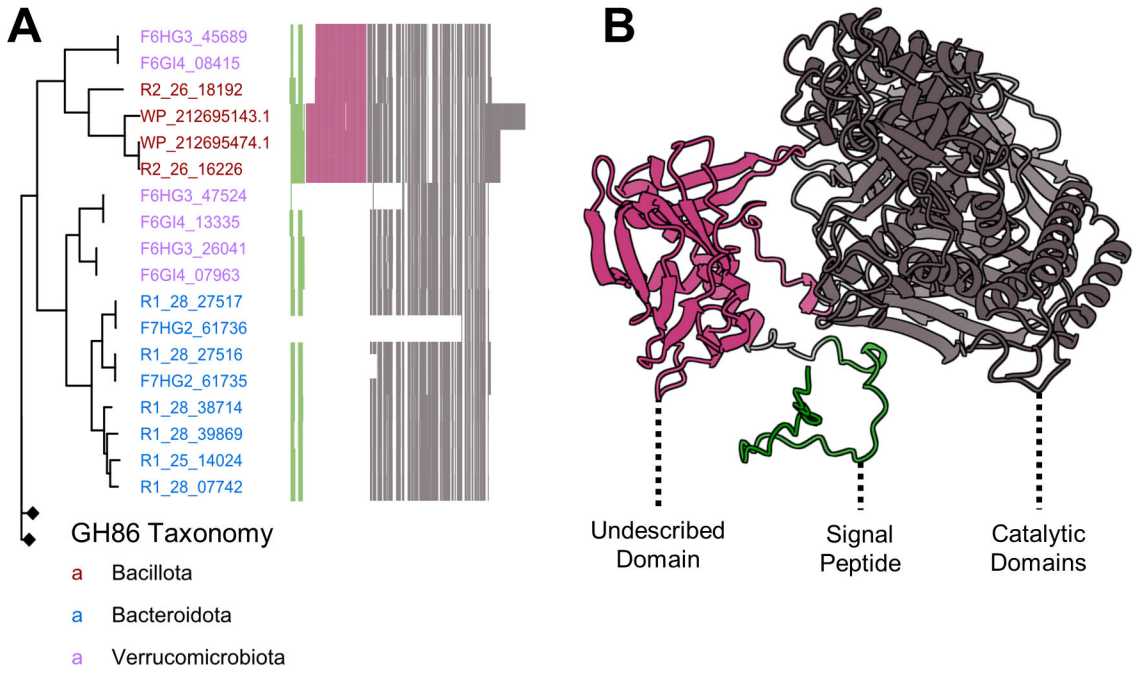

FIG 5 A β-agarase/β-porphyranase gene tree highlights an undescribed protein domain present in multiple phyla. (A) A gene tree of binned GH86 enzymes, with gene names colored by genome taxonomy. Nodes with black diamonds represent collapsed outgroup clades lacking the extra domain. A multiple sequence alignment is appended to the right of the tree, with colored vertical lines representing conserved amino acid positions and white vertical lines representing gaps. (B) The predicted protein structure of GH86 enzyme R2_26_16226, with conserved CAZy domains highlighted in gray, the predicted signal peptide in green, and the conserved new domain in pink. An uncollapsed version of the gene tree is included in Fig. S2, and a motif logo of the domain is presented in Fig. S3.

genes from the hydrothermal vent genome *Vallitalea pronyensis* (WP_212695143.1 and WP_212695474.1) (83) contained this pattern of approximately 168 amino acids. No other entries in the GenBank nr database contained sequences matching this region at greater than 50% amino acid identity (Fig. S2). Outside of the clade containing this novel domain, variability occurs primarily in the putative signal peptide region at the N-terminus of the protein, while the downstream porphyranase domain itself is far more conserved. Figure 5B displays the predicted three-dimensional structure of a *Kyphosus* symbiont GH86 enzyme, with the additional uncharacterized region positioned between the predicted signal peptide and annotated catalytic β-agarase and β-porphyranase domains. Although the function of this domain cannot be determined bioinformatically, it provides an interesting subject for further enzymatic characterization. Potentially novel properties might include modified substrate specificity, substrate concentration dependence, catalytic efficiency, and/or tolerance of different abiotic conditions.

MAG sequences were interrogated using antiSMASH BGC detection software to determine whether *Kyphosus* gut-associated microbes might encode any unusual secondary metabolites. The majority of *Bacillota*, *Bacteroidota*, *Verrucomicrobiota*, and *Gammaproteobacteria* MAGs from both fish gut inocula and bioreactor enrichments encoded BGCs typical of taxonomic relatives found in other vertebrate gut environments, such as lanthipeptides, beta-lactones, and arylpolyenes (84, 85). However, BGCs were not particularly abundant in our MAG catalog relative to other similar genomes. Our recovered *Gammaproteobacteria*, *Bacillota*, and *Bacteroidota* average fewer BGCs per genome than a random set of seawater MAGs representing each taxonomic group from the OceanDNA database. Thus, our host-associated MAGs may contain fewer BGCs per genome than their free-living relatives.

A total of 307 BGCs were annotated within our MAGs (Fig. 6). Twenty-three annotated BGCs were determined to be complete, meaning they were not located on contig edges, based on BiG-FAM analysis of antiSMASH predictions (61). Twenty BGCs represent

