Comparative metagenomics of tropical reef fishes show conserved core gut functions across hosts and diets with diet-related functional gene enrichments

Derek G. Wu[1], Cassandra R. Harris[1], Katie M. Kalis[1], Jennifer F. Biddle[1], Ibrahim F. Farag[1*]

[1]School of Marine Science and Policy, University of Delaware, Lewes DE 19958 USA

Running title: Fish gut metagenome

*corresponding author: faragif@udel.edu

Abstract

Fish gut microbial communities are important for the breakdown and energy harvesting of the host diet. Microbes within the fish gut are selected by environmental and evolutionary factors. To understand how fish gut microbial communities are shaped by diet, three tropical fish species (hawkfish, *Paracirrhites arcatus*; yellow tang, *Zebrasoma flavescens*; and triggerfish, *Rhinecanthus aculeatus*) were fed piscivorous (fish meal pellets), herbivorous (seaweed), and invertivorous (shrimp) diets, respectively. From fecal samples, a total of 43 metagenome assembled genomes (MAGs) were recovered from all fish diet treatments. Each host-diet treatment harbored distinct microbial communities based on taxonomy, with Proteobacteria, Bacteroidota, and Firmicutes being the most represented. Based on their metagenomes, microbial communities from all three host-diet treatments demonstrated a baseline ability to degrade proteinaceous, fatty acid, and simple carbohydrate inputs and carry out central carbon metabolism, lactate and formate fermentation, acetogenesis, nitrate respiration, and B vitamin

synthesis. The herbivorous yellow tang harbored a more functionally diverse microbial community with some complex polysaccharide degradation specialists, while the piscivorous hawkfish's gut community was more specialized for the degradation of proteins. The invertivorous triggerfish's gut microbiome lacked many carbohydrate degrading capabilities, resulting in a more specialized, functionally uniform community. Across all treatments, several MAGs were able to participate in only individual steps of the degradation of complex polysaccharides, suggestive of microbial community networks that degrade complex inputs. These data suggest the existence of a functional core microbiome that is common among fish species, although the specific taxonomic identities of the associated bacteria may differ.

Importance

The benefits of healthy microbiomes for vertebrate hosts include the breakdown of food into more readily usable forms and production of essential vitamins from their host's diet. Compositions of microbial communities in the guts of fish in response to diet have been studied, but lack a comprehensive understanding of the genome-based metabolic capabilities of specific microbes and how they support their hosts. Therefore, we assembled genomes of several gut microbes collected from the feces of three fish species that were being fed different diets to illustrate how individual microbes can carry out specific steps in the degradation and energy utilization of various food inputs and support their host. We found evidence that fish gut microbial communities share several core functions despite differences in microbial taxonomy. Herbivorous fish harbored a functionally diverse microbial community with plant matter degraders, while the piscivorous and invertivorous fish had microbiomes more specialized in protein degradation.

Introduction

There are many factors that affect the composition of gut microbes in fish, including the microbial composition of the surrounding water, diet, and environmental conditions such as salinity, temperature, and pH (Nayak, 2010; Yukgehnaish et al., 2020). Additionally, gut microbiomes are dynamic, where fish hatch with few microbes in their gut as larvae (Reid et al., 2009), which increase in quantity and diversity as they take in water and food from the surrounding environment (Nayak, 2010). The composition of the aquatic microbial environment has been shown to correlate with the gut microbiomes of the fishes in those aquatic environments (Arockiaraj et al., 2014; Léon-Zayas et al. 2020; Pan et al. 2022), while other studies have shown a divergence between gut and environmental microbial communities (Li et al., 2012). Gut microbiomes in individual fishes in the same environment often possess distinct microbial communities (Léon-Zayas et al. 2020), and these communities can remain relatively stable over time in the absence of disturbances (Pisaniello et al., 2023).

In marine fish, Proteobacteria, Fusobacteria, Firmicutes, Bacteroidetes, Actinobacteria and Verrucomicrobia are among the most commonly identified phyla in intestinal samples (Llewellyn et al., 2014). At lower taxonomic scales, *Vibrio*, *Pseudomonas*, *Achromobacter*, *Corynebacterium*, *Alteromonas*, *Flavobacterium*, and *Micrococcus* are considered to be the predominant gut colonizers (Nayak, 2010). Although the existence of baseline similarities in "core" gut microbial communities among congeneric fish hosts has been suggested (Shade & Handelsman, 2011; Givens et al., 2015; Le et al., 2019; Roeselers et al., 2011), there are also studies that show that microbiomes between congeneric hosts can diverge based on differences in environment and trophic level (Sullam et al., 2012). In a study of migratory rabbitfish, less than

half of the gut microbial community remained consistent across the fish's migratory path (Jones et al., 2018). There are also reported convergences of gut microbial communities between surgeonfish and nonsurgeonfish, suggesting that core microbiomes across fish species may exist (Miyake et al., 2014; Diwan et al., 2023).

Diet is an important factor that affects the gut microbiome. It has been shown that the midgut microbial community of convict surgeonfish reflects the microbial community of the algae that it was eating, suggesting a direct gut microbiome seeding related to diet (Parata et al., 2019). Yet, most diet-gut microbiome connections are more indirect. For example, larval diet supplementation provided to gilthead sea bream, European sea bass, and rainbow trout shifted their gut microbial communities in different ways, varying by host species (Antonopoulou et al., 2019). Similarly, changing the diet of sea bream from fishmeal to vegetables resulted in clear changes in microbial community composition, but not overall community diversity (Estruch et al., 2015).

Gut microbial communities also vary by feeding strategy and trophic level (Egerton et al., 2018; Pan et al. 2022). In the microbial community composition of surgeonfish, herbivorous and algivorous diets are most associated with distinct communities, while carnivorous and omnivorous diets tend to have dynamic and less distinct compositions (Miyake et al., 2014). Not only does diet affect the composition of microbes found in the gut, but it also affects the metabolism that occurs in the gut. Carnivorous diets are classified by the presence of pathways that relate to catabolism and biosynthesis of amino acids (Sanders et al., 2015). Herbivorous diets are classified by the pathways linked to pyruvate metabolism and a high abundance of short chain fatty acids, which are indicators of fermentation happening in the gut (Mountfort et al., 2002; Jones et al., 2018). Because herbivores do not intake an abundance of protein in their diet,

they go through the process of fermentation to produce amino acids (Sanders et al., 2015). To date, invertivorous diets have been less studied for gut microbiome functions.

Several studies have shown how individual gut microbes play putative roles in the degradation of key diet inputs (Ray et al., 2012). For example, in guts from carp fed grass diets, hundreds of bacterial isolates have been cultured that demonstrate cellulolytic activity (Li et al., 2014). A study of three herbivorous fishes found that the activity of amylase enzymes in the guts positively correlated with diets that were higher in rhodophytes, while laminarases were negatively correlated with phaephytic diets (Moran & Clements, 2005). Using a genomic content estimator, PICRUSt, based on 16S rRNA amplicon libraries, and enzymatic tests, herbivorous diets were associated with fish gut microbes that produced more cellulases, while carnivorous diets were correlated with gut microbes that produced more trypsins (Liu et al., 2016). In rabbitfish, seaweed diets resulted in diverse gut microbial communities that were able to produce nonstarch polysaccharide degrading enzymes (X. Zhang et al., 2017). However, these patterns likely vary by host species and diet, considering that there are studies that have found little functional ability of microbial communities to degrade diet inputs (McCauley et al., 2020). The majority of studies on fish gut microbiomes have utilized only small subunit ribosomal gene amplicon sequencing for analysis, and they have clearly shown that microbiomes vary based on diet (Liu et al., 2022; Antonopoulou et al., 2019; Jones et al., 2018; Kashinskaya et al., 2018; Estruch et al., 2015; Sullam et al., 2012). Few studies have examined in detail the metabolic pathways that exist across the gut microbiota of different fish species, which metagenomic analyses can provide unique insights into (Tarnecki et al., 2017; Stevenson et al. 2022). Current studies lack an approach to understanding how microbial community members may be working in concert to degrade macromolecules delivered via the host's diet. Without full genomes of

microbial community members across different host species and diets for examination, little can be known about the ways in which microbes interact with each other and work in community to degrade food materials.

As such, we examined the gut inhabitants from multiple coral reef fish to fully understand microbial interactions in fish guts. We examine three fish species in this study: yellow tang (*Zebrasoma flavescens*), humu humu Triggerfish (*Rhinecanthus aculeatus*), and arc-eye Hawkfish (*Paracirrhites arcatus*). Yellow tangs are herbivores that naturally feed on algae and marine plants; in the experiment they were fed seaweed, a natural herbivorous diet. Triggerfish are invertivores that naturally feed on shrimp, clams and snails; for this experiment these fish were fed frozen mysis shrimp. Hawkfish are generalist predators feeding naturally on small fishes and invertebrates; they were fed fish pellets for this study. The overarching goal of this study was to determine the extent to which fish gut microbial communities were structured according to selection pressures from the host diet using comparative metagenomic approaches. We hypothesized that the fish gut communities would primarily be determined by the biochemical capabilities of recovered metagenomes to degrade and harvest energy from the host's specific diet, thus the major energy harvesting pathways exhibited by the microbial communities would largely reflect the substrate inputs (i.e., fatty acid oxidation for high fat diets and mixed acid fermentation pathways for high polysaccharide diets). In this work, we instead show that the core metabolism of each fish gut microbiome is redundant across diets and species of fish, with greater accessory metabolisms for the herbivorous diet.

Methods

Fish Rearing, Diets and Sample Collection

The yellow tangs, triggerfish, and hawkfish used for this study were obtained from Live Aquaria® (Rhinelander, WI) and cared for according to IACUC approval through the University of Delaware (A1319). Fish were held in 37L aquaria on a large recirculating system at the UD Lewes Campus (Lewes, DE). Hawkfish (n=4) were kept individually in aquarium tanks due to aggressive behavior. Yellow tangs (n=5) and triggerfish (n=5) were kept in communal aquarium tanks with their same species. Hawkfish are piscivorous fish and were fed commercial fish meal pellets (Sustainable Aquatics Dry Hatchery Diet, Jefferson City TN). Yellow tangs are herbivorous fish and were fed commercially purchased nori seaweed (sushi wrapping, various vendors) as their diet. Triggerfish are invertivores and were fed commercially purchased mysis shrimp (San Francisco Bay Brand Frozen Mysis Shrimp) as their diet. We opted to collect gut microbiomes via fecal collection to preserve the fish for other experiments (Harris, 2019).

When collecting feces, the fish were placed in individual 2.5L aquaria. To minimize a seawater microbial signal, artificial seawater was created with Milliq-filtered water and Instant Ocean (Instant Ocean; Blacksburg, VA) to 33 ppm salinity. Fish were moved to these tanks 10 min after feeding and remained there overnight with air stones and no shared water flow to ensure that water quality remained high. After 12-15 hours, the fish were placed back in their 37L tanks. From each small tank, 1500 mL of water was sampled. The tank waters were filtered through an 8.0 micron cellulose filter (Merck Millipore; Burlington, MA). This filter size was selected to ensure fecal particles were preferentially collected instead of free-living bacteria in the water. The filters were labeled and placed in a 60mm x 15mm petri plate with a stackable lid and stored in a -80 freezer until used for DNA extractions.

Chemical Analysis of Food Sources

To determine major dietary differences between the fish-meal pellets and mysis shrimp, 20g of crushed, freeze dried mysis shrimp and 20g of fish-meal pellets were shipped to NP Analytics Laboratory (St. Louis, MO) for testing. Nutrition information for nori seaweed was retrieved from the commercial packaging.

DNA Extraction and Metagenomic Sequencing

DNA was extracted from the feces samples collected on the filter using the DNeasy PowerSoil Kit (Qiagen, Valencia CA). A quarter of the filter paper was used for the extractions. After each extraction, the amount of DNA in each sample was measured using the Qubit dsDNA HS assay kit with the Qubit fluorometer (Invitrogen; Waltham, MA). The DNA was stored in a -20C freezer. Samples with the highest DNA yield were selected from each species and sent to The University of Delaware DNA Sequencing and Genotyping Center at the Delaware Biotechnology Institute (Newark, DE) for library creation and sequencing on an Illumina HiSeq 2500.

Data Accessibility Statement

Raw data associated with these samples have been deposited in the NCBI database under BioProject 1112800.

Metagenome Quality Control, Assembly and Binning

Data processing, including reads quality checking and trimming, was performed using Nesoni (Victorian-Bioinformatics-Consortium, 2014) under default parameters and applying a quality score of Q20. Qualified reads for each dataset were assembled into contigs using IDBA-UD (Peng et al., 2012), metaSPAdes (Nurk et al., 2017), and Megahit v.1.1 (Li et al., 2016) with default settings and k-mer lengths varying from 27 to 117. Contigs longer than 1000 bp were binned using binning tools MaxBin2.0 (Wu et al., 2015), Metabat (Kang et al., 2015), and

Metabat2 (Kang et al., 2019). The resulting metagenome-assembled genomes (MAGs) were quality assessed using CheckM (Parks et al., 2015) with the "lineage_wf" option. Only bacterial MAGs, meeting criteria of ≥ 50% completion and ≤ 20% contamination, were retained. These MAGs were further refined by removing outlier contigs based on t-SNE signatures, GC content, taxonomic assignments, and differential coverages. Coverage calculations were performed by mapping trimmed reads against the contigs using BBMap v.37.61 (Bushnell, 2014). MAGs with < 10% contamination were deemed acceptable, while those with 10-20% contamination were manually cleaned again.

Phylogenetic Placement

The phylogenetic placements of the MAGs were estimated by creating a phylogenomic tree from concatenated ribosomal proteins, aligned using Muscle v3.8 (Edgar, 2004) and analyzed with FastTree v2.1 (Price et al., 2010). MAG identities were confirmed by calculating average nucleotide identity (ANI) and average amino acid identity (AAI) with custom reference genomes closely related to the screened lineages based on their single marker gene(s) closest matches. Taxonomic profiles were verified using GTDB-Tk v2.1.0 (Chaumeil et al., 2022).

Metabolic Reconstruction

Protein sequences from all MAGs were predicted using Prodigal v2.6.3 (Hyatt et al., 2010) using the default translation table 11. The collective metabolic pathways for all MAGs present in each depth were deduced via screening the predicted proteins against a custom hidden Markov model database composed of 126 key metabolic genes covering most of the substrate utilization, biosynthesis and energy metabolism related pathways. These pathways were further assessed for completion through querying the protein sequences of the potential MAGs against the Kyoto Encyclopedia of Genes and Genomes (KEGG) (https://www.genome.jp/kegg/)

(Kanehisa et al., 2023) database using the BlastKoala tool (Kanehisa et al., 2016). The pathway was deemed complete if the MAG encodes for at least 80% of the enzymes and other involved proteins of a given pathway.

