Diversity and Function of the Eastern Oyster (*Crassostrea virginica*) Microbiome

Zachary T. Pimentel^a, Keith Dufault-Thompson^a, Kayla T. Russo^a, Abigail K. Scro^b,
Roxanna M. Smolowitz^b, Marta Gomez-Chiarri^c, Ying Zhang^{a#}

^aDepartment of Cell and Molecular Biology, University of Rhode Island, Kingston, RI, USA

^bAquatic Diagnostic Laboratory, Roger Williams University, Bristol, RI, USA

^cDepartment of Fisheries, Animal and Veterinary Science, University of Rhode Island, Kingston,
RI, USA

Running Head: Eastern Oyster Microbiome Diversity and Function

Address correspondence to Ying Zhang, yingzhang@uri.edu

Abstract Word Count: 216

Article Word Count: 5,043

ABSTRACT

Marine invertebrate microbiomes play important roles in various host and ecological processes. However, a mechanistic understanding of host-microbe interactions is so far only available for a handful of model organisms. Here, an integrated taxonomic and functional analysis of the microbiome of the eastern oyster, *Crassostrea virginica*, was performed using 16S rRNA gene amplicon profiling, shotgun metagenomics, and genome-scale metabolic reconstruction. A relatively low number of amplicon sequence variants (ASVs) were observed in oyster tissues compared to water samples, while high variability was observed across individual oysters and among different tissue types. Targeted metagenomic sequencing of the gut microbiota led to further characterization of a dominant bacterial taxon, the class *Mollicutes*, which was captured by the reconstruction of a metagenome-assembled genome (MAG). Genome-scale metabolic reconstruction of the oyster *Mollicutes* MAG revealed a reduced set of metabolic functions and a high reliance on the uptake of host-derived nutrients. A chitin degradation and an arginine deiminase pathway were unique to the MAG as compared to other closely related *Mycoplasma* genomes, indicating a distinct mechanism of carbon and energy acquisition by the oyster-associated *Mollicutes*. A systematic reanalysis of public eastern oyster-derived microbiome data revealed the *Mollicutes* as a ubiquitous taxon among adult oysters despite their general absence in larvae and biodeposit samples, suggesting potential horizontal transmission via an unknown mechanism.

IMPORTANCE

Despite well-documented biological significance of invertebrate microbiomes, a detailed taxonomic and functional characterization is frequently missing from many non-model marine invertebrates. By using 16S rRNA gene-based community profiling, shotgun metagenomics, and genome-scale metabolic reconstruction, this study provides an integrated taxonomic and

functional analysis of the microbiome of the eastern oyster, *Crassostrea virginica*. Community profiling revealed a surprisingly low richness, as compared to surrounding seawater, and high variability among different tissue types and individuals. Reconstruction of a *Mollicutes* MAG enabled the phylogenomic positioning and functional characterization of the oyster-associated *Mollicutes*. Comparative analysis of the adult oyster gut, biodeposits, and oyster larvae samples indicated the potentially ubiquitous associations of the *Mollicutes* taxon with adult oysters. To the best of our knowledge, this study represented the first metagenomics derived functional inference of the eastern oyster microbiome. An integrated analytical procedure was developed for the functional characterization of microbiomes in other non-model host species.

INTRODUCTION

Oysters are filter feeding bivalve molluscs with great ecological and economic significance. They are considered ecosystem engineers due to their ability to form reefs that serve a variety of beneficial functions, including protecting shorelines from storm-related damage and providing habitat for other marine organisms [1,2]. Other important environmental functions provided by both wild and farmed oysters include biological remediation, sequestering of carbon through calcium carbonate shell productions, and alteration of biogeochemical cycles (such as the promotion of microbially mediated denitrification process) through their filter feeding lifestyle that involves deposition of feces and pseudofeces into the sediments [3,4]. For these reasons, efforts have been made to restore oyster populations around the United States. In addition, oysters represent a significant and growing portion of the aquaculture industry. In 2018, *Crassostrea* spp. of oysters represented almost one-third of the major species produced in world aquaculture [5].

Like other invertebrates, oysters are known to harbor a diverse range of microorganisms. Culture-based studies of oyster-associated microbes have revealed the presence of *Proteobacteria* (e.g. *Vibrio*, *Pseudomonas*, *Alteromonas*), *Actinobacteria* (e.g. *Micrococcus*), and *Bacteroidetes* (e.g. *Flavobacterium*) [6]. Some readily culturable members of the oyster microbiome, such as strains of the *Vibrio* genus, have been closely studied to profile their abundance [7], pathogenic potential [8], evolution and diversity [9], and inhibition of growth by probiotics [10,11]. Other well known, cultured microbes from oysters include the protozoan pathogen *Perkinsus marinus*, a causative agent of Dermo disease, which is one of the major diseases of adult eastern oysters [12]. Prior research has identified epizootiological factors involved with the prevalence and immunological consequences of the *P. marinus* infection [13,14]. However, little difference has been observed in the diversity and composition of the oyster-associated bacterial communities during *P. marinus* infection [15,16].

Culture-independent approaches, such as the profiling of amplicon libraries, lead to the detection of other previously uncultured taxa in the oyster microbiome, such as the *Chloroflexi*, *Firmicutes*, *Fusobacteria*, *Planctomycetes*, *Spirochaetes*, *Tenericutes*, and *Verrucomicrobia* [17–19]. A distinct microbiome is found in multiple oyster tissues (e.g. hemolymph, gill, mantle, and gut) when compared with microbial communities of the surrounding seawater, suggesting potential host selections that lead to the enrichment of specific groups [20]. Multiple factors, including changes in environmental conditions [15,17,21,22], diet [23], infection [24–26], and the use of probiotics [27], have been shown to influence the composition of oyster microbiomes during certain life stages and among different tissue types. All these studies set the stage for further investigating the taxonomic composition and functional potentials of the oyster microbiomes across different tissues.

In the absence of microbial isolates, shotgun metagenomics serves as a useful tool for gaining functional insights into uncultured members of a microbiome. Prior applications of metagenomics in marine invertebrates have revealed remarkable bacterial functions, including chemical defense mediated by secondary metabolites produced by the sponge microbiome [28] and the metabolic interactions between chemosynthetic symbionts and their hosts in deep-sea hydrothermal vents [29]. One potential challenge in the application of metagenomics to host tissues is the high abundance of host DNA that masks the signals from the tissue-associated microbiome [30]. This issue may not be easily resolved by targeting different sample types (e.g. using biodeposit samples to represent the gut microbiota), as a clear distinction is found between the fecal microbiome and the gut microbiome of filter feeding bivalves, e.g. the blue mussel, *Mytilus edulis* [31]. Thus, a successful application of metagenomics in oyster microbiome studies requires the development of customized protocols for the enrichment of microbial DNA from the oyster host tissues.

Here, an integrated microbiome analysis of the eastern oyster, *Crassostrea virginica*, was performed by combining 16S rRNA gene-based community profiling, shotgun metagenomics, and genome-scale metabolic reconstruction. Amplicon libraries of six distinct oyster tissue types were analyzed to compare the diversity and distribution of microbiomes among individual oysters. Metagenomic-based identification of oyster gut-associated bacteria was enabled with a specialized protocol that enriched microbes from host tissues. Functional characterization was performed through the application of a genome-scale metabolic reconstruction on a metagenome-assembled genome (MAG).

RESULTS

Sample Descriptions

Tissue samples were dissected from 60 farmed oysters collected from Ninigret Pond, Rhode Island, United States of America. In order to evaluate potential relationships between bacterial profiles and oyster health status, individual oysters were tested for potential infection by the most common causative agent of oyster disease in the region, the parasite *Perkinsus marinus*, using a real-time PCR assay (**Materials and Methods**). Overall, 14 of the 60 oysters were infected with *P. marinus* with a range of infection severities measured by the Mackin Index, an indicator of infection intensity based on cycle threshold values from the real-time PCR assay [32]. Subsets of infected and uninfected oysters were selected for further molecular profiling of the microbiome. Microbial community profiling was performed for 19 oysters on the gut, mantle, gill, and inner shell swab samples and for a subset of 9 oysters on the pallial fluid and hemolymph samples. As a comparison to the oyster tissues, samples were taken from the surrounding water column to profile the free-living (0.2-5 µm) and particle-associated (5-153 µm) microbial communities. Additionally, gut samples from 12 oysters were prepared for barcoded metagenomic sequencing, and pooled gut samples from another 10 oysters were prepared with an enrichment procedure for microbiome-enriched metagenomic sequencing (**Materials and Methods**). A summary of all oyster samples, the *P. marinus* Mackin Indices, and sequenced tissue types was provided in **Supplemental Table S1**.

Community Diversity of the Oyster Microbiome

The community richness, as measured by the Amplicon Sequence Variant (ASV) counts, was determined for samples of each tissue type (**Figure 1A**). The median number of ASVs per sample among the six tissue types ranged from 39 to 94, with the lowest counts in the gut and mantle and the highest counts in the inner shell swab (**Supplemental Table S2**). All oyster tissues had relatively lower richness compared to the surrounding seawater, as evidenced by the

sample-based rarefaction curves (**Supplemental Figure S1**). According to the rarefaction curves, the oyster tissue samples converged well before a sampling depth of 5,000 ASVs, with the majority of samples containing less than 200 distinct ASVs. In contrast, taxonomic profiling of the water samples barely converged with a sampling depth of over 20,000 ASVs, with 296 and 752 ASVs, respectively, identified from the free-living and particle-associated microbial fractions.

Despite the low number of ASVs across individual oyster tissue samples, relatively high variability was observed both among different tissue types and across different individuals. When comparing a tissue type across different individuals, only a small fraction of all observed ASVs were conserved (present in >80% of samples within a tissue type), with the shell swab representing the lowest percentage (1.3%) and the hemolymph representing the highest percentage (5.8%) of conserved ASVs (**Supplemental Table S2**). Similarly, when comparing the pooled ASVs from different oyster tissues, a total of 1,640 unique ASVs were identified among all the oyster samples, but only 50 ASVs were conserved among the six tissue types analyzed (**Supplemental Figure S2**). The inner shell swab samples were significantly more diverse when compared to the mantle samples based on richness (unique ASV counts) and Shannon index using the Tukey Honest Significant Differences test (TukeyHSD, p -value < 0.05) (**Figure 1A, 1B**). Similarly, a higher Shannon index was observed in the inner shell swab compared to the gut samples (**Figure 1B**). However, no statistical significance was detected among other tissue types. In addition, none of the tissue pairs showed any statistical difference based on the Simpson index (**Figure 1C**).

PERMANOVA analysis of the Bray-Curtis, Jaccard, Weighted UniFrac, and Unweighted UniFrac distances were performed to explore similarities and differences among samples of

different tissue types (**Supplemental Figure S3**). Analysis of all four distance matrices suggested that the gut samples were statistically different from the rest of the profiled tissues (p-value < 0.05). Similarly, significant differences were observed between mantle samples and all other tissue types based on Bray-Curtis, Jaccard, and Weighted UniFrac distances. However, the comparisons between gill, pallial fluid, hemolymph, and inner shell swab samples resulted in different estimations of the statistical significance depending on the distance matrix used. While the gill samples were different from all other samples based on the Jaccard or Unweighted Unifrac distances, the differences between gill and pallial fluid samples were not supported when Bray-Curtis or Weighted UniFrac distances were examined.

Taxonomic Identification of the Oyster Microbiome

Taxonomic assignment of the ASVs obtained from this study revealed several major bacterial classes that had a mean relative abundance of greater than 10% within at least one tissue type. These included *Gammaproteobacteria*, *Mollicutes*, *Fusobacteriia*, *Spirochaetia*, and *Chlamydiae*. Further, the *Gammaproteobacteria* was largely represented by the genus *Vibrio* in most tissue types (**Figure 2A,B**). The relative abundance of these major taxa was shown to vary among different tissue types based on the Tukey HSD test (p-value < 0.05). The genus *Vibrio* had a significantly lower relative abundance in gut samples as compared to other tissues (**Figure 2B**). In contrast, *Spirochaetia* had statistically higher presence in the mantle samples as compared to all other tissue types (**Figure 2D**). Meanwhile, the *Mollicutes* had statistically higher relative abundance in gut samples as compared to gill, mantle, and pallial fluid samples, but its higher mean abundance was not statistically significant when compared to the inner shell swab or hemolymph samples (**Figure 2E**). Another major taxon, *Fusobacteriia*, was relatively equally represented among all tissue types, with the mean relative abundances ranging from 11%

in the mantle to 21% in the pallial fluid (**Figure 2C**). Finally, *Chlamydiae* had a higher mean relative abundance in the gut samples as compared to other tissue samples. However, this higher mean relative abundance was largely attributed to four gut samples in which the relative abundance of *Chlamydiae* ranged from 38% to 86% (**Figure 2F**). Meanwhile, the *Chlamydiae* were absent in over one-third of the gut samples.

Compared to their presence in the oyster tissue samples, most of the major taxa had low relative abundances in the surrounding water column. For example, the relative abundance of *Mollicutes*, *Chlamydiae*, *Fusobacteriia*, and *Spirochaetia* were below 0.7% in both the free-living fraction and the particle-associated fraction of the seawater (**Supplemental Figure S4**), which were orders of magnitude lower than the mean abundance of these taxa in some oyster tissues (**Figure 2**). The *Gammaproteobacteria* had a higher relative abundance than the other taxa, with 11% in the free-living fraction and 13% in the particle-associated fraction of the seawater. However, the *Vibrio* was not a dominant *Gammaproteobacteria* in the seawater. Compared to a mean *Vibrio* abundance of above 25% across most oyster tissues (except for the gut), the *Vibrio* in seawater had a relative abundance of 0.1% in the free-living fraction and 0.3% in the particle-associated fraction (**Figure 2** and **Supplemental Figure S4**).

Metagenomic Sequencing of the Oyster Microbiome

Metagenomic sequencing was applied to the oyster gut samples in order to characterize the functional potential of the microbiome. Besides the barcoded sequencing of individual oyster samples, a procedure was adapted to enrich microbial signals and minimize genomic materials from the host (**Materials and Methods**). Overall, two rounds of metagenomic sequencing were performed, resulting in the collection of 13 metagenomes with a total of over 2.5 billion raw reads. Around 80% and 0.4% of the quality filtered reads were assigned to the oyster host and *P.*

marinus, respectively. Besides these, a total of 394,986,464 reads were used for the assembly and binning of MAGs across the 13 metagenomes.

A co-assembly of these reads resulted in 612,574 unique contigs. At least one complete 16S rRNA gene was detected for each of the major taxa, ranging from 1,396 to 1,562 base pairs in length. Six of the full-length 16S rRNA genes had identical sequences (with 100% coverage) to the ASVs (from the V4 region) from amplicon sequencing, with taxonomy mappings to the *Mollicutes* (*Mollicutes-1* and *Mollicutes-2*), *Chlamydiae*, *Spirochaetia* (*Spirochaetia-1* and *Spirochaetia-2*), and *Bacteroidia* (**Supplemental Table S3**). The relative abundance of these sequences was probed across different oyster tissues and between *P. marinus* infected and uninfected samples based on mappings to the ASV abundances (**Supplemental Figure S5**). The highest abundance was observed with ASVs mapped to the *Mollicutes-2* and *Spirochaetia-1*, with both presenting in different tissue types. The mean relative abundance of *Mollicutes-2* was around 10% in several tissue types and was significantly higher in the uninfected than the *P. marinus* infected oyster gill and pallial fluid samples. The mean relative abundance of *Spirochaetia-1* was around 20% in the mantle samples, while no significant differences were observed in its abundance between infected and uninfected oysters. Two sequences, *Mollicutes-1* and *Chlamydiae*, occurred mainly in the gut tissue, where the average relative abundance was around 5%. Both sequences had a visibly higher average relative abundance in uninfected gut samples as compared to the *P. marinus* infected tissues, but this difference was not statistically significant. Lastly, the *Spirochaetia-2* and *Bacteroidia* were present in only a few oyster tissues and had a low abundance throughout all samples.

Binning of the co-assembly produced two MAGs with completeness greater than 80% and contamination less than 2%. Each MAG contained a full-length 16S rRNA gene, with the

first from the class *Mollicutes* (*Mollicutes-1*) and the second from *Chlamydiae*. The *Mollicutes* MAG included 47 contigs with a total length of 0.62 Mb and a GC content of 28.7%. The estimated completeness and contamination was 97.4% and 1.8%, respectively. Similarly, the *Chlamydiae* MAG included 57 contigs with a total length of 1.08 Mb and a GC content of 41.0%, and it had a completeness of 86.2% and contamination of 0.4%. Besides the *Mollicutes* and *Chlamydiae* MAGs, a partial *Spirochetes* MAG was reconstructed, with a completeness of 43.2%, a contamination of 0.7%, and a full-length 16S rRNA gene (*Spirochaetia-1*) included in the MAG.

Phylogenomic reconstructions were performed based on conserved single-copy genes (CSCGs) identified between the two near-complete oyster MAGs and corresponding reference genomes in the *Mollicutes* and *Chlamydiae* taxa (**Materials and Methods**). In total, 34 CSCGs and 179 CSCGs, respectively, were used to build the *Mollicutes* and *Chlamydiae* phylogenies. Based on the phylogenomic reconstruction, the oyster *Mollicutes* MAG formed a basal branch to a group of marine *Mycoplasma*, with the nearest neighboring branches including *Mycoplasma todarodis* (isolated from a squid), *Mycoplasma marinum* (isolated from an octopus), and *Mycoplasmatales* bacterium DT_67 and DT_68, two MAGs obtained from deep-sea sinking particles (hereafter referred to as DT_67 and DT_68) [33] (**Figure 3A**). Meanwhile, the *Chlamydiae* MAG was found within the order *Parachlamydiales* most closely related to *Simkania negevensis*, an obligate intracellular bacteria with a broad host range from amoeba to animals [34,35] (**Supplemental Figure S6**).

