



Article

# The Purple Sea Urchin *Strongylocentrotus purpuratus* Demonstrates a Compartmentalization of Gut Bacterial Microbiota, Predictive Functional Attributes, and Taxonomic Co-Occurrence

Joseph A. Hakim <sup>1,\*</sup>, Julie B. Schram <sup>2</sup>, Aaron W. E. Galloway <sup>2</sup>, Casey D. Morrow <sup>3</sup>, Michael R. Crowley <sup>4</sup>, Stephen A. Watts <sup>1</sup> and Asim K. Bej <sup>1,\*</sup>

<sup>1</sup> Department of Biology, University of Alabama at Birmingham, 1300 University Blvd., Birmingham, AL 35294, USA; sawatts@uab.edu

<sup>2</sup> Oregon Institute of Marine Biology, University of Oregon, 63466 Boat Basin Rd, Charleston, OR 97420, USA; jschram@uoregon.edu (J.B.S.); agallow3@uoregon.edu (A.W.E.G.)

<sup>3</sup> Department of Cell, Developmental and Integrative Biology, University of Alabama at Birmingham, 1918 University Blvd., Birmingham, AL 35294, USA; caseym@uab.edu

<sup>4</sup> Department of Genetics, Heflin Center Genomics Core, School of Medicine, University of Alabama at Birmingham, 705 South 20th Street, Birmingham, AL 35294, USA; mcrowley@uab.edu

\* Correspondence: joe21@uab.edu (J.A.H.); abej@uab.edu (A.K.B.); Tel.: +1-(205)-934-9857 (A.K.B.)

Received: 8 November 2018; Accepted: 21 January 2019; Published: 26 January 2019



**Abstract:** The sea urchin *Strongylocentrotus purpuratus* (order Camarodonta, family Strongylocentrotidae) can be found dominating low intertidal pool biomass on the southern coast of Oregon, USA. In this case study, three adult sea urchins were collected from their shared intertidal pool, and the bacteriome of their pharynx, gut tissue, and gut digesta, including their tide pool water and algae, was determined using targeted high-throughput sequencing (HTS) of the 16S rRNA genes and bioinformatics tools. Overall, the gut tissue demonstrated *Arcobacter* and *Sulfurimonas* (Epsilonproteobacteria) to be abundant, whereas the gut digesta was dominated by *Psychromonas* (Gammaproteobacteria), *Propionigenium* (Fusobacteria), and Flavobacteriales (Bacteroidetes). Alpha and beta diversity analyses indicated low species richness and distinct microbial communities comprising the gut tissue and digesta, while the pharynx tissue had higher richness, more closely resembling the water microbiota. Predicted functional profiles showed Kyoto Encyclopedia of Genes and Genomes (KEGG) Level-2 categories of energy metabolism, membrane transport, cell motility, and signal transduction in the gut tissue, and the gut digesta represented amino acid, carbohydrate, vitamin and cofactor metabolisms, and replication and repair. Co-occurrence network analysis showed the potential relationships and key taxa, such as the highly abundant *Arcobacter* and *Propionigenium*, influencing population patterns and taxonomic organization between the gut tissue and digesta. These results demonstrate a trend of microbial community integration, allocation, predicted metabolic roles, and taxonomic co-occurrence patterns in the *S. purpuratus* gut ecosystem.

**Keywords:** Illumina; high-throughput sequencing (HTS), bacteriome; PICRUSt; 16S rRNA gene; PhyloToAST; CoNet; QIIME; LEfSe; KEGG

## 1. Introduction

The purple sea urchin *Strongylocentrotus purpuratus* (order Camarodonta, family Strongylocentrotidae) inhabits the rocky tide pools along the North-East Pacific from Alaska

to Baja Mexico. *S. purpuratus* is primarily herbivorous, which tempers the growth of marine vegetation and plays an important role in shaping the dynamic population patterns in their marine ecosystem [1–6]. The low intertidal tide pools on the southern Oregon coast are dominated by *S. purpuratus* and are interspersed with mosaics of tufted algae and invertebrate assemblages representing multiple phyla [7]. The microhabitats of these tide pools are influenced by the feeding activity of the inhabiting sea urchins [7]. The sea urchins present unique digestive physiology in a straightforward model and offer an evolutionary context to fundamental biological and physiological processes occurring in higher deuterostome organisms [8]. In general, the pharynx is enclosed within the Aristotle's Lantern, which is a pentamerally symmetric mastication apparatus of five tooth-like structures that assist in scraping and releasing intracellular nutrients from algae [9]. The pharynx tissue contains specialized mucus cells that contribute to the formation of a mucous envelope of ingested feed [10], forming individual pellets of gut digesta [11]. This gut digesta pellet formation has been considered an evolved digestive strategy of this organism, likely as a result of water flow dynamics in the gut lumen environment [12]. The role of gut bacteria in host health and digestion have been of interest beginning with the work of Lasker and Giese [9], who isolated gut bacteria from the gut digesta of *S. purpuratus*, showing the potential for these bacteria to digest polysaccharides from algal sources. In a separate study, bacteria isolated from the sea urchins *S. intermedius* and *S. nudus* demonstrated a similar algynolytic activity [13]. The importance of gut bacteria in sea urchin host health was further supported in *S. droebachiensis*, in which microbial suppression through antibiotics showed a reduced capacity for host incorporation of essential amino acids [14].

The microbial communities of the sea urchin gut pellets also play an important role in the biogeochemical cycles of the marine environment. In previous studies, it has been shown that the microbial community composition and their metabolic processes in the gut digesta remains stable following egestion into the environment [15–17]. For example, studies examining the chemical composition of *S. droebachiensis* egesta through flash combustion have shown increased in lipid, nitrogen, and organic carbon, and decreases in the carbon: nitrogen ratio, indicating the metabolic importance of the bacterial communities in the degradation and transformation of the contents within the pellets into a nutrient-rich food source for nearby marine organisms [17–20]. Additionally, it has been suggested that urchin gut microbiota are responsible for differences in algal digestion and synthesis of essential long chain fatty acids in both *S. purpuratus* and *S. droebachiensis* [20]. The pelleted egesta has also been considered as a mode for the dispersion of sea urchin gut microbiota into their environment [17,21]. Most of the early studies of the potential role of microbial communities in digestive processes of sea urchin gut ecosystem were conducted by culture-dependent methods [9]. However, recent advancement of the culture-independent method of high-throughput sequencing (HTS) of 16S rRNA genes from the metacommunity DNA has been shown to provide gut microbial community composition with high taxonomic coverage, including their potential metabolic functions [22,23]. Recently, the application of HTS on the V4 hypervariable segment of the 16S rRNA gene of the gut bacteriome of *Lytechinus variegatus* from the U.S. Gulf of Mexico [15,16,24] demonstrated distinct microbial community compositions between the gut tissue and gut digesta. Specifically, representative taxa from class Epsilonproteobacteria (assigned as *Arcobacter* / *Sulfuricum* through the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST)) dominated the gut tissue, whereas Gammaproteobacteria (namely *Vibrio*) were heightened in the digesta [15]. Additionally, predictive functional profiling of these compartmentalized microbial communities showed energy metabolisms such as oxidative phosphorylation, carbon fixation, nitrogen, methane, and sulfur metabolisms to be heightened in the gut tissue, compared to carbohydrate, amino acid, and lipid metabolisms in the digesta [16].