Substrate Degradation Analysis

Enzymes involved in the degradation of cellulose, hemicelluloses (xylan, mannan, and xyloglucan), and starch were identified as presented in Kumla et al. (2020). Enzymes for the degradation of chitin were identified as presented in De Tender et al. (2019). Corresponding CAZyme families for these enzymes were identified on the Carbohydrate-Active enZYmes Database (http://www.cazy.org/) (Drula et al., 2022). CAZyme family hits were isolated from the metagenome metabolic readout and reconstructed to follow the steps of macromolecular degradation. Exopeptidases were identified using the MEROPS Peptidase Database (https://www.ebi.ac.uk/merops/) (Rawlings et al., 2014). Enzymes critical for the beta oxidation of fatty acids, central carbon metabolic pathways, amino acid catabolism, and vitamin biosynthesis were identified via the KEGG database. Enzymes involved in lactate, acetate, and ethanol fermentation pathways were organized as described in the EcoCyc Database (https://ecocyc.org/) (Keseler et al., 2011). Formate fermentation enzymes were identified as described in Knappe & Sawers (1990). Hydrogenase functions were identified on the HydDB database (https://services.birc.au.dk/hyddb/) (Søndergaard et al., 2016).

Statistical Analysis

To visualize clustering patterns in substrate degradation enzymes in response to host diets, a principal component analysis of all screened substrate pathways was performed on all samples using the prcomp function in R Version 4.3.2 (R Core Team, 2023). The total number of gene copies for all enzymes involved in the degradation of proteins, cellulose, hemicellulose,

starch, and chitin normalized to genome completeness and a binary system for the presence/absence of fatty acid degradation were used for PCA processing. The loadings of the first two components are reported, explaining 76.7% of detected variance. PCA data were visualized using the fviz_pca_biplot function from the factoextra package (Kassambara & Mundt, 2020). To test if the centroids of diet treatments were different, a permutational multivariate analysis of variance (PERMANOVA) was performed using the adonis2 function from the vegan package (Oksanen et al., 2022). To assess the homogeneity of multivariate dispersions across diets, the betadisper and permutest functions were used to compute the distances from group centroids to individual samples with 999 permutations.

To determine if MAGs recovered from each treatment differed in the number of substrate degrading gene copies, the total gene copies for each degradation pathway normalized to MAG genome completeness were totaled. Then, a Kruskal Wallis test was performed to determine if there were significant differences among the three diets using the kruskal.test function. If significant, a post-hoc Wilcoxen test was performed to identify significantly different pairwise comparisons between diets.

Ethics Statement

The handling and care protocols of the fish used in this study were approved by the University of Delaware Institutional Animal Care and Use Committee (A1319), AUP #1319-2019-2.

Results

Diet Composition

The majority of both the fish meal pellet (56.1%) and shrimp (69.3%) diets was composed of protein (Supplemental Table 1). The fish meal pellets had a total fat content of 16.6%, while the shrimp diet was 6.87% fat (Supplemental Table 1). Of the fats, both the fish pellets and shrimp had a broad composition of fatty acids. From the fish pellet diet, the largest contributions of fats were C22:6n3 docosahexaenoic, C20:1n9 cis eicosenoic, and C16:1n7 palmitoleic acids (Supplemental Table 2). From the shrimp diet, the largest contributions of fats were C16:0 palmitic, C22:6n3 docosahexaenoic, and C20:5n3 eicosapentaenoic acids (Supplemental Table 2). Both diets had a small component of fiber (2.42% and 3.57% for fish pellets and shrimp, respectively) (Supplemental Table 1). Both diets could provide the gut microbial communities with biologically relevant minerals, such as potassium, calcium, phosphorus, and metals like iron, magnesium, zinc, manganese, and copper (Supplemental Table 1). The seaweed, fed as nori paper, contained almost equivalent protein and carbohydrates (5.7% and 5%, respectively), and minor components of fiber and sugar (both 0.4%). Seaweed had significant amounts of iodine, vitamins C, A and B12 and also contained calcium, potassium, iron and magnesium.

Metagenome-Assembled Genome (MAG) Recovery

A total of 18 MAGs were recovered from the piscivorous hawkfish treatment, with an average genome completeness and contamination of 91.44% and 4.47%, respectively (Table 1). These 18 MAGs contained 4 taxonomic classes: *Alphaproteobacteria* (8), *Bacteroidia* (3), *Gammaproteobacteria* (6), and *Gracilibacteria* (1) (Table 1). The gut microbiome showed more taxonomic diversity at subclass levels which indicates taxonomic niche diversification and specialization. For example, in *Gammaproteobacteria*, 4/6 are in the order *Enterobacteriales* and 2/6 are in the *Pseudomonadales*. All 3/3 *Bacteriodia* belong to the *Cytophagales* (Table 1).

275 From the herbivorous yellow tang treatment, a total of 18 MAGs were recovered, with an
276 average genome completeness and contamination of 94.19% and 3.63%, respectively (Table 1).
277 In contrast to the piscivorous gut community, the herbivorous community appeared to have more
278 taxonomic diversity at higher taxonomic ranks (e.g., phylum or order levels). The 18 MAGs
279 represented 7 taxonomic classes: *Alphaproteobacteria* (1), *Bacilli* (1), *Bacteroidia* (5), *Clostridia*
280 (3), *Gammaproteobacteria* (6), *Gracilibacteria* (1), and *Spirochaetia* (1) (Table 1). The
281 *Bacteroidia* were entirely comprised of the *Bacteroidales* order, and the Gammaproteobacteria
282 all belonged to the *Enterobacteriales*. The *Clostridia*, belonging to the Firmicutes phylum, were
283 represented by two closely related *Oscillospirales* MAGs and one *Lachnospirales* (Table 1).
284 A total of 7 MAGs were recovered from the invertivorous triggerfish treatment, with an
285 average genome completeness and contamination of 90.0% and 3.3%, respectively (Table 1).
286 The 7 MAGs represented 5 taxonomic classes: *Alphaproteobacteria* (2), *Bacilli* (1),
287 *Bacteriovoracia* (1), *Campylobacter* (1), and *Gammaproteobacteria* (2) (Table 1). Similar to the
288 piscivorous gut community, this community exhibited lower diversity at higher taxonomic levels
289 and more diversity at subclass levels, wherein all of the MAGs belonged to different orders.
290 *Diet-Based Analyses*
291 We investigated the gut microbiome composition in response to diet types by performing
292 principal component analysis on the copy numbers of enzymes that degrade substrates, which
293 were enriched in relation to each diet type. These data suggest that fish host diet is a major
294 driving selection factor for microbiome composition (PERMANOVA; $F = 6.1096$; $p = 0.003$)
295 (Figure 1). Hemicellulose, cellulose, chitin, and starch were forces loaded towards the
296 herbivorous diet whereas fatty acids and proteins were more predominant in the piscivorous and
297 invertivorous diets (Figure 1). The herbivorous gut community displayed a greater variance in

gene copies, while the invertivore gut community had the narrowest variation in substrate degradation gene compositions (homogeneity of dispersion; $F = 7.8859$; $p = 0.002$) (Figure 1). Although several of the MAGs collected from the herbivorous yellow tang demonstrated specialization for the degradation of complex polysaccharides, its microbiome also contained MAGs that were more specialized in protein and fatty acid degradation. In contrast, the invertivorous triggerfish microbiome had a narrower range of substrate degradation functions, suggestive of a more specialized community that is adapted to the high protein shrimp diet.

Carbohydrates

Overall genes encoding for carbohydrate degrading enzymes are more enriched in the MAGs recovered from herbivorous fish compared to fishes following other diets. For example, the degradation of cellulose could be completed in both the herbivorous and piscivorous diets (Figure 2a). MAGs recovered from the herbivorous yellow tang had significantly more gene copies for enzymes involved in the cellulose degradation pathway than both piscivorous and invertivorous hosts (Kruskal Wallis, $p = 0.01$) (Figure 3a). Within the piscivorous hawkfish gut, 7/18 MAGs [005, 009, 011, 022, 028, 030, and 033] could facilitate the hydrolysis of cellulose to glucose monomers (Figure 2a). Of these 7 MAGs, 4 are *Alphaproteobacteria,* 2 are *Bacteroidia*, and 1 is a *Gammaproteobacteria*. From the herbivorous diet, 9/18 MAGs [001, 002, 005, 006, 007, 009, 014.3, 015, and 016] could catalyze these reactions. Of these 9 MAGs, 3 are *Gammaproteobacteria*, 3 are *Bacteroidia*, 2 are *Clostridia*, and 1 is *Gracilibacteria*. In contrast, no MAGs recovered from the invertivorous triggerfish could degrade cellulose to simple monomers (Figure 2a). Additionally, in all treatments, there were several MAGs that could potentially catalyze at least one step of the larger cellulose degradation pathway but not the entire process, hereby referred to as “partial degraders” (Figure 2a). From the piscivorous treatment,

321 MAGs 001, 010, and 015 were partial degraders, while from the herbivorous treatment, MAGs
322 003, 011, 012, and 017 were partial degraders of cellulose. Most of these MAGs appear to be
323 opportunists that require other MAGs to initiate the breakdown of cellulose into simpler
324 substrates for them to use.

325 The MAGs recovered from the herbivorous yellow tang had significantly higher gene
326 copies of enzymes involved in hemicellulosic degradation than both the piscivorous and
327 invertivorous treatments (Kruskal Wallis, $p = 0.0075$) (Figure 3b). From the herbivorous
328 treatment, 5/18 MAGs [001, 006, 007, 011, and 015] could completely degrade at least one of the
329 three screened hemicelluloses (Figure 2b-d). Of these 5 MAGs, 2 are *Clostridia*, 2 are
330 *Bacteroidia*, and 1 is a *Gammaproteobacteria*. From the piscivorous treatment, only 2/18 [022
331 and 030] were able to fully degrade at least one hemicellulosic compound (Figure 2b-d). No
332 MAGs recovered from the invertivorous treatment could completely degrade hemicellulose
333 (Figure 2b-d). However, in all treatments, there were several MAGs that were partial degraders
334 of hemicellulose backbone or one of the side chains (Figure 2b-d). From the herbivorous
335 treatment, all MAGs except the complete degraders could be classified as partial degraders of
336 various hemicelluloses capable of performing one or more steps in the hemicellulose degradation
337 pathway. Similarly, all the MAGs recovered from the piscivorous could catalyze at least one step
338 of a hemicellulose degradation pathway; however, there was a lack of capacity to hydrolyze and
339 ferment many of the auxiliary sugars from the hemicellulose compounds.

340 The herbivorous gut community had significantly greater numbers of genes involved in
341 the degradation of chitin than both the invertivorous and piscivorous communities; 2.2 fold and
342 13.1 fold, respectively (Kruskal Wallis, $p = 0.0082$) (Figure 3c). In fact, the invertivorous gut
343 community showed an inability to catalyze the initial steps to either of the chitin degradation

344 pathways (Figure 2e). Meanwhile, 2/18 MAGs [009 and 030] from the piscivorous treatment and
345 5/18 MAGs [001, 003, 005, 006, and 015] from the herbivorous treatment could fully catalyze
346 the breakdown of chitin to simpler glucosamine units. The 2 MAGs from the piscivorous
347 treatment that can degrade chitin are both in the *Cytophagales* order within the *Bacteroidia* class.
348 Of the 5 chitin-degrading MAGs from the herbivorous treatment, 2 are *Bacteroidia*, 2 are
349 *Clostridia*, and 1 is a *Gammaproteobacteria*. All three gut communities produced amylases for
350 the degradation of starches. However, the herbivorous gut community members had significantly
351 more amylase gene copies than the piscivorous community members (Kruskal Wallis, $p = 0.05$)
352 (Figure 3d).

353 *Proteins*

354 All three hosts contained MAGs that produced exopeptidases for protein degradation,
355 allowing for basal levels of protein breakdown in the gut environment. The piscivorous gut
356 MAGs had 1.69 fold more exopeptidase gene copies per MAG than those of the herbivorous gut
357 community (Figure 3e), suggesting that protein degradation pathways were enriched in response
358 to the high protein fish pellet diet. From the piscivorous community, 16/18 MAGs had more than
359 10 exopeptidase genes, while 11/16 and 3/7 MAGs from the herbivorous and invertivorous guts
360 had more than 10 exopeptidase genes, respectively (Table 2). Additionally, MAG 005 from the
361 piscivorous gut (*Phaeomarinobacter* sp.) was the only MAG across all three treatments to have a
362 gene for collagenase, which could highlight a specialized niche for collagen degradation in
363 response to the fish pellet diet.

364 *Fatty Acids*

365 All three hosts had MAGs that could carry out complete beta oxidation of fatty acids. The
366 screen for fatty acid breakdown via beta oxidation included genes for acyl-CoA dehydrogenase,

367 enoyl-CoA hydratase, 3-hydroxyacyl-CoA dehydrogenase, and acetyl-CoA acyltransferase. In
368 the piscivorous gut community, 9/18 MAGs [003, 006, 007, 009, 012, 015, 018, 022, 030] were
369 able to carry out fatty acid oxidation to completion to produce acetyl CoA (Figure 4a). These
370 MAGs represent the *Gammaproteobacteria* (2), *Alphaproteobacteria* (3), and *Bacteroidia* (3)
371 classes. In contrast, only 4/18 MAGs [003, 004, 012, and 016] from the herbivorous community
372 and 3/7 MAGs [002, 005, and 006] from the invertivorous community could independently
373 complete all of fatty acid oxidation (Figure 4b-c). All four MAGs from the herbivorous yellow
374 tang microbiome that could complete beta oxidation were *Gammaproteobacteria*, and 3 of them
375 were in the *Enterobacterales* order. These data highlight a potential enrichment of fatty acid
376 oxidation capacities in the piscivorous hawkfish treatment, whose fish pellets contained high
377 amounts of various fatty acids.

378 *Energy Harvesting and Fermentation*

379 MAGs recovered from all three samples were analyzed for their metabolic capacities.
380 13/18, 14/18 and 5/7 MAGs from the piscivorous, herbivorous, and invertivorous treatments,
381 respectively, could carry all of glycolysis from glucose to pyruvate (Figure 4a-c). Following
382 glycolysis, 17/18, 17/18, and 7/7 MAGs from the piscivorous, herbivorous, and invertivorous gut
383 communities could carry out pyruvate oxidation, bridging glycolysis to the citric acid cycle
384 (Figure 4a-c). 13/18, 6/18, and 4/7 MAGs from the piscivorous, herbivorous, and invertivorous
385 guts, respectively, possessed more than 80% of the genes for the entire TCA cycle (Figure 4a-c).
386 Within each treatment, there are some MAGs that can complete individual energy-harvesting
387 steps within the TCA cycle, but not the complete process, which may be a product of genome
388 incompleteness (Table 2). Since the piscivorous hawkfish sample contained a higher proportion
389 of MAGs that could carry out the TCA cycle, particularly the second oxidation reactions, these

390 data could suggest that these MAGs are more predominantly metabolizing the amino acids from
391 their proteinaceous diet via central carbon metabolism.