To determine the similarity of the *Mollicutes* and *Chlamydiae* MAGs to genomes within their corresponding taxa, average nucleotide identity (ANI) and average amino acid identity (AAI) were calculated between the MAGs and all other species represented in their

corresponding CSCG phylogenies (**Supplemental Table S4, S5**). For the *Mollicutes* MAG, the highest pairwise ANI values were observed with *M. marinum* (66.8%), DT_67 (66.7%), DT_68 (66.4%), and *M. todarodis* (66.0%), all representing strains from marine sources, as well as *Mycoplasma mobile* (66.1%), a freshwater fish pathogen. Similarly, some of the highest pairwise AAI values were observed with *M. marinum* (51.9%) and *M. todarodis* (50.3%) with orthologs identified from 57.4% and 53.3%, respectively, of the protein-coding genes in the oyster *Mollicutes* MAG. While DT_67 and DT_68 had slightly higher AAI values of 54.4% and 53.6%, only a limited number of orthologs were identified, covering 22.5% and 22.2% of the protein-coding genes in the *Mollicutes* MAG (**Supplemental Table S4**). The oyster *Chlamydiae* MAG had an ANI between 62.8% and 64.8% to all reference genomes of the *Chlamydiae* phylum. The highest AAI value (47.9%) was observed with its nearest neighbor in the phylogenomic reconstruction, *Simkania negevensis*, with ortholog identifications for 53.6% of the protein-coding genes in the *Chlamydiae* MAG (**Supplemental Table S5**). Overall, both the *Mollicutes* and *Chlamydiae* MAGs had lower ANI and AAI values to the reference genomes than expected for strains of the same species (both generally $\geq 95\%$) [36,37].

Novel Functional Potentials of the Oyster *Mollicutes* MAG

A genome-scale metabolic reconstruction was performed to illustrate the metabolic capacity encoded by the oyster *Mollicutes* MAG (**Figure 3B**). Overall, the *Mollicutes* MAG encoded a highly reduced metabolism with few carbon utilization pathways and limited biosynthetic capability. Transport via the PTS system was predicted for glucose, fructose, mannose, and N-acetylglucosamine (GlcNAc). A complete glycolysis pathway and all subunits of ATP synthase were present, while genes of the tricarboxylic acid (TCA) cycle were largely missing. The ability to convert pyruvate to fumarate was predicted based on the presence of a malate dehydrogenase

and a fumarate hydratase. The conversion of acetyl-CoA to acetate was inferred by the identification of a phosphate acetyltransferase and an acetate kinase. An arginine deiminase (ADI) pathway was identified, potentially contributing to the production of ATP via the exchange of arginine and ornithine [38] and enabling the production of precursors for a *de novo* pyrimidine biosynthesis pathway (**Supplemental Table S6**). The *Mollicutes* MAG also encoded a putative chitinase with the catalytic domain homologous to an experimentally verified chitinase in *Pyrococcus furiosus* (PDB: 2DSK) and a chitin-binding domain at the C-terminus. It was noted that a mechanism of converting pyruvate to acetyl-CoA was missing in the MAG. However, putative subunits of a pyruvate dehydrogenase complex were identified from other contigs in the metagenomic co-assembly, with top BLAST hits to members of the *Mycoplasma* genus. Therefore, the potential conversion of pyruvate to acetyl-CoA was speculated in the metabolic reconstruction (**Figure 3B**).

A comparative genomic analysis was performed between the oyster *Mollicutes* MAG and genomes from closely related *Mycoplasma* isolates, including two marine species, *M. marinum* and *M. todarodis*, as well as one freshwater species, *M. mobile*. Genes in the oyster *Mollicutes* MAG can be largely classified into four categories based on their conservation with other species from the comparative genomic analysis (**Supplemental Figure S7**). Over 40% of the genes in the *Mollicutes* MAG were conserved among all reference genomes, encoding functions in central metabolism (e.g. ATP synthase, amino acid metabolism, nucleotide metabolism, *etc.*) and information storage and processing (e.g. replication, translation, transcription). Around 10% of genes were conserved between the *Mollicutes* MAG and the two marine isolates but were missing from *M. mobile*. These included genes belonging to a *de novo* pyrimidine biosynthesis pathway, multiple carbohydrate utilization genes, and an alpha-amylase domain-containing

protein. Only 18 genes (less than 3%) were uniquely conserved between the *Mollicutes* MAG and *M. mobile*, including an arginine deiminase and genes associated with DNA repair and the repair of enzymes under oxidative stress. Finally, around 36% of the genes were unique to the *Mollicutes* MAG. These included genes involved in pyruvate metabolism (e.g. malate dehydrogenase and fumarate hydratase) and the ADI pathway (e.g. carbamate kinase, ornithine carbamoyltransferase, and the arginine/ornithine antiporter). Overall, the oyster *Mollicutes* MAG maintained conserved functions with known *Mycoplasma* species while carrying unique metabolic capability in its genome.

Presence and Distribution of *Mollicutes* across Eastern Oyster Microbiome Studies

Following functional inferences obtained from the metabolic reconstruction of the oyster *Mollicutes* MAG, emphasis was placed on the composition and diversity of *Mollicutes* across different oyster microbiomes obtained from this and other studies. A total of three additional datasets were identified from public databases, including the community profiling of gut samples from the adult eastern oyster [39], the profiling of adult eastern oyster biodeposits [40], and the profiling of eastern oyster larvae homogenate [27]. Raw data from these prior studies were downloaded and reanalyzed with a standardized pipeline to minimize potential inconsistencies caused by variation in computational procedures (**Materials and Methods**). According to the comparative analysis, a consistently high relative abundance of *Mollicutes* was found in adult oyster gut samples, but this taxon was largely absent from larvae homogenate and adult biodeposits. This is in contrast to other major taxa examined. For example, the *Gammaproteobacteria*, dominated by the *Vibrio* genus, had significantly higher relative abundance in larvae homogenate compared to the adult gut or biodeposits (**Supplemental Figure S8**).

To further elucidate the taxonomic distribution of the *Mollicutes* ASVs recovered from oyster tissues and the surrounding seawater, a phylogenetic analysis was performed by placing the *Mollicutes* ASVs from this study and a previous study of the adult eastern oyster gut microbiome [39] into a reference tree of 16S rRNA gene sequences from the SILVA database (**Materials and Methods**). Branches in the tree were collapsed with labels indicating the environmental source of the SILVA references. A bubble plot was included on the right of the phylogeny to indicate the number of unique *Mollicutes* ASVs identified from different tissue or water samples (**Figure 4**).

The *Mollicutes* ASVs identified from oyster tissues had a largely distinct phylogenetic profile from the surrounding water samples. The tissue-associated ASVs consistently grouped around a clade dominated by references from diverse marine invertebrates, such as crabs, octopus, squid, and the Sydney rock oyster (*Saccostrea glomerata*) (marked with a star in **Figure 4**). In contrast, the water-associated ASVs grouped around a distant clade with references from water and sediments. Interestingly, a small number of ASVs from the oyster hemolymph, pallial fluid, and inner shell swab samples grouped within the clade of the water and sediment-associated *Mollicutes*, and the inner shell swab also included additional ASVs related to references in insects and marine invertebrates, such as coral, worms, and crustaceans. An additional clade, represented by only two references from the Sydney rock oyster and water, contained ASVs that are present in all the examined tissue types and some water samples. The phylogenetic profiling of *Mollicutes* ASVs in oyster tissues is well conserved between this study and the study by Pierce and Ward [39], where minor overlaps were found between ASVs from the oyster gut and the water samples. A detailed mapping of the *Mollicutes* references and ASVs to the different clades was shown in **Supplemental Table S7**.

DISCUSSION

The microbiome of marine invertebrates has potential significance in mediating biological and ecological functions of host organisms [41,42]. A comprehensive understanding of host-microbiome interactions relies on the taxonomic and functional characterization of the microbiome across different tissue types and individuals. Despite potential encounters with the highly diverse microbial communities in the surrounding water column through their filter feeding lifestyle, eastern oysters have been shown to generally contain low microbial diversity compared to the surrounding water column [20]. To the best of our knowledge, this study represents the first metagenome-derived functional inference of the oyster microbiome.

The amplicon-based 16S rRNA gene profiling revealed low richness in oyster tissues as compared to the surrounding seawater (**Figure 1** and **Supplemental Figure S1**), while high variability was observed among different individuals and tissue types (**Supplemental Figure S2**). Statistically significant differences have been identified between the composition of gut and mantle communities as compared to other tissue types, while less support has been found for the distinctions among gill, hemolymph, pallial fluid, and inner shell swab samples (**Supplemental Figure S3**). The open circulatory system and absence of anatomical structures that would further compartmentalize organs, such as a serous membrane or peritoneal cavity, may explain the overlap observed in the community profiles among the gill, hemolymph, pallial fluid, and inner shell swab samples. These features, for example, have previously been speculated to play a role in the spread of *P. marinus* among oyster tissues [12]. Interestingly, the gill and mantle samples demonstrated significant differences in the beta diversity measurements despite their close contact in the oyster, indicating potential mechanistic differences in the association of microbes with oyster gill and mantle tissues.

Five major bacterial taxa have been identified at the class level, with significant differences in relative abundance across tissues observed for the *Gammaproteobacteria* (dominated by the *Vibrio* genus), *Spirochaetia*, *Mollicutes*, and *Chlamydiae*. Meanwhile, a more consistent relative abundance of *Fusobacteriia* was observed across the six different tissue types examined (**Figure 2**). The presence of these major taxa has been further confirmed with metagenomic sequencing of the oyster gut microbiome, where representative full-length 16S rRNA genes were reconstructed for each taxon, and six of the 16S rRNA gene sequences had identical matches to ASVs targeting the V4 region. Two of the metagenome-assembled 16S rRNA genes were classified as *Mollicutes* (*Mollicutes-1* and *Mollicutes-2*). While the *Mollicutes-1* occurred primarily in the oyster gut and had a visibly higher but not statistically significant relative abundance in the uninfected than the *P. marinus* infected gut samples, the *Mollicutes-2* presented broadly among different tissue types and had a significantly higher relative abundance in the uninfected gill and pallial fluid samples compared to the *P. marinus* infected samples (**Supplemental Figure S5**). The differential abundance of *Mollicutes* sequences between uninfected and infected eastern oyster samples is similar to a prior study of the Sydney rock oyster (a Pacific species) where sequences related to *Mycoplasma* were present in the digestive gland of uninfected oysters but absent in oysters infected with the protozoan parasite, *Marteilia sydneyi* [24].

Additional insights into the oyster gut microbiome have been achieved with the reconstruction of MAGs. Overall, two MAGs of high completeness and low contamination have been identified for the *Mollicutes* and *Chlamydiae*, and one partial MAG has been identified from the *Spirochetes*. Specifically, phylogenomic assignment and the calculation of pairwise ANI and AAI values with reference genomes suggest the *Mollicutes* and *Chlamydiae* MAGs are

distinct from existing isolates (**Supplemental Table S4, S5**). Genome-scale metabolic reconstruction of the *Mollicutes* MAG has revealed a heavy reliance on host-derived nutrients, with several unique metabolic pathways identified in the *Mollicutes* MAG as compared to other neighboring strains in the phylogeny. One is a chitin utilization pathway, which supports the degradation of chitin for carbon and energy metabolism; another is a complete ADI pathway that could fuel ATP production through the utilization of arginine, an abundant amino acid in the eastern oyster [43]. Arginine has been previously implicated in infection dynamics of oysters. For example, *P. marinus* is speculated to sequester arginine to avoid host immune responses mediated by the production of nitric oxide, for which arginine is a precursor [44]. Interestingly, despite the presence of a GlcNAc utilization pathway in *M. marinum* and *M. todarodis* and the presence of an arginine deiminase gene in *M. mobile*, the complete ADI and chitin degradation pathways were only present in the *Mollicutes* MAG (**Figure 3B**).

The presence of the *Mollicutes* in the adult oyster gut has been observed in other 16S rRNA gene profiling studies [17,39,45]. A previous study also identified electron-dense bodies in eastern oyster gut goblet cells using transmission electron microscopy that resemble strains of the genus *Mycoplasma* [46]. Consistent with these prior studies, our analysis has provided additional support to the potentially universal presence of *Mollicutes* in the adult oyster gut. The absence of *Mollicutes* in adult biodeposits may also indicate a close association of the *Mollicutes* with digestive tissues, and this association may be mediated through a metabolic reliance of the *Mollicutes* on the host via the chitin utilization and arginine deiminase pathways (**Figure 3B**). Interestingly, the *Mollicutes* were also largely absent from samples of homogenized eastern oyster larvae, which was in contrast to the higher observed relative abundance of the *Vibrio* genus (**Supplemental Figure S8**). In addition, sequencing of surrounding water samples of the

adult oyster revealed a largely distant phylogenetic relationship between the water-associated and the oyster-associated *Mollicutes* populations (**Figure 4**), indicating a potential enrichment of specific *Mollicutes* strains in the host. Further study of the mechanisms of the *Mollicutes* acquisition, persistence, and physiology will begin to shed light on the nature of the relationship between the oyster host and *Mollicutes* (i.e. commensal, pathogen, mutualist). Overall, the integrated study of the taxonomic composition and metabolic potential of the eastern oyster microbiome will set the stage for future research on bacterial transmission dynamics, host range, and relative impacts on host health.

MATERIALS AND METHODS

Sample Collection and Dissection. Sixty live oysters were sampled from an aquaculture farm on Ninigret Pond, RI (USA) on 25 August 2017. Oysters were transported to the laboratory in coolers on ice. Upon arrival at the laboratory, the outer shells of individual oysters were scrubbed to remove visible sediments. For ten of the oysters, hemolymph and pallial fluid were collected by creating a small notch in the shell. Pallial fluid was drained through the notch prior to shucking, and the hemolymph was obtained with a sterile syringe from the pericardial cavity. Pallial fluid and hemolymph samples were centrifuged at 16,100 x g for 8 minutes, the supernatant removed, and the pellet was resuspended in 500 μ L of RNeasy lysis buffer (Qiagen, Carlsbad, CA). A sample of the gut, gills, mantle, and a swab of the inner shell (taken from the inner surface of the top and bottom shells using sterile cotton swabs) were stored individually in 1 mL of RNeasy lysis buffer (Qiagen, Carlsbad, CA). A mixed sample of mantle, gut, and rectum tissue was also preserved for pathogen quantification. In addition, 1 L of surface water collected at Ninigret Pond was pre-filtered through a 153 μ m mesh filter. Then, the water sample was filtered onto a 5 μ m pore size (diameter of 47 mm) filter (Sterlitech, Kent, WA) followed by a 0.2 μ m

pore size Sterivex filter (Millipore, Burlington, MA). All samples were stored in a -80°C freezer before further processing. Detailed tracking of all samples and their application in the 16S rRNA community profiling and metagenomic sequencing is provided in **Supplemental Table S1**.

***Perkinsus marinus* Quantification.** For each of the 60 oysters, DNA from mantle, gut, and rectum tissue samples was extracted using a previously described protocol with modifications [47]. First, 5 mg of archived oyster tissue was sterilely placed into a 1.5 mL microcentrifuge tube with 160 µL of preheated urea buffer and incubated at 60°C in a dry bath for one hour. Samples were vortexed intermittently throughout the incubation. After incubation, the tubes were vortexed again and placed into a 95°C dry bath for 15 minutes. Samples underwent centrifugation for 5 minutes at 15,000 x g, and 100 µL of the supernatant was removed and mixed with 1 µL of 100X TE buffer, 50 µL ammonium acetate, and 400 µL of 95% ethanol. After mixing, the samples were left at -20°C overnight to precipitate. The following day, samples underwent centrifugation for 20 minutes at 15,000 x g to pellet the precipitate and were washed three times with 70% ice-cold ethanol to remove any remaining impurities. DNA was eluted in 100 µL of buffer (10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0). Quantity and quality of the DNA were assessed with the NanoDrop 2000c (Thermo Scientific), and the samples were diluted to 100 ng in 3 µL to ensure consistency between qPCR reactions.

Subsequently, the DNA was used as a template in a multiplex, real-time PCR reaction testing for *P. marinus*, MSX, and SSO. *P. marinus* primers and dual-labeled probes from [48] were used in conjunction with MSX/SSO primers, MSX probe, and SSO probe (primer and probe sequences in **Supplemental Table S8**). Each 20 µL reaction was carried out using 300 nM *P. marinus* primers, 450 nM MSX/SSO primers and 75 nM probe with 3.44 µL water, 10 µL iQ Multiplex Powermix (BioRad, Hercules, CA), and 1 µL template DNA. The thermal cycler

protocol was as follows: 95°C for 60 seconds, followed by 40 cycles of 95°C for 15 seconds, then 60°C for 15 seconds.

DNA Extraction of Individual Tissue Samples. DNA extractions were performed with the Qiagen Allprep PowerFecal DNA/RNA Kit (Qiagen, Hilden, Germany) following the manufacturer's protocol with slight modifications. Modifications included the addition of 2 µL of Proteinase K (Fisher Scientific, Waltham, MA) and 13.5 µL of 20% SDS (Fisher Scientific, Waltham, MA) to the lysis buffer. Mechanical lysis was performed via two rounds of homogenization with a TissueLyser (Qiagen, Carlsbad, CA) for 2.5 minutes each at 30 Hz, and the samples were incubated for five minutes at room temperature in between.

16S rRNA Community Profiling. 16S rRNA gene profiling was performed using template DNA amplified with V4 primers 515F/806R with Illumina adapter overhangs (**Supplemental Table S8**) and 2x Phusion HF Master Mix (Thermo Fisher Scientific, Waltham, MA) [49,50]. The PCR was configured in a touchdown protocol and contained the following program: 3-minute initial denaturation at 94°C; followed by 10 cycles of a 45-second denaturation at 94°C, 60-second annealing at 60°C with every cycle decreasing 1°C, and 30 seconds of elongation at 72°C; followed by 25 cycles of a 45-second denaturation at 94°C, 60-second annealing at 50°C, and 30 seconds of elongation at 72°C; 10 minutes of final extension at 72°C, and an indefinite hold at 4°C. Sequencing libraries were prepared at the Rhode Island Genomics and Sequencing Center (RIGSC) using a standardized protocol as follows: 50 ng of Ampure XP cleaned PCR product was used as a template in a second PCR reaction (5 cycles) with 2x Phusion HF Master Mix in order to attach full indices and adapters with the Illumina Nextera Index Kit (Illumina, San Diego, CA). The resultant PCR products were cleaned with Ampure XP, profiled with an

Agilent BioAnalyzer DNA1000 chip (Agilent Technologies, Santa Clara, CA), quantified using a Qubit fluorometer (Invitrogen, Carlsbad, CA), and normalized for pooling. Then, the pooled library was quantified with qPCR with a Roche LightCycler 480 using the KAPA Biosystems Illumina Kit (KAPA Biosystems, Woburn, MA). The final pooled library underwent 2x250 bp sequencing on an Illumina MiSeq (Illumina, San Diego, CA).