Such HTS technology and bioinformatics analyses applied to the gut ecosystem of the naturally occurring sea urchin *S. purpuratus* can help establish a comprehensive microbial community composition and provide crucial information into the gut bacterial taxa and likely functions performed as they relate to host health and digestion. In this study, we have elaborated the microbial profiles of

the gut tissue, pharynx tissue, and mucous enveloped gut digesta of *S. purpuratus* collected from their natural rocky tide pool habitat on the coastline of Oregon, USA. In addition, samples of the tide pool seawater and adjacent algal community were collected and also analyzed for comparison with the gut tissue and digesta. We used HTS and bioinformatics tools to analyze the community composition, patterns of microbial taxa allocation in the gut environment, and the predicted metabolic functions of the bacterial microbiota in the gut ecosystem. These data were further refined using Phylogenetic Tools for Analysis of Species-level Taxa (PhyloToAST v1.4.0) [25] alongside Quantitative Insights into Microbial Ecology (QIIME v1.9.1) [26] to condense redundant taxonomic groups. This allowed increased resolution of microbial taxonomic groups to the species level and enhanced beta diversity inference. Additionally, the keystone taxa (herein “key” taxa) of the gut ecosystem were elaborated through topological analysis of Co-occurrence Network inferences (CoNet v1.1.1) [27–29] based on criteria described in Berry and Widder [30,31]. The results of this baseline case study demonstrate the microbial composition and associated functional capacity within the compartmentalized gut system of this evolutionarily and ecologically significant purple *S. purpuratus* sea urchin species.

## 2. Materials and Methods

### 2.1. Collection and Sample Preparation of *S. Purpuratus*

Adult *S. purpuratus* sea urchins (UR;  $n = 3$ ) were collected from within the same natural rocky tide pool habitat at Cape Arago, Oregon ( $43^{\circ}18'14.3''N$   $124^{\circ}24'05.1''W$ ) in September 2016, from within a  $1\text{ m}^2$  sampling plot (Figure 1), under permit is: #20366 (Oregon Department of Fish and Wildlife). Sea urchins were measured and sexed, which was followed by tissue dissection performed at the Oregon Institute of Marine Biology (OIMB) in Coos County, Oregon. For each sea urchin, an incision was made into the test area surrounding the Aristotle’s Lantern mastication structure using sterilized instruments, and the test was cut radially to expose the internal digestive tissue. The pharynx, which was enclosed by the Aristotle’s Lantern, was separated from the gut tissue and collected. The remaining digestive tissue (gut tissue) was gently rinsed with sterile phosphate buffered saline water (1x PBS, pH 7.4) (Fisher Scientific, Hampton, NH, USA), and the contents (gut digesta) were collected. The whole gut tissue was collected separately from the voided gut digesta. Replicate seawater samples (water;  $n = 3$ ) (1 L) from each tide pool was vacuum filtered separately through  $0.22\ \mu\text{m}$  filter paper (EMD Millipore Corporation, Danvers, MA, USA). The grazed-upon algal communities (algae;  $n = 3$ ) immediately surrounding the sea urchins were also collected as the general food source and used in this study. A total of 15 samples (pharynx,  $n = 3$ ; gut tissue,  $n = 3$ ; gut digesta,  $n = 3$ ; water,  $n = 3$ ; and algae,  $n = 3$ ) were placed into 95% (v/v) ethanol [32], flash frozen in liquid nitrogen, and shipped to the University of Alabama at Birmingham (UAB), where they were preserved at  $-20\ ^{\circ}\text{C}$  until used. Research performed under the Institutional Animal Care and Use Committee (IACUC-10043).



**Figure 1.** Sample collection site of *S. purpuratus* (purple sea urchins) from their natural rocky tide pool habitat along the coast of Oregon ( $43^{\circ}18'14.3''\text{N}$   $124^{\circ}24'05.1''\text{W}$ ). (A) Satellite image of the collection site (red marker) provided through Google Earth Pro (v.7.3.2.5491) (Data SIO, NOAA, US Navy NGA, GEBCO, Image Landsat/Copernicus; US Dept. of State Geographer; image date: December 2015). (B) Overview of the tide pool collection site (labeled as tide pool 1) showing naturally occurring sea urchins. (C) Sea urchin congregates with the algal food source in view. Photographs by J.B Schram.

## 2.2. Community DNA Extraction, Illumina MiSeq Sample Preparation, and High-Throughput Sequencing

The metacommunity DNA from each sample was purified using the Fecal DNA isolation kit (Zymo Research, Irvine, CA, USA; catalog no. D6010), and an amplicon library of the metacommunity V4 hypervariable region (V4) of the 16S rRNA gene was created using uniquely barcoded DNA oligonucleotide primers adapted from the Earth Microbiome Project ([www.earthmicrobiome.org](http://www.earthmicrobiome.org)) [33–35]. These primers consisted of the upstream nucleotide sequence for hybridization to the Illumina MiSeq flow-cell surface (underlined), a “pad” region (italicized), and a “linker” region (bolded). The forward primer (515F) for the V4 segment was: 5'-AATGATACGGCGACCACCGAGATCTACACTATGGTAATTGTGTGCCAGCMGCCGCGGTAA-3'. The reverse primer for the V4 segment (modified from 806R) also included a unique barcode (6 N's) region and was as follows: 5'-CAAGAGAAGACGGCATAACGAGATNNNNNNAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT-3' (Eurofins Genomics, Inc., Huntsville, AL, USA) [35,36]. Polymerase chain reaction (PCR) amplification was performed using the LongAmp Taq PCR Kit (New England Biolabs, Ipswich, MA, USA; catalog no. E5200S) at a total reaction volume of 50  $\mu\text{L}$  with the following reagents: 10  $\mu\text{L}$  of 5 $\times$  Reaction Buffer; 1.5  $\mu\text{L}$  of each dNTPs (200  $\mu\text{M}$ ); 2  $\mu\text{L}$  of each oligonucleotide primer (1.5  $\mu\text{M}$ ); 1.5  $\mu\text{L}$  of LongAmp<sup>®</sup>enzyme (5 U); 30  $\mu\text{L}$  of template DNA (2–5 ng/ $\mu\text{L}$ ); and 3  $\mu\text{L}$  of sterile H<sub>2</sub>O. The PCR proceeded with an initial denaturation at 94  $^{\circ}\text{C}$  for 1 min followed by 32 cycles of amplification of which each cycle consisted of denaturation at 94  $^{\circ}\text{C}$  for 30 sec, primer annealing at 50  $^{\circ}\text{C}$  for 1 min, and primer extension at 65  $^{\circ}\text{C}$  for 1 min, followed by the final extension at 65  $^{\circ}\text{C}$  for 3 min and a final hold at 4  $^{\circ}\text{C}$ . An amplicon fragment of approximately 380 bases was visualized through an ultraviolet (UV) transilluminator (Photodyne, Inc., Los Angeles, CA, USA) and excised with a sterile scalpel following electrophoresis through a 1.0% (w/v) Tris-borate-EDTA (TBE)/agarose gel [37]. The excised DNA fragments were purified using the QIAquick Gel Extraction Kit (Qiagen Inc., Venlo, Limburg; catalog no. 28704). PicoGreen dye (Life Technologies, Grand Island, NY, USA) was used to quantify each sample to adjust the concentration to 4 nM [35]. HTS was performed using the Illumina MiSeq platform [35,36], incorporating the 250 base paired-end kits from Illumina specific to the V4 region of the 16S rRNA gene.

## 2.3. Quality Assessment and Filtering

The raw sequence reads generated by HTS on the Illumina MiSeq platform were demultiplexed and converted to FASTQ format [38]. The read quality was evaluated using FastQC [39], and quality reads with 80% of bases at Q score >33 were retained for downstream analysis using the “fastx\_trimmer” command from the FASTX Toolkit [35,40]. Then, the paired-ends were merged using USEARCH [41],

and pairs with <50 base overlap and/or over 20 mismatching nucleotides were filtered. Read quality was again assessed after filtering using FASTQC, chimeric sequences were identified and removed using USEARCH [41]. Additionally, with the newly established bioinformatics techniques presented in the QIIME2 package (v2018.11) [42], an alternative approach to filtering and merging the paired-end sequence data was implemented. To do this, a “denoising” strategy was used based on the Poisson distribution through the Divisive Amplicon Denoising Algorithm program (DADA2, v1.10) [43,44]. This was performed utilizing the “qiime dada2 denoise-paired” module on the demultiplexed sequence data, with a truncation set at 250 bases for the forward and reverse reads. The sequence reads corresponding to each sample of this study have been deposited in NCBI SRA for public access (Bioproject number PRJNA504890).