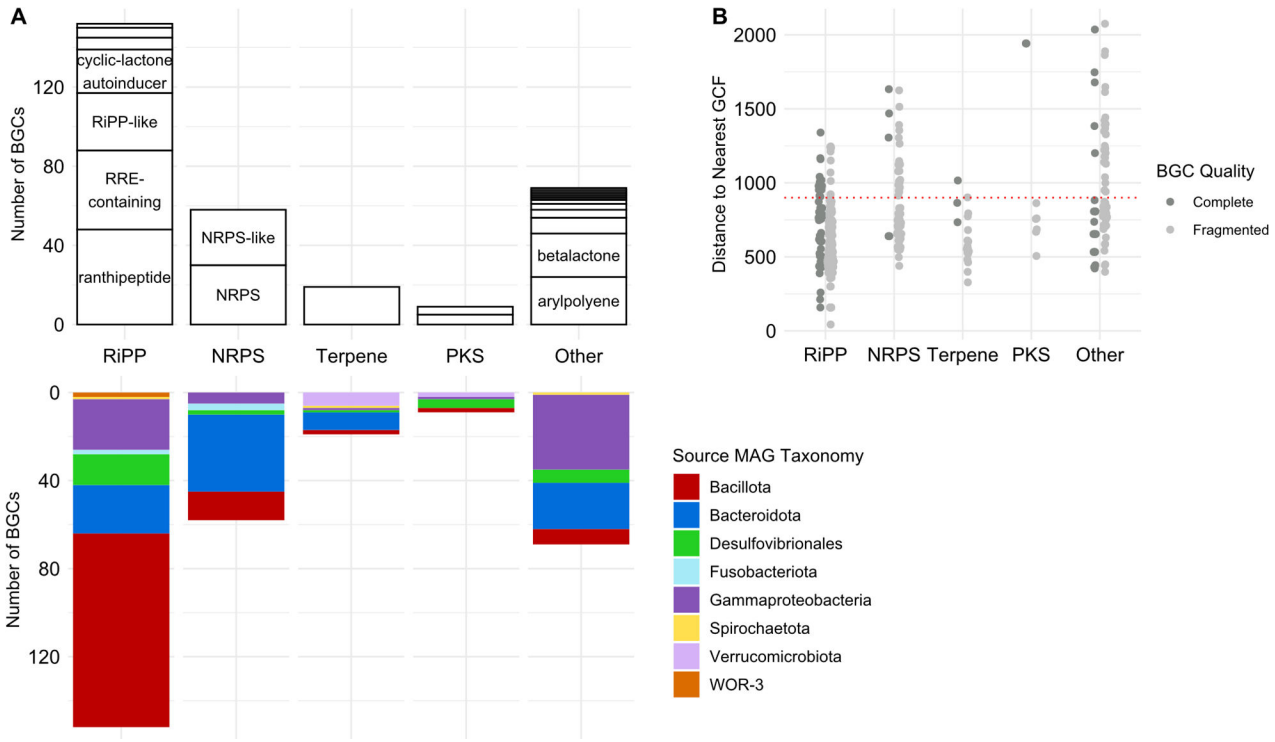

FIG 6 *Kyphosus* gut symbiont MAGs encode novel BGCs. (A) On the positive y-axis, counts of binned BGCs are grouped by BiG-SLiCE class and labeled by predicted product. On the negative y-axis, counts of binned BGCs are grouped by BiG-SLiCE class and colored by associated MAG taxonomy. (B) Distance of binned BGCs to the nearest gene cluster family as determined by BiG-FAM. A distance above 900, marked by a dashed red line, suggests novelty and divergence from previously described gene cluster families. BGCs are colored dark grey if they are annotated as complete by BiG-FAM. Abbreviations: RiPP, ribosomally synthesized and post-translationally modified peptide; RRE, RiPP recognition element; NRPS, non-ribosomal peptide synthetase; PKS, polyketide synthase.

putative novel gene cluster families, with BiG-FAM distances exceeding the standardized cutoff score of 900 (Fig. 6B). These novel gene cluster families may represent unique natural products or enzymes specialized to the *Kyphosus* gut environment. Complete biosynthetic gene cluster annotations, novelty assessment, and associated taxonomy are included in Table S6.

Community digestion of complex algal polysaccharides

Polysaccharide digestive capabilities vary among MAGs from different microbial taxa in the *Kyphosus* fish gut community, as shown in Fig. 7. Despite overall microbiome-wide diversity, the MAGs generated in this study show that few individual genomes contain all of the enzymes necessary to completely degrade even a single type of complex algal polysaccharide, let alone the huge variety of natural variants characteristic of marine macroalgae (86) that might be ingested by generalist herbivorous fishes. Each microbial genome instead contains a limited assortment of enzymes capable of partially degrading a selection of different carbohydrate moieties, including potentially incomplete breakdown products generated by other microbes. Combined pangenomic capabilities of several taxonomic groups appear to contain complementary collections of exported CAZymes that might facilitate adaptation to unpredictable variability in available polysaccharide content. Figure 7 summarizes predicted macroalgal digestion capabilities observed within individual MAGs for the most abundant taxonomic groups. The collaborative potential for all MAGs within each metagenomic sample is illustrated in Fig. S4.

Potential contributions to shared, community-wide degradation of algal polysaccharides through extracellular enzymes vary according to both microbial cell taxonomy and targeted substrate. More than 90% of CAZymes that target macroalgal polysaccharides