Microbial Cell Enrichment of Pooled Gut Samples. In order to enhance the metagenomic binning of co-assembled contigs, a microbial cell enrichment procedure was performed on the gut samples through the adaptation of a procedure previously applied to marine sponges [51]. The gut samples of 10 oysters were pooled and homogenized with 10 ml of ice-cold, sterile artificial seawater (28 ppt) in an autoclaved mortar and pestle and vortexed for 10 minutes. The homogenized sample was centrifuged at 100 x g at 4°C for 30 minutes. The supernatant was extracted and centrifuged again at 10,967 x g at 4°C for 30 minutes. After removing the supernatant, the resultant pellet was washed three times with ice-cold, sterile artificial seawater. The pellet from the final round of centrifugation was diluted with ice-cold, sterile artificial seawater and pushed through a 3 µm pore size (diameter of 47 mm) filter (Millipore, Burlington, MA) with a syringe. The flowthrough was subsequently pushed through a 0.2 µm pore size (diameter of 22 mm) filter (Sterlitech, Kent, WA). DNA extraction was then performed on the 0.2 µm filter using the Qiagen QIAamp PowerFecal DNA Kit following manufacturer protocols. Mechanical lysis was performed via two rounds of homogenization with a TissueLyser (Qiagen, Carlsbad, CA) for 1 minute each at 30 Hz, and the sample was incubated for five minutes at room temperature in between.

Shotgun Metagenomic Sequencing. Two rounds of metagenomic sequencing were performed. In the first round, 12 gut samples were selected from the samples in which 16S rRNA gene profiling was performed (**Supplemental Table S1**). In the second round, a single library was prepared using the microbial enriched DNA from pooled homogenate of oyster gut samples (described above). Preparation of metagenomic libraries was performed at the RIGSC using standardized protocols including DNA fragmentation with an S220 ultrasonicator (Covaris, Woburn, MA), quantification with a Qubit fluorometer (Invitrogen, Carlsbad, CA), library preparation (including end repair, A-tailing, adapter ligation, and size selection) on the Wafergen Apollo 324 with PrepX reagents and consumables (Takara Bio, Kusatsu, Japan), followed by quality control using an Agilent BioAnalyzer High Sensitivity chip (Agilent Technologies, Santa Clara, CA), and quantification of the final library with qPCR with a Roche LightCycler 480 using the KAPA Biosystems Illumina Kit (KAPA Biosystems, Woburn, MA). Sequencing of the metagenomic libraries was performed at Duke University's Sequencing and Genomic Technologies Core. For the first round, the pooled library was sequenced 2x150 bp across three lanes on an Illumina HiSeq 4000, while the second round was sequenced 2x150 bp on an Illumina NextSeq 500 (Mid-output).

Amplicon Sequence Analysis. Demultiplexed read pairs were analyzed using the Quantitative Insights Into Microbial Ecology 2 (QIIME2) software version 2018.4 [52]. Using the DADA2 plugin, samples were subjected to quality filtering, denoising, and chimera detection [53]. Following quality control and chimera removal of the sequencing data, samples with less than 5,000 total read pairs were dropped from further analysis due to the low sequencing depth (**Supplemental Table S2**). Taxonomic assignment was performed with QIIME2's classifier by mapping to the SILVA database release 132 [54]. Mitochondrial and chloroplast sequences were

removed following the taxonomic assignment. The resulting feature table and taxonomy data were exported and analyzed in R version 3.6.0. Alpha and beta diversity analyses were performed on the feature table with the vegan package in R [55]. Relative abundance of a taxon was calculated through dividing the ASV counts of that taxon in a given sample by the total number of ASVs in the same sample. The raw reads from three previous eastern oyster microbiome studies were retrieved from the NCBI Sequence Read Archive (SRA) linked to the following BioProject accession numbers: PRJNA518081 [27], PRJNA504404 [40], and PRJNA386685 [39] and subjected to the same analyses as specified above. Based on metadata from the SRA, samples that were not from the eastern oyster (water, mussels, etc.) were removed from the analysis after denoising was completed.

Metagenomic Assembly and Binning. The data from both rounds of metagenomic sequencing underwent trimming and adapter removal with Trimmomatic (leading and trailing bases below a quality score of 3 were removed, a 4 bp sliding window was used with an average quality score of 15, and a minimum read length of 130 bp was required) and the last 10bp of all reads were removed with the CutAdapt software [56,57]. Due to differences in the chemistry associated with the Illumina NextSeq and HiSeq platforms, the reads from the two rounds of metagenomic sequencing underwent slightly different quality control measures. The reads from the sample that was sequenced on the Illumina NextSeq 500 also underwent additional steps in Trimmomatic and CutAdapt in order to remove PolyG. In Trimmomatic, a sequence composed of 50 guanine residues was added to the adapter sequences, and a sequence of 10 guanine residues was used as an adapter sequence in CutAdapt (allowing for an error rate of 20%).

The quality filtered reads were then mapped with BBMap to a combined reference database containing the *C. virginica* genome (RefSeq Assembly Accession GCF_002022765.2),

the *P. marinus* genome (RefSeq Assembly Accession GCF_000006405.1), all complete bacterial genomes downloaded on 18 April 2018, and a collection of manually curated draft genomes of bacteria isolated from oysters (strain names and NCBI RefSeq accessions in **Supplemental Table S9**). Reads mapped to a bacterial genome or unmapped to any genomes in the reference database were used for the metagenomic co-assembly and binning. Paired reads were considered as being mapped to a bacterial genome if at least one read from the pair was mapped. Metagenomic co-assembly was performed using MEGAHIT with default parameters [58]. Binning of the co-assembled contigs was performed with default parameters in MetaBat2 [59] using read mappings performed with BBMap, with the resulting sam files converted to bam format and sorted with SAMtools [60]. Completeness and contamination of the MAGs reconstructed from the oyster microbiome were assessed using CheckM. The *ssu_finder* function in CheckM was also used to identify 16S rRNA genes in the metagenomic co-assembly [61]. Candidate 16S rRNA genes greater than 1,000 bp in length were retained.

Phylogenetic Positioning of *Mollicutes* ASVs. The ASVs taxonomically assigned to the class *Mollicutes* from this and prior studies [39], two full-length *Mollicutes* 16S rRNA genes assembled from the metagenomes collected in this study, and the full-length 16S rRNA genes obtained from four *Firmicutes* outgroups (**Supplemental Table S7**), were placed into a reference tree with SILVA's ACT server, using the SINA aligner [62]. The minimum identity was set at 85% and the number of neighbors per query set at 100. The resultant phylogeny was collapsed based on the environment where the reference SILVA sequences were derived in iTol [63].

Phylogenomic Analyses of *Mollicutes* and *Chlamydiae* MAGs. Phylogenomic reconstructions were performed on the *Mollicutes* MAG and the *Chlamydiae* MAG with reference genomes from

the two corresponding taxa (**Supplemental Tables S4, S5**). To ensure consistency in gene calling, all genomes were first analyzed using Prodigal with genetic code 4 for the Mollicutes [64] and standard genetic code for the *Chlamydiae*. Conserved single-copy genes (CSCGs) were identified through the analysis of bidirectional best BLAST hits as described in a previous study [65]. The identified CSCGs were individually aligned with MUSCLE, the alignments were concatenated, and a phylogeny was reconstructed with RAxML using the JTT substitution model and the GAMMA model of rate heterogeneity. Pairwise average nucleotide identity (ANI) values were computed with OrthANI [66]. The average amino acid identity (AAI) was computed by the mean protein identity values of all bidirectional best BLAST hits identified based on the following thresholds: e-value less than or equal to 1×10^{-3} ; sequence identity greater than or equal to 30%; coverage of 70% or higher to both sequences in the alignment.

Metabolic Reconstruction of the *Mollicutes* MAG. A metabolic reconstruction of the *Mollicutes* MAG was developed based on an initial mapping of the proteome to the KEGG database through the KAAS server [67] and the annotation of transporters was based on homology searches to the Transporter Classification Database [68]. The draft reconstruction underwent manual curations through comparisons to three *Mycoplasma* isolates closely related to the *Mollicutes* MAG (i.e. *M. mobile*, *M. todarodis*, and *M. marinum*) and using reference annotations from an existing metabolic model of *Mycoplasma pneumoniae* [69]. The metabolic reconstruction was incorporated into a genome-scale model using PSAMM [70]. Metabolic pathway gaps were identified using the PSAMM *gapcheck* function, which in turn were used to guide the manual curation of gene functions from the MAG and from other contigs identified from the metagenomic co-assembly. For example, the pyruvate dehydrogenase complex was initially not identified in the *Mollicutes* MAG, but its subunits were found on other contigs in the

co-assembly with top BLAST hits (in the NCBI non-redundant protein database) to members of the *Mycoplasma* genus. In this case, the pyruvate dehydrogenase reaction was included in the metabolic model to highlight the potential presence of this function. Some gap reactions, such as the transport of acetate, were included in the metabolic network to represent the potential export of acetate as a metabolic product.

DATA AVAILABILITY

Raw sequencing data is available from the NCBI Sequence Read Archive under the BioProject accession PRJNA658576.

ACKNOWLEDGEMENTS

We would like to thank Alexa Sterling, Ashley Hamilton, Benjamin Korry, Dina Proestou, Evelyn Takyi, Kathryn Markey Lundgren, Rebecca Stevick, and Samuel Hughes for advice and/or assistance related to sampling and processing of oysters. Thanks to Janet Atoyan of the Rhode Island Genomics and Sequencing Center for advice and assistance related to sample preparation for sequencing.

FUNDING

This work was supported by the National Science Foundation (NSF) grant #OIA-1929078. Metabolic reconstruction was supported by NSF grant #DBI-1553211, and student fellowships were partially supported by grant #OIA-1655221. Sample collection, parasite quantification, and community and metagenomic sequencing were supported by a Rhode Island Science and Technology Advisory Council Collaborative Research Grant in 2017. Microbial community sequencing and metagenomic library preparation were conducted at a Rhode Island NSF EPSCoR research facility, the Genomics and Sequencing Center, supported in part by the

National Science Foundation EPSCoR Cooperative Agreement #OIA-1655221. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agencies.

AUTHOR CONTRIBUTIONS

Research Design & Planning: YZ, ZTP, MGC; Sample Collection: ZTP, MGC, KDT, YZ; *P. marinus* Quantification: AKS, RMS; Molecular Sequencing: ZTP, YZ; Data Analysis: ZP, YZ; Metabolic Reconstruction: KDT, KTR, ZTP, YZ; Manuscript Writing: ZP, YZ; Manuscript Editing and Approval: All authors.

REFERENCES

1. Peterson CH, Grabowski JH, Powers SP. Estimated enhancement of fish production resulting from restoring oyster reef habitat: quantitative valuation. *Mar Ecol Prog Ser.* 2003;264: 249–264.
2. Piazza BP, Banks PD, La Peyre MK. The potential for created oyster shell reefs as a sustainable shoreline protection strategy in Louisiana. *Restor Ecol.* 2005. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1526-100X.2005.00062.x>
3. Kellogg ML, Cornwell JC, Owens MS, Paynter KT. Denitrification and nutrient assimilation on a restored oyster reef. *Mar Ecol Prog Ser.* 2013;480: 1–19.
4. Olivier AS, Jones L, Le Vay L, Christie M, Wilson J, Malham SK. A global review of the ecosystem services provided by bivalve aquaculture. *Reviews in Aquaculture.* 2020. pp. 3–25. doi:10.1111/raq.12301
5. FAO. The State of World Fisheries and Aquaculture 2020: Sustainability in action. Food and Agriculture Organization of the United Nations; 2020.
6. Pierce ML, Ward JE. Microbial Ecology of the Bivalvia, with an Emphasis on the Family Ostreidae. *J Shellfish Res.* 2018;37: 793–806.
7. Blackstone GM, Nordstrom JL, Vickery MCL, Bowen MD, Meyer RF, DePaola A. Detection of pathogenic *Vibrio parahaemolyticus* in oyster enrichments by real time PCR. *J Microbiol Methods.* 2003;53: 149–155.
8. Wang R, Zhong Y, Gu X, Yuan J, Saeed AF, Wang S. The pathogenesis, detection, and prevention of *Vibrio parahaemolyticus*. *Frontiers in Microbiology.* 2015.

doi:10.3389/fmicb.2015.00144

9. Wegner KM, Mathias Wegner K, Piel D, Bruto M, John U, Mao Z, et al. Molecular Targets for Coevolutionary Interactions Between Pacific Oyster Larvae and Their Sympatric Vibrios. *Frontiers in Microbiology*. 2019. doi:10.3389/fmicb.2019.02067
10. Karim M, Zhao W, Rowley D, Nelson D, Gomez-Chiarri M. Probiotic Strains for Shellfish Aquaculture: Protection of Eastern Oyster, *Crassostrea virginica*, Larvae and Juveniles Against Bacterial Challenge. *J Shellfish Res*. 2013;32: 401–408.
11. Lim HJ, Kapareiko D, Schott EJ, Hanif A, Wikfors GH. Isolation and Evaluation of New Probiotic Bacteria for use in Shellfish Hatcheries: I. Isolation and Screening for Bioactivity. *Journal of Shellfish Research*. 2011. pp. 609–615. doi:10.2983/035.030.0303
12. Smolowitz R. A review of current state of knowledge concerning *Perkinsus marinus* effects on *Crassostrea virginica* (Gmelin) (the eastern oyster). *Vet Pathol*. 2013;50: 404–411.
13. Ford SE, Smolowitz R. Infection dynamics of an oyster parasite in its newly expanded range. *Marine Biology*. 2007. pp. 119–133. doi:10.1007/s00227-006-0454-6
14. Proestou DA, Sullivan ME. Variation in global transcriptomic response to *Perkinsus marinus* infection among eastern oyster families highlights potential mechanisms of disease resistance. *Fish Shellfish Immunol*. 2020;96: 141–151.
15. Pierce ML, Evan Ward J, Holohan BA, Zhao X, Hicks RE. The influence of site and season on the gut and pallial fluid microbial communities of the eastern oyster, *Crassostrea virginica* (Bivalvia, Ostreidae): community-level physiological profiling and genetic structure. *Hydrobiologia*. 2016. pp. 97–113. doi:10.1007/s10750-015-2405-z
16. Sakowski EG. The microbiome of the eastern oyster, *Crassostrea virginica*, in health and disease. University of Delaware. 2015. Available: <http://udspace.udel.edu/handle/19716/17503>
17. King GM, Judd C, Kuske CR, Smith C. Analysis of stomach and gut microbiomes of the eastern oyster (*Crassostrea virginica*) from coastal Louisiana, USA. *PLoS One*. 2012;7: e51475.
18. Fernandez-Piquer J, Bowman JP, Ross T, Tamplin ML. Molecular analysis of the bacterial communities in the live Pacific oyster (*Crassostrea gigas*) and the influence of postharvest temperature on its structure. *J Appl Microbiol*. 2012;112: 1134–1143.
19. Roterman YR, Benayahu Y, Reshef L, Gophna U. The gill microbiota of invasive and indigenous *Spondylus* oysters from the Mediterranean Sea and northern Red Sea. *Environmental Microbiology Reports*. 2015. pp. 860–867. doi:10.1111/1758-2229.12315
20. Lokmer A, Kuenzel S, Baines JF, Wegner KM. The role of tissue-specific microbiota in initial establishment success of Pacific oysters. *Environ Microbiol*. 2016;18: 970–987.

- 713 21. Lokmer A, Mathias Wegner K. Hemolymph microbiome of Pacific oysters in response to
714 temperature, temperature stress and infection. *ISME J.* 2015;9: 670–682.
- 715 22. Britt A, Bernini M, McSweeney B, Dalapati S, Duchin S, Cavanna K, et al. The effects of
716 atrazine on the microbiome of the eastern oyster: *Crassostrea virginica*. *Sci Rep.* 2020;10:
717 11088.
- 718 23. Simons AL, Churches N, Nuzhdin S. High turnover of faecal microbiome from algal
719 feedstock experimental manipulations in the Pacific oyster (*Crassostrea gigas*). *Microb*
720 *Biotechnol.* 2018;11: 848–858.
- 721 24. Green TJ, Barnes AC. Bacterial diversity of the digestive gland of Sydney rock oysters,
722 *Saccostrea glomerata* infected with the paramyxean parasite, *Marteilia sydneyi*. *J Appl*
723 *Microbiol.* 2010;109: 613–622.
- 724 25. King WL, Jenkins C, Seymour JR, Labbate M. Oyster disease in a changing environment:
725 Decrypting the link between pathogen, microbiome and environment. *Marine*
726 *Environmental Research.* 2019. pp. 124–140. doi:10.1016/j.marenvres.2018.11.007
- 727 26. King WL, Siboni N, Williams NLR, Kahlke T, Nguyen KV, Jenkins C, et al. Variability in
728 the Composition of Pacific Oyster Microbiomes Across Oyster Families Exhibiting
729 Different Levels of Susceptibility to OsHV-1 μ var Disease. *Front Microbiol.* 2019;10: 473.
- 730 27. Stevick RJ, Sohn S, Modak TH, Nelson DR, Rowley DC, Tammi K, et al. Bacterial
731 Community Dynamics in an Oyster Hatchery in Response to Probiotic Treatment. *Frontiers*
732 *in Microbiology.* 2019. doi:10.3389/fmicb.2019.01060
- 733 28. Pita L, Rix L, Slaby BM, Franke A, Hentschel U. The sponge holobiont in a changing
734 ocean: from microbes to ecosystems. *Microbiome.* 2018;6: 46.
- 735 29. Dubilier N, Bergin C, Lott C. Symbiotic diversity in marine animals: the art of harnessing
736 chemosynthesis. *Nat Rev Microbiol.* 2008;6: 725–740.
- 737 30. Lloyd-Price J, Mahurkar A, Rahnavard G, Crabtree J, Orvis J, Hall AB, et al. Strains,
738 functions and dynamics in the expanded Human Microbiome Project. *Nature.* 2017;550:
739 61–66.
- 740 31. Griffin TW, Baer JG, Ward JE. Direct Comparison of Fecal and Gut Microbiota in the Blue
741 Mussel (*Mytilus edulis*) Discourages Fecal Sampling as a Proxy for Resident Gut
742 Community. *Microb Ecol.* 2020. doi:10.1007/s00248-020-01553-2
- 743 32. Mackin JG. Oyster Disease Caused by *Dermocystidium Marinum* and Other
744 Microorganisms in Louisiana. 1962.
- 745 33. Ramírez AS, Vega-Orellana OM, Viver T, Poveda JB, Rosales RS, Poveda CG, et al. First
746 description of two moderately halophilic and psychrotolerant *Mycoplasma* species isolated
747 from cephalopods and proposal of *Mycoplasma marinum* sp. nov. and *Mycoplasma*
748 *todarodis* sp. nov. *Syst Appl Microbiol.* 2019;42: 457–467.