#### 2.4. Taxonomic Distribution

The resultant quality assessed sequence files were processed using QIIME (v1.9.1) [26] along with PhylotoAST (v1.4.0) [25] to condense redundant operational taxonomic units (OTUs) [45]. First, OTUs were selected at a 97% sequence similarity threshold using the default UCLUST algorithm option in QIIME (v1.9.1) [41]. Representative OTU sequences were then selected using the “most\_abundant” option, and taxonomy was assigned to the representative sequences at a 60% confidence threshold using the Ribosomal Database Project (RDP) classifier [46], trained with the GreenGenes reference database (v13.8) [47,48]. At this stage, OTUs occurring at less than 0.0005% average abundance across all samples in the study were filtered [49–53]. Then by using the PhyloToAST (v1.4.0) workflow, the species-level resolution was enhanced using the “assign\_taxonomy\_by\_blast\_result.py” command to assign taxonomy through BLAST [45] to the GreenGenes (v13.8) database, and redundant OTUs were merged through the “condense\_workflow.py” command [25]. Variation in the read-depth was accounted for by subsampling of the condensed OTU table using both the median and minimum read count values across all samples as described in de Carcer et al. [54] through the “single\_rarefaction.py” command in QIIME (v1.9.1), and both subsampled OTU tables were assessed for downstream analysis. Additionally, for the top 100 taxa determined in the rarefied OTU table, the representative sequences were extracted and aligned to multiple databases using the SILVA ACT: Alignment, Classification and Tree Service ([www.arb-silva.de/aligner](http://www.arb-silva.de/aligner)) [55]. For this analysis, the SSU (Small Sub-Unit) category was selected, and a minimum similarity identity was set to 0.9, with 20 neighbors per query sequence. Sequences below an identity threshold of 70% were discarded. For taxonomic identification, the least common ancestor (LCA) method was used, and the databases selected included GreenGenes [47,48], Ribosomal Database Project (RDP) [56], and SILVA [57]. Lastly, for the alternative QIIME2 (v2018.11) method, the denoised and merged sequence data was used to generate representative sequences with the “qiime feature-table tabulate-seqs” command. These representative sequences, herein referred as amplicon sequence variants (ASVs), were assigned taxonomic identities through the “qiime feature-classifier” command utilizing the “classify-sklearn” option [58] against the GreenGenes (v13.8) database.

#### 2.5. Alpha Diversity

The PhyloToAST (v1.4.0) condensed and minimum-count subsampled OTU table (herein, rarefied OTU table) was used to determine the taxonomic distribution and alpha diversity metrics of each sample. The rarefied OTU table was merged according to biological replicates and used to create the relative abundance graph of phyla represented at >1% abundant, as well as the top 100 most resolved taxonomic identities across all sample groups using Microsoft Excel Software (Seattle, WA, USA). Taxa represented at >1% in the gut system (gut tissue and digesta) were also visualized, with standard deviations calculated through STAMP (v2.1.3) [59]. Shannon [60–62], and Simpson [61,63] diversity measurements were determined through the “alpha\_diversity.py” command in QIIME (v1.9.1). These values were plotted as a kernel density estimator-smoothed histogram using the “diversity.py” command through PhyloToAST (v1.4.0), to show both the diversity value and the range

of underlying data points (density) for each sample group. Kruskal–Wallis H-tests were performed for the five groups to show the alpha diversity variation between groups at a significance value of  $p = 0.1$  [64].

## 2.6. Beta Diversity

For beta diversity, the rarefied OTU table was used to determine the Bray–Curtis distance matrix values [65]. These values were also used to calculate significant grouping among biological replicates ( $n = 3$ ) through an analysis of similarity (ANOSIM) and multivariate analysis of variance (Adonis) of groups, both set at 999 permutations, determined through the QIIME (v.1.9.1) “compare\_categories.py” module utilizing the Vegan (v2.4.3) R package implementation of the statistical methods [66–68]. Additionally, this ANOSIM and Adonis analysis was performed on each OTU table generated in this study (5 total), which included the unfiltered OTU table, filtered OTU table (<0.0005%), PhyloToAST (v1.4.0) condensed OTU table, condensed median, and condensed minimum subsampled OTU tables. Visualization of beta diversity trends was performed using Plymouth Routines in Multivariate Ecological Research (PRIMER-6) software (Primer-E Ltd, Plymouth Marine Laboratory, Plymouth UK, v6.1.2) [67]. In PRIMER-6, a 2D multidimensional scale (MDS) plot was generated using the Bray–Curtis distance matrices, to show variation between each sample, along with an overlay of Bray–Curtis similarity values [69]. A dendrogram was also generated based on clustering by group average [69]. A 2D MDS and dendrogram cluster analysis was also performed on the top 100 OTUs and the remaining rare OTUs, to show the contributions of both the heightened and rare taxa to the observed sample community diversity and cluster patterns. In addition to the Bray–Curtis based analyses described above, the rarefied OTU table was used to determine the weighted and unweighted Unifrac distances [70], which was calculated through the “beta\_diversity\_through\_plots.py” module of QIIME (v1.9.1), and used to generate the 3D principle coordinates analysis (PCoA) plots through the “PCoA.py” command in PhyloToAST (v1.4.0). These values were also used to calculate the ANOSIM and Adonis metrics for group analyses as previously described, and uploaded into PRIMER-6 to generate the dendrogram based on group average. Heatmap analysis was performed using the rarefied OTU table in R (v3.3.2), incorporating the heatmap.2 function from gplots (v3.0.1) package [71]. In brief, the associated sample group dendrogram was created through the Vegan (v2.4.3) package [68] using the Bray–Curtis distance metric of the grouped biological replicate count data and clustered according to the group average algorithm. Microbial taxa represented at <1% of the total dataset were filtered from the heatmap. A color palette was selected using the RColorBrewer package [72], and the relative abundances were shown for each taxon across all sample groups (black bar lines). Linear discriminant analysis (LDA) effect size (LEfSe) analysis was used to determine the taxa contributing to the effect size between the compartmentalized gut microbial communities of the gut tissue ( $n = 3$ ) and digesta ( $n = 3$ ) [73]. This analysis was performed through the Hutlab Galaxy web application ([huttenhower.sph.harvard.edu/galaxy/](http://huttenhower.sph.harvard.edu/galaxy/)), and incorporated the non-parametric Kruskal-Wallis sum-rank test for significant differential abundance set at a significance of  $p = 0.05$  [64], followed by LDA to estimate effect size at  $\log(10)$  values [73,74]. The results were plotted to show those taxa that demonstrated an LDA of  $\pm 3$  for effect size.

## 2.7. Predicted Functional Analysis

The functional capacity associated with the microbial communities of the gut tissue and digesta was determined using the Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt v1.1.2) package [75] and analyzed in STAMP (v2.1.3) [59]. For this analysis, an OTU table was constructed by the “pick\_closed\_reference\_otus.py” strategy to ensure representative taxonomic information through the GreenGenes (v13.8) database as suggested in PICRUSt [47,48,75]. The resultant OTU table was normalized by copy number, and Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthology (KO) Ids were predicted along with the weighted Nearest Sequenced Taxon Index (NSTI) values for the confidence of predictions using the “predict\_metagenomes.py”

command. The assigned functional categories were then collapsed into levels (KEGG-Level-2, and 3) through the “categorize\_by\_function.py” command. The KEGG-Level-2 and 3 profiles were uploaded into STAMP (v2.1.3) for two-group scatter plot analysis, to determine the metabolic categories that are preferentially enriched in each group at two levels of hierarchical classification. Additionally, LEfSe analysis [73] was performed on the KO Ids using an LDA score of  $\pm 2.4$ , again utilizing the non-parametric Kruskal–Wallis sum-rank test for significant differential abundance set at a significance of  $p = 0.05$  [64] and LDA for effect size using  $\log(10)$  values [73,74], to demonstrate those categories contributing most to functional profile dissimilarity.