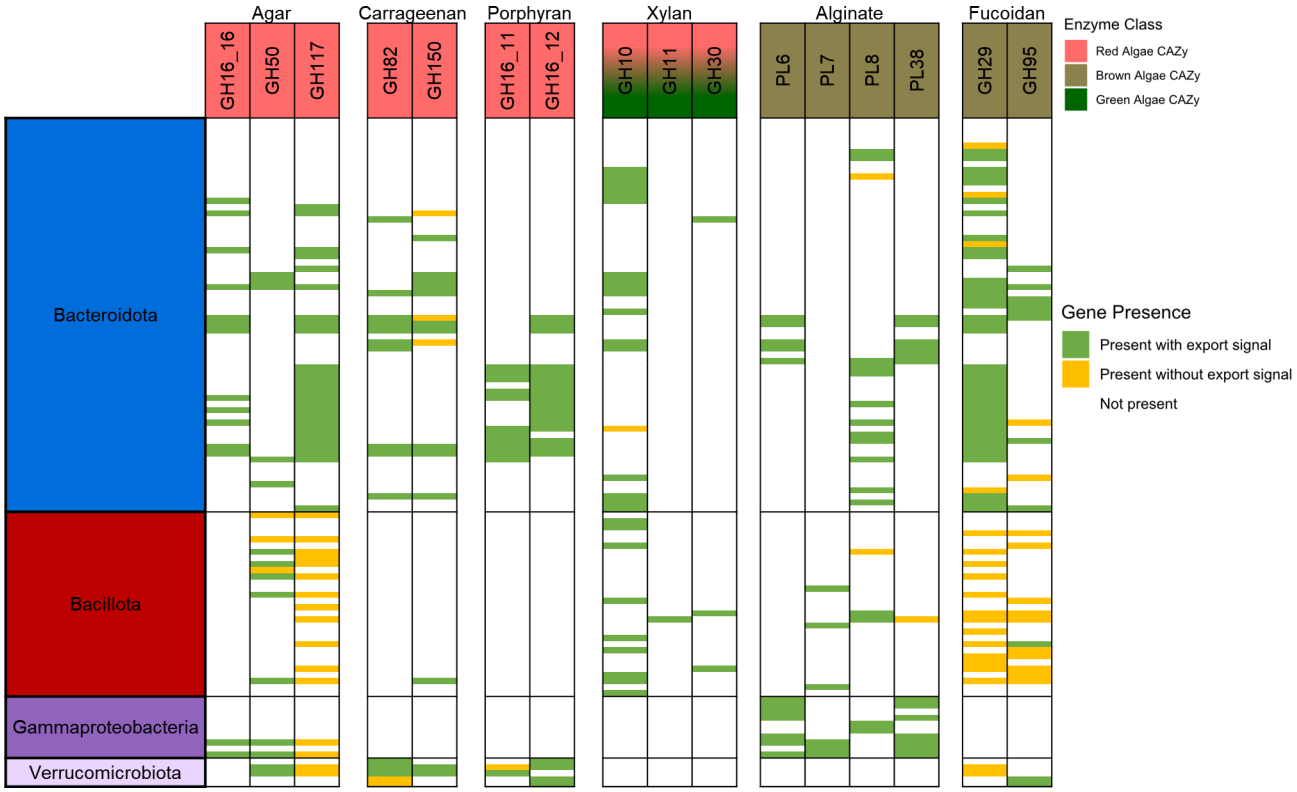

FIG 7 *Kyphosus* gut symbiont MAGs encode the capacity to degrade various algal polysaccharides collaboratively but not solitarily. Each row represents a single MAG from the annotated taxonomic lineage. Only MAGs from the four lineages with the highest concentration of CAZymes (*Bacillota*, *Bacteroidota, Gammaproteobacteria,* and *Verrucomicrobiota*) are shown. MAGs with no applicable CAZy classes are not shown, and CAZy classes not associated with a single substrate or not found in any MAG are not shown. Green bars denote a signal peptide annotated to at least one of the appropriate CAZyme in a single MAG, while yellow bars mark the absence of a signal peptide on all appropriate CAZyme candidates within a MAG.

from *Bacteroidota* MAGs contain signal peptides that indicate export or integration into the cellular membrane. In contrast, CAZymes in *Bacillota* MAGs largely lack these signal peptides in enzymes predicted to degrade fucoidan and agar but have more abundant signal peptides within the smaller set of CAZymes targeting xylan and alginates. Few *Bacillota* MAGs contain all of the enzymes required to fully degrade complex algal polysaccharides such as porphyran, suggesting that cells from this taxonomic group might scavenge partial breakdown products degraded extracellularly by other taxa.

Verrucomicrobiota polysaccharide digestion enzymes appear to be more specialized toward red algae, with genomes consistently containing CAZymes predicted to digest agar, carrageenan, and porphyran. However, MAGs from this phylum seem to be lacking enzymes predicted to target green or brown algal polysaccharides. *Gammaproteobacteria* MAGs appear to have more enzymes involved in the digestion of non-sulfated polysaccharides, such as alginate, and occasionally enzymes involved in agar degradation. Thus, the *Gammaproteobacteria* symbionts analyzed here may have specialized in polysaccharide types that are easier to digest.

DISCUSSION

The recovery and characterization of 211 MAGs from *Kyphosus* gut and enrichment metagenomes connect detailed taxonomic classification with the potential of the major microbial contributors to digest complex algal polysaccharides. Algal polysaccharide-targeting enzymes from this study are divergent in sequence from previously sequenced and characterized representatives from other environments, clarifying prior assumptions about the metabolic capacities of this system using 16S rRNA or community

composition. This study confirms and expands upon earlier work showing that certain members of the *Bacillota* and *Verrucomicrobiota* lineages are unexpectedly richer in some CAZyme and sulfatase enzyme classes than their respective taxonomic relatives (20). These CAZyme-rich MAGs provide the first genomic evidence supporting prior observations of laminarin, carrageenan, and alginate degradation in *Kyphosus* guts (87, 88). Differences between source inocula and the metagenomes of bioreactor enrichments starting from *Kyphosus* gut bacteria highlight potential challenges in harnessing this microbiota for bioenergy preprocessing of macroalgal feedstocks.

This study describes specific genes encoding SCFA production pathways in the genomes of fish gut microbiota. Microbial fatty acids serve as a key metabolite in gut-brain communication (89) and are a major source of available carbon for the host (90). SCFA pathway diversity is unexpectedly low for a system previously shown to contain high SCFA concentrations *in vivo* (16). However, this observation is consistent with a few dominant lineages, primarily the *Bacteroidota*, producing high amounts of SCFAs from the breakdown products of algal polysaccharides. Prior chemical work has observed that propanoate is more abundant than butyrate in *Kyphosus* guts (16), and our pathway enzyme abundance information at the genome level supports these observations (Fig. 3). Metabolic capacities in our *Kyphosus* metagenomes also match previous observations that bony fishes with carbohydrate-rich diets consistently lack branched SCFAs and have low rates of protein fermentation by gut bacteria (91).

Mannitol has been suggested as a major source of fermentation substrate in some algivorous fishes, based on the large percentage of mannitol in some brown macroalgae, observed degradation of mannitol by *Kyphosus* guts (92), and the relative accessibility of this compound compared to complex sulfated algal polysaccharides. The presence of both mannitol 2-dehydrogenase (EC 1.1.1.67) and D-mannitol-1-phosphate dehydrogenase (EC 1.1.1.17) genes in our MAGs, the latter of which was initially proposed by Seeto et al. (93) but not found in recent metagenomic investigations of *Kyphosus* guts by Stevenson et al. (19), suggests that mannitol utilization may differ more between individual fish and algivorous species than previously thought.

Other observations in prior work on *Kyphosus* (16) noted rates of sulfate reduction higher than methanogenesis, although both processes were negligible compared to SCFA production. This aligns with the low abundance of *Desulfovibrionales* and the near complete absence of Archaea in our metagenomes, consistent with observations that dietary red macroalgae inhibit methanogenesis and thus the success of gut Archaea (33). Both sulfate reduction and methanogenesis appear to be minor sources of energy available for Kyphosid host absorption, compared to fermentation by *Bacteroidota* and *Bacillota*.