392 The central carbon pathways observed in these microbial communities produce reduced
393 electron carriers that need to be oxidized in order to sustain the catabolic pathways. From the
394 piscivorous treatment, 4/18 MAGs [001, 010, 012, and 030] could perform dissimilatory nitrate
395 reduction to ammonia as a mechanism for anaerobic respiration (Figure 4a). 5/18 MAGs [003,
396 004, 007, 012, and 016] from the herbivorous treatment are dissimilatory nitrate respirers, while
397 2/7 MAGs [002 and 008] from the invertivorous treatment can carry out nitrate respiration
398 (Figure 4b-c). Lactate fermentation was rare among all three treatments, with only 2/18, 1/18,
399 and 0/7 MAGs from the piscivorous, herbivorous, and invertivorous diets, respectively, being
400 able to produce a L-lactate dehydrogenase (Figure 4a-c).

401 All three gut communities demonstrated an ability to perform the latter half of the
402 acetogenesis pathway to reduce acetyl-CoA to acetate. From the piscivorous treatment, 6/18
403 MAGs had genes for both phosphate acetyltransferase and acetate kinase, synthesizing ATP
404 (Figure 4a). Similarly, 11/18 MAGs from the herbivorous diet and 4/7 MAGs from the
405 invertivorous diet could produce these enzymes (Figure 4b-c). Using the acetate products, 5/18,
406 7/18, and 3/7 MAGs from the piscivorous, herbivorous, and invertivorous communities,
407 respectively, could produce alcohol dehydrogenase homologs for the production of ethanol and
408 consumption of reduced NADH (Figure 4a-c).

409 Alternatively, all three gut communities had genes for formate acetyltransferase and
410 formate dehydrogenase, which are involved in the fermentation of formate to carbon dioxide and
411 hydrogen gas. From the piscivorous treatment, 3/18 MAGs could produce formate
412 acetyltransferase while 8/18 MAGs could produce formate dehydrogenase (Figure 4a). In

413 contrast, 11/18 MAGs from the herbivorous diet could produce formate acetyltransferase, while
414 only 3/18 MAGs could produce formate dehydrogenase (Figure 4b). From the invertivorous diet,
415 4/7 MAGs possessed formate acetyltransferase genes, while 3/7 MAGs possessed genes for
416 formate dehydrogenase (Figure 4c). None of the MAGs in this study could perform dissimilatory
417 sulfate respiration or methanogenesis (Figure 4).

418 *Vitamin and Hydrogenase Production*

419 All communities had MAGs that could produce vitamins B1, B2, B7, B9, and B12.
420 Thiamine (B1) could be produced by 7/18, 10/18, and 3/7 MAGs from the piscivorous,
421 herbivorous, and invertivorous communities, respectively (Table 2). Riboflavin (B2) could be
422 synthesized by 4/18, 6/18, and 2/7 MAGs from the piscivorous, herbivorous, and invertivorous
423 hosts guts, respectively (Table 2). Biotin (B7) could be synthesized by 12/18, 10/18, and 3/7
424 MAGs from the piscivorous, herbivorous, and invertivorous diet communities, respectively,
425 using one of the various Bio pathways (Table 2). Hydrotetrafolate, a precursor to folic acid (B9),
426 could be produced by 14/18, 9/18, and 3/7 MAGs from the piscivorous, herbivorous, and
427 invertivorous hosts guts, respectively (Table 2). Cobalamin (B12) could be made by 7/18, 4/18,
428 and 2/7 MAGs from the piscivorous, herbivorous, and invertivorous hosts guts, respectively
429 (Table 2).

430 From the piscivorous gut community, the only MAG capable of producing a component
431 of a nickel-iron hydrogenase is 001 (*Photobacterium aphoticum*), being able to produce a
432 periplasmic Group 1 hydrogenase, which is generally responsible for catalyzing steps in
433 anaerobic respiration using metals, nitrates, or sulfates as terminal electron acceptors
434 (Supplemental Figure 1). Similarly only 1 MAG [016] from the invertivorous treatment produces
435 a group A1 and A3 iron-iron hydrogenase, which produces fermentative hydrogen gas

436 (Supplemental Figure 1). Also from the invertivorous diet, 2/7 MAGs [002 and 008] from this
437 diet produce group 1 nickel-iron hydrogenases., which also act as respiratory electron receptors
438 (Supplemental Figure 1).

439 In contrast, the herbivorous treatment microbial community synthesizes various iron and
440 nickel hydrogenases. 6/18 MAGs [005, 010, 015, 017, 023, and 006] produce group A1 and A3
441 iron-iron hydrogenases (Supplemental Figure 1). 7/18 MAGs [001, 005, 010, 015, 017, 023, and
442 006] produce group B iron-iron hydrogenases involved in hydrogen gas fermentation
443 (Supplemental Figure 1). 4/18 MAGs [010, 015, 017, and 006] produce group C1 iron-iron
444 hydrogenases (Supplemental Figure 1). Only 2/18 MAGs [017 and 006] produce group C2 iron-
445 iron hydrogenases, which are transcriptional regulators (Supplemental Figure 1). 4/18 MAGs
446 [010, 015, 023, and 006] produce group C3 iron-iron hydrogenases (Supplemental Figure 1).
447 3/18 MAGs [001, 004, and 005] produce group 1 nickel-iron hydrogenases, also transcriptional
448 regulators (Supplemental Figure 1). Within the herbivorous yellow tang gut, MAGs 006, 010,
449 015, and 017 produce the greatest variety of hydrogenases, and 3 of them belong to the
450 *Clostridia* class within Firmicutes.

451 *Gut Models*

452 Based on the metagenomic results presented in the current study, models of the gut
453 microbiomes were constructed to illustrate the concerted effort of the gut microbiota to degrade
454 complex macromolecular diet inputs and harvest energy from them to highlight the differences
455 seen across the samples (Figures 5-7). The metabolic reconstructions highlight the potential for
456 individual gut microbes to perform individual steps in either degradation or energy utilization
457 pathways, creating products that are then usable by other members of the microbial community.
458

Discussion

Despite a number of publications that describe species-specific gut communities in fish, in this study using metagenome-assembled genomes, we found evidence of a core microbiome in terms of function, but not taxonomy (which was seen in Givens et al. (2015), Le et al. (2019), and Roeselers et al. (2011)), across three different fish species with different diets. Our study benefited from the aquarium-based approach we used, as we had defined food sources that the fish were fed for weeks in captivity, compared to the uncontrolled diet that may be found in nature. Core functions include protein degradation, fatty acid oxidation, central carbon metabolism, nitrate respiration, acetogenesis, and formate oxidation. Considering our sampling design and dataset limits, additional functions may be possible in fish guts and more taxa should be closely examined; however, our results show that previous work based on small subunit ribosomal sequences or the interpretation of metabolism based solely on taxonomic data is a limited view of a fish gut microbiome.

Despite the number of shared functions creating the core of the microbial community, diet types also showed selective trends. The microbial communities associated with the piscivorous and invertivorous diets were the most functionally similar, as expected, considering the compositions of the fish pellets and shrimp diets, consistent with the model of diet-based microbiomes (Wong & Rawls, 2012). Both diets were high in protein and fat content, and the MAGs recovered from these diets showed high gene copy numbers for exopeptidases and abilities to carry out beta oxidation of fatty acids. In comparing these two diet treatments, the piscivorous gut community showed much more variation in the metabolic capabilities of the recovered MAGs. For example, the piscivorous gut community, although specialized for the degradation of proteins and fatty acids, could also degrade plant matter. It could be that plant

482 material is indirectly ingested during fish consumption. In contrast, the invertivorous triggerfish
483 is primarily digesting shrimp, which likely do not contain substantial amounts of complex plant
484 polysaccharides in their intestines. Therefore, the triggerfish microbial community specializes in
485 digesting proteins and fats with some versatility for simple sugars.

486 As expected, the herbivorous yellow tang gut microbial community demonstrated a
487 strong ability to degrade complex plant polysaccharides, likely due to their nori seaweed diet.
488 Unexpectedly, the MAGs recovered from the invertivorous triggerfish diet demonstrated an
489 inability to degrade chitinous compounds. This suggests that the gut microbial community is not
490 responsible for the degradation of chitin that would be introduced into the gut system as shrimp
491 exoskeletons. This contrasts with the results which found that rainbow trout being fed
492 chitinaceous insect exuviae had gut microbiomes that were enriched in chitinolytic bacterial
493 genera; however, these results were based on predictive metagenomes based solely on 16S rRNA
494 gene amplicons (Rimoldi et al. 2023). Exochitinase production has been observed in various fish
495 species being fed insect diets, although the source of these enzymes (host versus gut
496 microbiome) was not specified (German et al., 2010; Eggink et al., 2022). Possible explanations
497 for this discrepancy in the gut microbial community's inability to degrade chitin could relate to
498 the host physiology. For example, the triggerfish may be masticating and pre-digesting the
499 chitinous shrimp shells in such a way that allows the microbial community to utilize enzymes
500 with broad substrates that could catabolize these compounds. These host forces may work in
501 concert with the gut community to carry out digestion.

502 With the higher fatty acid diet of the piscivorous hawkfish and the invertivorous
503 triggerfish, their microbiomes carried a strong signature of beta oxidation. The beta oxidation of
504 fatty acids is a reductive cycle that rapidly degrades long chain fatty acids to directly produce

ATP and acetyl-CoA. This high energy-yielding process explains its abundance in MAGs recovered from these treatments, while the herbivorous yellow tang gut microbial community had fewer MAGs capable of fatty acid degradation. Instead of relying on this high yield process, our data suggest that the herbivorous gut community relies more on mixed acid fermentation and acetate-producing pathways for the production of ATP and recycling of reduced electron carriers. This establishes a major difference in the proteinaceous and fatty acid diet economy expected in the piscivorous and invertivorous treatments and the polysaccharide diet economy expected in the herbivorous treatment.

Further supporting the polysaccharide economy expected in the herbivorous yellow tang gut, the MAGs collected from this treatment possessed a deep redundancy and breadth of hydrogenases. These hydrogenases varied in function, including being final electron acceptors in anaerobic respiration and hydrogen gas production at the end of fermentation. Since the herbivorous treatment is associated with higher concentrations of fibrous substrates, the constituent sugars would rely on electron mediated processes to degrade the substrates and recycle cofactors. The higher proportion of MAGs that produce hydrogenases supports this expectation, with these hydrogenases playing a putative role in balancing electron budgets and gradients to support respiration and fermentation pathways.

A limitation of this study that is worth noting is the possibility for there to be host-specific factors such as gut physiology and chemistry that could also be selecting for taxon and function-specific MAGs. While the present study found substantial evidence for the linkage between diet inputs and community composition, it has been shown that host species may harbor distinct microbial communities (Li et al., 2012). Future research should study model fish species that have adaptable diets or whose diets change with development (Benavides et al. 1994). Then,

diets could be shifted and changes in the microbial community composition could be followed as a function of time and treatment. The methodology used here shows that fish under treatment can have their microbiome analyzed while remaining viable for longer treatments of study. If these diet changes are considered in the context of ecological disturbances, these studies would be indicative of the likelihood of adjusting to rapidly changing environments, which is important to understand given the cascade of environmental changes being mediated by anthropogenic activity.

We recognize that keeping fish viable for this study meant that internal shifts of microbes along gut structures were unable to be observed (Zhang et al. 2017; Leon-Zayas et al. 2020; Pardesi et al. 2022). As such, these data serve as a basis for further studies whose experimental design could include multiple scenarios for sampling. However, our data does show strongly that single-gene taxonomic analyses fall short of determining differences in fish gut microbiomes. The ability to sample the full metagenome and build MAG-based models of metabolism shows that fish guts are more similar than expected for functional potential. Future studies that use advanced omics approaches to gain insights into host-microbe interactions and functions in gut systems could focus on the expression of substrate degradation genes in these environments. The present study provides insight into the probable interactions between the host diet and the genome compositions of the gut microbial communities, but in order to get insight into the functions of these MAGs within these microbiomes, transcriptomics or proteomics can provide knowledge of how these MAGs are behaving and regulating the expression of their genomes.