### 2.8. Co-Occurrence Analysis of Microbial Taxa

Significant co-occurrence patterns occurring between the microbial communities of the gut tissue and the gut digesta were determined using Co-occurrence Network inference (CoNet v1.1.1) [27–29]. To do this, the rarefied OTU data were uploaded into Cytoscape (v3.6.0) [76] through the CoNet (v1.1.1) plugin with taxa assigned to sample type (gut tissue and gut digesta). Links between higher level taxa were not explored and a parent-child exclusion was applied. Taxonomic entries with a cumulative group sum of 200 and at least 2/3 of samples containing non-zero values were kept [27–29,77]. Significant co-occurrences between taxa were determined by utilizing the Pearson [78,79], Spearman [80], Bray–Curtis [65], Kullback–Leibler [81], and mutual information similarity [82], with a  $10^{-8}$  pseudo-count [27–29,77]. The 200 highest (most positive) and lowest (most negative) edges were selected and merged by the union approach using the mean value [77]. The multi-edge scores were shuffled row-wise at 100 permutations (for null distributions), followed by bootstrapping at 100 permutations (for randomizations). The  $p$ -values of the multi-edges assigned to node pairs were merged using the Brown method [83], with unstable edges filtered out, and the corrected significance value (q-value) was determined with a threshold set at  $p < 0.05$  for significance [27–29].

The final network was constructed in Cytoscape (v3.6.0) using the radial layout algorithm in the yFiles plugin (v1.0) [84], and topological parameters were determined by NetworkAnalyzer (v2.7) [85] using an undirected approach. Node sizes were scaled to their group abundance, colored according to phylum (class for Proteobacteria), and assigned a shape according to group membership (circle for gut tissue, “-gut”; diamond for gut digesta; “-dig”). The edges were scaled by the q-value and colored according to their positive (co-presence; green) and negative (co-exclusion; red) association. Based on the topological features determined through NetworkAnalyzer (v2.7), those nodes tending to have a high degree (number of edges), closeness centrality, and low betweenness centrality have been referred to as key taxa as described by Berry and Widder et al. [30] [77,86,87]. These features were plotted as a scatter plot ( $y =$  closeness centrality;  $x =$  betweenness centrality; node size is scaled to degree) through Microsoft Excel Software (Seattle, WA, USA). The top 10 nodes based on their closeness centrality values were selected as likely key taxa.

## 3. Results

### 3.1. Environmental Conditions and Sea Urchin Measurements

The tide pool location sea water conditions were determined to have a salinity of 30.19 ppt with a 7.7 pH. The dissolved oxygen content of the tide pool water was determined to be 59% and the temperature was determined to be 13.1 °C. The sea urchins of this study weighed between 38.98–53.35 g and had a mean diameter of 5.1–5.6 cm, a height of 2.6–2.8 cm, and a spheroid volume of 38.4–41.6 cm<sup>3</sup>. The sexes of the sea urchins for the study were determined as UR1 = F, UR2 = F, and UR3 = M.

### 3.2. Quality Assessment and Sample Statistics

The total sequences generated through Illumina MiSeq-based HTS of the bacterial 16S rRNA gene of the 15 samples of the study generated a total of 1,714,746 forward and reverse reads (Table 1).

Quality checking and trimming using the FASTX Toolkit, followed by merging of the forward and reverse sequences resulted in 1,249,827 total reads. Grouping of biological replicate data ( $n = 3$ ) showed the following total sequence read counts: algae (340,438), gut digesta (221,684), gut tissue (231,854), pharynx (194,640), and water (261,211). Clustering of sequences into OTUs and taxa assignment revealed a total of 44,664 unique assignments. Filtering of rare OTUs occurring at less than 0.0005% reduced the number of OTUs to 4290 unique observations across all samples. Condensing of redundant taxonomic IDs through PhyloToAST (v1.4.0) showed a total of 776 OTUs. Rarefaction of the condensed OTU table to the median read count value (77,806) and the minimum read count value (49,641) both maintained 776 unique observations. Through the alternative strategy utilizing DADA2 (v1.10) implemented in QIIME2 (v2018.11), a total of 1134 unique features were determined (ASVs) representing a total of 467,866 reads across all samples, which were subsequently assigned to 371 taxonomic identities when collapsed to the species level (data not shown).

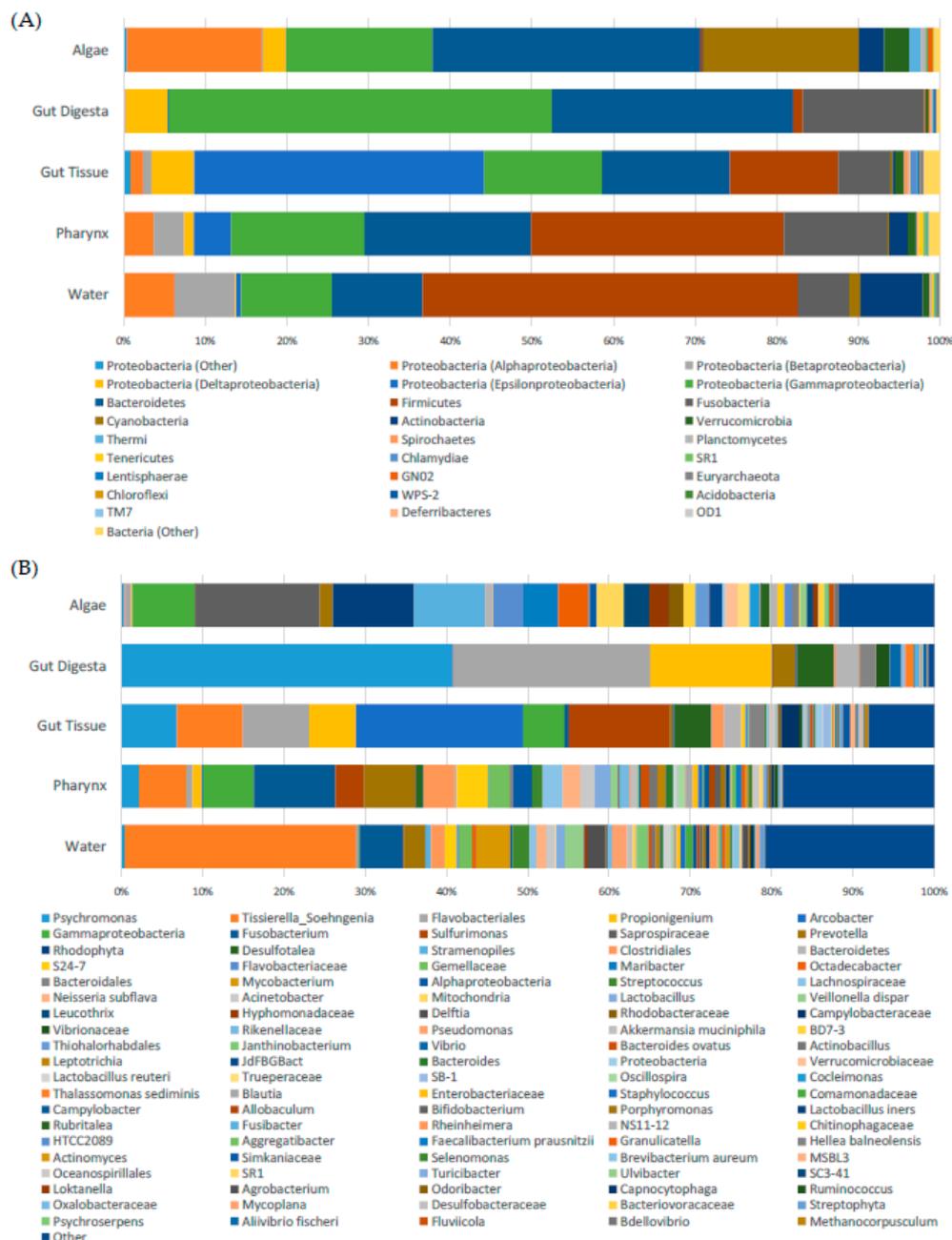
**Table 1.** Sequence reads, operational taxonomic unit (OTU) count, and alpha diversity of each sample of the study. The table shows the sequence read count (1) before and (2) after quality checking and filtering based on low quality reads using FASTX Toolkit, the (3) unique unfiltered (Unfilt.) OTU observations following chimera removal, (4) the filtered (Filt.) OTUs after removal of rare OTUs (<0.0005% abundant in all samples), (5) the resultant condensed (Cond.) OTU count following the merging of redundant taxonomic information through Phylogenetic Tools for Analysis of Species-level Taxa (PhyloToAST) (v1.4.0), (6) the OTU count following subsampling to the median (Med.) value (77,806), and (7) the minimum (Min.) value (49,641). Also shown are the (8) Shannon and (9) Simpson diversity indices corresponding to each sample determined using the condensed OTU table data.