34. Michel R, Müller K-D, Zoeller L, Walochnik J, Hartmann M, Schmid E-N. Free-living amoebae serve as a host for the Chlamydia-like bacterium *Simkania negevensis*. *Acta Protozool.* 2005;44: 113–121.
35. Lieberman D, Kahane S, Lieberman D, Friedman MG. Pneumonia with serological evidence of acute infection with the Chlamydia-like microorganism “Z.” *Am J Respir Crit Care Med.* 1997;156: 578–582.
36. Konstantinidis KT, Tiedje JM. Towards a genome-based taxonomy for prokaryotes. *J Bacteriol.* 2005;187: 6258–6264.
37. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol.* 2007;57: 81–91.
38. Noens EEE, Lolkema JS. Convergent evolution of the arginine deiminase pathway: the ArcD and ArcE arginine/ornithine exchangers. *Microbiologyopen.* 2017;6. doi:10.1002/mbo3.412
39. Pierce ML, Evan Ward J. Gut Microbiomes of the Eastern Oyster (*Crassostrea virginica*) and the Blue Mussel (*Mytilus edulis*): Temporal Variation and the Influence of Marine Aggregate-Associated Microbial Communities. *mSphere.* 2019. doi:10.1128/msphere.00730-19
40. Murphy AE, Kolkmeier R, Song B, Anderson IC, Bowen J. Bioreactivity and Microbiome of Biodeposits from Filter-Feeding Bivalves. *Microb Ecol.* 2019;77: 343–357.
41. Jones BW, Nishiguchi MK. Counterillumination in the Hawaiian bobtail squid, *Euprymna scolopes* Berry (Mollusca: Cephalopoda). *Marine Biology.* 2004. pp. 1151–1155. doi:10.1007/s00227-003-1285-3
42. van Oppen MJH, Blackall LL. Coral microbiome dynamics, functions and design in a changing world. *Nature Reviews Microbiology.* 2019. pp. 557–567. doi:10.1038/s41579-019-0223-4
43. Tikunov AP, Johnson CB, Lee H, Stoskopf MK, Macdonald JM. Metabolomic Investigations of American Oysters Using ¹H-NMR Spectroscopy. *Marine Drugs.* 2010. pp. 2578–2596. doi:10.3390/md8102578
44. Villamil L, Gómez-León J, Gómez-Chiarri M. Role of nitric oxide in the defenses of *Crassostrea virginica* to experimental infection with the protozoan parasite *Perkinsus marinus*. *Dev Comp Immunol.* 2007;31: 968–977.
45. Arfken A, Song B, Bowman JS, Piehler M. Denitrification potential of the eastern oyster microbiome using a 16S rRNA gene based metabolic inference approach. *PLoS One.* 2017;12: e0185071.
46. Harshbarger JC, Chang SC. Chlamydiae (with phages), mycoplasmas, and richettsiae in

785 Chesapeake Bay bivalves. *Science*. 1977;196: 666–668.

786 47. Aranishi F, Okimoto T. A simple and reliable method for DNA extraction from bivalve
787 mantle. *J Appl Genet*. 2006;47: 251–254.

788 48. Faveri JD, De Faveri J, Smolowitz RM, Roberts SB. Development and Validation of a Real-
789 Time Quantitative PCR Assay for the Detection and Quantification of *Perkinsus marinus* in
790 the Eastern Oyster, *Crassostrea virginica*. *Journal of Shellfish Research*. 2009. pp. 459–464.
791 doi:10.2983/035.028.0306

792 49. Parada AE, Needham DM, Fuhrman JA. Every base matters: assessing small subunit rRNA
793 primers for marine microbiomes with mock communities, time series and global field
794 samples. *Environ Microbiol*. 2016;18: 1403–1414.

795 50. Apprill A, McNally S, Parsons R, Weber L. Minor revision to V4 region SSU rRNA 806R
796 gene primer greatly increases detection of SAR11 bacterioplankton. *Aquat Microb Ecol*.
797 2015;75: 129–137.

798 51. Fieseler L, Horn M, Wagner M, Hentschel U. Discovery of the Novel Candidate Phylum
799 “Poribacteria” in Marine Sponges. *Appl Environ Microbiol*. 2004;70: 3724–3732.

800 52. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al.
801 Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2.
802 *Nat Biotechnol*. 2019;37: 852–857.

803 53. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2:
804 High-resolution sample inference from Illumina amplicon data. *Nat Methods*. 2016;13:
805 581–583.

806 54. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal
807 RNA gene database project: improved data processing and web-based tools. *Nucleic Acids*
808 *Res*. 2013;41: D590–6.

809 55. Dixon P. VEGAN, a package of R functions for community ecology. *Journal of Vegetation*
810 *Science*. 2003. pp. 927–930. doi:10.1111/j.1654-1103.2003.tb02228.x

811 56. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence
812 data. *Bioinformatics*. 2014;30: 2114–2120.

813 57. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads.
814 *EMBnet.journal*. 2011. p. 10. doi:10.14806/ej.17.1.200

815 58. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. MEGAHIT: an ultra-fast single-node
816 solution for large and complex metagenomics assembly via succinct de Bruijn graph.
817 *Bioinformatics*. 2015;31: 1674–1676.

818 59. Kang D, Li F, Kirton ES, Thomas A, Egan RS, An H, et al. MetaBAT 2: an adaptive
819 binning algorithm for robust and efficient genome reconstruction from metagenome

- assemblies. PeerJ Preprints; 2019 Feb. Report No.: e27522v1.
doi:10.7287/peerj.preprints.27522v1
60. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25: 2078–2079.
61. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res*. 2015;25: 1043–1055.
62. Pruesse E, Peplies J, Glöckner FO. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics*. 2012;28: 1823–1829.
63. Letunic I, Bork P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics*. 2007;23: 127–128.
64. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*. 2010;11: 119.
65. Zhang Y, Sievert SM. Pan-genome analyses identify lineage- and niche-specific markers of evolution and adaptation in Epsilonproteobacteria. *Front Microbiol*. 2014;5: 110.
66. Lee I, Ouk Kim Y, Park S-C, Chun J. OrthoANI: An improved algorithm and software for calculating average nucleotide identity. *Int J Syst Evol Microbiol*. 2016;66: 1100–1103.
67. Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res*. 2007;35: W182–5.
68. Saier MH Jr, Reddy VS, Tsu BV, Ahmed MS, Li C, Moreno-Hagelsieb G. The Transporter Classification Database (TCDB): recent advances. *Nucleic Acids Res*. 2016;44: D372–9.
69. Wodke JAH, Puchalka J, Lluch-Senar M, Marcos J, Yus E, Godinho M, et al. Dissecting the energy metabolism in *Mycoplasma pneumoniae* through genome-scale metabolic modeling. *Molecular Systems Biology*. 2013. p. 653. doi:10.1038/msb.2013.6
70. Steffensen JL, Dufault-Thompson K, Zhang Y. PSAMM: A Portable System for the Analysis of Metabolic Models. *PLoS Comput Biol*. 2016;12: e1004732.

LEGENDS:

FIGURES:

Figure 1. Diversity of eastern oyster microbiome samples as measured by (A) richness, based on

the unique ASV count in each sample, (B) Shannon Index, and (C) Simpson Index. Samples were grouped by different tissue types, with acronyms as follows: Gil (Gill), Gut (Gut), Hmp (Hemolymph), Mtl (Mantle), Pfd (Pallial Fluid), and Swb (Inner Shell Swab). Pairwise statistical significance was assessed with Tukey's Honest Significant Differences test: * p-value < 0.05, ** p-value < 0.01.

Figure 2. Boxplot showing the relative abundance of major taxa in individual samples grouped by tissue type: *Gammaproteobacteria* (A), *Vibrio* (B), *Fusobacteriia* (C), *Spirochaetia* (D), *Mollicutes* (E), and *Chlamydiae* (F). Each tissue type was given a three letter code as follows: Gil (Gill), Gut (Gut), Hmp (Hemolymph), Mtl (Mantle), Pfd (Pallial Fluid), and Swb (Inner Shell Swab). Pairwise statistical significance was assessed with Tukey's Honest Significant Differences test: * p-value < 0.05, ** p-value < 0.01, *** p-value < 0.001.

Figure 3. Phylogenomic and functional analysis of the oyster *Mollicutes* metagenome-assembled genome (MAG) identified from this study. (A) Maximum likelihood phylogenomic reconstruction based on conserved single-copy genes. Members of the *Firmicutes* were selected as outgroups in the phylogeny (**Supplemental Table S4**). Support values were based on 100 iterations of bootstrapping. (B) Visualization of the central metabolic pathways reconstructed from the oyster *Mollicutes* MAG (created with BioRender). Metabolites are connected with directed edges indicating biochemical conversions or transport and diffusion processes. The edges were coded by circled numbers, with further details represented in **Supplemental Table S6**. Additional coding was provided by the edge colors to indicate conservations between the oyster *Mollicutes* MAG and reference genomes, including the marine host-associated

Mycoplasma marinum and *Mycoplasma todarodis*, and the freshwater host-associated *Mycoplasma mobile*. Green—conserved in all four genomes; blue—conserved between the *Mollicutes* MAG and the marine *Mycoplasma* (*M. marinum* and *M. todarodis*) but absent in *M. mobile*; purple—conserved between the *Mollicutes* MAG and the freshwater *M. mobile*, but absent from the marine *Mycoplasma*; magenta—unique functions in the *Mollicutes* MAG. The conversion from pyruvate to acetyl-CoA was marked as black because the function of a pyruvate dehydrogenase was identified outside of the MAG from unbinned contigs that had top BLAST hits to members of the *Mycoplasma*.

Figure 4. Phylogenetic positioning of oyster-associated *Mollicutes* ASVs. The branches were collapsed to summarize the environments represented by reference 16S rRNA genes in the SILVA database. The number of SILVA reference sequences (n_{REF}) and unique ASV sequences (n_{ASV}) were marked in each collapsed clade. Corresponding counts of unique ASVs in the different sample types were shown as a bubble plot for each clade on the right of the phylogeny, with filled circles sized according to the ASV counts and colored based on sample types. Data from this study and a prior study [39] was included in the analysis. Sample types included seawater the following oyster tissues: Gil (Gill), Gut (Gut), Hmp (Hemolymph), Mtl (Mantle), Pfd (Pallial Fluid), and Swb (Inner Shell Swab). A given ASV may appear in multiple sample types, and thus, the sum of values in each row of the bubble plot is not necessarily the number of ASVs (n_{ASV}) in the collapsed branch. The clade labeled as Marine Invertebrates is marked with an orange star to indicate the primary clade of oyster-associated *Mollicutes*, which includes the two metagenome-assembled full-length 16S rRNA genes, *Mollicutes-1* and *Mollicutes-2*, and their corresponding ASVs (**Supplemental Table S3**).

SUPPLEMENTAL FIGURES:

Figure S1. Rarefaction curves for the gut (A), gill (B), mantle (C), hemolymph (D), pallial fluid (E), inner shell swab (F), and water samples (G).

Figure S2. Venn diagram showing the overlap of ASVs identified from different oyster tissues. ASVs observed in at least one sample of a tissue type were counted as present in that tissue type. The number of samples in each tissue type can be found in **Supplemental Table S2**.

Figure S3. Statistical significance in the comparison of beta diversity between tissues based on four distance metrics: Bray-Curtis (A), Jaccard (B), Weighted UniFrac (C), and Unweighted UniFrac distances (D). Statistically significant differences were marked with a yellow background to the corresponding p-values based on the permutational multivariate analysis of variance (PERMANOVA) test (p-value < 0.05).

Figure S4. Relative abundance of the major taxa (*Gammaproteobacteria/Vibrio*, *Fusobacteriia*, *Spirochaetia*, *Mollicutes*, and *Chlamydiae*) in the free-living fraction (0.2-5 µm) and the particle-associated fraction (5-153 µm) of seawater samples. The relative abundance of the *Gammaproteobacteria* (in green), given their high value, are assigned to the primary (left) y-axis. The other taxa (in blue), are assigned to the secondary (right) y-axis.

Figure S5. Relative abundance of 16S rRNA gene sequences from the metagenomic assembly as measured by their corresponding ASVs (**Supplemental Table S3**). The abundance

measurements from different sample types were grouped based on the detection of *Perkinsus marinus* infections from individual oysters. Tissues that showed a significant difference between uninfected and *P. marinus* infected samples were marked: * p-value < 0.05.

Figure S6. Maximum likelihood phylogenomic reconstruction of the oyster *Chlamydiae* MAG based on conserved single copy genes identified from the MAG and reference *Chlamydiae* genomes (**Supplemental Table S5**). Species of the *Chlamydiales* order were collapsed into a single clade and formed a neighboring clade to the *Parachlamydiales* clade that includes the oyster *Chlamydiae* MAG. Members of the *Verrucomicrobia* were used as outgroups in the phylogeny. Support values were based on 100 iterations of bootstrapping.

Figure S7. Distribution of COG functions among the conserved and unique genes of the oyster *Mollicutes* MAG. The COG functions were profiled among four distinct gene groups based on conservation between the oyster *Mollicutes* MAG and reference genomes, including the marine host-associated *Mycoplasma marinum* and *Mycoplasma todarodis*, and the freshwater host-associated *Mycoplasma mobile*: All 4—conserved in all four genomes; *Mollicutes* MAG Only—unique functions in the *Mollicutes* MAG; *Mollicutes* MAG & Marine Isolates—conserved between the *Mollicutes* MAG and the marine *Mycoplasma* (*M. marinum* and *M. todarodis*) but absent in *M. mobile*; *Mollicutes* MAG & *M. mobile*—conserved between the *Mollicutes* MAG and the freshwater *M. mobile*, but absent from the marine *Mycoplasma*.

Figure S8. Comparison of relative abundances of the major taxa among different eastern oyster microbiome studies. Four distinct eastern oyster microbiome datasets were included in the

comparison, including the adult gut samples from this study and from [39], larvae homogenate samples [27], and biodeposit samples [40]. Pairwise statistical significance was assessed with Tukey's Honest Significant Differences test. Significant tissue pairs are marked as follows: * p-value < 0.05, ** p-value < 0.01, *** p-value < 0.001.

SUPPLEMENTAL TABLES:

Table S1. Summary of the oyster samples analyzed in this study. Sample identifiers for the amplicon sequencing were shown for each tissue sample of each individual oyster. Samples marked with a green background were removed from further analysis due to low sequencing depth.

Table S2. Number of samples obtained from different oyster tissues (before and after removing samples based on quality control) and summary statistics of the ASVs identified in each tissue type.

Table S3. Summary of the metagenome-assembled full-length 16S rRNA genes. Six of the gene sequences were mapped to ASV sequences with 100% identity throughout the entire length of the ASVs. Taxonomic assignments were shown based on comparison to reference genes in the SILVA database.

Table S4. Average nucleotide identity (ANI) and average amino acid identity (AAI) of the oyster *Mollicutes* MAG compared to reference genomes and outgroups analyzed in the phylogenomic reconstruction (**Figure 3A**).

Table S5. Average nucleotide identity (ANI) and average amino acid identity (AAI) of the oyster *Chlamydiae* MAG compared to reference genomes and outgroups analyzed in the phylogenomic reconstruction (**Supplemental Figure S6**).

Table S6. Metabolic functions corresponding to the visualization of central metabolism reconstructed from the oyster *Mollicutes* MAG. Each number in the Reaction Number column refers to a metabolic function correspondingly labeled in **Figure 3**. Gene identifiers corresponding to each function were included under the four genomes: *M. mobile*, *M. marinum*, *M. todarodis*, and oyster *Mollicutes* MAG. Each entry was classified in the Type column based on conservation of genes between the *Mollicutes* MAG and other genomes: All4–functions conserved in all four genomes analyzed; MAG+Marine–functions conserved between the MAG and the marine species *M. marinum* and *M. todarodis*; MAG+Mobile–functions conserved between MAG and *M. mobile*; MAG Only–functions unique to the MAG. Additional information includes the reaction name, metabolic pathway(s), Enzyme Commission Numbers (EC Number), COG classification, and eggNOG annotations were provided.

Table S7. Metadata showing the references and ASVs included in the phylogeny of **Figure 4**. Reference genes were labeled as identifiers in the SILVA reference sequence database. ASVs were labeled as their unique identifiers. The Clade assignments indicate the positioning of a reference or ASV to a specific clade in **Figure 4**. More specific information about the source environment of each reference gene was represented in the Environment column. The source samples (this study vs. [39]) for each ASV were indicated in the Sample Types column.

990

991 **Table S8.** Summary of primers and probes used in this study. The locus-specific part of the
 992 primers 515F w/ Overhang and 806R w/ Overhang are underlined while the rest of the sequence
 993 contains a sequencing adapter.

994

995 **Table S9.** Strain names and NCBI RefSeq assembly accessions for a set of manually curated
 996 draft genomes of oyster derived isolates.

997

Figure 2. Boxplot showing the relative abundance of major taxa in individual samples grouped by tissue type: *Gammaproteobacteria* (A), *Vibrio* (B), *Fusobacteriia* (C), *Spirochaetia* (D), *Mollicutes* (E), and *Chlamydiae* (F). Each tissue type was given a three letter code as follows: Gil (Gill), Gut (Gut), Hmp (Hemolymph), Mtl (Mantle), Pfd (Pallial Fluid), and Swb (Inner Shell Swab). Pairwise statistical significance was assessed with Tukey's Honest Significant Differences test: * p-value < 0.05, ** p-value < 0.01, *** p-value < 0.001.