Although *Kiritimatiellales* MAGs recovered from *K. cinerascens* fish guts contain more enzymes targeting algal polysaccharides than other members of their phyla, these taxa were not recovered from enrichment metagenomes. However, this should not be problematic for enrichment processing if the dominant *Bacteroidota* contain CAZymes with overlapping specificities for the same substrates, as suggested in Fig. 7. Additional work comparing MAGs from lumen and mucosal samples may provide additional insights into metabolic capacities that might be more abundant in the transient vs permanent resident fraction of the microbiome (19). Future enzyme-focused work will be needed to characterize sample-specific polysaccharide degradative chemistry in order to parse the specific roles of each taxa. *Vallitalea* and *Verrucomicrobiota* enzymes may encode some unique functionalities, as suggested by the extra domain present in their β-porphyranase sequences (Fig. 5). Isolation and *in vitro* characterization of bioinformatically predicted enzyme activities will be necessary to fully integrate these discoveries into aquaculture and bioenergy applications.

Metagenomic data from the MAGs in this study suggest that few individual cells have the genomic potential to independently degrade all of the complex sulfated polysaccharide substrates present in marine macroalgae. However, secreted and extracellularly exposed transmembrane CAZymes may enable collaborative interactions between fish

gut microbes to facilitate complete digestion of these molecules, without the high metabolic cost of encoding a complete, independent repertoire in every genome. A division of labor strategy cannot be fully confirmed without *in vitro* tests (94), although the first condition of genomically encoded functional complementarity appears to hold true between *Kyphosus* symbionts based on bioinformatic criteria. In one similar study, gene-based observations of complementarity for marine lignocellulose-degrading bacteria align with *in vitro* observations that support a division of labor hypothesis (95). Future work involving cultured representatives and enriched microcosms will be required to pin down the ecological strategies used by symbionts in this system.

This study provides a new baseline for *Kyphosus* microbiota at the genome level but begets many new questions requiring additional experimentation. Further work that connects enrichment composition, feedstock polysaccharide composition, and physical configuration parameters to chemical measurements of degraded polysaccharides will help determine which phyla are required for complete polysaccharide breakdown. The incorporation of novel enzyme sequences identified here may warrant the creation of new subclasses, based on classification techniques such as sequence similarity networks (96). Isolation and characterization of divergent proteins with unexpected new domains may reveal new enzymatic properties unique to this system. Metatranscriptomic analyses utilizing the genome catalogs presented here will enable detailed analysis of substrate-specific metabolic pathway expression and species collaboration. *Kyphosus* digestive systems have long been studied as models for herbivorous fish gut fermentation and can now be explored further using these additional techniques to deliver a deeper understanding of their degradative and fermentative capabilities.

Conclusion

The new metagenome-assembled genomes recovered from herbivorous fish guts and corresponding bioreactors described here provide a genomic catalog of *Kyphosus* gut symbionts highlighting untapped diversity in enzymatic and collaborative potential in the degradation of algal polysaccharides. The extensive sequence divergence of enzymes encoded within these genomes from previously characterized CAZyme family examples supports the promise of herbivorous fish guts as a source of novel and industrially relevant enzymes. Expansion of these discoveries will not only clarify ecological interactions but have the potential to improve the applicability of macroalgae in the bioenergy and aquaculture sectors.

ACKNOWLEDGMENTS

This publication includes data generated at the UC San Diego IGM Genomics Center utilizing an Illumina NovaSeq 6000 that was purchased with funding from a National Institutes of Health SIG grant (#S10 OD026929). Computational analyses were performed using the San Diego Supercomputer Center's Triton Shared Computing Cluster.

This work was funded by the United States Government, Department of Energy, Advanced Research Projects Agency—Energy grant ARPA-E DE-FOA-0001858 to R.S.N. and L.M.L.L. with contracts to C.E.N., L.W.K., and E.E.A., National Science Foundation grants OCE-1837116 and EF-2025217 to E.E.A., and National Institutes of Health NIEHS grant R01-ES030316 to E.E.A.

AUTHOR AFFILIATIONS

[1]Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California San Diego, La Jolla, California, USA
[2]Marine Biology Research Division, Scripps Institution of Oceanography, University of California San Diego, La Jolla, California, USA
[3]Daniel K. Inouye Center for Microbial Oceanography: Research and Education, School of Ocean and Earth Science and Technology, University of Hawai'i at Mānoa, Honolulu, Hawaii, USA

[4]National Renewable Energy Laboratory, Golden, Colorado, USA
[5]Ocean-Era Inc., Kailua-Kona, Hawaii, USA
[6]Center for Microbiome Innovation, University of California San Diego, La Jolla, California, USA
[7]Department of Molecular Biology, School of Biological Sciences, University of California San Diego, La Jolla, California, USA

AUTHOR ORCIDs

Aaron Oliver http://orcid.org/0000-0002-0410-8284
Sheila Podell http://orcid.org/0000-0001-7073-5190
Wesley J. Sparagon http://orcid.org/0000-0002-9627-4174
Eric E. Allen http://orcid.org/0000-0002-1229-8794

FUNDING

Funder	Grant(s)	Author(s)
DOE | Advanced Research Projects Agency - Energy (ARPA-E)	DE-FOA-000185	Robert S. Nelson Lieve M. L. Laurens
National Science Foundation (NSF)	OCE-1837116, EF-2025217	Eric E. Allen
HHS | NIH | National Institute of Environmental Health Sciences (NIEHS)	R01-ES030316	Eric E. Allen

AUTHOR CONTRIBUTIONS

Aaron Oliver, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing – original draft, Writing – review and editing | Sheila Podell, Conceptualization, Investigation, Methodology, Resources, Supervision, Writing – review and editing | Linda Wegley Kelly, Funding acquisition, Project administration, Supervision, Writing – review and editing | Wesley J. Sparagon, Investigation, Writing – review and editing | Alvaro M. Plominsky, Investigation, Methodology, Writing – review and editing | Robert S. Nelson, Funding acquisition, Investigation | Lieve M. L. Laurens, Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review and editing | Simona Augyte, Investigation, Methodology | Neil A. Sims, Conceptualization, Project administration, Supervision | Craig E. Nelson, Conceptualization, Funding acquisition, Project administration, Writing – review and editing | Eric E. Allen, Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review and editing

DATA AVAILABILITY

All custom code used for data analysis and visualization are available at https://github.com/AaronAOliver/KyphosusMAGs. Sequence reads are available under SRA BioProject numbers PRJNA819194 and PRJNA1023379. Complete MAG sequences and predicted proteins are available on Zenodo (https://zenodo.org) under DOI no. 10.5281/zenodo.8277654.

ADDITIONAL FILES

The following material is available online.