| Sample           | Raw Reads         | Trimmed Reads     | Unfilt. OTUs   | Filt. OTUs   | Cond. OTUs  | Cond. OTUs Med. | Cond. OTUs Min. | Shannon Diversity | Simpson Diversity |
|------------------|-------------------|-------------------|----------------|--------------|-------------|-----------------|-----------------|-------------------|-------------------|
| Algae 1          | 154,561           | 121,543           | 4985           | 1345         | 268         | 259             | 243             | 5.3590            | 0.9376            |
| Algae 2          | 137,323           | 103,116           | 2973           | 855          | 226         | 220             | 204             | 4.7960            | 0.9340            |
| Algae 3          | 160,926           | 115,779           | 4745           | 1383         | 329         | 325             | 302             | 4.9626            | 0.9156            |
| Gut Digesta 1    | 76,329            | 60,871            | 2133           | 580          | 155         | 155             | 151             | 3.1936            | 0.8260            |
| Gut Digesta 2    | 74,249            | 61,267            | 2144           | 599          | 119         | 119             | 113             | 2.4760            | 0.5991            |
| Gut Digesta 3    | 123,640           | 99,546            | 3056           | 611          | 126         | 121             | 106             | 2.9289            | 0.7891            |
| Gut Tissue 1     | 128,539           | 90,412            | 4384           | 1094         | 328         | 322             | 301             | 4.1236            | 0.8440            |
| Gut Tissue 2     | 68,644            | 51,895            | 2311           | 898          | 273         | 273             | 273             | 3.8787            | 0.8436            |
| Gut Tissue 3     | 123,735           | 89,547            | 4765           | 1418         | 403         | 397             | 368             | 4.8038            | 0.9250            |
| Pharynx Tissue 1 | 107,276           | 72,954            | 5663           | 1558         | 430         | 430             | 418             | 6.0719            | 0.9642            |
| Pharynx Tissue 2 | 86,515            | 58,281            | 4417           | 1488         | 431         | 431             | 430             | 5.9389            | 0.9632            |
| Pharynx Tissue 3 | 92,987            | 63,405            | 4918           | 1489         | 402         | 402             | 397             | 6.2246            | 0.9697            |
| Water 1          | 123,154           | 81,885            | 4725           | 1679         | 504         | 504             | 487             | 5.2838            | 0.8797            |
| Water 2          | 137,044           | 93,713            | 6127           | 1087         | 403         | 400             | 386             | 5.7289            | 0.9546            |
| Water 3          | 119,824           | 85,613            | 5608           | 1406         | 400         | 400             | 377             | 4.2286            | 0.8211            |
| Summary          | total = 1,714,746 | total = 1,249,827 | total = 44,664 | total = 4290 | total = 776 | total = 776     | total = 776     | avg. = 4.6666     | avg. = 0.8778     |

### 3.3. Taxonomic Distribution across Samples

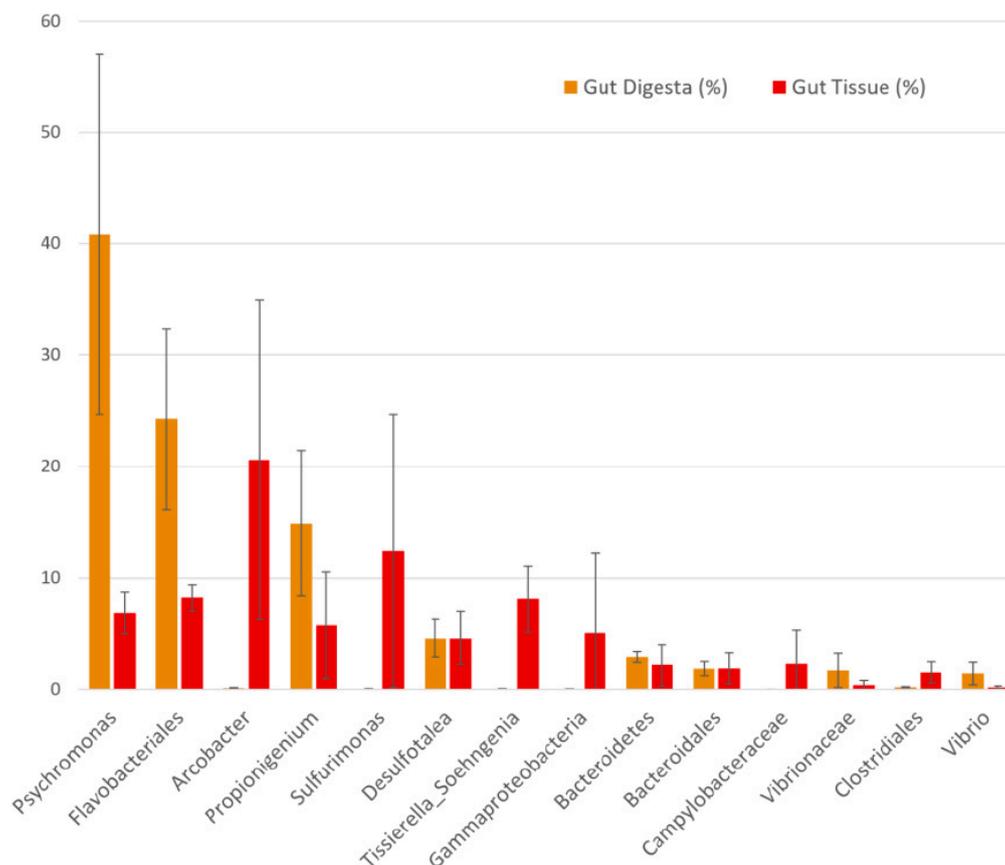
Taxonomic distribution across all samples showed the gut tissue represented a uniquely heightened amount of Epsilonproteobacteria in the order of Campylobacterales, namely family Campylobacteraceae (*Arcobacter*) (~20%) and Helicobacteraceae (*Sulfurimonas*) (~12%) as compared to the other samples of the study (Figure 2A,B). Members of Firmicutes (*Tissierella Soehngenii*) were observed in the gut tissue and appeared to be present in the pharynx tissue and the environmental samples, particularly the water. Also observed were Bacteroidetes (Flavobacteriales, ~2%), as well as

Gammaproteobacteria (*Psychromonas*, ~7%), Deltaproteobacteria (*Desulfotalea*, ~5%), and Fusobacteria (*Propionigenium*, ~5%), to lesser degrees of abundance.