References:

Antonopoulou, E., Nikouli, E., Piccolo, G., Gasco, L., Gai, F., Chatzifotis, S., Mente, E., & Kormas, K. A. (2019). Reshaping gut bacterial communities after dietary tenebrio molitor larvae meal supplementation in three fish species. *Aquaculture*, 503, 628–635. https://doi.org/10.1016/j.aquaculture.2018.12.013

Arockiaraj, J., Kumaresan, V., Chaurasia, M. K., Bhatt, P., Palanisamy, R., Pasupuleti, M., Gnanam, A. J., & Kasi, M. (2014). Molecular Characterization of a Novel Cathepsin B from Striped Murrel *Channa striatus*: Bioinformatics Analysis, Gene Expression, Synthesis of Peptide and Antimicrobial Property. *Turkish Journal of Fisheries and Aquatic Sciences*, 14(2). https://doi.org/10.4194/1303-2712-v14_2_08

Benavides, A. G., Cancino, J. M., & Ojeda, F. P. (1994). Ontogenetic change in the Diet Of *Aplodactylus punctatus* (Pisces: Aplodactylidae): An ecophysiological explanation. *Marine Biology*, 118(1), 1–5. https://doi.org/10.1007/bf00699213

Bushnell, B. (2014). BBMap: A Fast, Accurate, Splice-Aware Aligner. *Lawrence Berkeley National Laboratory*. LBNL Report #: LBNL-7065E. Retrieved from https://escholarship.org/uc/item/1h3515gn

Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P., & Parks, D. H. (2022). GTDB-TK V2: Memory friendly classification with the genome taxonomy database. *Bioinformatics*, 38(23), 5315–5316. https://doi.org/10.1093/bioinformatics/btac672

De Tender, C., Mesuere, B., Van der Jeugt, F., Haegeman, A., Ruttink, T., Vandecasteele, B., Dawyndt, P., Debode, J., & Kuramae, E. E. (2019). Peat substrate amended with chitin modulates the N-cycle, siderophore and chitinase responses in the lettuce rhizobiome. *Scientific Reports*, 9(1), Article 1. https://doi.org/10.1038/s41598-019-46106-x

Diwan, A. D., Harke, S. N., & Panche, A. N. (2023). Host-microbiome interaction in fish and shellfish: An overview. *Fish and Shellfish Immunology Reports*, 4, 100091. https://doi.org/10.1016/j.fsirep.2023.100091

Drula, E., Garron, M.-L., Dogan, S., Lombard, V., Henrissat, B., & Terrapon, N. (2022). The carbohydrate-active enzyme database: Functions and literature. *Nucleic Acids Research*, 50(D1), D571–D577. https://doi.org/10.1093/nar/gkab1045

Edgar, R. C. (2004). Muscle: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5), 1792–1797. https://doi.org/10.1093/nar/gkh340

Egerton, S., Culloty, S., Whooley, J., Stanton, C., & Ross, R. P. (2018). The gut microbiota of marine fish. *Frontiers in Microbiology*, 9. https://doi.org/10.3389/fmicb.2018.00873

Eggink, K. M., Pedersen, P. B., Lund, I., & Dalsgaard, J. (2022). Chitin digestibility and intestinal exochitinase activity in Nile tilapia and rainbow trout fed different black soldier fly larvae meal size fractions. *Aquaculture Research*, 53(16), 5536–5546. https://doi.org/10.1111/are.16035

Estruch, G., Collado, M. C., Peñaranda, D. S., Tomás Vidal, A., Jover Cerdá, M., Pérez Martínez, G., & Martinez-Llorens, S. (2015). Impact of fishmeal replacement in diets for Gilthead Sea Bream (*Sparus aurata*) on the gastrointestinal microbiota determined by pyrosequencing the 16S rRNA gene. *PLOS ONE*, 10(8). https://doi.org/10.1371/journal.pone.0136389

German, D. P., Nagle, B. C., Villeda, J. M., Ruiz, A. M., Thomson, A. W., Contreras Balderas, S., & Evans, D. H. (2010). Evolution of herbivory in a carnivorous clade of minnows

(Teleostei: Cyprinidae): Effects on gut size and digestive physiology. *Physiological and Biochemical Zoology*, 83(1), 1–18. https://doi.org/10.1086/648510

Givens, C., Ransom, B., Bano, N., & Hollibaugh, J. (2015). Comparison of the gut microbiomes of 12 bony fish and 3 shark species. *Marine Ecology Progress Series*, 518, 209–223. https://doi.org/10.3354/meps11034

Harris, C. R. (2019) Interactions between predators, diets, and the gut microbiome of tropical reef fish [Unpublished master's thesis]. University of Delaware.

Hyatt, D., Chen, G.-L., LoCascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*, 11(1). https://doi.org/10.1186/1471-2105-11-119

Jones, J., DiBattista, J. D., Stat, M., Bunce, M., Boyce, M. C., Fairclough, D. V., Travers, M. J., & Huggett, M. J. (2018). The microbiome of the gastrointestinal tract of a range-shifting marine herbivorous fish. *Frontiers in Microbiology*, 9. https://doi.org/10.3389/fmicb.2018.02000

Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M., & Ishiguro-Watanabe, M. (2023). KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Research*, 51(D1), D587–D592. https://doi.org/10.1093/nar/gkac963

Kanehisa, M., Sato, Y., & Morishima, K. (2016). Blastkoala and Ghostkoala: KEGG tools for functional characterization of genome and metagenome sequences. *Journal of Molecular Biology*, 428(4), 726–731. https://doi.org/10.1016/j.jmb.2015.11.006

Kang, D. D., Froula, J., Egan, R., & Wang, Z. (2015). MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ*, 3. https://doi.org/10.7717/peerj.1165

Kang, D. D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., & Wang, Z. (2019). MetaBAT 2: An adaptive binning algorithm for robust and efficient genome reconstruction from Metagenome Assemblies. *PeerJ*, 7. https://doi.org/10.7717/peerj.7359

Kashinskaya, E. N., Simonov, E. P., Kabilov, M. R., Izvekova, G. I., Andree, K. B., & Solovyev, M. M. (2018). Diet and other environmental factors shape the bacterial communities of Fish Gut in an eutrophic lake. *Journal of Applied Microbiology*, 125(6), 1626–1641. https://doi.org/10.1111/jam.14064

Kassambara, A. & Mundt, F. (2020). factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R package version 1.0.7.999. http://www.sthda.com/english/rpkgs/factoextra

Keseler, I. M., Collado-Vides, J., Santos-Zavaleta, A., Peralta-Gil, M., Gama-Castro, S., Muñiz-Rascado, L., Bonavides-Martinez, C., Paley, S., Krummenacker, M., Altman, T., Kaipa, P., Spaulding, A., Pacheco, J., Latendresse, M., Fulcher, C., Sarker, M., Shearer, A. G., Mackie, A., Paulsen, I., … Karp, P. D. (2011). EcoCyc: A comprehensive database of Escherichia coli biology. *Nucleic Acids Research*, 39(Database issue), D583–D590. https://doi.org/10.1093/nar/gkq1143

Knappe, J., & Sawers, G. (1990). A radical-chemical route to acetyl-CoA: The anaerobically induced pyruvate formate-lyase system of *Escherichia coli*. *FEMS Microbiology Reviews*, 6(4), 383–398. https://doi.org/10.1111/j.1574-6968.1990.tb04108.x

Kumla, J., Suwannarach, N., Sujarit, K., Penkhrue, W., Kakumyan, P., Jatuwong, K., Vadthanarat, S., & Lumyong, S. (2020). Cultivation of Mushrooms and Their Lignocellulolytic Enzyme Production Through the Utilization of Agro-Industrial Waste. *Molecules*, 25(12), Article 12. https://doi.org/10.3390/molecules25122811

Le, D., Nguyen, P., Nguyen, D., Dierckens, K., Boon, N., Lacoere, T., Kerckhof, F.-M., De Vrieze, J., Vadstein, O., & Bossier, P. (2019). Gut microbiota of migrating wild rabbit fish (*Siganus guttatus*) larvae have low spatial and temporal variability. *Microbial Ecology*, 79(3), 539–551. https://doi.org/10.1007/s00248-019-01436-1

León-Zayas, R., McCargar, M., Drew, J. A., & Biddle, J. F. (2020). Microbiomes of fish, sediment and seagrass suggest connectivity of coral reef microbial populations. *PeerJ*, 8. https://doi.org/10.7717/peerj.10026

Li, D., Luo, R., Liu, C.-M., Leung, C.-M., Ting, H.-F., Sadakane, K., Yamashita, H., & Lam, T.-W. (2016). Megahit v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods*, 102, 3–11. https://doi.org/10.1016/j.ymeth.2016.02.020

Li, H., Wu, S., Wirth, S., Hao, Y., Wang, W., Zou, H., Li, W., & Wang, G. (2014). Diversity and activity of cellulolytic bacteria, isolated from the gut contents of grass carp (*Ctenopharyngodon idellus*) (Valenciennes) fed on Sudan grass (*Sorghum sudanense*) or artificial feedstuffs. *Aquaculture Research*, 47(1), 153–164. https://doi.org/10.1111/are.12478

Li, X., Yu, Y., Feng, W., Yan, Q., & Gong, Y. (2012). Host species as a strong determinant of the intestinal microbiota of fish larvae. *The Journal of Microbiology*, 50(1), 29–37. https://doi.org/10.1007/s12275-012-1340-1

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. *Scientific Reports*, 6(1). https://doi.org/10.1038/srep24340

Liu, Y., Li, X., Li, Y., Li, J., & Zhu, S. (2022). Gut microbiomes of cyprinid fish exhibit host-species symbiosis along gut trait and diet. *Frontiers in Microbiology*, 13. https://doi.org/10.3389/fmicb.2022.936601

Llewellyn, M. S., Boutin, S., Hoseinifar, S. H., & Derome, N. (2014). Teleost microbiomes: The State of the art in their characterization, manipulation and importance in aquaculture and fisheries. *Frontiers in Microbiology,* 5. https://doi.org/10.3389/fmicb.2014.00207

McCauley, M., German, D. P., Lujan, N. K., & Jackson, C. R. (2020). Gut microbiomes of sympatric Amazonian wood-eating catfishes (Loricariidae) reflect host identity and little role in wood digestion. *Ecology and Evolution*, 10(14), 7117–7128. https://doi.org/10.1002/ece3.6413

Miyake, S., Ngugi, D. K., & Stingl, U. (2014). Diet strongly influences the gut microbiota of surgeonfishes. *Molecular Ecology*, 24(3), 656–672. https://doi.org/10.1111/mec.13050

Moran, D., & Clements, K. D. (2002). Diet and endogenous carbohydrases in the temperate marine herbivorous fish *Kyphosus sydneyanus*. *Journal of Fish Biology,* 60(5), 1190–1203. https://doi.org/10.1111/j.1095-8649.2002.tb01714.x

Mountfort, D. O., Campbell, J., & Clements, K. D. (2002). Hindgut fermentation in three species of marine herbivorous fish. *Applied and Environmental Microbiology*, 68(3), 1374–1380. https://doi.org/10.1128/aem.68.3.1374-1380.2002

Nayak, S.K. (2010). Role of gastrointestinal microbiota in fish. *Aquaculture Research*, 41(11), pp.1553-1573. https://doi.org/10.1111/j.1365-2109.2010.02546.x

Nurk, S., Meleshko, D., Korobeynikov, A., & Pevzner, P. A. (2017). Metaspades: A new versatile metagenomic assembler. *Genome Research*, 27(5), 824–834. https://doi.org/10.1101/gr.213959.116

Oksanen, J., Simpson, G., Blanchet, F., Kindt, R., Legendre, P., Minchin, P., O'Hara, R., Solymos, P., Stevens, M., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., De Caceres, M., Durand, S., Evangelista, H., FitzJohn, R., Friendly, M., Furneaux, B., Hannigan, G., Hill, M., Lahti, L., McGlinn, D., Ouellette, M., Ribeiro Cunha, E., Smith, T., Stier, A., Ter Braak, C., & Weedon, J. (2022). vegan: Community Ecology Package. R package version 2.6-4. https://CRAN.R-project.org/package=vegan

Pan, B., Han, X., Yu, K., Sun, H., Mu, R., Lian C-A. (2022) Geographical distance, host evolutionary history and diet drive gut microbiome diversity of fish across the Yellow River. *Molecular Ecology* 32(5): 1183-1196. https://onlinelibrary.wiley.com/doi/full/10.1111/mec.16812

Parata, L., Nielsen, S., Xing, X., Thomas, T., Egan, S., & Vergés, A. (2019). Age, gut location and diet impact the gut microbiome of a tropical herbivorous surgeonfish. *FEMS Microbiology Ecology*. https://doi.org/10.1093/femsec/fiz179

Pardesi, B., Roberton, A. M., Lee, K. C., Angert, E. R., Rosendale, D. I., Boycheva, S., White, W. L., & Clements, K. D. (2022). Distinct microbiota composition and fermentation products indicate functional compartmentalization in the hindgut of a marine herbivorous fish. *Molecular Ecology*, 31(8), 2494–2509. https://doi.org/10.1111/mec.16394

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., & Tyson, G. W. (2015). Checkm: Assessing the quality of microbial genomes recovered from isolates, single cells, and Metagenomes. *Genome Research*, 25(7), 1043–1055. https://doi.org/10.1101/gr.186072.114

Peng, Y., Leung, H. C., Yiu, S. M., & Chin, F. Y. (2012). IDBA-Ud: A de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics*, 28(11), 1420–1428. https://doi.org/10.1093/bioinformatics/bts174

Pisaniello, A., Handley, K. M., White, W. L., Angert, E. R., Boey, J. S., & Clements, K. D. (2023). Host individual and gut location are more important in gut microbiota community composition than temporal variation in the marine herbivorous fish *Kyphosus sydneyanus*. *BMC Microbiology*, 23(1). https://doi.org/10.1186/s12866-023-03025-2

Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS ONE*, 5(3). https://doi.org/10.1371/journal.pone.0009490

R Core Team. (2023). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Rawlings, N. D., Waller, M., Barrett, A. J., & Bateman, A. (2014). MEROPS: The database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Research*, 42(D1), D503–D509. https://doi.org/10.1093/nar/gkt953

Ray, A. K., Ghosh, K., & Ringø, E. (2012). Enzyme-producing bacteria isolated from fish gut: A Review. *Aquaculture Nutrition*, 18(5), 465–492. https://doi.org/10.1111/j.1365-2095.2012.00943.x

Reid, H. I., Treasurer, J. W., Adam, B., & Birkbeck, T. H. (2009). Analysis of bacterial populations in the gut of developing cod larvae and identification of *Vibrio logei*, *Vibrio anguillarum* and *Vibrio splendidus* as pathogens of cod larvae. *Aquaculture*, 288(1–2), 36–43. https://doi.org/10.1016/j.aquaculture.2008.11.022

Rimoldi, S., Ceccotti, C., Brambilla, F., Faccenda, F., Antonini, M., & Terova, G. (2023). Potential of shrimp waste meal and insect exuviae as sustainable sources of chitin for fish feeds. *Aquaculture*, 567, 739256. https://doi.org/10.1016/j.aquaculture.2023.739256

Roeselers, G., Mittge, E. K., Stephens, W. Z., Parichy, D. M., Cavanaugh, C. M., Guillemin, K., & Rawls, J. F. (2011). Evidence for a core gut microbiota in the zebrafish. *The ISME Journal*, 5(10), 1595–1608. https://doi.org/10.1038/ismej.2011.38

Sanders, J.G., Beichman, A.C., Roman, J., Scott, J.J., Emerson, D., McCarthy, J.J. and Girguis, P.R. (2015). Baleen whales host a unique gut microbiome with similarities to both carnivores and herbivores. *Nature Communications*, 6, p.8285. https://doi.org/10.1038/ncomms9285

Shade, A., & Handelsman, J. (2011). Beyond the Venn diagram: The Hunt for a core microbiome. *Environmental Microbiology*, 14(1), 4–12. https://doi.org/10.1111/j.1462-2920.2011.02585.x

Søndergaard, D., Pedersen, C. N. S., & Greening, C. (2016). HydDB: A web tool for hydrogenase classification and analysis. *Scientific Reports*, 6(1), Article 1. https://doi.org/10.1038/srep34212

Stevenson, S. J. R., Lee, K. C., Handley, K. M., Angert, E. R., White, W. L., & Clements, K. D. (2022). Substrate degradation pathways, conserved functions and community composition of the hindgut microbiota in the herbivorous marine fish *Kyphosus sydneyanus*. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 272, 111283. https://doi.org/10.1016/j.cbpa.2022.111283

Sullam, K.E., Essinger, S.D., Lozupone, C.A., O'Connor, M.P., Rosen, G.L., Knight, R.O.B., Kilham, S.S. and Russell, J.A. (2012). Environmental and ecological factors that shape

the gut bacterial communities of fish: a meta-analysis. *Molecular Ecology*, 21(13), pp.3363-3378. https://doi.org/10.1111/j.1365-294X.2012.05552.x

Tarnecki, A. M., Burgos, F. A., Ray, C. L., & Arias, C. R. (2017). Fish intestinal microbiome: Diversity and symbiosis unravelled by metagenomics. *Journal of Applied Microbiology*, 123(1), 2–17. https://doi.org/10.1111/jam.13415

Victorian-Bioinformatics-Consortium. *Nesoni*, https://github.com/Victorian-Bioinformatics-Consortium/nesoni (2014).