Supplemental Material

Supplemental figures (mBio00496-24-s0001.pdf). Fig. S1 to S4.
Table S1 (mBio00496-24-s0002.xlsx). Enrichment sample information.
Table S2 (mBio00496-24-s0003.xlsx). Assembly statistics for enrichment metagenomes.

Table S3 (mBio00496-24-s0004.xlsx). MIMAG-compliant completeness and lineage information for recovered MAGs.
Table S4 (mBio00496-24-s0005.xlsx). Predicted viral contigs and auxiliary metabolic genes.
Table S5 (mBio00496-24-s0006.xlsx). Mannitol utilization genes for MAGs and metagenomes.
Table S6 (mBio00496-24-s0007.xlsx). Detailed BGC annotations and nearest gene cluster families.

REFERENCES

1. Wu P, Wang T, Liu Y, Li C, Xiao Y, Xu S, Han T, Lin L, Quan Q. 2022. Differences of macroalgal consumption by eight herbivorous coral reef fishes from the Xisha Islands, China. Front Mar Sci 9. https://doi.org/10.3389/fmars.2022.882196
2. Knudsen SW, Clements KD. 2016. World-wide species distributions in the family Kyphosidae (Teleostei: Perciformes). Mol Phylogenet Evol 101:252–266. https://doi.org/10.1016/j.ympev.2016.04.037
3. Dell CLA, Longo GO, Burkepile DE, Manfrino C. 2020 Few herbivore species consume dominant macroalgae on a Caribbean coral reef. Front Mar Sci 7. https://doi.org/10.3389/fmars.2020.00676
4. Pillans RD, Babcock RC, Thomson DP, Haywood MDE, Downie RA, Vanderklift MA, Rochester WA. 2017. Habitat effects on home range and schooling behaviour in a herbivorous fish (*Kyphosus bigibbus*) revealed by acoustic tracking. Mar Freshwater Res 68:1454. https://doi.org/10.1071/MF16199
5. Augyte S, Sims NA, Martin K, Van Wychen S, Panczak B, Alt H, Nelson R, Laurens LML. 2023. Tropical red macroalgae cultivation with a focus on compositional analysis. Plants (Basel) 12:3524. https://doi.org/10.3390/plants12203524
6. Patel AK, Vadrale AP, Singhania RR, Michaud P, Pandey A, Chen S-J, Chen C-W, Dong C-D. 2023. Algal polysaccharides: current status and future prospects. Phytochem Rev 22:1167–1196. https://doi.org/10.1007/s11101-021-09799-5
7. Popper ZA, Michel G, Hervé C, Domozych DS, Willats WGT, Tuohy MG, Kloareg B, Stengel DB. 2011. Evolution and diversity of plant cell walls: from algae to flowering plants. Annu Rev Plant Biol 62:567–590. https://doi.org/10.1146/annurev-arplant-042110-103809
8. Helbert W. 2017. Marine polysaccharide sulfatases. Front Mar Sci 4. https://doi.org/10.3389/fmars.2017.00006
9. Bäumgen M, Dutschei T, Bornscheuer UT. 2021. Marine polysaccharides: occurrence, enzymatic degradation and utilization. Chembiochem 22:2247–2256. https://doi.org/10.1002/cbic.202100078
10. Jagtap AS, Manohar CS. 2021. Overview on microbial enzymatic production of algal oligosaccharides for nutraceutical applications. Mar Biotechnol (NY) 23:159–176. https://doi.org/10.1007/s10126-021-10027-6
11. Drula E, Garron ML, Dogan S, Lombard V, Henrissat B, Terrapon N. 2022. The carbohydrate-active enzyme database: functions and literature. Nucleic Acids Res 50:D571–D577. https://doi.org/10.1093/nar/gkab1045
12. Li J, He Z, Liang Y, Peng T, Hu Z. 2022. Insights into algal polysaccharides: a review of their structure, depolymerases, and metabolic pathways. J Agric Food Chem 70:1749–1765. https://doi.org/10.1021/acs.jafc.1c05365
13. Barbeyron T, Brillet-Guéguen L, Carré W, Carrière C, Caron C, Czjzek M, Hoebeke M, Michel G. 2016. Matching the diversity of sulfated biomolecules: creation of a classification database for sulfatases reflecting their substrate specificity. PLoS One 11:e0164846. https://doi.org/10.1371/journal.pone.0164846
14. Rimmer DW, Wiebe WJ. 1987. Fermentative microbial digestion in herbivorous fishes. J Fish Biol 31:229–236. https://doi.org/10.1111/j.1095-8649.1987.tb05228.x