Figure 3. Phylogenomic and functional analysis of the oyster *Mollicutes* metagenome-assembled genome (MAG) identified from this study. **(A)** Maximum likelihood phylogenomic reconstruction based on conserved single-copy genes. Members of the *Firmicutes* were selected as outgroups in the phylogeny (**Supplemental Table S4**). Support values were based on 100 iterations of bootstrapping. **(B)** Visualization of the central metabolic pathways reconstructed from the oyster *Mollicutes* MAG (created with BioRender). Metabolites are connected with directed edges indicating biochemical conversions or transport and diffusion processes. The edges were coded by circled numbers, with further details represented in **Supplemental Table S6**. Additional coding was provided by the edge colors to indicate conservations between the oyster *Mollicutes* MAG and reference genomes, including the marine host-associated *Mycoplasma marinum* and *Mycoplasma todarodis*, and the freshwater host-associated *Mycoplasma mobile*. Green—conserved in all four genomes; blue—conserved between the *Mollicutes* MAG and the marine *Mycoplasma* (*M. marinum* and *M. todarodis*) but absent in *M. mobile*; purple—conserved between the *Mollicutes* MAG and the freshwater *M. mobile*, but absent from the marine *Mycoplasma*; magenta—unique functions in the *Mollicutes* MAG. The conversion from pyruvate to acetyl-CoA was marked as black because the function of a pyruvate dehydrogenase was identified outside of the MAG from unbinned contigs that had top BLAST hits to members of the *Mycoplasma*.

Figure 4. Phylogenetic positioning of oyster-associated *Mollicutes* ASVs. The branches were collapsed to summarize the environments represented by reference 16S rRNA genes in the SILVA database. The number of SILVA reference sequences (n_{REF}) and unique ASV sequences (n_{ASV}) were marked in each collapsed clade. Corresponding counts of unique ASVs in the different sample types were shown as a bubble plot for each clade on the right of the phylogeny, with filled circles sized according to the ASV counts and colored based on sample types. Data from this study and a prior study [39] was included in the analysis. Sample types included seawater the following oyster tissues: Gil (Gill), Gut (Gut), Hmp (Hemolymph), Mtl (Mantle), Pfd (Pallial Fluid), and Swb (Inner Shell Swab). A given ASV may appear in multiple sample types, and thus, the sum of values in each row of the bubble plot is not necessarily the number of ASVs (n_{ASV}) in the collapsed branch. The clade labeled as Marine Invertebrates is marked with an orange star to indicate the primary clade of oyster-associated *Mollicutes*, which includes the two metagenome-assembled full-length 16S rRNA genes, *Mollicutes-1* and *Mollicutes-2*, and their corresponding ASVs (**Supplemental Table S3**).

Figure S1. Rarefaction curves for the gut (A), gill (B), mantle (C), hemolymph (D), pallial fluid (E), inner shell swab (F), and water samples (G).

Figure S2. Venn diagram showing the overlap of ASVs identified from different oyster tissues. ASVs observed in at least one sample of a tissue type were counted as present in that tissue type. The number of samples in each tissue type can be found in **Supplemental Table S2.**

Figure S3. Statistical significance in the comparison of beta diversity between tissues based on four distance metrics: Bray-Curtis (**A**), Jaccard (**B**), Weighted UniFrac (**C**), and Unweighted UniFrac distances (**D**). Statistically significant differences were marked with a yellow background to the corresponding p-values based on the permutational multivariate analysis of variance (PERMANOVA) test (p-value < 0.05).

A.		Bray-Curtis PERMANOVA				
		Gut	Gill	Hemolymph	Pallial Fluid	Shell Swab
Mantle		0.001	0.002	0.001	0.003	0.001
Gut			0.001	0.001	0.001	0.001
Gill				0.014	0.084	0.001
Hemolymph					0.899	0.423
Pallial Fluid						0.235

B.		Jaccard PERMANOVA				
		Gut	Gill	Hemolymph	Pallial Fluid	Shell Swab
Mantle		0.001	0.003	0.002	0.044	0.001
Gut			0.001	0.001	0.001	0.001
Gill				0.002	0.005	0.001
Hemolymph					0.552	0.01
Pallial Fluid						0.028

C.		Weighted UniFrac PERMANOVA				
		Gut	Gill	Hemolymph	Pallial Fluid	Shell Swab
Mantle		0.001	0.001	0.001	0.002	0.001
Gut			0.001	0.001	0.001	0.001
Gill				0.015	0.117	0.001
Hemolymph					0.69	0.23
Pallial Fluid						0.424

D.		Unweighted UniFrac PERMANOVA				
		Gut	Gill	Hemolymph	Pallial Fluid	Shell Swab
Mantle		0.001	0.037	0.004	0.08	0.001
Gut			0.001	0.006	0.005	0.002
Gill				0.004	0.007	0.001
Hemolymph					0.173	0.016
Pallial Fluid						0.02

Figure S4. Relative abundance of the major taxa (*Gammaproteobacteria/Vibrio*, *Fusobacteriia*, *Spirochaetia*, *Mollicutes*, and *Chlamydiae*) in the free-living fraction (0.2-5 μm) and the particle-associated fraction (5-153 μm) of seawater samples. The relative abundance of the *Gammaproteobacteria* (in green), given their high value, are assigned to the primary (left) y-axis. The other taxa (in blue), are assigned to the secondary (right) y-axis.

Figure S5. Relative abundance of 16S rRNA gene sequences from the metagenomic assembly as measured by their corresponding ASVs (**Supplemental Table S3**). The abundance measurements from different sample types were grouped based on the detection of *Perkinsus marinus* infections from individual oysters. Tissues that showed a significant difference between uninfected and *P. marinus* infected samples were marked:

* p-value < 0.05.

Figure S6. Maximum likelihood phylogenomic reconstruction of the oyster *Chlamydiae* MAG based on conserved single copy genes identified from the MAG and reference *Chlamydiae* genomes (**Supplemental Table S5**). Species of the *Chlamydiales* order were collapsed into a single clade and formed a neighboring clade to the *Parachlamydiales* clade that includes the oyster *Chlamydiae* MAG. Members of the *Verrucomicrobia* were used as outgroups in the phylogeny. Support values were based on 100 iterations of bootstrapping.

Figure S7. Distribution of COG functions among the conserved and unique genes of the oyster *Mollicutes* MAG. The COG functions were profiled among four distinct gene groups based on conservation between the oyster *Mollicutes* MAG and reference genomes, including the marine host-associated *Mycoplasma marinum* and *Mycoplasma todarodis*, and the freshwater host-associated *Mycoplasma mobile*: All 4—conserved in all four genomes; *Mollicutes* MAG Only—unique functions in the *Mollicutes* MAG; *Mollicutes* MAG & Marine Isolates—conserved between the *Mollicutes* MAG and the marine *Mycoplasma* (*M. marinum* and *M. todarodis*) but absent in *M. mobile*; *Mollicutes* MAG & *M. mobile*—conserved between the *Mollicutes* MAG and the freshwater *M. mobile*, but absent from the marine *Mycoplasma*.

Figure S8. Comparison of relative abundances of the major taxa among different eastern oyster microbiome studies. Four distinct eastern oyster microbiome datasets were included in the comparison, including the adult gut samples from this study and from [39], larvae homogenate samples [27], and biodeposit samples [40]. Pairwise statistical significance was assessed with Tukey's Honest Significant Differences test. Significant tissue pairs are marked as follows: * p-value < 0.05, ** p-value < 0.01, *** p-value < 0.001.

Table S1. Summary of the oyster samples analyzed in this study. Sample identifiers for the amplicon sequencing were shown for each tissue sample of each individual oyster. Samples marked with a green background were removed from further analysis due to low sequencing depth.

Oyster ID	<i>P. marinus</i> Infection Mackin Index	16S rRNA Gene Sequencing	Hemolymph and Pallial Fluid 16S rRNA Gene Sequenced	Individual Digestive Tract Metagenome Sequenced (Round I)	Digestive Tract Included in the Pooled Metagenomic Sequencing (Round II)	Sample IDs					
						Gut	Gill	Mtl	Swb	Hmp	Pfd
1	0										
2	0	X	X			ZP58	ZP20	ZP1	ZP39	ZP77	ZP86
3	0	X	X	ZTP1		ZP59	ZP21	ZP2	ZP40	ZP78	ZP87
4	0	X	X	ZTP2		ZP60	ZP22	ZP3	ZP41	ZP79	ZP88
5	0	X	X			ZP61	ZP23	ZP4	ZP42	ZP80	ZP89
6	3	X	X	ZTP8		ZP62	ZP24	ZP5	ZP43	ZP81	ZP90
7	1	X	X			ZP63	ZP25	ZP6	ZP44	ZP82	ZP91
8	0	X	X	ZTP3		ZP64	ZP26	ZP7	ZP45	ZP83	ZP92
9	5	X	X	ZTP4		ZP65	ZP27	ZP8	ZP46	ZP84	ZP93
10	2	X	X			ZP66	ZP28	ZP9	ZP47	ZP85	ZP94
11	0				X						
12	0.5	X		ZTP5		ZP67	ZP29	ZP10	ZP48		
13	0				X						
14	0										
15	0										
16	0				X						
17	0										
18	3	X		ZTP6		ZP68	ZP30	ZP11	ZP49		
19	0										
20	0										
21	0										
22	0										
23	0										
24	0										
25	0										
26	2	X		ZTP7		ZP69	ZP31	ZP12	ZP50		
27	0.5	X				ZP70	ZP32	ZP13	ZP51		
28	1	X				ZP71	ZP33	ZP14	ZP52		
29	0										
30	0										
31	0.5	X				ZP72	ZP34	ZP15	ZP53		
32	0				X						
33	0				X						
34	0				X						
35	0				X						
36	0				X						
37	0				X						
38	0				X						
39	0										
40	0										
41	0										
42	0										
43	0										
44	0										
45	0										
46	0										
47	0										
48	0										
49	0										
50	3	X		ZTP9		ZP732	ZP35	ZP16	ZP54		

[illegible]

Table S2. Number of samples obtained from different oyster tissues (before and after removing samples based on quality control) and summary statistics of the ASVs identified in each tissue type.

Sample Type	Samples Sequenced	Samples Retained After QC	Median ASV Count Per Sample	25th Percentile ASV Count Per Sample	75th Percentile ASV Count Per Sample	Percentage of ASVs in > 80% Samples	% ASVs That Overlap With Water Samples
Gut	19	17	39	23	64	1.65%	9.90%
Gill	19	16	53	36.75	86.5	2.78%	10.90%
Mantle	19	19	39	26	47.5	1.45%	12.50%
Shell Swab	19	18	94.5	57.75	119	1.29%	10.00%
Pallial Fluid	9	9	46	40	70	3.44%	22.50%
Hemolymph	9	9	54	50	72	5.77%	18.90%

Table S3. Summary of the metagenome-assembled full-length 16S rRNA genes. Six of the gene sequences were mapped to ASV sequences with 100% identity throughout the entire length of the ASVs. Taxonomic assignments were shown based on comparison to reference genes in the SILVA database.

[illegible]

Table S4. Average nucleotide identity (ANI) and average amino acid identity (AAI) of the oyster *Mollicutes* MAG compared to reference genomes and outgroups analyzed in the phylogenomic reconstruction (**Figure 3A**).

Strain Name	Type	NCBI RefSeq Accession	ANI Value with <i>Mollicutes</i> MAG (%)	AAI Value with <i>Mollicutes</i> MAG (%)	Gene Overlap with <i>Mollicutes</i> MAG	Gene Overlap with <i>Mollicutes</i> MAG (%)
<i>Staphylococcus aureus</i> subsp. <i>aureus</i> NCTC 8325	Outgroup	GCF_000013425.1	63.24	43.89	252	39.94
<i>Enterococcus faecalis</i> V583	Outgroup	GCF_000007785.1	63.15	43.89	244	38.67
<i>Listeria monocytogenes</i> EGD-e	Outgroup	GCF_000196035.1	62.15	43.85	253	40.10
<i>Lactobacillus plantarum</i> WCFS1	Outgroup	GCF_000203855.3	62.12	43.36	230	36.45
<i>Mycoplasmatales bacterium</i> DT_67	Reference	GCA_013215715.1	66.75	54.40	142	22.50
<i>Mycoplasmatales bacterium</i> DT_68	Reference	GCA_013215685.1	66.35	53.55	140	22.19
<i>Mycoplasma marinum</i>	Reference	GCF_004335975.1	66.80	51.94	362	57.37
<i>Mycoplasma todarodis</i>	Reference	GCF_004335995.1	66.04	50.32	336	53.25
<i>Mycoplasma crocodyli</i> MP145	Reference	GCF_000025845.1	65.49	49.70	301	47.70
<i>Mycoplasma alligatoris</i> A21JP2	Reference	GCF_000178375.1	65.39	49.41	292	46.28
<i>Mycoplasma phocicerebrale</i>	Reference	GCF_003383595.3	65.72	49.11	261	41.36
<i>Mycoplasma hyorhinis</i>	Reference	GCF_001705605.1	65.50	49.06	270	42.79
<i>Mycoplasma mobile</i> 163K	Reference	GCF_000008365.1	66.08	48.99	311	49.29
<i>Mycoplasma hominis</i>	Reference	GCF_000759375.2	65.60	48.80	260	41.20
<i>Mycoplasma hyopneumoniae</i>	Reference	GCF_002257505.1	63.78	48.67	194	30.74
<i>Mycoplasma</i> sp. Bg1	Reference	GCA_001641205.1	64.11	46.84	91	14.42
<i>Mycoplasma</i> sp. Bg2	Reference	GCA_001641225.1	64.21	45.89	84	13.31
<i>Mycoplasma capricolum</i> subsp. <i>capripneumoniae</i> 87001	Reference	GCF_000835085.1	65.39	45.86	238	37.72
<i>Mycoplasma mycoides</i> subsp. <i>capri</i>	Reference	GCF_900489525.1	65.35	45.36	249	39.46
<i>Mycoplasma</i> sp. (ex <i>Biomphalaria glabrata</i>)	Reference	GCF_001484045.1	63.69	44.14	219	34.71
<i>Candidatus Hepatoplasma crinochetorum</i> Av	Reference	GCF_000582535.1	64.07	43.53	227	35.97
<i>Mycoplasma gallisepticum</i>	Reference	GCF_001676495.1	63.01	43.48	213	33.76
<i>Candidatus Mycoplasma liparidae</i>	Reference	GCA_009884515.1	63.56	43.10	210	33.28
<i>Mycoplasma pneumoniae</i> M129	Reference	GCF_000027345.1	61.05	42.75	190	30.11

Table S5. Average nucleotide identity (ANI) and average amino acid identity (AAI) of the oyster *Chlamydiae* MAG compared to reference genomes and outgroups analyzed in the phylogenomic reconstruction (**Supplemental Figure S6**).

Strain Name	Type	NCBI RefSeq Accession	ANI Value with Chlamydiae MAG (%)	AAI Value with Chlamydiae MAG (%)	Gene Overlap with Chlamydiae MAG	Gene Overlap with Chlamydiae MAG (%)
Nibricoccus aquaticus	Outgroup	GCF_002310495.1	64.84	40.98	236	24.33
Opiritatus terrae PB90-1	Outgroup	GCF_000019965.1	63.07	41.12	241	24.85
Lacunisphaera limnophila	Outgroup	GCF_001746835.1	63.37	40.94	244	25.15
Chlamydia psittaci 84/55	Reference	GCF_000298375.2	63.75	46.18	365	37.63
Chlamydia avium 10DC88	Reference	GCF_000583875.1	63.78	46.34	366	37.73
Chlamydia psittaci VS225	Reference	GCF_000298455.2	63.73	46.27	368	37.94
Chlamydia psittaci Mat116	Reference	GCF_000338695.1	63.89	46.33	370	38.14
Chlamydia psittaci CP3	Reference	GCF_000298535.2	63.90	46.30	372	38.35
Chlamydia muridarum str. Nigg isolate CM007	Reference	GCF_002997955.1	63.54	45.80	380	39.18
Chlamydia sp. H15-1957-10C I	Reference	GCF_900239945.1	63.62	46.15	380	39.18
Chlamydia pneumoniae J138	Reference	GCF_000011165.1	64.17	46.32	381	39.28
Chlamydia gallinacea strain JX-1	Reference	GCF_002007725.1	63.78	46.11	382	39.38
Chlamydia pneumoniae CWL011	Reference	GCF_001007065.1	63.83	46.30	382	39.38
Chlamydia pneumoniae CWL029	Reference	GCF_000008745.1	63.78	46.30	382	39.38
Chlamydia pneumoniae K7	Reference	GCF_001007045.1	63.80	46.30	382	39.38
Chlamydia pneumoniae TW-183	Reference	GCF_000007205.1	63.90	46.30	382	39.38
Chlamydia pneumoniae Wien3	Reference	GCF_001007085.1	63.84	46.30	382	39.38
Chlamydia pneumoniae Wien1	Reference	GCF_001007025.1	63.69	46.31	382	39.38
Chlamydia pneumoniae CM1	Reference	GCF_001007125.1	63.87	46.35	382	39.38
Chlamydia pneumoniae PB2	Reference	GCF_001007145.1	63.69	46.36	382	39.38
Chlamydia pneumoniae CV14	Reference	GCF_001007105.1	63.69	46.36	382	39.38
Chlamydia suis isolate 9-1b I	Reference	GCF_900149505.2	63.89	46.15	383	39.48
Chlamydia pneumoniae AR39	Reference	GCF_000091085.2	63.72	46.32	383	39.48
Chlamydia gallinacea 08-1274/3	Reference	GCF_000471025.2	63.93	46.11	384	39.59
Chlamydia abortus strain GN6	Reference	GCF_002205755.1	63.72	46.29	384	39.59
Chlamydia muridarum str. Nigg isolate CM021	Reference	GCF_002997735.1	63.62	45.89	385	39.69
Chlamydia psittaci WS/RT/E30	Reference	GCF_000298475.2	63.72	46.14	385	39.69
Chlamydia pecorum PV3056/3	Reference	GCF_000470765.1	64.12	46.43	385	39.69
Chlamydia pecorum W73	Reference	GCF_000470805.1	63.57	46.44	385	39.69
Chlamydia muridarum str. Nigg isolate CM001	Reference	GCF_002998115.1	63.62	45.92	386	39.79
Chlamydia psittaci MN	Reference	GCF_000298435.2	63.80	46.08	386	39.79
Chlamydia psittaci GR9	Reference	GCF_000298415.1	64.05	46.09	386	39.79
Chlamydia psittaci WC	Reference	GCF_000298515.2	63.99	46.17	386	39.79
Chlamydia abortus strain 15-70d24	Reference	GCF_900416725.1	63.69	46.21	386	39.79
Chlamydia pneumoniae LPCoLN	Reference	GCF_000024145.1	63.80	46.25	386	39.79
Chlamydia pecorum P787	Reference	GCF_000470825.1	63.50	46.53	386	39.79
Chlamydia trachomatis L2b/UCH-1/proctitis	Reference	GCF_000068525.2	63.59	45.94	387	39.90
Chlamydia trachomatis strain L2b/CS19/08	Reference	GCF_000971705.1	63.59	45.94	387	39.90
Chlamydia trachomatis L2b/795	Reference	GCF_000318985.1	63.55	45.94	387	39.90
Chlamydia trachomatis L2b/8200/07	Reference	GCF_000318905.1	63.11	45.94	387	39.90
Chlamydia trachomatis L2b/UCH-2	Reference	GCF_000318925.1	63.11	45.94	387	39.90
Chlamydia trachomatis strain L2b/CS784/08	Reference	GCF_000971725.1	63.59	45.95	387	39.90
Chlamydia trachomatis L2/434/Bu(i)	Reference	GCF_000364765.1	63.52	45.96	387	39.90
Chlamydia trachomatis L2/434/Bu(f)	Reference	GCF_000364785.1	63.52	45.96	387	39.90
Chlamydia trachomatis RC-L2/55	Reference	GCF_000441815.1	63.80	45.97	387	39.90