Wong, S., & Rawls, J. F. (2012). Intestinal microbiota composition in fishes is influenced by host ecology and environment. *Molecular Ecology*, 21(13), 3100–3102. https://doi.org/10.1111/j.1365-294x.2012.05646.x

Wu, Y.-W., Simmons, B. A., & Singer, S. W. (2015). MaxBin 2.0: An automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics*, 32(4), 605–607. https://doi.org/10.1093/bioinformatics/btv638

Yukgehnaish, K., Kumar, P., Sivachandran, P., Marimuthu, K., Arshad, A., Paray, B. A., & Arockiaraj, J. (2020). Gut microbiota metagenomics in aquaculture: Factors influencing gut microbiome and its physiological role in fish. *Reviews in Aquaculture*, 12(3), 1903–1927. https://doi.org/10.1111/raq.12416

Zhang, X., Wu, H., Li, Z., Li, Y., Wang, S., Zhu, D., Wen, X., & Li, S. (2017). Effects of dietary supplementation of *Ulva pertusa* and non-starch polysaccharide enzymes on gut microbiota of *Siganus canaliculatus*. *Journal of Oceanology and Limnology*, 36(2), 438–449. https://doi.org/10.1007/s00343-017-6235-x

Zhang, Z., Li, D., Refaey, M. M., & Xu, W. (2017). High spatial and temporal variations of microbial community along the southern catfish gastrointestinal tract: Insights into

Dynamic Food Digestion. *Frontiers in Microbiology*, 8. https://doi.org/10.3389/fmicb.2017.01531

Acknowledgements:

We would like to thank our colleagues at UD in Lewes for assistance with fish care and sampling. This work was supported by The Mindlin Foundation, funding to CR via the UD Graduate Scholars Fellowship, funding to KK and JFB from the WM Keck Foundation and funding to IF and JFB by ExxonMobil Research and Engineering Company. DW was supported as a UD School of Marine Science & Policy scholar. Support from the University of Delaware CBCB Bioinformatics Data Science Core Facility (RRID:SCR_017696), including use of the BIOMIX and BioStore computational resources was made possible through funding from Delaware INBRE (NIGMS P20GM103446), NIH Shared Instrumentation Grant (S10OD028725) the State of Delaware, and the Delaware Biotechnology Institute.

Figures

Figure 1: Principal component analysis of normalized enzyme gene copy numbers across all MAGs recovered from the three diet treatments. The total numbers of enzymes of each degradation pathway (proteins, cellulose, hemicellulose, starch, and chitin) normalized to the MAG's genome completeness were used to generate components. Fatty acids were included as a binary where 1 designated that the MAG could complete all four steps of beta oxidation, while 0 designated that it could not. Ellipses represent t distributions around the centroids. Ellipses are labeled for clarity. Individual points represent individual MAGs. The larger shapes in the ellipses are group centroids.

Figure 2: Polysaccharide degradation pathway to monosaccharide products. Shown are a) cellulose, b) xylan, c) mannan, d) xyloglucan, and e) chitin. Enzymes and their CAZyme IDs are noted above arrows connecting substrates and products. MAGs that possess genes for each enzyme are denoted beneath the arrow, where dark brown (left), green (center), and pink (right) font represent the piscivorous, herbivorous, and invertivorous treatments, respectively.

Figure 3: Number of substrate degradation genes per MAG normalized to genome completeness. Shown are a) cellulose, b) hemicellulose, c) chitin, d) starch, and e) proteins in each diet treatment. Kruskal Wallis values above bars represent group comparisons. Pairwise comparison values denoted by brackets are the result of post-hoc Wilcoxon tests between diets.

Figure 4: Heatmaps of metabolic functions that can be performed by recovered MAGs. Samples shown are a) piscivorous hawkfish, b) herbivorous yellow tang, and c) invertivorous triggerfish. Red indicates the ability to perform a metabolic pathway, while blue indicates inability.

Figure 5: Digital illustration of simplified piscivorous hawkfish gut microbiome community interactions. Diet inputs enter the gut from the left and leave to the right following bacteria-facilitated degradation. The top 8 MAGs with the highest genome completeness and *Phaeomarinobacter* sp. recovered from this diet treatment are shown. Not all metabolic functions are depicted. Substrates such as fatty acids, proteins, and polysaccharides overlayed behind or in the vicinity of MAGs indicates that the MAG is able to degrade them. Arrows connecting fermentation or respiration products and substrates intersect with MAGs that can carry out that specific conversion. Figure was created on BioRender (https://BioRender.com/).

Figure 6: Digital illustration of simplified herbivorous yellow tang gut microbiome community interactions. Diet inputs enter the gut from the left and leave to the right following

821 bacteria-facilitated degradation. The top 8 MAGs with the highest genome completeness
822 recovered from this diet treatment are shown. Not all metabolic functions are depicted.
823 Substrates and arrows follow as in Figure 5.
824 **Figure 7: Digital illustration of simplified invertivorous triggerfish gut microbiome**
825 **community interactions.** Diet inputs enter the gut from the left and leave to the right following
826 bacteria-facilitated degradation. All 7 MAGs recovered from this diet treatment are shown. Not
827 all metabolic functions are depicted. Substrates and arrows follow as in Figure 5.

Table 1: MAG taxonomic assignments, completeness and contamination

MAG ID[1]	Phylum	Class	Order	Family	Genus	Species	Genome Completeness (%)	Contamination (%)
H1N1.001	Proteobacteria	Gammaproteobacteria	Enterobacterales	Vibrionaceae	*Photobacterium*	*aphoticum*	97.58	2.55
H1N1.003	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Alcanivoracaceae	*Alcanivorax*	sp.	100	0.74
H1N1.005	Proteobacteria	Alphaproteobacteria	Parvibaculales	Phaeomarinobacteraceae	*Phaeomarinobacter*	sp.	86.52	5.77
H1N1.006	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Phaeobacter*	*italicus*	99.66	6.79
H1N1.007	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Thalassospiraceae	*Thalassospira*	sp.	100	1
H1N1.009	Bacteroidota	Bacteroidia	Cytophagales	Cyclobacteriaceae	*Fabibacter*	*pacificus*	98.73	2.89
H1N1.010	Proteobacteria	Gammaproteobacteria	Enterobacterales	Vibrionaceae	*Vibrio*	*fortis*	74.03	9.2
H1N1.011	Patescibacteria	Gracilibacteria	UBA1369	GCA-2718715	-	-	73.06	0.93
H1N1.012	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Ruegeria*	sp.	98.27	3.69
H1N1.015	Bacteroidota	Bacteroidia	Cytophagales	Cyclobacteriaceae	*Fabibacter*	*spongicola*	98.01	0.6
H1N1.018	Proteobacteria	Alphaproteobacteria	Minwuiales	Minwuiaceae	*Minwuia*	sp.	98.91	3.77
H1N1.021	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Porticoccaceae	*Porticoccus*	sp.	89.33	3.83
H1N1.022	Proteobacteria	Gammaproteobacteria	Enterobacterales	Alteromonadaceae	*Alteromonas*	sp.	92.49	9.21
H1N1.025	Proteobacteria	Gammaproteobacteria	Enterobacterales	Alteromonadaceae	*Pseudoalteromonas*	*spongiae*	70.69	6.03
H1N1.026	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	*Erythrobacter*	*flavus*	94.33	0.86
H1N1.028	Proteobacteria	Alphaproteobacteria	Caulobacterales	Maricaulaceae	*Maricaulis*	sp.	93.44	1.68
H1N1.030	Bacteroidota	Bacteroidia	Cytophagales	Flammeovirgaceae	*Sediminitomix*	sp.	93.49	8.17
H1N1.033	Proteobacteria	Alphaproteobacteria	UBA8366	GCA-2696645	GRI0909	sp.	87.3	12.75
T4LTN.001	Bacteroidota	Bacteroidia	Bacteroidales	Bacteroidaceae	*Bacteroides*	sp.	98.51	0.5
T4LTN.002	Bacteroidota	Bacteroidia	Bacteroidales	Paludibacteraceae	-	-	85.3	5.67
T4LTN.003	Proteobacteria	Gammaproteobacteria	Enterobacterales	Vibrionaceae	*Vibrio*	*owensii*	92.93	1.51
T4LTN.004	Proteobacteria	Gammaproteobacteria	Enterobacterales	Shewanellaceae	*Shewanella*	sp.	94.4	3.15
T4LTN.005	Bacteroidota	Bacteroidia	Bacteroidales	Tannerellaceae	*Parabacteroides*	sp.	96.92	7.31
T4LTN.006	Firmicutes	Clostridia	Lachnospirales	Lachnospiraceae	*Lachnoclostridium*	sp.	99.11	4.83
T4LTN.007	Proteobacteria	Gammaproteobacteria	Enterobacterales	Alteromonadaceae	*Psychrosphaera*	sp.	91.46	10.06
T4LTN.009	Proteobacteria	Gammaproteobacteria	Enterobacterales	Alteromonadaceae	*Psychrosphaera*	sp.	92.4	3.7
T4LTN.010	Spirochaetota	Spirochaetia	Sphaerochaetales	Sphaerochaetaceae	*Spirochaeta*	sp.	100	2.3
T4LTN.011	Bacteroidota	Bacteroidia	Bacteroidales	UBA1402	UBA8389	sp.	100	2.02

T4LTN.012	Proteobacteria	Gammaproteobacteria	Enterobacterales	Vibrionaceae	*Enterovibrio*	sp.	97.84	0.9
T4LTN.014.2	Firmicutes	Bacilli	RF39	CAG-822	UBA5364	sp.	96.63	0.61
T4LTN.014.3	Patescibacteria	Gracilibacteria	BD1-5	UBA6164	MAAV01	sp.	83.15	1.69
T4LTN.015	Firmicutes	Clostridia	Oscillospirales	Ruminococcaceae	UMGS363	sp.	97.32	6.64
T4LTN.016	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Saccharospirillaceae	*Oleispira*	sp.	87.84	3.6
T4LTN.017	Firmicutes	Clostridia	Oscillospirales	Ruminococcaceae	-	-	99.32	0.91
T4LTN.022	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	*Erythrobacter*	*flavus*	95.9	7.03
T4LTN.023	Bacteroidota	Bacteroidia	Bacteroidales	P3	-	-	86.31	2.89
Tr3LTN.002	Proteobacteria	Gammaproteobacteria	Enterobacterales	Shewanellaceae	*Shewanella*	sp.	98.87	1.08
Tr3LTN.005	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Epibacterium*	sp.	98.94	3.55
Tr3LTN.006	Proteobacteria	Gammaproteobacteria	Psuedomonadales	Cellvibrionaceae	GCA-2707785	sp.	95.52	1.92
Tr3LTN.008	Campylobacterota	Campylobacteria	Campylobacterales	Arcobacteraceae	UBA4036	sp.	99.19	2.37
Tr3LTN.016	Firmicutes	Bacilli	-	-	-	-	84.8	7.11
Tr3LTN.017	Proteobacteria	Alphaproteobacteria	UBA1280	UBA12223	UBA12223	-	81.93	2.99
Tr3LTN.026	Bdellovibrionota	Bacteriovoracia	Bacteriovoracales	Bacteriovoracaceae	-	-	70.65	4.1

[1] Naming: H1N1, Hawfish; T4LTN, Yellow tang; Tr3LTN, Triggerfish

Table 2: Summary of biochemical functions that could be performed by all recovered MAGs.