15. Pardesi B, Roberton AM, Lee KC, Angert ER, Rosendale DI, Boycheva S, White WL, Clements KD. 2022. Distinct microbiota composition and fermentation products indicate functional compartmentalization in the hindgut of a marine herbivorous fish. Mol Ecol 31:2494–2509. https://doi.org/10.1111/mec.16394
16. Mountfort DO, Campbell J, Clements KD. 2002. Hindgut fermentation in three species of marine herbivorous fish. Appl Environ Microbiol 68:1374–1380. https://doi.org/10.1128/AEM.68.3.1374-1380.2002
17. Perry WB, Lindsay E, Payne CJ, Brodie C, Kazlauskaite R. 2020. The role of the gut microbiome in sustainable teleost aquaculture. Proc Biol Sci 287:20200184. https://doi.org/10.1098/rspb.2020.0184
18. Sparagon WJ, Gentry EC, Minich JJ, Vollbrecht L, Laurens LML, Allen EE, Sims NA, Dorrestein PC, Kelly LW, Nelson CE. 2022. Fine scale transitions of the microbiota and metabolome along the gastrointestinal tract of herbivorous fishes. Anim Microbiome 4:33. https://doi.org/10.1186/s42523-022-00182-z
19. Stevenson SJR, Lee KC, Handley KM, Angert ER, White WL, Clements KD. 2022. Substrate degradation pathways, conserved functions and community composition of the hindgut microbiota in the herbivorous marine fish *Kyphosus sydneyanus*. Comp Biochem Physiol A Mol Integr Physiol 272:111283. https://doi.org/10.1016/j.cbpa.2022.111283
20. Podell S, Oliver A, Kelly LW, Sparagon WJ, Plominsky AM, Nelson RS, Laurens LML, Augyte S, Sims NA, Nelson CE, Allen EE. 2023. Herbivorous fish microbiome adaptations to sulfated dietary polysaccharides. Appl Environ Microbiol 89:e0215422. https://doi.org/10.1128/aem.02154-22
21. Serag AM, Abdel-Sabour MS, El-Hadidi M, Maged M, Magdy M, Ramadan MF, Refaat MH. 2022. Comparative 16S metabarcoding of Nile tilapia gut microbiota from the northern lakes of Egypt. Appl Biochem Biotechnol 194:2168–2182. https://doi.org/10.1007/s12010-021-03750-2
22. Liu H, Guo X, Gooneratne R, Lai R, Zeng C, Zhan F, Wang W. 2016. The gut microbiome and degradation enzyme activity of wild freshwater fishes influenced by their trophic levels. Sci Rep 6:24340. https://doi.org/10.1038/srep24340
23. Hu Z, Tong Q, Chang J, Yu J, Li S, Niu H, Ma D. 2021. Gut bacterial communities in the freshwater snail *Planorbella trivolvis* and their modification by a non-herbivorous diet. PeerJ 9:e10716. https://doi.org/10.7717/peerj.10716
24. Yang G, Jian SQ, Cao H, Wen C, Hu B, Peng M, Peng L, Yuan J, Liang L. 2019. Changes in microbiota along the intestine of grass carp (*Ctenopharyngodon idella*): community, interspecific interactions, and functions. Aquaculture 498:151–161. https://doi.org/10.1016/j.aquaculture.2018.08.062
25. Pudlo NA, Pereira GV, Parnami J, Cid M, Markert S, Tingley JP, Unfried F, Ali A, Varghese NJ, Kim KS, Campbell A, Urs K, Xiao Y, Adams R, Martin D, Bolam DN, Becher D, Eloe-Fadrosh EA, Schmidt TM, Abbott DW, Schweder T, Hehemann JH, Martens EC. 2022. Diverse events have transferred genes for edible seaweed digestion from marine to human gut bacteria. Cell Host Microbe 30:314–328. https://doi.org/10.1016/j.chom.2022.02.001
26. Hehemann JH, Correc G, Barbeyron T, Helbert W, Czjzek M, Michel G. 2010. Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. Nature 464:908–912. https://doi.org/10.1038/nature08937
27. Abe K, Nomura N, Suzuki S. 2020. Biofilms: hot spots of horizontal gene transfer (HGT) in aquatic environments, with a focus on a new HGT mechanism. FEMS Microbiol Ecol 96:fiaa031. https://doi.org/10.1093/femsec/fiaa031
28. Romero Marcia AD, Yao T, Chen M-H, Oles RE, Lindemann SR. 2021. Fine carbohydrate structure of dietary resistant glucans governs the structure and function of human gut microbiota. Nutrients 13:2924. https://doi.org/10.3390/nu13092924
29. Lindemann SR. 2020. A piece of the pie: engineering microbiomes by exploiting division of labor in complex polysaccharide consumption.