**Figure 2.** Taxonomic distribution of microbial communities in the gut ecosystem and rocky tide pool environment of the sea urchin *S. purpuratus*. (A) The relative abundance of phyla (class for Proteobacteria) represented at >1% are shown, with phyla <1% grouped as “Other.” (B) The top 100 taxa at the most resolvable level across all samples were also visualized, and taxa not included as the top 100 were assigned as “Other.” OTUs were picked at 97% similarity threshold, filtered at <0.0005%, condensed using PhyloToAST (v1.4.0), and subsampled to minimum OTU count (rarefied OTU table). Taxonomic identities were determined by using the GreenGenes (v13.8) database, and the color code corresponds to each taxon observed across the gut and environmental samples. Grouping of biological replicates ( $n = 3$ ) was supported by an analysis of similarity (ANOSIM) and multivariate analysis of variance (Adonis) ( $p < 0.001$ ). Relative abundance plot was created through Microsoft Excel Software (Seattle, WA, USA).

The dominant microbial taxa in the gut digesta were observed to be members of *Psychromonas* (~40%), *Propionigenium* (~15%), and class Flavobacteriales (~25%). Compared to the gut tissue, these taxa comprised a large relative abundance (~80%) of the bacterial microbiota observed in the gut digesta. The gut digesta also included members of phylum Bacteroidetes (3%), class Gammaproteobacteria identified as Vibrionaceae and *Vibrio* (~1 and ~2% respectively), and *Desulfotalea* (~5%) at noticeable relative abundances (Figure 3).



**Figure 3.** Comparison of the observed taxa between the gut tissue ( $n = 3$ ) and gut digesta ( $n = 3$ ) using the rarefied OTU table data. Taxa observed at <1% were filtered from the graph. Standard deviation and relative abundances were determined through STAMP (v2.1.3), and the graph was generated through Microsoft Excel Software (Seattle, WA, USA).

The pharynx tissue presented many of the same bacterial taxa observed in the water and to a lesser extent the algae samples. Of these shared taxa, *Fusobacterium* (~10%) was observed at the highest abundance, which was followed by the families S27-7 (Bacteroidales) and Gemellaceae, and genus *Prevotella*. The presence of these bacteria in the gut tissue and gut digesta were negligible (<1%). However, the pharynx tissue also included members of *Tissierella\_Soehngenia* (~6%), *Sulfurimonas* (~3%), and *Desulfotalea* (~1%) which were observed in the gut tissue and digesta.

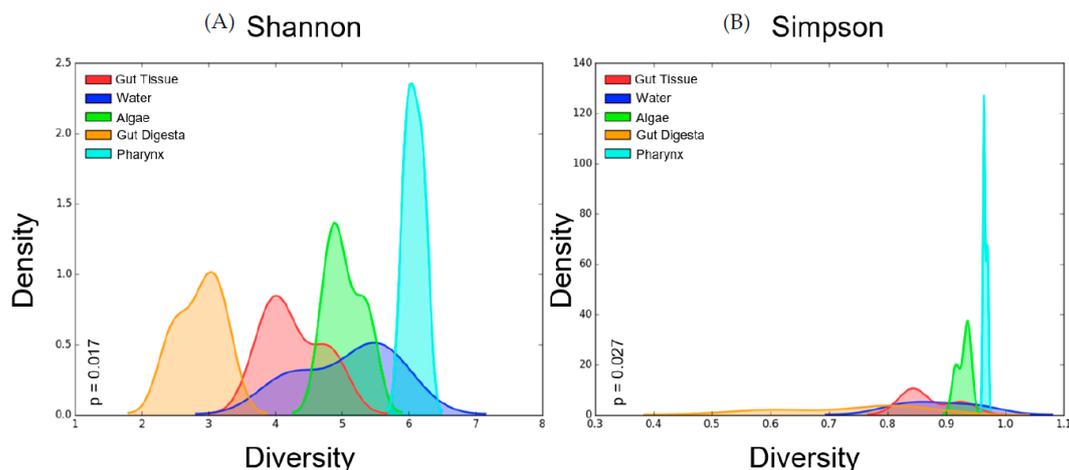
The algae samples showed Saprospiraceae (~15%), Rhodophyta (~10%), and *Stramenopiles* (~9%) to be heightened, the presence of which was negligible in the other samples in the present study. We also observed the genera *Maribacter* and *Octadecabacter* at equal capacities (~4%). All microbial taxa at their most resolved level, including their group abundance identified in the rarefied OTU table used in this study, have been elaborated (Table S1). Additionally, the unfiltered OTU table generated in this study has also been elaborated (Table S2).

From the alternative strategy utilizing ASVs, the taxonomic distribution was in concert with the OTU picking strategy, with only slight variations in relative abundance (Table S3). In the gut tissue,

a slightly higher relative abundance of *Arcobacter* (~22%) and *Sulfurimonas* (~14%) was determined through the ASV method compared to the OTU picking strategy. A variation in relative abundances of heightened taxa was also observed in the gut digesta, where *Psychromonas* was more highly represented at ~50%, whereas *Propionigenium* (~13%) and Flavobacteriales (~14%) were marginally less abundant. The pharynx tissue was also consistent between the two strategies, with a slightly higher abundance of *Fusobacterium* (12%) and *Sulfurimonas* (~7%). The ASV method was able to resolve an extra phylogenetic level of one heightened feature in the pharynx tissue, which was classified to phylum Gammaproteobacteria (~6%) through the OTU picking method, but was determined as order Legionellales (~10%) through the ASV approach. For the water samples, *Tissierella\_Soengenia* (~32%) comprised the highest abundance, followed by *Fusobacterium* (~7%), and for the algae samples, Saprospiraceae, Rhodophyta, and *Stramenopiles* were confirmed through the ASV method, at slightly higher relative abundances compared to the OTU picking method. Additionally, the results of the alignment of the representative sequences corresponding to the top 100 taxa determined in the rarefied OTU table through SILVA ACT: Alignment, Classification and Tree Service ([www.arb-silva.de/aligner](http://www.arb-silva.de/aligner)) have been elaborated (Table S4).