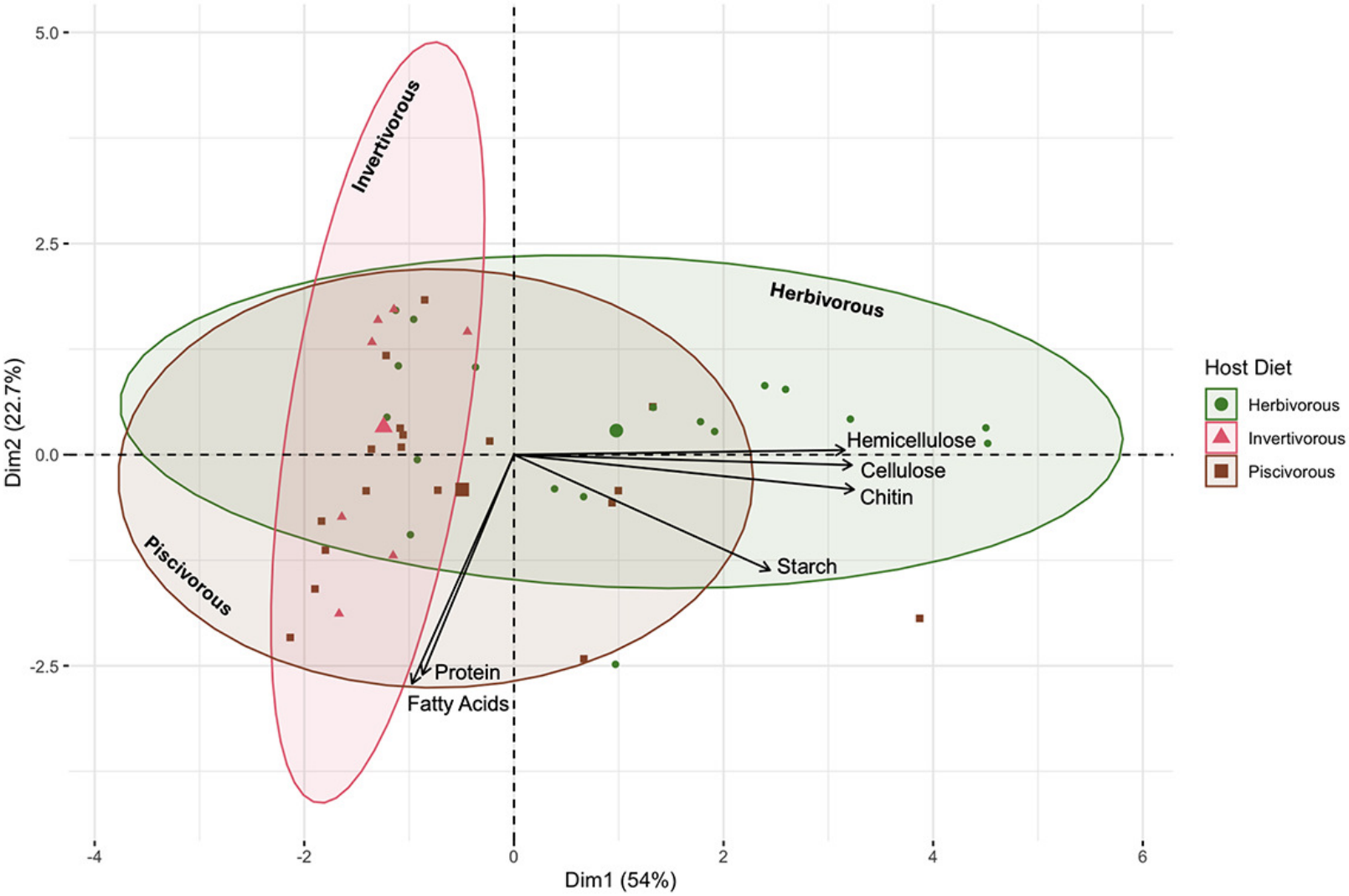
MAG ID[1]	Family	Host Diet	Substrate Utilization[2]	Energy Production[3]	Vitamin Production	Benefits to Host
H1N1.001	Vibrionaceae	Piscivorous	Proteins, Starch	Glycolysis, Pyruvate Oxidation, TCA Cycle, Nitrate reduction, Formate fermentation, Acetogenesis	B1, B2, B7, B9, B12	Protein Degradation, Vitamin Production
H1N1.003	Alcanivoracaceae	Piscivorous	Proteins, Fatty Acids	Glycolysis, Pyruvate Oxidation, TCA Cycle, Beta Oxidation, Acetogenesis (p)	B2, B7, B9, B12	Protein Degradation, Vitamin Production
H1N1.005	Phaeomarinobacteraceae	Piscivorous	Collagen, Proteins, Cellulose	Glycolysis, TCA Cycle (p), Pyruvate Oxidation	B7	Vitamin Production
H1N1.006	Rhodobacteraceae	Piscivorous	Proteins, Fatty Acids	Glycolysis, Pyruvate Oxidation, TCA Cycle (p), Beta Oxidation, Acetogenesis	B7, B9, B12	Protein Degradation, Vitamin Production
H1N1.007	Thalassospiraceae	Piscivorous	Proteins, Fatty Acids	Glycolysis, Pyruvate Oxidation, TCA Cycle, Beta Oxidation, Lactate fermentation, Acetogenesis (p)	B1, B9, B12	Protein Degradation, Vitamin Production
H1N1.009	Cyclobacteriaceae	Piscivorous	Proteins, Cellulose, Starch, Chitin	Glycolysis, Pyruvate Oxidation, TCA Cycle (p), Acetogenesis (p)	B7, B9	Protein Degradation, Vitamin Production
H1N1.010	Vibrionaceae	Piscivorous	Proteins, Starch	Pyruvate Oxidation, TCA Cycle (p), Nitrate reduction, Acetogenesis, Formate fermentation	B1, B7, B9, B12	Protein Degradation, Vitamin Production
H1N1.011	GCA-2718715	Piscivorous	Cellulose	Glycolysis (p), Lactate fermentation, Acetogenesis (p)	-	-
H1N1.012	Rhodobacteraceae	Piscivorous	Proteins, Fatty Acids	Glycolysis, Pyruvate Oxidation, TCA Cycle (p), Beta Oxidation, Nitrate reduction, Acetogenesis	B7, B9, B12	Protein Degradation, Vitamin Production
H1N1.015	Cyclobacteriaceae	Piscivorous	Fatty Acids	Glycolysis, Pyruvate Oxidation, TCA Cycle (p), Beta Oxidation, Acetogenesis (p)	B7, B9	Vitamin Production
H1N1.018	Minwuiaceae	Piscivorous	Proteins, Fatty Acids	Glycolysis, Pyruvate Oxidation, TCA Cycle, Beta Oxidation, Acetogenesis(p)	B12	Protein Degradation, Vitamin Production
H1N1.021	Porticoccaceae	Piscivorous	Proteins	Glycolysis (p), Pyruvate Oxidation, Acetogenesis (p)	B9	Protein Degradation, Vitamin Production
H1N1.022	Alteromonadaceae	Piscivorous	Proteins, Fatty Acids, Cellulose, Hemicellulose, Starch	Glycolysis, Pyruvate Oxidation, TCA Cycle, Beta Oxidation, Acetogenesis (p)	B1, B2, B7, B9	Protein Degradation, Vitamin Production
H1N1.025	Alteromonadaceae	Piscivorous	Proteins, Starch	Glycolysis, Pyruvate Oxidation, Acetogenesis	B2, B9	Protein Degradation, Vitamin Production
H1N1.026	Sphingomonadaceae	Piscivorous	Proteins	Glycolysis, Pyruvate Oxidation, TCA Cycle (p), Acetogenesis (p)	B1, B7	Protein Degradation, Vitamin Production
H1N1.028	Maricaulaceae	Piscivorous	Proteins, Cellulose	Glycolysis (p), Pyruvate Oxidation, Acetogenesis (p)	B1, B7, B9	Protein Degradation, Vitamin Production
H1N1.030	Flammeovirgaceae	Piscivorous	Proteins, Fatty Acids, Cellulose, Hemicellulose, Starch, Chitin	Glycolysis, Pyruvate Oxidation, TCA Cycle (p), Beta Oxidation, Acetogenesis	B1, B7, B9	Protein Degradation, Vitamin Production
H1N1.033	GCA-2696645	Piscivorous	Proteins, Cellulose	Glycolysis (p), Pyruvate Oxidation, TCA Cycle (p), Acetogenesis (p)	-	Protein Degradation
T4LTN.001	Bacteroidaceae	Herbivorous	Cellulose, Hemicellulose, Starch, Proteins, Chitin	Glycolysis, Pyruvate Oxidation, TCA Cycle (p), H2-Producing Fermentation, Acetogenesis (p)	B1, B2, B7, B9	Plant Matter Degradation, Vitamin Production
T4LTN.002	Paludibacteraceae	Herbivorous	Cellulose, Proteins	Glycolysis, Pyruvate Oxidation, Acetogenesis (p)	-	Plant Matter Degradation
T4LTN.003	Vibrionaceae	Herbivorous	Starch, Proteins, Fatty Acids, Chitin	Glycolysis, Pyruvate Oxidation, TCA Cycle (p), Nitrate reduction, Acetogenesis, Formate fermentation, Beta Oxidation	B1, B2, B7, B9	Plant Matter Degradation, Vitamin Production
T4LTN.004	Shewanellaceae	Herbivorous	Starch, Proteins, Fatty Acids	Glycolysis, Pyruvate Oxidation, TCA Cycle, Nitrate reduction, Acetogenesis, Formate fermentation, Beta Oxidation	B1, B2, B7, B9, B12	Plant Matter Degradation, Vitamin Production
T4LTN.005	Tannerellaceae	Herbivorous	Cellulose, Hemicellulose, Starch, Proteins, Chitin	Glycolysis, Pyruvate Oxidation, TCA Cycle (p), H_2-Producing Fermentation, Acetogenesis (p)	B1, B2, B7, B9, B12	Plant Matter Degradation, Vitamin Production
T4LTN.006	Lachnospiraceae	Herbivorous	Cellulose, Hemicellulose, Starch, Proteins, Chitin	Glycolysis, Pyruvate Oxidation, Lactate fermentation, H_2-Producing Fermentation, Acetogenesis (p)	B7, B12	Plant Matter Degradation, Vitamin Production
T4LTN.007	Alteromonadaceae	Herbivorous	Cellulose, Hemicellulose, Starch	Glycolysis, Pyruvate Oxidation, TCA Cycle (p), Nitrate reduction, Acetogenesis	B1, B2, B7, B9	Plant Matter Degradation, Vitamin Production
T4LTN.009	Alteromonadaceae	Herbivorous	Cellulose, Starch	Glycolysis, Pyruvate Oxidation, TCA Cycle (p), Acetogenesis	B1, B7	Plant Matter Degradation, Vitamin Production
T4LTN.010	Sphaerochaetaceae	Herbivorous	Proteins	Glycolysis, Pyruvate Oxidation, H_2-Producing Fermentation, Acetogenesis (p)	-	-
T4LTN.011	UBA1402	Herbivorous	Hemicellulose	Glycolysis, Pyruvate Oxidation, TCA Cycle (p), Acetogenesis (p)	B9	Plant Matter Degradation, Vitamin Production
T4LTN.012	Vibrionaceae	Herbivorous	Starch, Proteins, Fatty Acids	Glycolysis, Pyruvate Oxidation, TCA Cycle, Nitrate reduction, Acetogenesis, Formate fermentation, Beta Oxidation	B1, B2, B7, B9, B12	Plant Matter Degradation, Vitamin Production
T4LTN.014.2	CAG-822	Herbivorous	Starch	Glycolysis (p), Pyruvate Oxidation, Acetogenesis (p)	-	Plant Matter Degradation
T4LTN.014.3	UBA6164	Herbivorous	Simple Saccharides	-	-	-
T4LTN.015	Ruminococcaceae	Herbivorous	Cellulose, Hemicellulose, Chitin	Glycolysis, Pyruvate Oxidation, H_2-Producing Fermentation, Acetogenesis (p)	-	Plant Matter Degradation

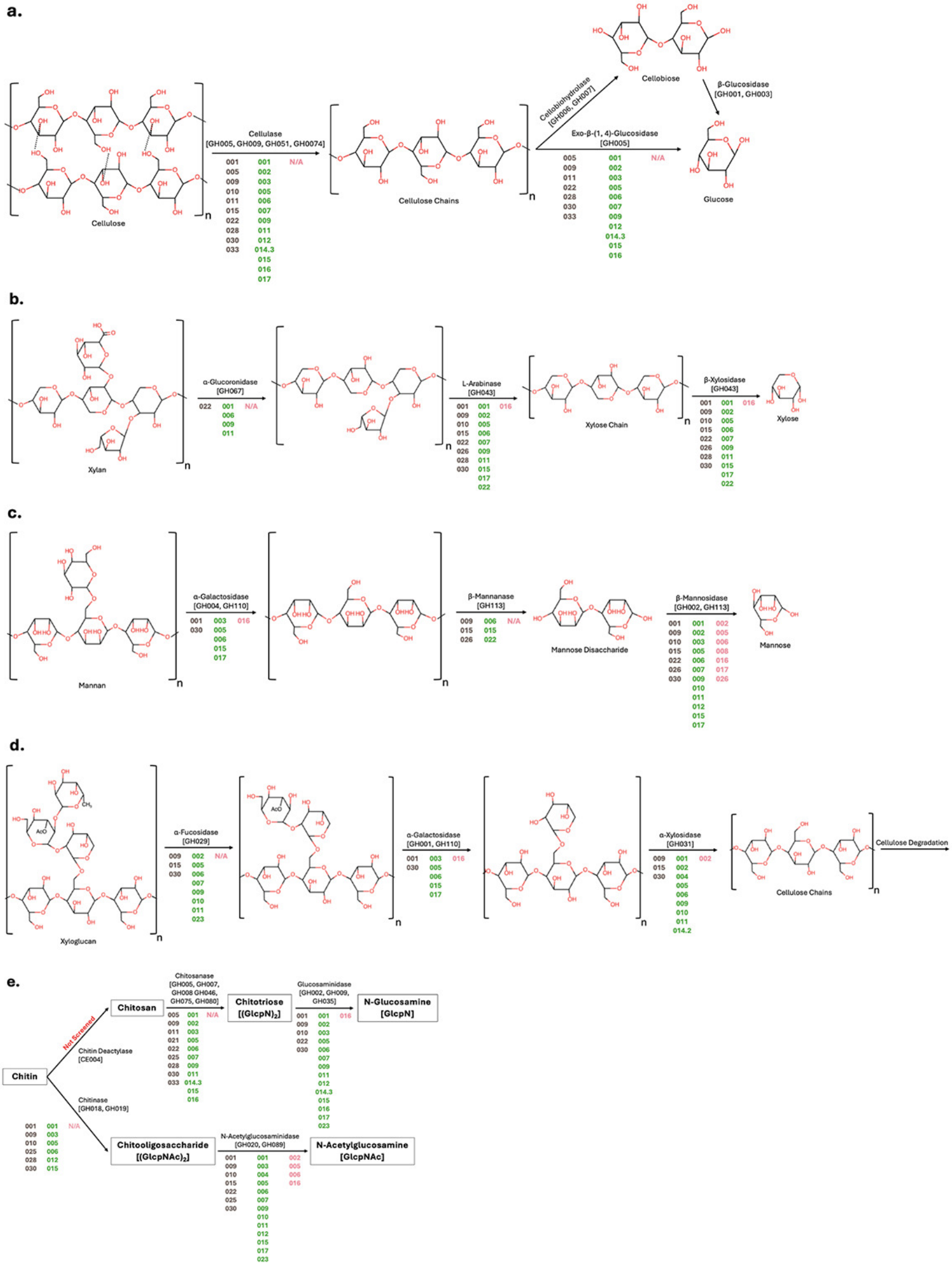
T4LTN.016	Saccharospirillaceae	Herbivorous	Cellulose, Proteins, Fatty Acids	Pyruvate Oxidation, Nitrate reduction, Acetogenesis, Beta Oxidation	B1, B7, B9	Plant Matter Degradation, Vitamin Production
T4LTN.017	Ruminococcaceae	Herbivorous	Simple Saccharides	Glycolysis, Pyruvate Oxidation, H_2-Producing Fermentation, Acetogenesis (p)	B7	Vitamin Production
T4LTN.022	Sphingomonadaceae	Herbivorous	Proteins	Glycolysis, Pyruvate Oxidation, TCA Cycle (p), Acetogenesis (p)	B1	Vitamin Production
T4LTN.023	P3	Herbivorous	Proteins	Glycolysis, Pyruvate Oxidation, H_2-Producing Fermentation, Acetogenesis (p)	B1	Vitamin Production
Tr3LTN.002	Shewanellaceae	Invertivorous	Proteins, Fatty Acids	Glycolysis, Pyruvate Oxidation, TCA Cycle, Nitrate reduction, Acetogenesis, Formate fermentation, Beta Oxidation	B1, B2, B7, B9, B12	Protein Degradation, Vitamin Production
Tr3LTN.005	Rhodobacteraceae	Invertivorous	Proteins, Fatty Acids	Glycolysis, Pyruvate Oxidation, TCA Cycle (p), Beta Oxidation	B2, B9, B12	Protein Degradation, Vitamin Production
Tr3LTN.006	Cellvibrionaceae	Invertivorous	Proteins, Fatty Acids	Glycolysis, Pyruvate Oxidation, Beta Oxidation	B1, B7, B9	Protein Degradation, Vitamin Production
Tr3LTN.008	Arcobacteraceae	Invertivorous	Simple Saccharides	Glycolysis, Pyruvate Oxidation, TCA Cycle (p)	B1, B7	Vitamin Production
Tr3LTN.016	Class: Bacilli	Invertivorous	Simple Saccharides	Glycolysis, Pyruvate Oxidation, H_2-Producing Fermentation	-	-
Tr3LTN.017	UBA12223	Invertivorous	Simple Saccharides	Glycolysis (p), Pyruvate Oxidation, TCA Cycle (p)	-	-
Tr3LTN.026	Bacteriovoracaceae	Invertivorous	Simple Saccharides	Glycolysis (p), Pyruvate Oxidation	-	-