Curr Opin Chem Eng 30:96–102. https://doi.org/10.1016/j.coche.2020.08.004
30. Sofia F. 2018 The state of world fisheries and aquaculture 2018-meeting the sustainable development goals. Food and Agriculture Organization of the United Nations, Rome.
31. Cottrell RS, Metian M, Froehlich HE, Blanchard JL, Sand Jacobsen N, McIntyre PB, Nash KL, Williams DR, Bouwman L, Gephart JA, Kuempel CD, Moran DD, Troell M, Halpern BS. 2021. Time to rethink trophic levels in aquaculture policy. Rev Aquac 13:1583–1593. https://doi.org/10.1111/raq.12535
32. Wan AHL, Davies SJ, Soler - Vila A, Fitzgerald R, Johnson MP. 2019. Macroalgae as a sustainable aquafeed ingredient. Rev Aquac 11:458–492. https://doi.org/10.1111/raq.12241
33. Sofyan A, Irawan A, Herdian H, Harahap MA, Sakti AA, Suryani AE, Novianty H, Kurniawan T, Darma ING, Windarsih A, Jayanegara A. 2022. Effects of various macroalgae species on methane production, rumen fermentation, and ruminant production: a meta-analysis from *in vitro* and *in vivo* experiments. Anim Feed Sci Technol 294:115503. https://doi.org/10.1016/j.anifeedsci.2022.115503
34. Pardesi B, Roberton AM, Wollmuth EM, Angert ER, Rosendale DI, White WL, Clements KD. 2022. *Tannockella kyphosi* gen. nov., sp. nov., a member of the family *Erysipelotrichaceae*, isolated from the hindgut of the marine herbivorous fish *Kyphosus sydneyanus*. Int J Syst Evol Microbiol 72. https://doi.org/10.1099/ijsem.0.005374
35. Maneein S, Milledge JJ, Nielsen BV, Harvey PJ. 2018. A review of seaweed pre-treatment methods for enhanced biofuel production by anaerobic digestion or fermentation. Fermentation 4:100. https://doi.org/10.3390/fermentation4040100
36. González-Gloria KD, Rodríguez-Jasso RM, Aparicio E, Chávez González ML, Kostas ET, Ruiz HA. 2021. Macroalgal biomass in terms of third-generation biorefinery concept: current status and techno-economic analysis – a review. Bioresour Technol Rep 16:100863. https://doi.org/10.1016/j.biteb.2021.100863
37. Sutherland AD, Varela JC. 2014. Comparison of various microbial inocula for the efficient anaerobic digestion of *Laminaria hyperborea*. BMC Biotechnol 14:7. https://doi.org/10.1186/1472-6750-14-7
38. Milledge JJ, Nielsen BV, Maneein S, Harvey PJ. 2019. A brief review of anaerobic digestion of algae for bioenergy. Energies 12:1166. https://doi.org/10.3390/en12061166
39. Marshall WS, Grosell M. 2006. Ionic and solute transport, p 179–196. In Evans DH, Clairborne JB (ed), The physiology of fishes, 3rd ed
40. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170
41. Wood DE, Lu J, Langmead B. 2019. Improved metagenomic analysis with Kraken 2. Genome Biol 20:257. https://doi.org/10.1186/s13059-019-1891-0
42. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, Connor R, Funk K, Kelly C, Kim S, Madej T, Marchler-Bauer A, Lanczycki C, Lathrop S, Lu Z, Thibaud-Nissen F, Murphy T, Phan L, Skripchenko Y, Tse T, Wang J, Williams R, Trawick BW, Pruitt KD, Sherry ST. 2022. Database resources of the national center for biotechnology information. Nucleic Acids Res 50:D20–D26. https://doi.org/10.1093/nar/gkab1112
43. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. Genome Res 27:824–834. https://doi.org/10.1101/gr.213959.116
44. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11:119. https://doi.org/10.1186/1471-2105-11-119
45. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153
46. Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, Busk PK, Xu Y, Yin Y. 2018. dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. Nucleic Acids Res 46:W95–W101. https://doi.org/10.1093/nar/gky418
47. Teufel F, Almagro Armenteros JJ, Johansen AR, Gíslason MH, Pihl SI, Tsirigos KD, Winther O, Brunak S, von Heijne G, Nielsen H. 2022. SignalP 6.0 predicts all five types of signal peptides using protein language models. Nat Biotechnol 40:1023–1025. https://doi.org/10.1038/s41587-021-01156-3
48. Aramaki T, Blanc-Mathieu R, Endo H, Ohkubo K, Kanehisa M, Goto S, Ogata H. 2020. KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. Bioinformatics 36:2251–2252. https://doi.org/10.1093/bioinformatics/btz859
49. Buchfink B, Reuter K, Drost HG. 2021. Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat Methods 18:366–368. https://doi.org/10.1038/s41592-021-01101-x
50. López-Mondéjar R, Tláskal V, da Rocha UN, Baldrian P. 2022. Global distribution of carbohydrate utilization potential in the prokaryotic tree of life. mSystems 7:e0082922. https://doi.org/10.1128/msystems.00829-22
51. Nishimura Y, Yoshizawa S. 2022. The OceanDNA MAG catalog contains over 50,000 prokaryotic genomes originated from various marine environments. Sci Data 9:305. https://doi.org/10.1038/s41597-022-01392-5
52. Uritskiy GV, DiRuggiero J, Taylor J. 2018. MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. Microbiome 6:158. https://doi.org/10.1186/s40168-018-0541-1
53. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055. https://doi.org/10.1101/gr.186072.114
54. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. 2019. GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. Bioinformatics 36:1925–1927. https://doi.org/10.1093/bioinformatics/btz848
55. Parks DH, Chuvochina M, Rinke C, Mussig AJ, Chaumeil PA, Hugenholtz P. 2022. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. Nucleic Acids Res 50:D785–D794. https://doi.org/10.1093/nar/gkab776
56. Ren J, Song K, Deng C, Ahlgren NA, Fuhrman JA, Li Y, Xie X, Poplin R, Sun F. 2020. Identifying viruses from metagenomic data using deep learning. Quant Biol 8:64–77. https://doi.org/10.1007/s40484-019-0187-4
57. Nayfach S, Camargo AP, Schulz F, Eloe-Fadrosh E, Roux S, Kyrpides NC. 2021. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. Nat Biotechnol 39:578–585. https://doi.org/10.1038/s41587-020-00774-7
58. Pons JC, Paez-Espino D, Riera G, Ivanova N, Kyrpides NC, Llabrés M. 2021. VPF-class: taxonomic assignment and host prediction of uncultivated viruses based on viral protein families. Bioinformatics 37:1805–1813. https://doi.org/10.1093/bioinformatics/btab026
59. Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, van Wezel GP, Medema MH, Weber T. 2021. antiSMASH 6.0: improving cluster detection and comparison capabilities. Nucleic Acids Res 49:W29–W35. https://doi.org/10.1093/nar/gkab335
60. Kautsar SA, van der Hooft JJJ, de Ridder D, Medema MH. 2021. BiG-SLiCE: a highly scalable tool maps the diversity of 1.2 million biosynthetic gene clusters. Gigascience 10:giaa154. https://doi.org/10.1093/gigascience/giaa154
61. Kautsar SA, Blin K, Shaw S, Weber T, Medema MH. 2021. BiG-FAM: the biosynthetic gene cluster families database. Nucleic Acids Res 49:D490–D497. https://doi.org/10.1093/nar/gkaa812
62. Sánchez-Navarro R, Nuhamunada M, Mohite OS, Wasmund K, Albertsen M, Gram L, Nielsen PH, Weber T, Singleton CM. 2022. Long-read metagenome-assembled genomes improve identification of novel complete biosynthetic gene clusters in a complex microbial activated sludge ecosystem. mSystems 7:e0063222. https://doi.org/10.1128/msystems.00632-22
63. Huang R, Wang Y, Liu D, Wang S, Lv H, Yan Z. 2023. Long-read metagenomics of marine microbes reveals diversely expressed secondary metabolites. Microbiol Spectr 11:e0150123. https://doi.org/10.1128/spectrum.01501-23
64. Pascal Andreu V, Roel-Touris J, Dodd D, Fischbach MA, Medema MH. 2021. The gutSMASH web server: automated identification of primary metabolic gene clusters from the gut microbiota. Nucleic Acids Res 49:W263–W270. https://doi.org/10.1093/nar/gkab353
65. Asnicar F, Thomas AM, Beghini F, Mengoni C, Manara S, Manghi P, Zhu Q, Bolzan M, Cumbo F, May U, Sanders JG, Zolfo M, Kopylova E, Pasolli E,