### 3.4. Alpha Diversity

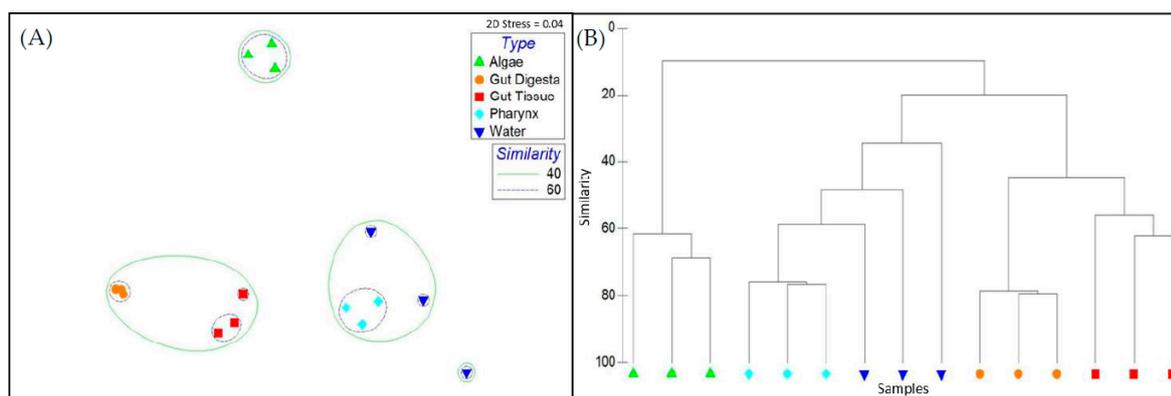
Alpha diversity measures performed on the rarefied OTU table showed the highest Shannon diversity in the pharynx tissues (avg =  $6.08 \pm 0.08$  SEM), followed by the water ( $5.08 \pm 0.44$  SEM) and algae ( $5.04 \pm 0.17$  SEM), with the gut tissue (avg =  $4.27 \pm 0.28$  SEM) and gut digesta (avg =  $2.87 \pm 0.21$  SEM) showing the lowest diversity. Simpson diversity showed a similar trend in the sea urchin gut samples, with the pharynx tissue showing the highest diversity (avg =  $0.976 \pm 0.002$  SEM), and the gut tissue ( $0.871 \pm 0.027$  SEM) and gut digesta ( $0.738 \pm 0.070$  SEM) showing the lowest. The water ( $0.885 \pm 0.039$  SEM) and algae ( $0.929 \pm 0.007$  SEM) microbial profiles were more diverse than the gut tissue and digesta, but less diverse than the pharynx tissue (Table 1). Kruskal–Wallis H-test analysis through PhyloToAST (v1.4.0) showed significant differences between the Shannon ( $p = 0.017$ ) and Simpson ( $p = 0.027$ ) alpha diversity values across the 5 groups of the study. The kernel density smoothed histograms plotted using PhyloToAST (v1.4.0) visualized the alpha diversity values, with density representing the intra-sample variation in the group (high density = low variation between the diversity of samples in the group). For both Shannon (Figure 4A) and Simpson (Figure 4B) diversity, the highest density corresponded to the pharynx tissue samples, indicating low intra-sample variation in the group. Shannon diversity showed the broadest diversity value range in the water (Figure 4A), and Simpson showed the broadest range in the gut digesta (Figure 4B). For Shannon diversity, the gut digesta and gut tissue samples had distinct histogram peaks. For Simpson diversity, the range of alpha diversity measures for the gut digesta was broader, indicating a minimal density peak, and the gut tissue showed a low but noticeable peak (Figure 4A,B).



**Figure 4.** Per-group alpha diversity measurements calculated across all samples in the study. (A) Shannon and (B) Simpson alpha diversity histograms were smoothed by kernel density estimation. The Kruskal–Wallis H-tests were performed for the five groups and showed a significance value of  $p = 0.017$  for the Shannon and  $p = 0.027$  for Simpson diversity measurements, indicating significant differences between each group’s alpha diversity. The X-axis shows the diversity value of Shannon (values much greater than 0 are more diverse) and Simpson (values closer to 1 are more diverse). The histogram values of each sample were smoothed through kernel estimation to show the range of sample data points within each group. The Y-axis depicts the density function, which denotes the distribution of data points falling within this range (higher peak represents more clustered data points). Relevant p-values are listed in each graph. Plots were generated using the “diversity.py” command through PhyloToAST (v1.4.0).

### 3.5. Beta Diversity

Microbial taxonomic distribution patterns determined through Bray–Curtis metrics across all samples revealed the gut tissues to cluster together at >50% and the gut digesta >60% (Figure 5A). These two sample groups maintained a Bray–Curtis similarity >40%. The pharynx tissues group demonstrated low intrasample variation (Bray–Curtis similarity >60%) and had a microbial community structure that was the most similar to the water samples (Bray–Curtis similarity >40%, Figure 5B). The algae samples clustered together at a value >60% but were least similar to the other samples of the study. The observed cluster patterns were strengthened when only the top 100 OTUs were plotted through 2D MDS (Figure S1A,B). For the rare taxa (the taxa not included in the top 100 OTUs), although low intrasample variation was observed, the gut tissue group was clustered nearer the pharynx and water samples (Figure S1C,D). Although the general trends of within- and between-group sample similarity were supported through the weighted (Figure S2A,B) and unweighted (Figure S3A,B) Unifrac 3D PCoA and dendrogram analysis, there were slight variations in cluster patterns between the two Unifrac approaches. Through the weighted Unifrac method, the gut tissue and gut digesta microbial communities clustered closer together as compared to the unweighted method, which showed the gut tissue samples to more closely resemble the pharynx. However, in concert with the Bray–Curtis method, both methods demonstrated the pharynx tissue to more closely resemble the water samples, and the algae samples maintained a divergent cluster pattern away from the other samples of the study.



**Figure 5.** Beta diversity analysis of microbial communities observed across all samples in the study using Bray–Curtis similarity metrics determined for the rarefied OTU table. (A) A 2D multidimensional scale (MDS) plot analysis was performed to show sample cluster patterns based on observed OTUs, with a 40% and 60% Bray–Curtis similarity overlay, and the stress value (2D Stress = 0.04) was indicated. (B) Dendrogram analysis was also performed and each sample’s cluster patterns were based on group average. The OTU table was pretreated via standardization by the total and log transformation prior to Bray–Curtis analysis. Figure legends are shown in the 2D MDS plot. Data was generated and plotted through PRIMER-6 software (Primer-E Ltd, Plymouth Marine Laboratory, Plymouth UK, v6.1.2).

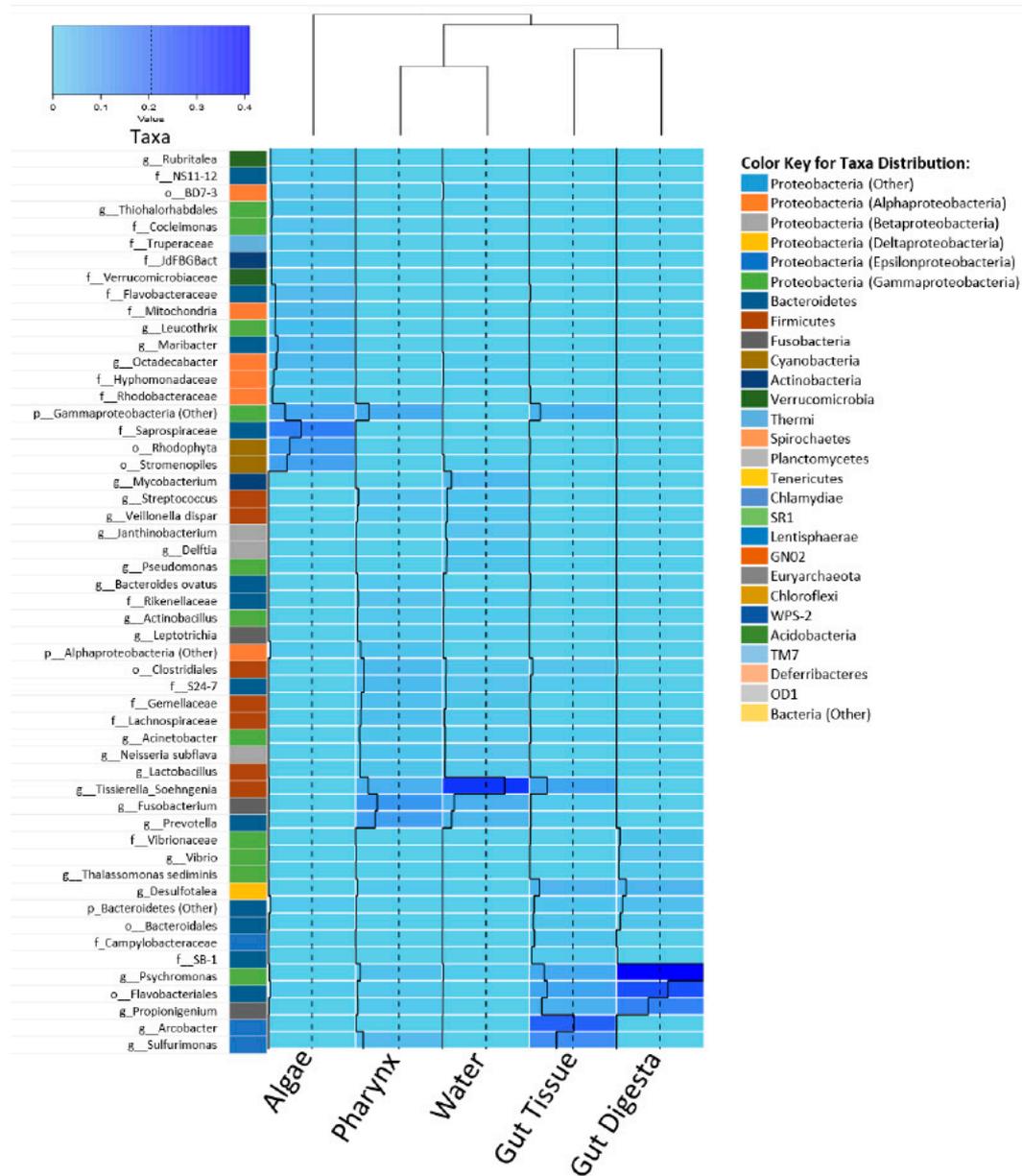
ANOSIM and Adonis analysis supported biological replicate grouping and demonstrated the highest grouping similarity corresponding to the OTU data that was rarefied to the minimum count value (ANOSIM  $R = 0.94$ ; Adonis  $R^2 = 0.76$ ). However, significant grouping patterns of biological replicate groups were observed when the analysis was performed on each OTU table generated in this study (unfiltered, filtered, condensed, median rarefied, and minimum rarefied OTU tables) with  $p = 0.001$  for ANOSIM and Adonis (Table 2). Additionally, ANOSIM and Adonis performed on the rarefied OTU table using the weighted (ANOSIM  $R = 0.92$ ; Adonis  $R^2 = 0.77$ ) and unweighted (ANOSIM  $R = 0.88$ ; Adonis  $R^2 = 0.73$ ) Unifrac statistic also supported significant biological replicate grouping, with  $p = 0.001$  determined using both methods, with the weighted method showing a slightly higher value for both statistics.