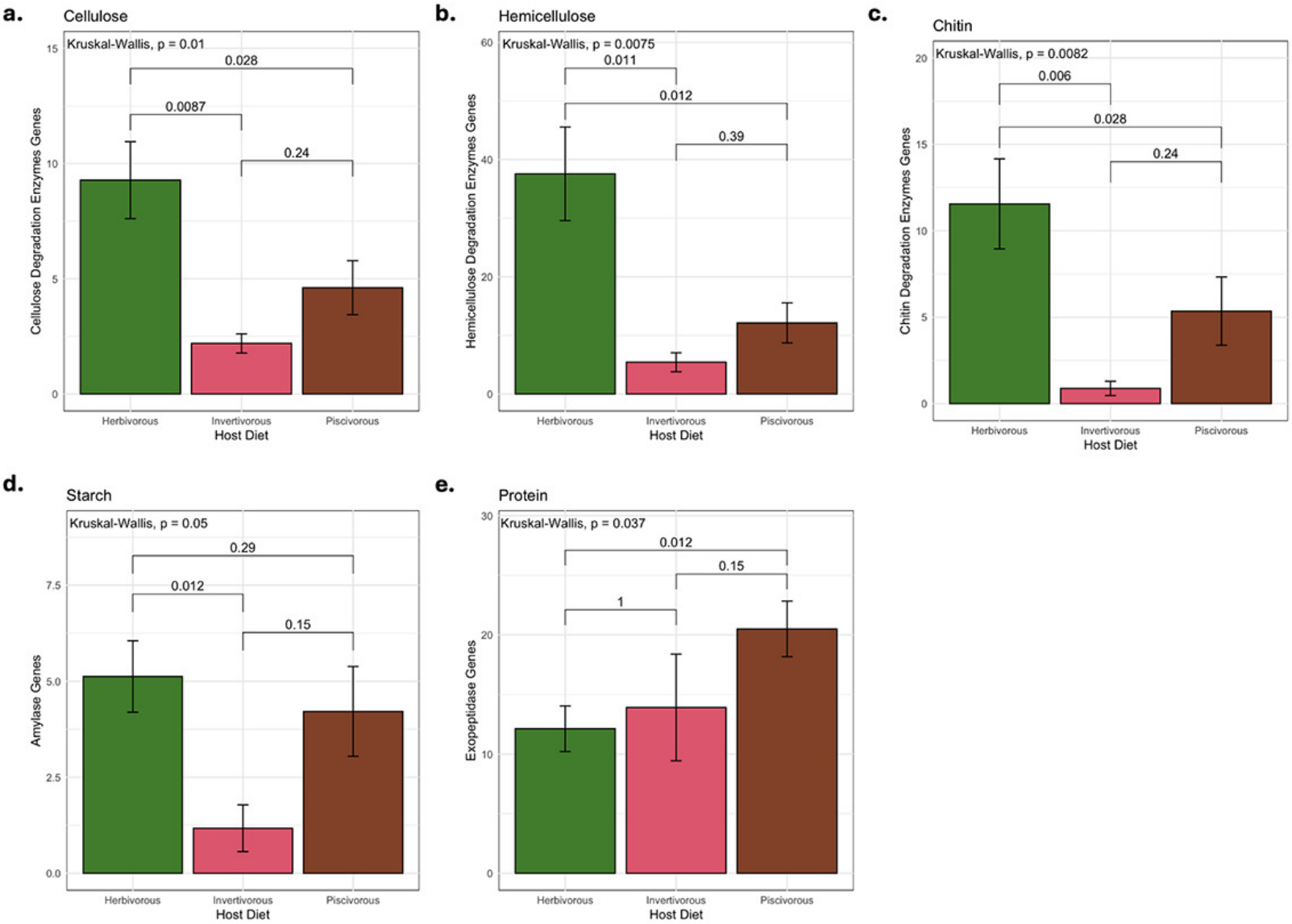
[1] Naming: H1N1, Hawfish; T4LTN, Yellow tang; Tr3LTN, Triggerfish

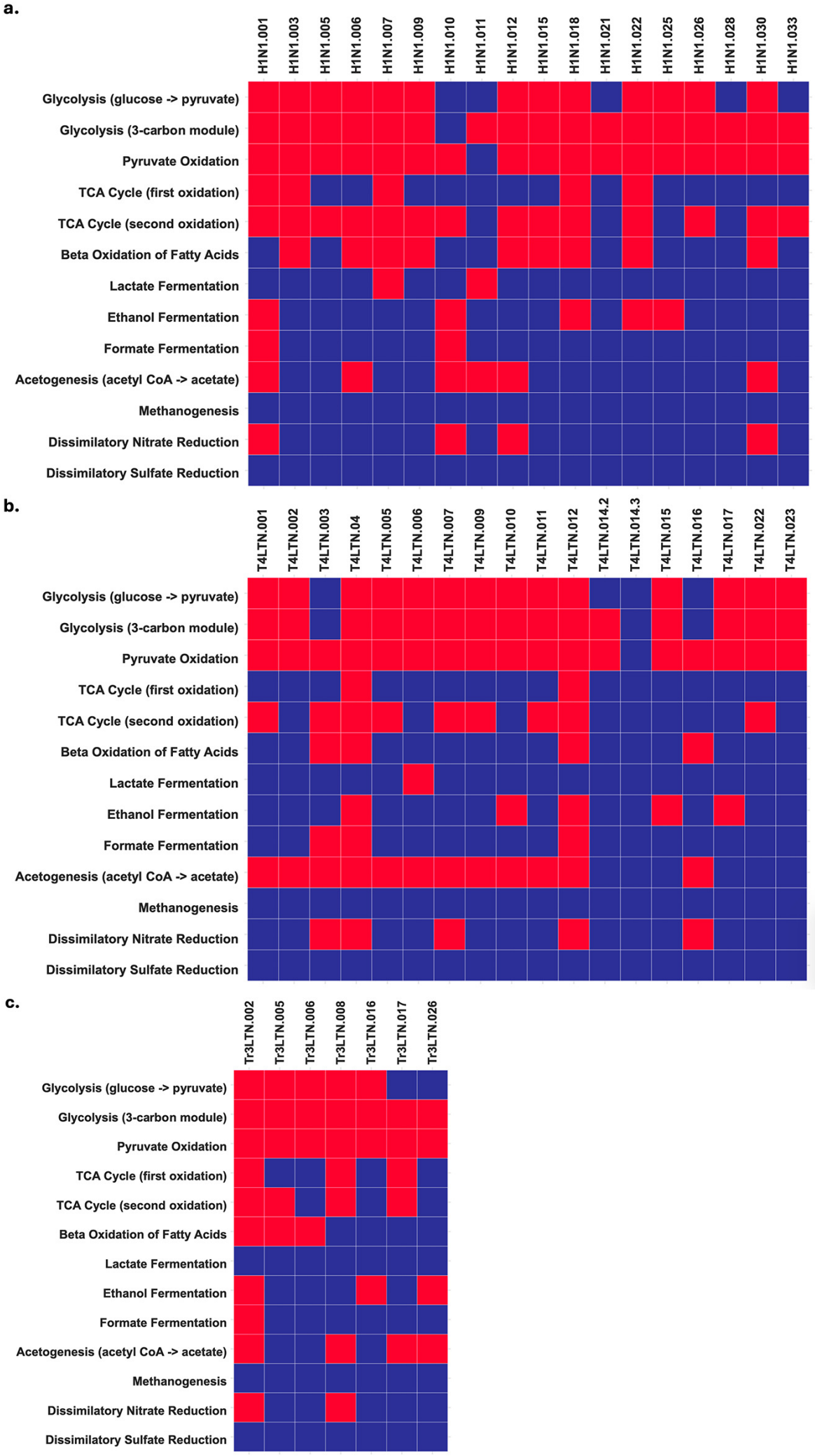
[2]Carbohydrate substrates are included if the MAG was capable of completely degrading it from start to finish. Starch is listed if the MAG possesses more than 5 amylase genes. Proteins are listed if the MAG possesses more than 10 exopeptidase genes. If the MAG could only produce β-glucosidase, simple saccharides is listed.

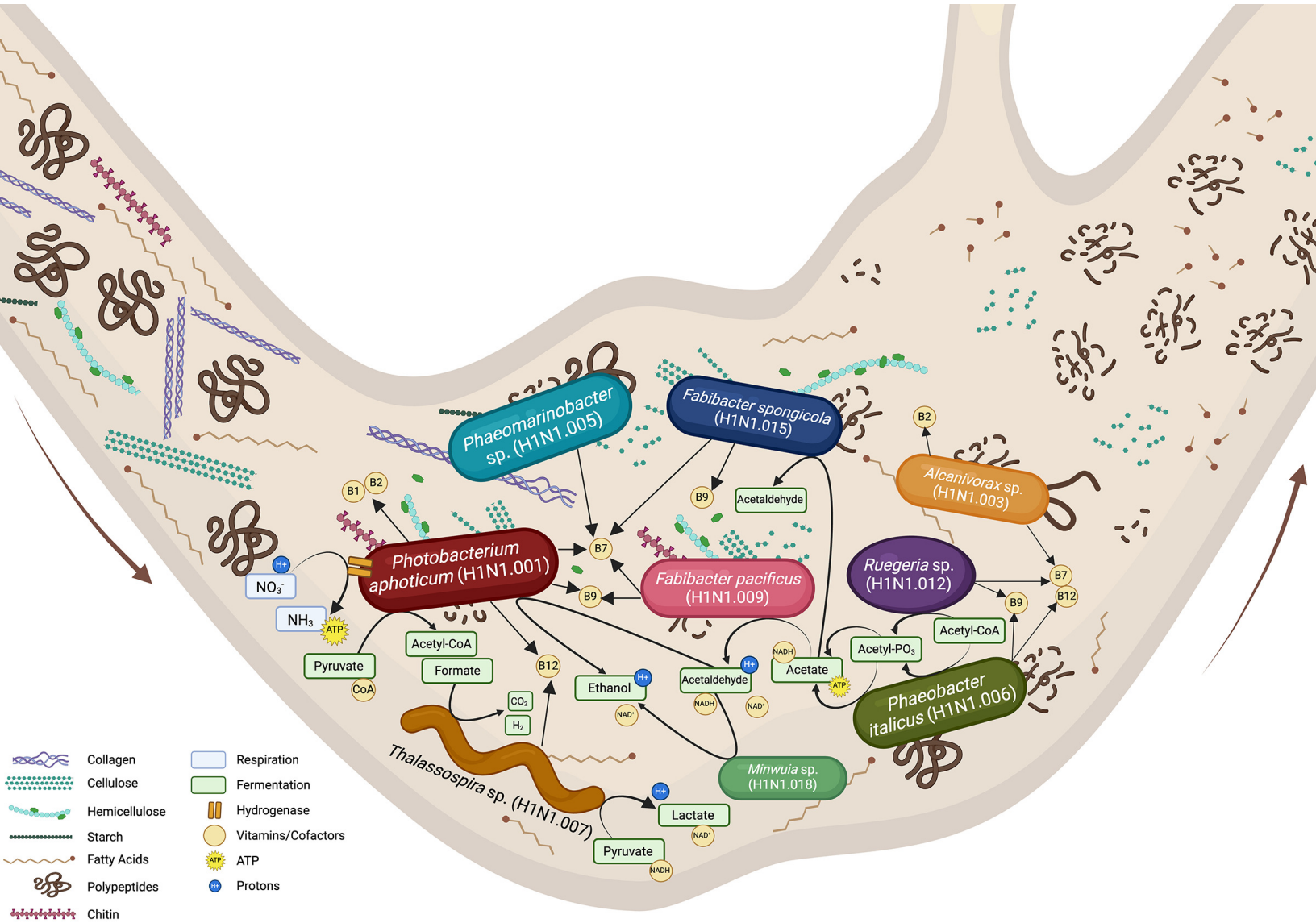
[3]For energy production, the pathway is included if the MAG can complete it entirely. If the MAG can complete an energy yielding/harvesting step of a pathway but not the entire process, the function is listed with a “p” (partial).