Knight R, Mirarab S, Huttenhower C, Segata N. 2020. Precise phylogenetic analysis of microbial isolates and genomes from metagenomes using PhyloPhlAn 3.0. Nat Commun 11:2500. https://doi.org/10.1038/s41467-020-16366-7
66. Segata N, Börnigen D, Morgan XC, Huttenhower C. 2013. PhyloPhlAn is a new method for improved phylogenetic and taxonomic placement of microbes. Nat Commun 4:2304. https://doi.org/10.1038/ncomms3304
67. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772–780. https://doi.org/10.1093/molbev/mst010
68. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313. https://doi.org/10.1093/bioinformatics/btu033
69. R Core Team. 2022. R: a language and environment for statistical computing. Vienna, Austria. Available from: https://www.R-project.org/
70. Wang L-G, Lam T-Y, Xu S, Dai Z, Zhou L, Feng T, Guo P, Dunn CW, Jones BR, Bradley T, Zhu H, Guan Y, Jiang Y, Yu G. 2020. Treeio: an R package for phylogenetic tree input and output with richly annotated and associated data. Mol Biol Evol 37:599–603. https://doi.org/10.1093/molbev/msz240
71. Yu G. 2020. Using ggtree to visualize data on tree-like structures. Curr Protoc Bioinformatics 69:e96. https://doi.org/10.1002/cpbi.96
72. Xu S, Dai Z, Guo P, Fu X, Liu S, Zhou L, Tang W, Feng T, Chen M, Zhan L, Wu T, Hu E, Jiang Y, Bo X, Yu G. 2021. ggtreeExtra: compact visualization of richly annotated phylogenetic data. Mol Biol Evol 38:4039–4042. https://doi.org/10.1093/molbev/msab166
73. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797. https://doi.org/10.1093/nar/gkh340
74. Zhou L, ggmsa YG. 2022. ggmsa: plot multiple sequence alignment using "ggplot2" Available from: http://yulab-smu.top/ggmsa/
75. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2 – approximately maximum-likelihood trees for large alignments. PLoS One 5:e9490. https://doi.org/10.1371/journal.pone.0009490
76. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH. 2015. CDD: NCBI's conserved domain database. Nucleic Acids Res 43:D222–D226. https://doi.org/10.1093/nar/gku1221
77. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. 2022. ColabFold: making protein folding accessible to all. Nat Methods 19:679–682. https://doi.org/10.1038/s41592-022-01488-1
78. Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, Morris JH, Ferrin TE. 2021. UCSF ChimeraX: structure visualization for researchers, educators, and developers. Protein Sci 30:70–82. https://doi.org/10.1002/pro.3943
79. Crooks GE, Hon G, Chandonia JM, Brenner SE. 2004. WebLogo: a sequence logo generator. Genome Res 14:1188–1190. https://doi.org/10.1101/gr.849004
80. Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, Schulz F, Jarett J, Rivers AR, Eloe-Fadrosh EA, et al. 2017. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. Nat Biotechnol 35:725–731. https://doi.org/10.1038/nbt.3893
81. Pascal Andreu V, Augustijn HE, Chen L, Zhernakova A, Fu J, Fischbach MA, Dodd D, Medema MH. 2023. gutSMASH predicts specialized primary metabolic pathways from the human gut microbiota. Nat Biotechnol 41:1416–1423. https://doi.org/10.1038/s41587-023-01675-1
82. Hehemann JH, Kelly AG, Pudlo NA, Martens EC, Boraston AB. 2012. Bacteria of the human gut microbiome catabolize red seaweed glycans with carbohydrate-active enzyme updates from extrinsic microbes. Proc Natl Acad Sci USA 109:19786–19791. https://doi.org/10.1073/pnas.1211002109
83. Ben Aissa F, Postec A, Erauso G, Payri C, Pelletier B, Hamdi M, Ollivier B, Fardeau M-L. 2014. *Vallitalea pronyensis* sp. nov., isolated from a marine alkaline hydrothermal chimney. Int J Syst Evol Microbiol 64:1160–1165. https://doi.org/10.1099/ijs.0.055756-0
84. Donia MS, Fischbach MA. 2015. Small molecules from the human microbiota. Science 349:1254766. https://doi.org/10.1126/science.1254766
85. Zhang X-X, Lv Q-B, Yan Q-L, Zhang Y, Guo R-C, Meng J-X, Ma H, Qin S-Y, Zhu Q-H, Li C-Q, Liu R, Liu G, Li S-H, Sun D-B, Ni H-B. 2022. A catalog of over 5,000 metagenome-assembled microbial genomes from the caprinae gut microbiota. Microbiol Spectr 10:e0221122. https://doi.org/10.1128/spectrum.02211-22
86. Arnosti C, Wietz M, Brinkhoff T, Hehemann J-H, Probandt D, Zeugner L, Amann R. 2021. The biogeochemistry of marine polysaccharides: sources, inventories, and bacterial drivers of the carbohydrate cycle. Ann Rev Mar Sci 13:81–108. https://doi.org/10.1146/annurev-marine-032020-012810
87. Moran D, Clements KD. 2002. Diet and endogenous carbohydrases in the temperate marine herbivorous fish *Kyphosus sydneyanus*. J Fish Biol 60:1190–1203. https://doi.org/10.1111/j.1095-8649.2002.tb01714.x
88. Skea GL, Mountfort DO, Clements KD. 2005. Gut carbohydrases from the New Zealand marine herbivorous fishes *Kyphosus sydneyanus* (Kyphosidae), *Aplodactylus arctidens* (Aplodactylidae) and *Odax pullus* (Labridae). Comp Biochem Physiol B Biochem Mol Biol 140:259–269. https://doi.org/10.1016/j.cbpc.2004.10.008
89. Majumdar A, Siva Venkatesh IP, Basu A. 2023. Short-chain fatty acids in the microbiota–gut–brain axis: role in neurodegenerative disorders and viral infections. ACS Chem Neurosci 14:1045–1062. https://doi.org/10.1021/acschemneuro.2c00803
90. Morrison DJ, Preston T. 2016. Formation of short chain fatty acids by the gut microbiota and their impact on human metabolism. Gut Microbes 7:189–200. https://doi.org/10.1080/19490976.2015.1134082
91. Clements KD, German DP, Piché J, Tribollet A, Choat JH. 2016. Integrating ecological roles and trophic diversification on coral reefs: multiple lines of evidence identify parrotfishes as microphages. Biol J Linn Soc. https://doi.org/10.1111/bij.12914
92. White WL, Coveny AH, Robertson J, Clements KD. 2010. Utilisation of mannitol by temperate marine herbivorous fishes. J Exp Mar Biol Ecol 391:50–56. https://doi.org/10.1016/j.jembe.2010.06.007
93. Seeto GS, Veivers PC, Clements KD, Slaytor M. 1996. Carbohydrate utilisation by microbial symbionts in the marine herbivorous fishes *Odax cyanomelas* and *Crinodus lophodon*. J Comp Physiol B 165:571–579. https://doi.org/10.1007/BF00387519
94. Giri S, Waschina S, Kaleta C, Kost C. 2019. Defining division of labor in microbial communities. J Mol Biol 431:4712–4731. https://doi.org/10.1016/j.jmb.2019.06.023
95. Peng Q, Lin L, Tu Q, Wang X, Zhou Y, Chen J, Jiao N, Zhou J. 2023. Unraveling the roles of coastal bacterial consortia in degradation of various lignocellulosic substrates. mSystems 8:e0128322. https://doi.org/10.1128/msystems.01283-22
96. Viborg AH, Terrapon N, Lombard V, Michel G, Czjzek M, Henrissat B, Brumer H. 2019. A subfamily roadmap of the evolutionarily diverse glycoside hydrolase family 16 (GH16). J Biol Chem 294:15973–15986. https://doi.org/10.1074/jbc.RA119.010619