Chlamydia trachomatis L2c	Reference	GCF_000220105.1	63.37	45.97	387	39.90
Chlamydia trachomatis 434/Bu	Reference	GCF_000068585.1	63.52	45.97	387	39.90
Chlamydia trachomatis L2/25667R	Reference	GCF_000318885.1	63.15	45.97	387	39.90
Chlamydia trachomatis strain tet9	Reference	GCF_004135145.1	63.79	45.98	387	39.90
Chlamydia pecorum E58	Reference	GCF_000204135.1	63.90	46.36	387	39.90
Chlamydia muridarum str. Nigg isolate CM006	Reference	GCF_002997975.1	63.64	45.89	388	40.00
Chlamydia trachomatis RC-J/953	Reference	GCF_000441675.1	63.50	45.90	388	40.00
Chlamydia trachomatis RC-J/971	Reference	GCF_000441795.1	63.42	45.90	388	40.00
Chlamydia trachomatis RC-L2(s)/46	Reference	GCF_000441615.1	63.58	45.93	388	40.00
Chlamydia trachomatis RC-J/943	Reference	GCF_000441655.1	63.69	45.94	388	40.00
Chlamydia trachomatis RC-J/966	Reference	GCF_000441755.1	63.70	45.94	388	40.00
Chlamydia abortus strain 15-49d3	Reference	GCF_900416705.1	64.18	46.23	388	40.00
Chlamydia abortus strain 84/2334	Reference	GCF_008329805.1	64.25	46.25	388	40.00
Chlamydia trachomatis strain SQ15	Reference	GCF_002777135.1	63.65	45.80	389	40.10
Chlamydia trachomatis strain SQ19	Reference	GCF_002777155.1	63.62	45.81	389	40.10
Chlamydia trachomatis strain QH111L	Reference	GCF_001885175.1	63.21	45.84	389	40.10
Chlamydia trachomatis A/HAR-13	Reference	GCF_000012125.1	63.46	45.85	389	40.10
Chlamydia trachomatis strain SQ12	Reference	GCF_002776995.1	63.39	45.87	389	40.10
Chlamydia trachomatis strain chxRP4F1	Reference	GCF_002088315.1	63.31	45.87	389	40.10
Chlamydia trachomatis L1/440/LN	Reference	GCF_000318825.1	63.20	45.87	389	40.10
Chlamydia muridarum str. Nigg	Reference	GCF_000006685.1	63.63	45.88	389	40.10
Chlamydia muridarum str. Nigg isolate CM012	Reference	GCF_002998175.1	63.64	45.88	389	40.10
Chlamydia muridarum strain Nigg3	Reference	GCF_000772145.1	63.63	45.88	389	40.10
Chlamydia muridarum str. Nigg3 CMUT3-5 strain Nigg3	Reference	GCF_000830945.1	63.63	45.88	389	40.10
Chlamydia muridarum str. Nigg 2 MCR	Reference	GCF_000709455.1	63.63	45.88	389	40.10
Chlamydia trachomatis RC-L2(s)/3	Reference	GCF_000441695.1	63.59	45.89	389	40.10
Chlamydia muridarum str. Nigg isolate CM009	Reference	GCF_002997915.1	63.63	45.89	389	40.10
Chlamydia trachomatis G/11222	Reference	GCF_000092705.1	63.26	45.89	389	40.10
Chlamydia muridarum str. Nigg isolate CM002	Reference	GCF_002998035.1	63.64	45.89	389	40.10
Chlamydia muridarum str. Nigg CM972	Reference	GCF_000830965.1	63.63	45.89	389	40.10
Chlamydia muridarum str. Nigg isolate CM004	Reference	GCF_002998235.1	63.64	45.89	389	40.10
Chlamydia muridarum str. Nigg isolate CM003	Reference	GCF_002998015.1	63.64	45.89	389	40.10
Chlamydia muridarum str. Nigg isolate CM005	Reference	GCF_002997995.1	63.64	45.89	389	40.10
Chlamydia muridarum str. Nigg isolate CM008	Reference	GCF_002997935.1	63.63	45.89	389	40.10
Chlamydia muridarum strain Nigg3 clone G0.1.1	Reference	GCF_000767405.1	63.63	45.89	389	40.10
Chlamydia muridarum strain Nigg3 clone G28.38.1	Reference	GCF_000767445.1	63.63	45.89	389	40.10
Chlamydia muridarum str. Nigg isolate CM010	Reference	GCF_002997815.1	63.63	45.89	389	40.10
Chlamydia muridarum strain Nigg3-28	Reference	GCF_003433255.1	63.63	45.89	389	40.10
Chlamydia trachomatis L3/404/LN	Reference	GCF_000319105.1	63.09	45.89	389	40.10
Chlamydia trachomatis L1/115	Reference	GCF_000318845.1	63.07	45.89	389	40.10
Chlamydia trachomatis strain chxRP4D10	Reference	GCF_002088335.1	63.32	45.89	389	40.10
Chlamydia trachomatis strain SQ10	Reference	GCF_002776975.1	63.46	45.89	389	40.10
Chlamydia trachomatis D/CS637/11	Reference	GCF_001183765.1	63.41	45.89	389	40.10
Chlamydia trachomatis strain SQ14	Reference	GCF_002777015.1	63.47	45.89	389	40.10
Chlamydia trachomatis Ia/Sotonla3	Reference	GCF_000318785.1	63.46	45.90	389	40.10

Chlamydia trachomatis la/Sotonla1	Reference	GCF_000318765.1	63.42	45.90	389	40.10
Chlamydia trachomatis L1/224	Reference	GCF_000318865.1	63.61	45.90	389	40.10
Chlamydia trachomatis strain SQ20	Reference	GCF_002776845.1	63.16	45.90	389	40.10
Chlamydia trachomatis strain SQ24	Reference	GCF_002776885.1	63.16	45.90	389	40.10
Chlamydia trachomatis strain SQ09	Reference	GCF_002776955.1	63.42	45.90	389	40.10
Chlamydia trachomatis D/UW-3/CX	Reference	GCF_000008725.1	63.20	45.91	389	40.10
Chlamydia abortus strain 1H 1	Reference	GCF_900002505.1	63.87	46.23	389	40.10
Chlamydia abortus strain 1B_covac 1	Reference	GCF_900002515.1	63.87	46.23	389	40.10
Chlamydia abortus strain 162STDY5437294 1	Reference	GCF_000952935.1	63.87	46.23	389	40.10
Chlamydia trachomatis RC-F/69	Reference	GCF_000441585.1	63.62	45.80	390	40.21
Chlamydia trachomatis RC-J(s)/122	Reference	GCF_000441735.1	63.50	45.84	390	40.21
Chlamydia trachomatis strain chxRP1H1	Reference	GCF_002088355.1	63.34	45.84	390	40.21
Chlamydia trachomatis G/9301	Reference	GCF_000092805.1	63.39	45.84	390	40.21
Chlamydia trachomatis G/11074	Reference	GCF_000092725.1	63.46	45.85	390	40.21
Chlamydia trachomatis strain SQ02	Reference	GCF_002777055.1	63.24	45.85	390	40.21
Chlamydia trachomatis J/6276tet1	Reference	GCF_000441775.1	63.77	45.85	390	40.21
Chlamydia trachomatis strain SQ01	Reference	GCF_002777035.1	63.27	45.85	390	40.21
Chlamydia trachomatis G/9768	Reference	GCF_000092685.1	63.39	45.86	390	40.21
Chlamydia trachomatis D-LC	Reference	GCF_000093005.1	63.20	45.86	390	40.21
Chlamydia trachomatis	Reference	GCF_000590755.1	63.66	45.86	390	40.21
Chlamydia trachomatis D-EC	Reference	GCF_000092485.1	63.20	45.86	390	40.21
Chlamydia trachomatis strain la/CS190/96	Reference	GCF_001183845.1	63.47	45.87	390	40.21
Chlamydia trachomatis strain SQ05	Reference	GCF_002777075.1	63.27	45.87	390	40.21
Chlamydia psittaci M56	Reference	GCF_000298495.2	63.91	46.09	390	40.21
Chlamydia abortus strain GIMC 2006: CabB577	Reference	GCF_002895085.1	64.08	46.15	390	40.21
Chlamydia abortus S26/3	Reference	GCF_000026025.1	63.68	46.15	390	40.21
Chlamydia psittaci 6BC	Reference	GCF_000204255.1	63.64	46.16	390	40.21
Chlamydia psittaci strain GIMC 2003: Cps25SM	Reference	GCF_002752695.1	64.17	46.16	390	40.21
Chlamydia psittaci 02DC15	Reference	GCF_000270425.1	63.94	46.22	390	40.21
Chlamydia trachomatis	Reference	GCF_000590615.1	62.86	45.75	391	40.31
Chlamydia trachomatis strain SQ25	Reference	GCF_002777095.1	63.65	45.77	391	40.31
Chlamydia trachomatis strain F-6068	Reference	GCF_001655575.1	62.80	45.81	391	40.31
Chlamydia trachomatis strain E-103	Reference	GCF_001655335.1	63.68	45.81	391	40.31
Chlamydia suis isolate 4-29b I	Reference	GCF_900149385.2	63.54	45.84	391	40.31
Chlamydia suis isolate SWA-2 I	Reference	GCF_900169085.1	63.60	45.88	391	40.31
Chlamydia felis Fe/C-56	Reference	GCF_000009945.1	63.54	45.97	391	40.31
Chlamydia psittaci NJ1	Reference	GCF_000298555.2	63.78	46.14	391	40.31
Chlamydia psittaci 6BC	Reference	GCF_000191925.1	63.58	46.18	391	40.31
Chlamydia psittaci strain GIMC 2005: CpsCP1	Reference	GCF_002752735.1	63.93	46.19	391	40.31
Chlamydia psittaci strain GIMC 2004: CpsAP23	Reference	GCF_002752715.1	63.87	46.19	391	40.31
Chlamydia psittaci 01DC11	Reference	GCF_000270405.1	64.05	46.19	391	40.31
Chlamydia psittaci C19/98	Reference	GCF_000270385.1	63.81	46.19	391	40.31
Chlamydia psittaci 08DC60	Reference	GCF_000270445.1	63.61	46.20	391	40.31
Chlamydia caviae GPIC	Reference	GCF_000007605.1	63.66	46.26	391	40.31
Chlamydia trachomatis	Reference	GCF_000590575.1	63.05	45.75	392	40.41
Chlamydia trachomatis	Reference	GCF_000590675.1	62.88	45.76	392	40.41
Chlamydia trachomatis E/Bour	Reference	GCF_000318645.1	63.66	45.76	392	40.41
Chlamydia trachomatis strain E-32931	Reference	GCF_001655495.1	63.27	45.76	392	40.41
Chlamydia trachomatis strain E-160	Reference	GCF_001655375.1	63.67	45.76	392	40.41
Chlamydia trachomatis strain E-547	Reference	GCF_001655415.1	63.64	45.76	392	40.41

Chlamydia trachomatis strain E/CS1025/11	Reference	GCF_001183805.1	63.07	45.76	392	40.41
Chlamydia trachomatis E/SW3	Reference	GCF_000304495.1	63.20	45.76	392	40.41
Chlamydia trachomatis Sweden2	Reference	GCF_000210495.1	63.35	45.76	392	40.41
Chlamydia trachomatis F/SW5	Reference	GCF_000304535.1	63.88	45.76	392	40.41
Chlamydia trachomatis strain E-8873	Reference	GCF_001655455.1	63.11	45.76	392	40.41
Chlamydia trachomatis strain F/CS847/08	Reference	GCF_001183825.1	63.06	45.77	392	40.41
Chlamydia trachomatis F/11-96	Reference	GCF_000590635.1	63.21	45.77	392	40.41
Chlamydia trachomatis strain E-DK-20	Reference	GCF_001655535.1	63.43	45.77	392	40.41
Chlamydia trachomatis E/11023	Reference	GCF_000092745.1	63.60	45.77	392	40.41
Chlamydia trachomatis F/SW4	Reference	GCF_000304515.1	63.80	45.77	392	40.41
Chlamydia trachomatis E/150	Reference	GCF_000092665.1	63.64	45.77	392	40.41
Chlamydia trachomatis strain SQ29	Reference	GCF_002776795.1	63.32	45.78	392	40.41
Chlamydia trachomatis strain SQ32	Reference	GCF_002776745.1	63.31	45.78	392	40.41
Chlamydia trachomatis C/TW-3	Reference	GCF_000507225.1	63.29	45.79	392	40.41
Chlamydia trachomatis A/7249	Reference	GCF_000318565.1	63.30	45.81	392	40.41
Chlamydia trachomatis A2497	Reference	GCF_000226605.1	63.30	45.82	392	40.41
Chlamydia trachomatis B/TZ1A828/OT	Reference	GCF_000026905.1	63.23	45.82	392	40.41
Chlamydia suis isolate 3-25b I	Reference	GCF_900149465.2	63.36	45.82	392	40.41
Chlamydia trachomatis A/363	Reference	GCF_000318525.1	63.07	45.82	392	40.41
Chlamydia trachomatis A2497 serovar A	Reference	GCF_000284475.1	63.30	45.82	392	40.41
Chlamydia suis isolate 5-27b I	Reference	GCF_900149625.2	63.57	45.85	392	40.41
Chlamydia psittaci strain Rostinovo-70	Reference	GCF_006385615.1	64.16	46.11	392	40.41
Chlamydia psittaci strain AMK	Reference	GCF_009857155.1	64.00	46.11	392	40.41
Chlamydia psittaci strain Ful127	Reference	GCF_003999235.1	63.96	46.19	392	40.41
Chlamydia trachomatis RC-F(s)/342	Reference	GCF_000441715.1	63.55	45.72	393	40.52
Chlamydia trachomatis B/Jali20/OT	Reference	GCF_000026925.1	63.56	45.77	393	40.52
Chlamydia trachomatis RC-F(s)/852	Reference	GCF_000441635.1	63.45	45.69	394	40.62
Chlamydia sp. S15-834C isolate S15-834K I	Reference	GCF_900239975.1	64.04	45.98	394	40.62
Neochlamydia sp. S13	Reference	GCF_000648235.2	63.67	47.33	451	46.49
Waddlia chondrophila WSU 86-1044	Reference	GCF_000092785.1	63.32	47.67	457	47.11
Parachlamydia acanthamoebae UV-7	Reference	GCF_000253035.1	63.39	47.74	465	47.94
Candidatus Protochlamydia naegleriophila strain KNic cPNK	Reference	GCF_001499655.1	63.07	47.50	474	48.87
Simkania negevensis Z gsn.131	Reference	GCF_000237205.1	62.98	47.91	520	53.61

Table S6. Metabolic functions corresponding to the visualization of central metabolism reconstructed from the oyster *Molluscus* MAG. Each number in the Reaction Number column refers to a metabolic function correspondingly labeled in **Figure 3**. Gene identifiers corresponding to each function were included under the four genomes: *M. mobile*, *M. marinum*, *M. todarodis*, and oyster *Molluscus* MAG. Each entry was classified in the Type column based on conservation of genes between the *Molluscus* MAG and other genomes: All4—functions conserved in all four genomes analyzed; MAG+Marine—functions conserved between the MAG and the marine species *M. marinum* and *M. todarodis*; MAG+Mobile—functions conserved between MAG and *M. mobile*; MAG Only—functions unique to the MAG. Additional information includes the reaction name, metabolic pathway(s), Enzyme Commission Numbers (EC Number), COG classification, and eggNOG annotations were provided.