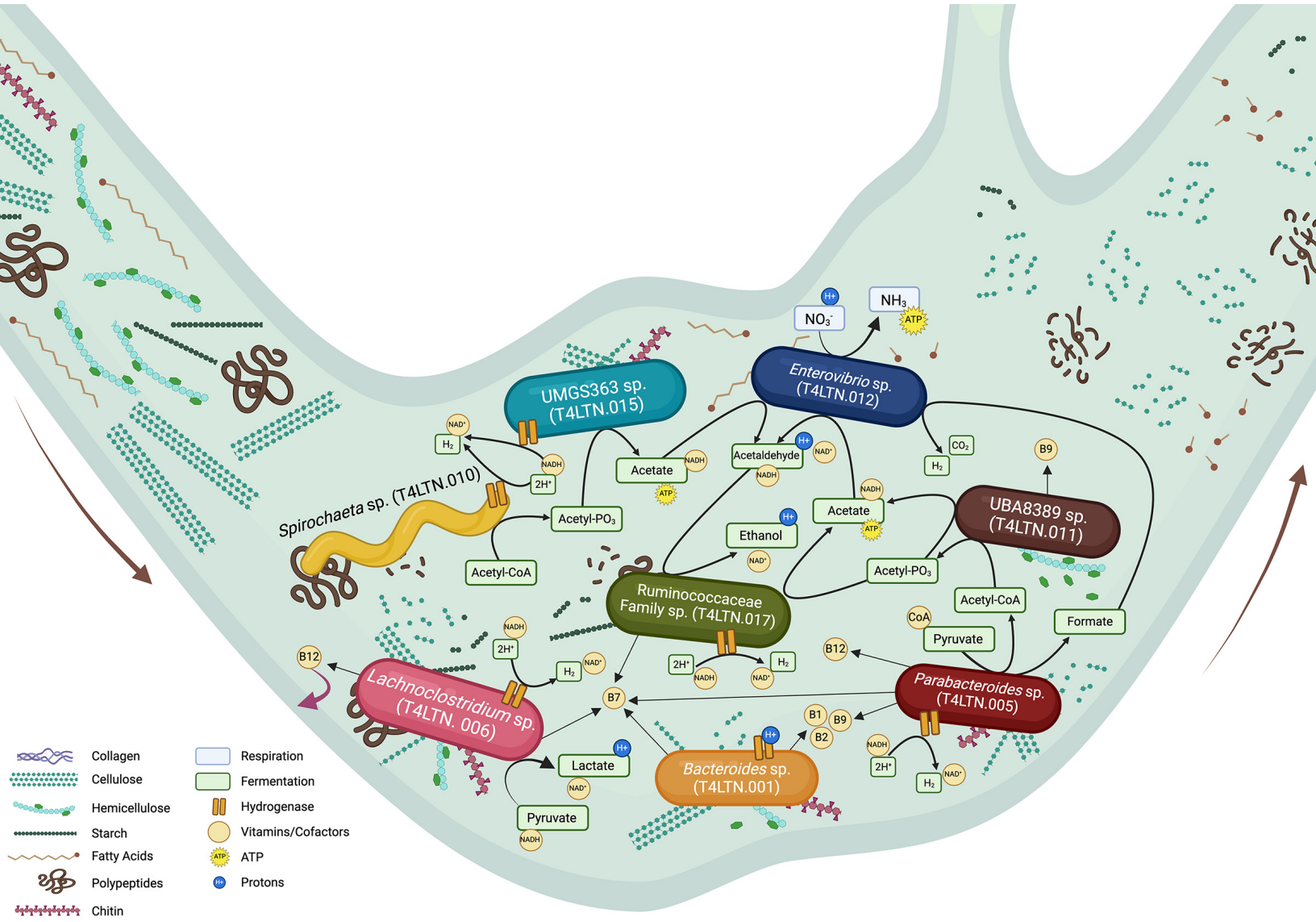

Invertivorous
Herbivorous
Piscivorous
Hemicellulose
Cellulose
Chitin
Starch
Protein
Fatty Acids
Dim1 (54%)
Dim2 (22.7%)
5.0
2.5
0.0
-2.5
-4
-2
0
2
4
6
Host Diet
Herbivorous
Invertivorous
Piscivorous

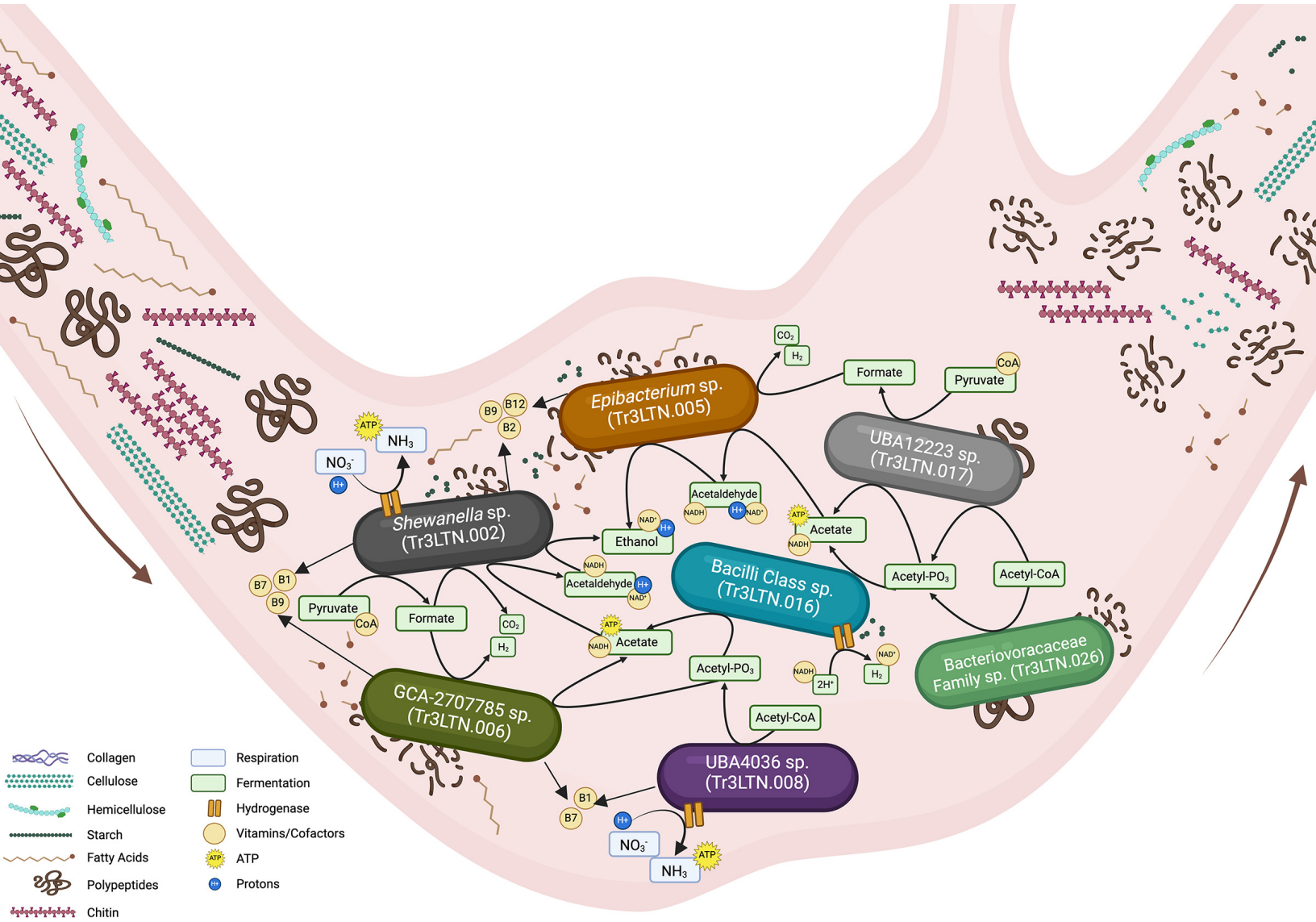

a.
Cellulose
Cellulase
[GH005, GH009, GH051, GH0074]
001 005 009 010 011 015 022 028 030 033
001 002 003 005 006 007 009 011 012 014.3 015 016 017
N/A
Cellulose Chains
Cellobiohydrolase
[GH006, GH007]
Cellobiose
β-Glucosidase
[GH001, GH003]
Exo-β-(1, 4)-Glucosidase
[GH005]
005 009 011 022 028 030 033
001 002 003 005 006 007 009 012 014.3 015 016
N/A
Glucose
b.
Xylan
α-Glucoronidase
[GH067]
022
001 006 009 011
N/A
L-Arabinase
[GH043]
001 009 010 015 022 026 028 030
001 002 005 006 007 009 011 015 017 022
016
Xylose Chain
β-Xylosidase
[GH043]
001 009 010 015 022 026 028 030
001 002 005 006 007 009 011 015 017 022
016
Xylose
c.
Mannan
α-Galactosidase
[GH004, GH110]
001 030
003 005 006 015 017
016
β-Mannanase
[GH113]
009 015 026
006 015 022
N/A
Mannose Disaccharide
β-Mannosidase
[GH002, GH113]
001 009 010 015 022 026 030
001 002 003 005 006 007 009 010 011 012 015 017
002 005 006 008 016 017 026
Mannose
d.
Xyloglucan
α-Fucosidase
[GH029]
009 015 030
002 005 006 007 009 010 011 023
N/A
α-Galactosidase
[GH001, GH110]
001 030
003 005 006 015 017
016
α-Xylosidase
[GH031]
009 015 030
001 002 004 005 006 009 010 011 014.2
002
Cellulose Chains
Cellulose Degradation
e.
Chitin
Not Screened
Chitin Deactylase
[CE004]
Chitosan
Chitosanase
[GH005, GH007, GH008 GH046, GH075, GH080]
005 009 011 021 022 025 028 030 033
001 002 003 005 006 007 009 011 014.3 015 016
N/A
Chitotriose
[(GlcpN)2]
Glucosaminidase
[GH002, GH009, GH035]
001 009 010 022 030
001 002 003 005 006 007 009 011 012 014.3 015 016 017 023
016
N-Glucosamine
[GlcpN]
Chitinase
[GH018, GH019]
001 009 010 025 028 030
001 003 005 006 012 015
N/A
Chitooligosaccharide
[(GlcpNAc)2]
N-Acetylglucosaminidase
[GH020, GH089]
001 009 010 015 022 025 030
001 003 004 005 006 007 009 010 011 012 015 017 023
002 005 006 016
N-Acetylglucosamine
[GlcpNAc]


a.
Cellulose
Kruskal-Wallis, p = 0.01
0.028
0.0087
0.24
Cellulose Degradation Enzymes Genes
Herbivorous
Invertivorous
Piscivorous
Host Diet
b.
Hemicellulose
Kruskal-Wallis, p = 0.0075
0.011
0.012
0.39
Hemicellulose Degradation Enzymes Genes
Herbivorous
Invertivorous
Piscivorous
Host Diet
c.
Chitin
Kruskal-Wallis, p = 0.0082
0.006
0.028
0.24
Chitin Degradation Enzymes Genes
Herbivorous
Invertivorous
Piscivorous
Host Diet
d.
Starch
Kruskal-Wallis, p = 0.05
0.29
0.012
0.15
Amylase Genes
Herbivorous
Invertivorous
Piscivorous
Host Diet
e.
Protein
Kruskal-Wallis, p = 0.037
0.012
0.15
1
Exopeptidase Genes
Herbivorous
Invertivorous
Piscivorous
Host Diet


a.
H1N1.001
H1N1.003
H1N1.005
H1N1.006
H1N1.007
H1N1.009
H1N1.010
H1N1.011
H1N1.012
H1N1.015
H1N1.018
H1N1.021
H1N1.022
H1N1.025
H1N1.026
H1N1.028
H1N1.030
H1N1.033
Glycolysis (glucose -> pyruvate)
Glycolysis (3-carbon module)
Pyruvate Oxidation
TCA Cycle (first oxidation)
TCA Cycle (second oxidation)
Beta Oxidation of Fatty Acids
Lactate Fermentation
Ethanol Fermentation
Formate Fermentation
Acetogenesis (acetyl CoA -> acetate)
Methanogenesis
Dissimilatory Nitrate Reduction
Dissimilatory Sulfate Reduction
b.
T4LTN.001
T4LTN.002
T4LTN.003
T4LTN.04
T4LTN.005
T4LTN.006
T4LTN.007
T4LTN.009
T4LTN.010
T4LTN.011
T4LTN.012
T4LTN.014.2
T4LTN.014.3
T4LTN.015
T4LTN.016
T4LTN.017
T4LTN.022
T4LTN.023
Glycolysis (glucose -> pyruvate)
Glycolysis (3-carbon module)
Pyruvate Oxidation
TCA Cycle (first oxidation)
TCA Cycle (second oxidation)
Beta Oxidation of Fatty Acids
Lactate Fermentation
Ethanol Fermentation
Formate Fermentation
Acetogenesis (acetyl CoA -> acetate)
Methanogenesis
Dissimilatory Nitrate Reduction
Dissimilatory Sulfate Reduction
c.
Tr3LTN.002
Tr3LTN.005
Tr3LTN.006
Tr3LTN.008
Tr3LTN.016
Tr3LTN.017
Tr3LTN.026
Glycolysis (glucose -> pyruvate)
Glycolysis (3-carbon module)
Pyruvate Oxidation
TCA Cycle (first oxidation)
TCA Cycle (second oxidation)
Beta Oxidation of Fatty Acids
Lactate Fermentation
Ethanol Fermentation
Formate Fermentation
Acetogenesis (acetyl CoA -> acetate)
Methanogenesis
Dissimilatory Nitrate Reduction
Dissimilatory Sulfate Reduction


Phaeomarinobacter sp. (H1N1.005)
Fabibacter spongicola (H1N1.015)
Alcanivorax sp. (H1N1.003)
Photobacterium aphoticum (H1N1.001)
Fabibacter pacificus (H1N1.009)
Ruegeria sp. (H1N1.012)
Phaeobacter italicus (H1N1.006)
Minwuia sp. (H1N1.018)
Thalassospira sp. (H1N1.007)
B1
B2
B7
B9
B12
Acetaldehyde
Acetyl-CoA
Formate
Pyruvate
CoA
ATP
NH3
NO3-
H+
CO2
H2
Ethanol
NAD+
NADH
Acetate
Acetyl-PO3
Lactate
Collagen
Cellulose
Hemicellulose
Starch
Fatty Acids
Polypeptides
Chitin
Respiration
Fermentation
Hydrogenase
Vitamins/Cofactors
ATP
Protons


UMGS363 sp. (T4LTN.015)
Enterovibrio sp. (T4LTN.012)
Spirochaeta sp. (T4LTN.010)
Ruminococcaceae Family sp. (T4LTN.017)
UBA8389 sp. (T4LTN.011)
Lachnoclostridium sp. (T4LTN. 006)
Parabacteroides sp. (T4LTN.005)
Bacteroides sp. (T4LTN.001)
NO3-
NH3
ATP
H+
H2
NAD+
NADH
2H+
Acetate
Acetaldehyde
CO2
B9
Acetyl-PO3
Acetyl-CoA
Ethanol
Formate
CoA
Pyruvate
B12
B7
B1
B2
Lactate
Collagen
Cellulose
Hemicellulose
Starch
Fatty Acids
Polypeptides
Chitin
Respiration
Fermentation
Hydrogenase
Vitamins/Cofactors
ATP
Protons


Epibacterium sp. (Tr3LTN.005)
Shewanella sp. (Tr3LTN.002)
UBA12223 sp. (Tr3LTN.017)
Bacilli Class sp. (Tr3LTN.016)
Bacteriovoracaceae Family sp. (Tr3LTN.026)
GCA-2707785 sp. (Tr3LTN.006)
UBA4036 sp. (Tr3LTN.008)
Formate
Pyruvate
CoA
CO2
H2
Acetaldehyde
Ethanol
Acetate
Acetyl-PO3
Acetyl-CoA
NH3
NO3-
ATP
NADH
NAD+
H+
2H+
B9
B12
B2
B7
B1
Collagen
Cellulose
Hemicellulose
Starch
Fatty Acids
Polypeptides
Chitin
Respiration
Fermentation
Hydrogenase
Vitamins/Cofactors
ATP
Protons