Reaction Number	<i>Mycoplasmamobile</i>	<i>Mycoplasmamarinum</i>	<i>Mycoplasmatodarodis</i>	<i>MolluscusMAG</i>	Type	Reaction(s)	Pathway(s)	EC Number(s)	COG	EggnoG Annotation Description
1	NC_006908.1_570	NZ_PSZO01000018.1_8	NZ_PSZP01000066.1_2	k141_135469_7	All 4	alpha-D-Glucose 6-phosphate ketol-isomerase, alpha-D-Glucose 6-phosphate ketol-isomerase, beta-D-Glucose 6-phosphate ketol-isomerase	Glycolysis / Gluconeogenesis	5.1.3.15	G	Glucose-6-phosphate isomerase
2	-	NZ_PSZO01000009.1_11	NZ_PSZP01000028.1_5	k141_45383_31	MAG+Marine	D-mannose-6-phosphate aldose-ketose-isomerase	Fructose and mannose metabolism	5.3.1.8	G	Isomerase
3	-	NZ_PSZO01000024.1_8	NZ_PSZP01000006.1_3	k141_194231_4	MAG+Marine	N-Acetyl-D-glucosamine-6-phosphate amidohydrolase	Amino sugar and nucleotide sugar metabolism	3.5.1.25	G	Belongs to the metallo-dependent hydrolases superfamily. NAG4 family
4	-	NZ_PSZO01000024.1_9	NZ_PSZP01000006.1_4	k141_194231_4_3	MAG+Marine	D-glucosamine-6-phosphate aminohydrolase (ketol isomerizing)	Amino sugar and nucleotide sugar metabolism	3.5.99.6	G	Catalyzes the reversible isomerization-deamination of glucosamine 6-phosphate (GlcN6P) to form fructose 6-phosphate (Fru6P) and ammonium ion
5	NC_006908.1_329	NZ_PSZO01000003.1_100	NZ_PSZP01000041.1_2	k141_194231_3_2	All 4	ATP-D-tagatose-6-phosphate 1-phosphotransferase, ATP-D-fructose-6-phosphate 1-phosphotransferase, ATP: D-fructose-6-phosphate 1-phosphotransferase, CTP: D-tagatose 6-phosphate 1-phosphotransferase, UTP: D-tagatose 6-phosphate 1-phosphotransferase, ITP: D-tagatose 6-phosphate 1-phosphotransferase	Galactose metabolism, Methane metabolism, Glycolysis / Gluconeogenesis	2.7.1.11, 2.7.1.5, 6.2.7.1.144	G	Catalyzes the phosphorylation of D-fructose 6-phosphate to fructose 1,6-bisphosphate by ATP, the first committing step of glycolysis
6	NC_006908.1_131	NZ_PSZO01000018.1_17	NZ_PSZP01000034.1_13	k141_135469_2	All 4	D-fructose-1,6-bisphosphate D-glyceraldehyde-3-phosphate-lyase (glycerone-phosphate-forming), sedoheptulose 1,7-bisphosphate D-glyceraldehyde-3-phosphate-lyase, D-fructose 1-phosphate D-glyceraldehyde-3-phosphate-lyase, beta-D-fructose-1,6-bisphosphate D-glyceraldehyde-3-phosphate-lyase (glycerone-phosphate-forming)	Carbon metabolism, Fructose and mannose metabolism, Glycolysis / Gluconeogenesis	4.1.2.13	G	Fructose-bisphosphate aldolase class-II
7	NC_006908.1_430	NZ_PSZO01000001.1_4	NZ_PSZP01000055.1_3	k141_194231_2_1	All 4	D-glyceraldehyde-3-phosphate aldose-ketose-isomerase	Glycolysis / Gluconeogenesis	5.3.1.1	G	Involved in the gluconeogenesis. Catalyzes stereospecifically the conversion of dihydroxyacetone phosphate (DHAP) to D-glyceraldehyde-3-phosphate (G3P)
8	NC_006908.1_194	NZ_PSZO01000010.1_27	NZ_PSZP01000029.1_4	k141_392341_3	All 4	D-glyceraldehyde-3-phosphate: NAD+ oxidoreductase (phosphorylating)	Glycolysis / Gluconeogenesis	1.2.1.12, 1.2.1.5, 9	G	Belongs to the glyceraldehyde-3-phosphate dehydrogenase family
9	NC_006908.1_467	NZ_PSZO01000001.1_10	NZ_PSZP01000055.1_1	k141_194231_2_4	All 4	ATP: 3-phospho-D-glycerate 1-phosphotransferase	Glycolysis / Gluconeogenesis	2.7.2.3	F	Belongs to the phosphoglycerate kinase family
10	NC_006908.1_340	NZ_PSZO01000001.1_3	NZ_PSZP01000023.1_5	k141_194231_2_0	All 4	D-phosphoglycerate 2,3-phosphomutase	Glycolysis / Gluconeogenesis	5.4.2.11, 5.4.2.1, 2	G	Catalyzes the interconversion of 2-phosphoglycerate and 3-phosphoglycerate
11	NC_006908.1_187	NZ_PSZO01000024.1_6	NZ_PSZP01000033.1_9	k141_194231_4_6	All 4	2-phospho-D-glycerate hydro-lyase (phosphoenolpyruvate-forming)	Glycolysis / Gluconeogenesis	4.2.1.11	F	Catalyzes the reversible conversion of 2-phosphoglycerate into phosphoenolpyruvate. It is essential for the degradation of carbohydrates via glycolysis
12	NC_006908.1_330	NZ_PSZO01000003.1_99	NZ_PSZP01000041.1_3	k141_194231_3_1	All 4	ATP: pyruvate 2-O-phosphotransferase, GTP: pyruvate 2-O-phosphotransferase, dATP: pyruvate 2-O-phosphotransferase, dGTP: pyruvate 2-O-phosphotransferase	Glycolysis / Gluconeogenesis, Purine metabolism	2.7.1.40	G	Belongs to the pyruvate kinase family
13	NC_006908.1_174	NZ_PSZO01000001.1_48	NZ_PSZP01000047.1_3	k141_29341_24	All 4	propanoyl-CoA: phosphate propanoyltransferase, acetyl-CoA: phosphate acetyltransferase	Propanoate metabolism, Taurine and hypotaurine metabolism	2.3.1.8, 2.3.1.22, 2	C	Phosphate acetyltransferase
14	NC_006908.1_175	NZ_PSZO01000002.1_9	NZ_PSZP01000023.1_6	k141_29341_25	All 4	ATP: propanoate phosphotransferase	Propanoate metabolism	2.7.2.1, 2.7.2.15	F	Catalyzes the formation of acetyl phosphate from acetate and ATP. Can also catalyze the reverse reaction
15	-	-	-	k141_347318_1	MAG Only	(S)-malate: NAD+ oxidoreductase (decarboxylating)	Pyruvate metabolism	1.1.1.38, 1.1.1.3, 9	C	Psort location Cytoplasmic, score
16	-	-	-	k141_347318_2	MAG Only	(S)-malate hydro-lyase (fumarate-forming)	Citrate cycle (TCA cycle)	4.2.1.2	C	Involved in the TCA cycle. Catalyzes the stereospecific interconversion of fumarate to L-malate
17	-	NZ_PSZO01000001.1_35	NZ_PSZP01000021.1_2	k141_305077_3	MAG+Marine	N6-(1,2-dicarboxyethyl)AMP AMP-lyase (fumarate-forming), 1-(5-Phosphoribosyl)-5-amino-4-(N-succinocarboxamide)-imidazole AMP-lyase	Purine metabolism	4.3.2.2	F	Belongs to the lyase 1 family. Adenylosuccinate lyase subfamily
18	-	NZ_PSZO01000001.1_36	NZ_PSZP01000021.1_3	k141_305077_7	MAG+Marine	IMP-L-aspartate ligase (GDP-forming)	Purine metabolism	6.3.4.4	F	Plays an important role in the de novo pathway of purine nucleotide biosynthesis. Catalyzes the first committed step in the biosynthesis of AMP from IMP
19	NC_006908.1_231	-	-	k141_135469_1_0	MAG+Mobile	L-Arginine iminohydrolase	Arginine biosynthesis	3.5.3.6	E	Arginine deiminase
20	-	-	-	k141_135469_1_1	MAG Only	Carbamoyl-phosphate: L-ornithine carbamoyltransferase	Arginine biosynthesis	2.1.3.3	E	Reversibly catalyzes the transfer of the carbamoyl group from carbamoyl phosphate (CP) to the N (epsilon) atom of ornithine (ORN) to produce L-citrulline
21	-	-	-	k141_375758_7	MAG Only	ATP: carbamate phosphotransferase, ATP: carbamate phosphotransferase	Nitrogen metabolism, Arginine biosynthesis	2.7.2.2, 6.3.5.5	E	Belongs to the carbamate kinase family
22	-	NZ_PSZO01000008.1_15	NZ_PSZP01000008.1_19	k141_245899_1_1	MAG+Marine	carbamoyl-phosphate: L-aspartate carbamoyltransferase	Pyrimidine metabolism	2.1.3.2	F	Belongs to the ATCase OTCase family
23	-	NZ_PSZO01000006.1_12	NZ_PSZP01000008.1_17	k141_245899_1_2	MAG+Marine	(S)-dihydroorotate amidohydrolase	Pyrimidine metabolism	3.5.2.3	F	Belongs to the metallo-dependent hydrolases superfamily. DHOase family. Class I DHOase subfamily

24	-	NZ_PSZO0100 0006.1_8	NZ_PSP0100 0008.1_14	k141_245899_8	MAG+Marine	(S)-dihydroorotate:NAD+ oxidoreductase	Pyrimidine metabolism	1.3.1.14	F	Belongs to the dihydroorotate dehydrogenase family. Type 1 subfamily
25	-	NZ_PSZO0100 0006.1_6	NZ_PSP0100 0008.1_12	k141_245899_1 0	MAG+Marine	Orotidine-5'-phosphate:diphosphate phospho-alpha-D-ribosyl-transferase	Pyrimidine metabolism	2.4.2.10	F	Catalyzes the transfer of a ribosyl phosphate group from 5-phosphoribose 1-diphosphate to orotate, leading to the formation of orotidine monophosphate (OMP)
26	-	NZ_PSZO0100 0006.1_7	NZ_PSP0100 0008.1_13	k141_245899_9	MAG+Marine	orotidine-5'-phosphate carboxy-lyase (UMP-forming)	Pyrimidine metabolism	4.1.1.23	F	Catalyzes the decarboxylation of orotidine 5'- monophosphate (OMP) to uridine 5'- monophosphate (UMP)
27	-	-	-	k141_198499_1 0	MAG Only	chitinase	Amino sugar and nucleotide sugar metabolism	3.2.1.14	G	Chitin-binding domain type 3
28	NC_006908. 1_628	NZ_PSZO0100 0016.1_18	NZ_PSP0100 0028.1_4	k141_232514_2 4	All 4	protein-N(pi)-phosphohistidine:D-fructose 1-phosphotransferase	Transport	2.7.1.202	GT	Phosphotransferase system
29	-	NZ_PSZO0100 0024.1_11	NZ_PSP0100 0006.1_6	k141_194231_4 7	MAG+Marine	N-Acetyl-D-glucosamine transport via PEP:Pyr PTS,N-Acetyl-D-glucosamine transport via PEP:Pyr PTS (periplasm), Glucose IICBA Two Transporter, protein-N(pi)-phosphohistidine:D-glucose 6-phosphotransferase	Transport	2.7.1.199	G	Phosphotransferase system, EIIb
30	-	NZ_PSZO0100 0024.1_4	NZ_PSP0100 0033.1_7	k141_194231_3 9	MAG+Marine	Glucose IICBA Two Transporter,protein-N(pi)-phosphohistidine:D-glucose 6-phosphotransferase	Transport	2.7.1.199	G	PTS system, glucose-specific
31	-	-	-	k141_135469_1 2	MAG Only	Arginine Three Transporter,Arginine Four Transporter,Arginine Two Transporter	Transport	N/A	S	C4-dicarboxylate anaerobic carrier

Table S7. Metadata showing the references and ASVs included in the phylogeny shown in **Figure 4**. Reference genes were labeled as identifiers in the SILVA reference sequence database. ASVs were labeled as their unique identifiers. The Clade assignments indicate the positioning of a reference or ASV to a specific clade in **Figure 4**. More specific information about the source environment of each reference gene was represented in the Environment column. The source samples (this study vs. [39]) for each ASV were indicated in the Sample Types column.

Leaf ID	Clade	Environment	Type	Sample Types
JX576039.1.1346	Water & Sediment	Pit mud	SILVA Ref	
HF558631.1.1496	Water & Sediment	Mine tailing material	SILVA Ref	
2a671e73e3622f97f2d0506a1ffec705	Water & Sediment		ASV	This study-Water
KC358307.1.1332	Water & Sediment	Soda lake	SILVA Ref	
KC358310.1.1332	Water & Sediment	Soda lake	SILVA Ref	
HM127395.1.1483	Water & Sediment	Qinghai Lake water	SILVA Ref	
KC358220.1.1332	Water & Sediment	Soda lake	SILVA Ref	
65fd578fb9a7c4b9bc79b6ea339c38fc	Water & Sediment		ASV	Pierce-Water
KX014170.1.1521	Water & Sediment	Surface water from mangrove ecosystem	SILVA Ref	
KX014436.1.1528	Water & Sediment	Surface water from mangrove ecosystem	SILVA Ref	
KX097092.1.1497	Water & Sediment	Deep-sea sediment	SILVA Ref	
AB013835.1.1512	Water & Sediment	Deep-sea sediment	SILVA Ref	
1add5752236c30af9424263706ff85fa	Water & Sediment		ASV	This study-Inner Shell
6175df30f759d85441ba6a9ab018a2cb	Water & Sediment		ASV	This study-Water
EU925865.1.1487	Water & Sediment	Bering Sea sediment	SILVA Ref	
AB807258.1.1459	Water & Sediment	Ocean drilling core sample	SILVA Ref	
6fb57c651d2ac434efdf261dee0bda26	Water & Sediment		ASV	Pierce-Water
JN442952.1.1411	Water & Sediment	Hypersaline microbial mat	SILVA Ref	
4eefe7161bb97f73bf050dc2437969cf	Water & Sediment		ASV	This study-Water
2281c31ca8d25739554114a0be807c03	Water & Sediment		ASV	Pierce-Water
d54a5f521da2a7c148ece6171a430a51	Water & Sediment		ASV	Pierce-Water
KM270399.1.1479	Water & Sediment	Saline lake	SILVA Ref	
EU050916.1.1403	Water & Sediment	Arctic sediment	SILVA Ref	
KJ176755.1.1384	Water & Sediment	Lake sediment core	SILVA Ref	
74162fd6ab90302ffede29d48b70263	Water & Sediment		ASV	Pierce-Water
JN448389.1.1444	Water & Sediment	Hypersaline microbial mat	SILVA Ref	
af20f6246749430b77897825f858b3c6	Water & Sediment		ASV	This study-Inner Shell
080faf1408120dfc448a1fdec1085be3	Water & Sediment		ASV	This study-Water
EU142042.1.1496	Water & Sediment	Methane seep sediment	SILVA Ref	
CP009415.4513.6037	Water & Sediment	Methane Seep/Candidatus Izimaplasma sp. HR1	SILVA Ref	
3aa78e3f425d18d5af6a66b1e5780c54	Water & Sediment		ASV	Pierce-Water
FJ717180.1.1531	Water & Sediment	Marine sediment	SILVA Ref	
cbd3328aa47abdca0f8dae3421633aae	Water & Sediment		ASV	Pierce-Water
JQ245525.1.1389	Water & Sediment	Mud volcano	SILVA Ref	
FJ814754.1.1361	Water & Sediment	Submarine canyon	SILVA Ref	

JRFF01000022.152793.154320	Water & Sediment	Ocean Sediment Core/Candidatus Izimaplasma sp. HR2	SILVA Ref	
51856eea0ee2651043ff1a278bf36525	Water & Sediment		ASV	Pierce-Water
KC682425.1.1494	Water & Sediment	Juan de Fuca Ridge	SILVA Ref	
HQ716421.1.1366	Water & Sediment	Pig feces - Sus scrofa	SILVA Ref	
da2cd984ec6db03128e36b7485235ec9	Water & Sediment		ASV	This study-Water
KR068293.1.1531	Water & Sediment	Yak rumen - Bos grunniens	SILVA Ref	
GQ135550.1.1410	Water & Sediment	Swine waste in anaerobic reactor	SILVA Ref	
JQ191012.1.1337	Water & Sediment	Human stool - Homo sapiens	SILVA Ref	
GQ449141.1.1407	Water & Sediment	Cow feces - Bos taurus	SILVA Ref	
KF842498.1.1414	Water & Sediment	Human stool - Homo sapiens	SILVA Ref	
JQ083746.1.1520	Water & Sediment	Mouse colon tissue - Mus musculus	SILVA Ref	
89fc24ad31bddb4925c52f16e155738	Water & Sediment		ASV	Pierce-Gut
JN680638.1.1493	Water & Sediment	Cockroach gut - Shelfordella lateralis	SILVA Ref	
KX672408.1.1394	Water & Sediment	Fermentation pitmud	SILVA Ref	
JN173097.1.1543	Water & Sediment	Anaerobic bioreactor	SILVA Ref	
JQ191065.1.1326	Water & Sediment	Human stool - Homo sapiens	SILVA Ref	
JQ084329.1.1533	Water & Sediment	Mouse colon tissue - Mus musculus	SILVA Ref	
2c4ea112630127a462e26f509c2d904e	Water & Sediment		ASV	Pierce-Gut
EF406725.1.1538	Water & Sediment	Mouse colon - Mus musculus	SILVA Ref	
HM992539.1.1324	Water & Sediment	Oil sands tailings	SILVA Ref	
KJ176751.1.1377	Water & Sediment	Lake sediment core	SILVA Ref	
EU522660.1.1371	Water & Sediment	Oil sands tailings	SILVA Ref	
AB546068.1.1499	Water & Sediment	Subsurface crude oil deposits	SILVA Ref	
MG264277.1.1360	Water & Sediment	Desert oasis	SILVA Ref	
GU196243.1.1503	Water & Sediment	Anaerobic digester	SILVA Ref	
6497c9a76544b602ad7854da49f5f256	Water & Sediment		ASV	Pierce-Water
2bfc705b6717b30f7fdc24e6eb431124	Water & Sediment		ASV	Pierce-Water
FJ676076.1.1359	Water & Sediment	Cow feces - Bos taurus	SILVA Ref	
5e943fee671c8ccd15085626ccb384db	Water & Sediment		ASV	Pierce-Gut
AB606300.1.1464	Water & Sediment	Mouse feces - Mus musculus	SILVA Ref	
EF406813.1.1467	Water & Sediment	Mouse colon - Mus musculus	SILVA Ref	
JQ084129.1.1453	Water & Sediment	Mouse colon tissue - Mus musculus	SILVA Ref	
EU508626.1.1365	Water & Sediment	Mouse colon - Mus musculus	SILVA Ref	
HQ703848.1.1440	Water & Sediment	Qinghai lake sediment	SILVA Ref	

CEVG01027442.2481.3988	Water & Sediment	Marine water	SILVA Ref	
EU010153.1.1456	Water & Sediment	Marine water sample	SILVA Ref	
KP410318.1.1478	Water & Sediment	Marine water	SILVA Ref	
KC358313.1.1292	Water & Sediment	Soda lake	SILVA Ref	
FJ745224.1.1347	Water & Sediment	Marine surface water	SILVA Ref	
62a777b4962a35b697cf1c8a1112861a	Water & Sediment		ASV	This study-Water
e0084d8ea3189efa7f1d8cd888c9d78f	Water & Sediment		ASV	This study-Hemolymph, This study-Pallial Fluid, This study-Water
a11eebcc9b928c5977cb791b507b3a30	Fish & Crustaceans		ASV	Pierce-Gut
HG792199.1.1509	Fish & Crustaceans	Chinese mitten crab gut - Eriocheir sinensis	SILVA Ref	
HG792215.1.1493	Fish & Crustaceans	Chinese mitten crab gut - Eriocheir sinensis	SILVA Ref	
HG792192.1.1512	Fish & Crustaceans	Chinese mitten crab gut - Eriocheir sinensis	SILVA Ref	
HG792194.1.1517	Fish & Crustaceans	Chinese mitten crab gut - Eriocheir sinensis	SILVA Ref	
HG792244.1.1511	Fish & Crustaceans	Chinese mitten crab gut - Eriocheir sinensis	SILVA Ref	
AB980093.1.1479	Fish & Crustaceans	Deep-sea crab gut - Shinkaia crosnieri	SILVA Ref	
676e960911a82dc423bebe63fe04eec4	Fish & Crustaceans		ASV	Pierce-Gut
AB980083.1.1478	Fish & Crustaceans	Deep-sea crab gut - Shinkaia crosnieri	SILVA Ref	
975050216a47a7bda5485e656e5de095	Fish & Crustaceans		ASV	Pierce-Gut
9c5b4d7085c8a0346c7382e24b2afdc2	Fish & Crustaceans		ASV	Pierce-Gut
HG792201.1.1524	Fish & Crustaceans	Chinese mitten crab gut - Eriocheir sinensis	SILVA Ref	
f997d299ea3b6dc494abcdcfb459b71d	Fish & Crustaceans		ASV	Pierce-Water
FM165585.1.1291	Fish & Crustaceans	Earthworm - Lumbricus rubellus	SILVA Ref	
LC124566.1.1487	Fish & Crustaceans	Lake sediment	SILVA Ref	
FPLS01044383.16.1558	Fish & Crustaceans	Soil	SILVA Ref	
MF195079.1.1418	Fish & Crustaceans	Shrimp gut - Litopenaeus vannamei	SILVA Ref	
9242d1724a55496e2341aa79b6eaddfb	Fish & Crustaceans		ASV	Pierce-Water
KC120656.1.1519	Fish & Crustaceans	Turbot gut - Scophthalmus maximus	SILVA Ref	
GQ866067.1.1510	Fish & Crustaceans	Lobster gut - Nephrops norvegicus	SILVA Ref	
HE610324.1.1554	Fish & Crustaceans	Mud crab gut - Scylla paramamosain	SILVA Ref	
HE610323.1.1554	Fish & Crustaceans	Mud crab gut - Scylla paramamosain	SILVA Ref	

FM995177.1.1509	Sydney Rock Oyster & Water	Sydney rock oyster gut - <i>Saccostrea glomerata</i>	SILVA Ref	
KC189779.1.1496	Sydney Rock Oyster & Water	Freshwater spring	SILVA Ref	
49fb9cdc0de782007ba86481a9dda8a8	Sydney Rock Oyster & Water		ASV	This study-Gut, This study-Gill, This study-Mantle, This study-Hemolymph, This study-Pallial Fluid, This study-Inner Shell
9a5b0c85c23610c16b6c1e9600ccc205	Sydney Rock Oyster & Water		ASV	Pierce-Gut
a098a400272df20a41b671bb780ee5f0	Sydney Rock Oyster & Water		ASV	Pierce-Gut, Pierce-Water
6704131fb97e7403ce010d105346ead2	Sydney Rock Oyster & Water		ASV	Pierce-Gut, Pierce-Water
3d30a792cd5d7b9f6aed623763a06559	Sydney Rock Oyster & Water		ASV	Pierce-Gut
ff86244cf432bc769c451eae73709079	Sydney Rock Oyster & Water		ASV	Pierce-Gut
28b75f71e71b48f18e8ed3181fa76da1	Sydney Rock Oyster & Water		ASV	Pierce-Gut, Pierce-Water
GQ275102.1.1395	Insects & Marine Invertebrates	Ant gut - <i>Lasius alienus</i>	SILVA Ref	
AJ515719.1.1452	Insects & Marine Invertebrates	Deep-sea shrimp gut - <i>Rimicaris exoculata</i>	SILVA Ref	
AJ515720.1.1485	Insects & Marine Invertebrates	Deep-sea shrimp gut - <i>Rimicaris exoculata</i>	SILVA Ref	
4183f35d6c771afd415bf68ead2d7bf8	Insects & Marine Invertebrates		ASV	This study-Inner Shell
KC669119.1.1488	Insects & Marine Invertebrates	Coral - <i>Stylophora pistillata</i>	SILVA Ref	
5bada3cc83169941170aae8fb04e69e5	Insects & Marine Invertebrates		ASV	This study-Water
DQ890427.1.1476	Insects & Marine Invertebrates	Mudshrimp midgut - <i>Pestarella tyrrhena</i>	SILVA Ref	
9384a1a63342d6ceaec7ce2790952a08	Worms & Crustaceans		ASV	Pierce-Water
AB990245.1.1443	Worms & Crustaceans	Glacier ice worm - <i>Mesenchytraeus solifugus</i>	SILVA Ref	
AB990194.1.1443	Worms & Crustaceans	Glacier ice worm - <i>Mesenchytraeus solifugus</i>	SILVA Ref	
HE610322.1.1517	Worms & Crustaceans	Mud crab gut - <i>Scylla paramamosain</i>	SILVA Ref	
GAFS01005998.1.1256	Worms & Crustaceans	Crayfish - <i>Astacus leptodactylus</i>	SILVA Ref	
EU246820.1.1408	Worms & Crustaceans	Hermit crab gut - <i>Calcinus obscurus</i>	SILVA Ref	
EU646197.1.1267	Worms & Crustaceans	Woodlouse midgut hepatopancreas - <i>Tylos europaeus</i>	SILVA Ref	
04cce8908ae1d6087c5aa6a5f55015e2	Worms & Crustaceans		ASV	This study-Inner Shell
AB980074.1.1450	Worms & Crustaceans	Deep-sea crab gut - <i>Shinkaia crosnieri</i>	SILVA Ref	
KM213010.1.1485	Marine Invertebrates	Gastropod gills - <i>Cyathernia naticoides</i>	SILVA Ref	
42a6a72d9479e34df095e842db7e8e85	Marine Invertebrates		ASV	This study-Gut, This study-Hemolymph
2feb36895c21f6b050a7c39618152daa	Marine Invertebrates		ASV	This study-Gut
83e4cd8399a192c9493c0fdb75f4adfd	Marine Invertebrates		ASV	Pierce-Gut
d457c3782498adda04571e6bb87612a9	Marine Invertebrates		ASV	Pierce-Gut
0b56c037731484c9a0e20df61475122c	Marine Invertebrates		ASV	Pierce-Gut

LT716015.1.1523	Marine Invertebrates	Mycoplasma todarodis (Todarodes sagittatus)	SILVA Ref	
LT716014.1.1525	Marine Invertebrates	Mycoplasma marinum (Octopus vulgaris)	SILVA Ref	
c2a9e7814ff595604b5a3faccbe0eb13	Marine Invertebrates		ASV	Pierce-Gut
DQ340199.1.1438	Marine Invertebrates	Mudsucker gut - Gillichthys mirabilis	SILVA Ref	
KX431272.1.1480	Marine Invertebrates	Fish gut - Nibea sp.	SILVA Ref	
MG563382.1.1425	Marine Invertebrates	Bearded dragon lung - Pogona vitticeps	SILVA Ref	
L24105.1.1435	Marine Invertebrates	Mycoplasma gallinarum	SILVA Ref	
MH539050.1.1383	Marine Invertebrates	Mycoplasma gallinarum	SILVA Ref	
MH538983.1.1424	Marine Invertebrates	Mycoplasma gallinarum	SILVA Ref	
GBAQ01032348.1.1419	Marine Invertebrates	Sea hare hepatopancreas - Aplysia californica	SILVA Ref	
1c60fe87d8fc2f0c54354f62954d6c96	Marine Invertebrates		ASV	Pierce-Gut
GU070687.1.1438	Marine Invertebrates	Abalone gut - Haliotis diversicolor	SILVA Ref	
JF521605.1.1439	Marine Invertebrates	Green alga - Bryopsis sp.	SILVA Ref	
HE648941.1.1402	Marine Invertebrates	Green alga - Bryopsis sp.	SILVA Ref	
MH424685.1.1506	Marine Invertebrates	Abalone gut - Haliotis diversicolor	SILVA Ref	
HQ393436.1.1436	Marine Invertebrates	Abalone gut - Haliotis diversicolor	SILVA Ref	
8e2fb29be45701af1755f8854ab60c1b	Marine Invertebrates		ASV	This study-Inner Shell
21924c609714c24361cf08ec1f8dc3ec	Marine Invertebrates		ASV	Pierce-Gut
HQ393430.1.1439	Marine Invertebrates	Abalone gut - Haliotis diversicolor	SILVA Ref	
HQ393431.1.1439	Marine Invertebrates	Abalone gut - Haliotis diversicolor	SILVA Ref	
HQ393422.1.1439	Marine Invertebrates	Abalone gut - Haliotis diversicolor	SILVA Ref	
HQ393434.1.1439	Marine Invertebrates	Abalone gut - Haliotis diversicolor	SILVA Ref	
KY048453.1.1299	Marine Invertebrates	Gastropod gut - Benedictia baicalensis	SILVA Ref	
4f4ec76f8a81f1ae284e21008885e2c9	Marine Invertebrates		ASV	This study-Gut, This study-Gill, This study-Hemolymph, This study-Inner Shell
2cbf7629d429c942302155869e9c7b42	Marine Invertebrates		ASV	This study-Gut, This study-Gill, This study-Mantle, This study-Hemolymph, This study-Pallial Fluid, This study-Inner Shell

362e4d31693840baffae2a07eece3fcd	Marine Invertebrates		ASV	This study-Gut, This study-Gill, This study-Mantle, This study-Hemolymph, This study-Pallial Fluid, This study-Inner Shell
9a3d3bbd5f3e5ff9c29498585132c05b	Marine Invertebrates		ASV	This study-Gut
eb07e8dd3a79ae10dfb68c23dea60b03	Marine Invertebrates		ASV	Pierce-Gut
8960af3413fbfecc97f49105df5a63d7	Marine Invertebrates		ASV	Pierce-Gut
129b093af14309adefb5f152a8446235	Marine Invertebrates		ASV	Pierce-Gut
47647d3d724f76e01f9572f6dfb8a27c	Marine Invertebrates		ASV	Pierce-Gut
e22a8bdafdeb9b7cc2fa03484e9ebe8c	Marine Invertebrates		ASV	Pierce-Gut
e245e1af8ee7d14d044c9c84e58bebea	Marine Invertebrates		ASV	This study-Gut, This study-Gill, This study-Hemolymph
8f7c0d43f4fb3a063e80feae5b722698	Marine Invertebrates		ASV	This study-Gut, This study-Gill, This study-Mantle, This study-Hemolymph, This study-Pallial Fluid, This study-Inner Shell
6b4094dfedd3b8cc53f6927d2d8b7aed	Marine Invertebrates		ASV	This study-Gut, This study-Gill, This study-Mantle, This study-Hemolymph, This study-Pallial Fluid, This study-Inner Shell
af330ba2b893323765865d09486b21b9	Marine Invertebrates		ASV	This study-Gut
34b4a4b6c7f1262d395df27972ba3c18	Marine Invertebrates		ASV	This study-Gut, This study-Gill, This study-Hemolymph, This study-Inner Shell
dd496bf86e610c9c63b05f3f48fe7606	Marine Invertebrates		ASV	This study-Gut, This study-Gill, This study-Hemolymph, This study-Pallial Fluid, This study-Inner Shell
d3820332a5bbe58bdccfd5f33b51446	Marine Invertebrates		ASV	This study-Gut, This study-Gill, This study-Mantle, This study-Hemolymph, This study-Pallial Fluid, This study-Inner Shell
e046e9603a2c6c7c1b2e86f0093384a1	Marine Invertebrates		ASV	This study-Mantle
f4d7fc831aa034de18f61eaf4c7dcada	Marine Invertebrates		ASV	This study-Gut
0b1446b54dde1f608788e18fdaaa3052	Marine Invertebrates		ASV	This study-Inner Shell
1ce770523b3a4cf7164a5b4802cd39f	Marine Invertebrates		ASV	Pierce-Gut
a5d3b17bdc2815835a0ce577ab421ec	Marine Invertebrates		ASV	Pierce-Gut
4bcd22e587c7872c5368e3b4f910be21	Marine Invertebrates		ASV	Pierce-Gut
b7dba329a50781a9f23132c25d8708b9	Marine Invertebrates		ASV	Pierce-Gut
b7931f775f4cbf41d69b7b14ee21ba13	Marine Invertebrates		ASV	Pierce-Gut
8d79337500acff30feacab4a23dbc840	Marine Invertebrates		ASV	Pierce-Gut
FM995178.1.1493	Marine Invertebrates	Sydney rock oyster gut - <i>Saccostrea glomerata</i>	SILVA Ref	
72e3d7eb5eea3444aa6bf645a4a05d71	Marine Invertebrates		ASV	This study-Gut

8950b83f17c010cbb4d8d58d6ac1900d	Marine Invertebrates		ASV	This study-Gut, This study-Gill, This study-Mantle, This study-Hemolymph, This study-Pallial Fluid, This study-Inner Shell
4435068650d63043eede5a03c52f91d0	Marine Invertebrates		ASV	This study-Gut, This study-Gill, This study-Hemolymph, This study-Pallial Fluid, This study-Inner Shell
ed2d558fd100a22e357759594baacab8	Marine Invertebrates		ASV	This study-Gut, This study-Gill, This study-Mantle, This study-Pallial Fluid, This study-Inner Shell
59b1fe0f695032a98908b04bc70ac9f1	Marine Invertebrates		ASV	Pierce-Water
d69f165cf236785419047f1e1b9fd9e4	Marine Invertebrates		ASV	Pierce-Gut
7798bda0bcf5c98b8540624e48d3e878	Marine Invertebrates		ASV	Pierce-Gut
6f59d2901a8dca363123ed13163692d6	Marine Invertebrates		ASV	Pierce-Gut
JQ269320.1.1479	Marine Invertebrates	Biofilm on antifouling paint	SILVA Ref	
243095f5b3e79906a7b7fb20345a4ee6	Marine Invertebrates		ASV	Pierce-Water
DQ917875.1.1417	Marine Invertebrates	Coral - Muricea elongata	SILVA Ref	
DQ917865.1.1457	Marine Invertebrates	Coral - Muricea elongata	SILVA Ref	
DQ917898.1.1513	Marine Invertebrates	Coral - Muricea elongata	SILVA Ref	
DQ395544.1.1401	Marine Invertebrates	Deep-sea octacoral	SILVA Ref	
DQ395564.1.1519	Marine Invertebrates	Deep-sea octacoral	SILVA Ref	
DQ395752.1.1309	Marine Invertebrates	Deep-sea octacoral	SILVA Ref	
DQ395829.1.1528	Marine Invertebrates	Deep-sea octacoral	SILVA Ref	
DQ395851.1.1513	Marine Invertebrates	Deep-sea octacoral	SILVA Ref	
NC_004567.2:502863-504433	Outgroup (Firmicutes)	Lactobacillus plantarum WCFS1	Outgroup	
NC_007795.1:448585-450606	Outgroup (Firmicutes)	Staphylococcus aureus subsp. aureus NCTC 8325	Outgroup	
NC_004668.1:248466-249987	Outgroup (Firmicutes)	Enterococcus faecalis V583	Outgroup	
NC_003210.1:243556-245041	Outgroup (Firmicutes)	Listeria monocytogenes EGD-e	Outgroup	

Table S8. Summary of primers and probes used in this study. The locus-specific part of the primers 515F w/ Overhang and 806R w/ Overhang are underlined while the rest of the sequence contains a sequencing adapter.

Name	Type	Sequence
515F w/ Overhang	Primer	5' TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGTGYCAGCMGCCGCGGTAA
806R w/ Overhang	Primer	5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGGACTACNVGGGTWCTAAT
Perkinsus_F	Primer	5' CGCCTGTGAGTATCTCTCGA
Perkinsus_R	Primer	5' GTTGAAGAGAAGAATCGCGTGAT
MSX/SSO_F	Primer	5' TSGRGATTACCYSGCCTTC
MSX/SSO_R	Primer	5' ACAGGTCAGTGATGCCCTTAG
MSX	Probe	5' SHEX/TTGCACACGAGTTCAACCTTGCCTG/3BHQ1
SSO	Probe	5' 5Cy5/AATGACCCAGTCAGCGGGCCGA/3BHQ1
Perkinsus	Probe	5' 56-FAM/CGCAAACCTCGACTGTGTTGTGGTG/3BHQ1

Table S9. Strain names and NCBI RefSeq assembly accessions for a set of manually curated draft genomes of oyster derived isolates.

Strain	RefSeq Assembly Accession
<i>Vibrio tubiashii</i> subsp. <i>europaeus</i>	GCF_001695575.1
<i>Vibrio bivalvicida</i> 605T	GCF_001399455.2
<i>Aliiroseovarius crassostreae</i> CV919-312	GCF_001307765.1
<i>Vibrio coralliilyticus</i> RE22	GCF_001297935.1
<i>Bacillus pumilus</i> RI06-95	GCF_001183525.1
<i>Vibrio</i> sp. J2-17	GCF_001244075.1
<i>Vibrio</i> sp. J2-31	GCF_001243385.1
<i>Vibrio crassostreae</i> LGP107	GCF_001048575.1
<i>Vibrio</i> sp. J2-4	GCF_001243675.1
<i>Vibrio</i> sp. J2-29	GCF_001048655.1
<i>Stenotrophomonas maltophilia</i> MF89	GCF_000455685.1
<i>Pseudomonas stutzeri</i> MF28	GCF_000455665.1
<i>Vibrio</i> sp. J2-12	GCF_001243825.1
<i>Vibrio</i> sp. J2-15	GCF_001244175.1
<i>Vibrio</i> sp. J2-1	GCF_001048595.1
<i>Vibrio crassostreae</i> LGP7	GCF_001048535.1
<i>Vibrio crassostreae</i> J5-4	GCF_001368875.1
<i>Vibrio</i> sp. J2-3	GCF_001243545.1
<i>Vibrio</i> sp. J2-26	GCF_001262155.1
<i>Vibrio</i> sp. J2-8	GCF_001048615.1
<i>Vibrio</i> sp. J2-6	GCF_001243735.1
<i>Vibrio crassostreae</i> J5-20	GCF_001048515.1
<i>Vibrio crassostreae</i> J5-19	GCF_001048495.1
<i>Vibrio crassostreae</i> J5-5	GCF_001368895.1
<i>Vibrio crassostreae</i> LGP8	GCF_001048555.1
<i>Microbacterium maritipicum</i> MF109	GCF_000455825.1
<i>Vibrio crassostreae</i> J5-15	GCF_001048475.1
<i>Pseudomonas aeruginosa</i> WC55	GCF_000455705.1