**Table 2.** Grouping statistics performed on each OTU table generated in the study. Results show both ANOSIM and Adonis measurements of the initial OTU table (unfiltered; Unfilt.), followed by filtering of those taxa represented at  $<0.0005\%$  in the study (filtered; Filt.), condensing (Cond.) using PhyloToAST (v1.4.0) and rarefying to the median (Med.) and minimum (Min.) of the condensed OTU table file. Analysis was performed on grouped biological sample replicates (pharynx,  $n = 3$ ; gut tissue,  $n = 3$ ; gut digesta,  $n = 3$ ; water,  $n = 3$ ; and algae,  $n = 3$ ).

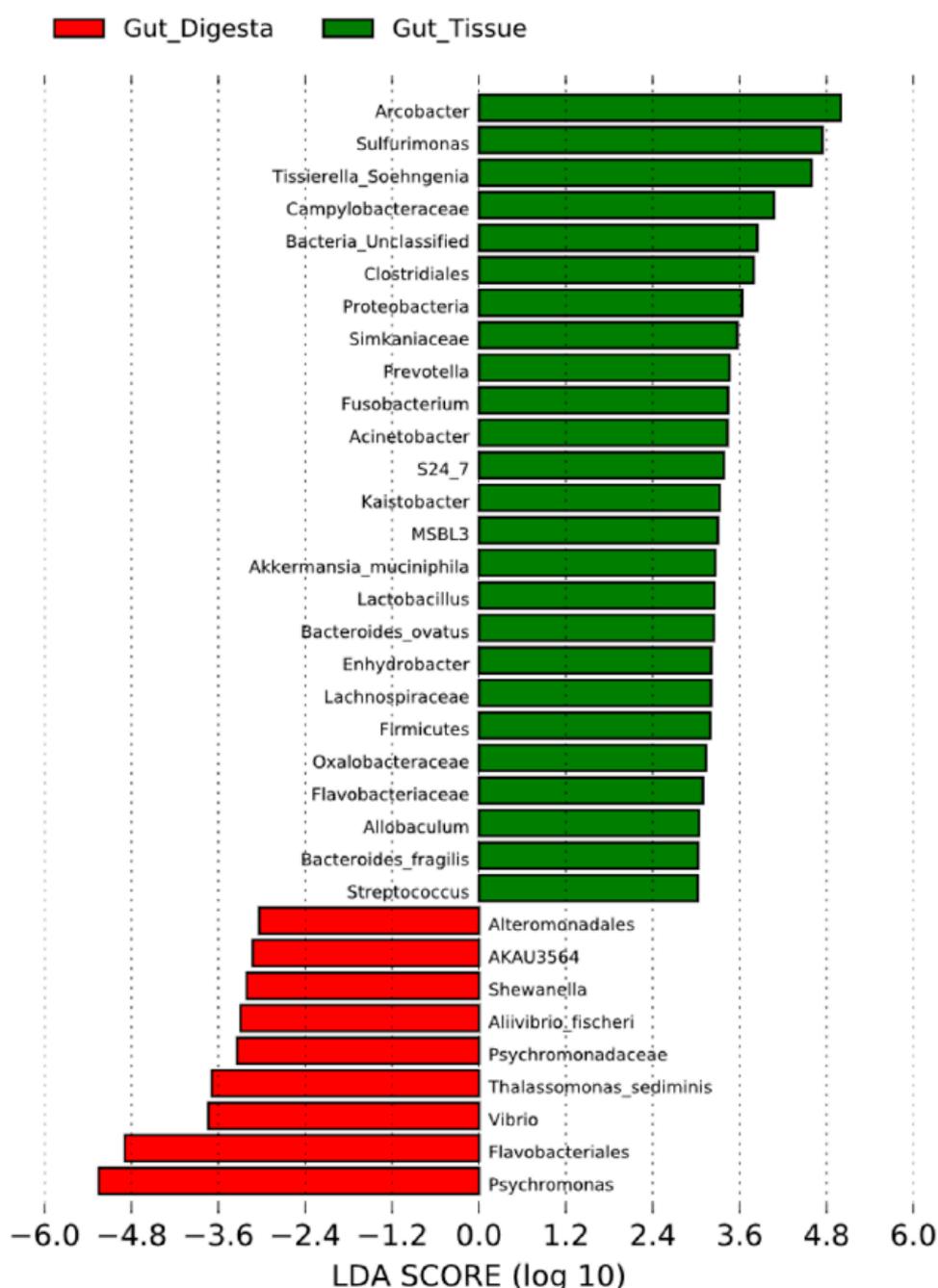
| Diversity Measure | Unfilt. OTUs    | Filt. OTUs      | Cond. OTUs      | Cond. OTUs Med. | Cond. OTUs Min  |
|-------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| ANOSIM            | $R = 0.93185$   | $R = 0.93185$   | $R = 0.94074$   | $R = 0.94074$   | $R = 0.94222$   |
| Adonis            | $R^2 = 0.69518$ | $R^2 = 0.71145$ | $R^2 = 0.74894$ | $R^2 = 0.74951$ | $R^2 = 0.75688$ |

Similarity trends between sample groups were elaborated by heatmap analysis, which also depicted the relative abundance associated with each taxon contributing to the group diversity across the grouped biological replicate samples (Figure 6). Dendrogram analysis across grouped biological replicates for the heatmap analysis revealed the gut digesta and gut tissue to have a more similar microbial ecology composition, and likewise for the water and pharynx groups. The algal food source was the least similar to the other samples of the study, confirming the similarity trends observed in the 2D MDS plot analysis (Figure 5A). LEfSe analysis of the gut tissue and digesta at an LDA score of  $\pm 3$  showed the taxa contributing most to the dissimilarity (effect size) of the gut tissue to be *Arcobacter*, *Sulfurimonas*, and *Tissierella\_Soehngenia*, whereas the gut digesta revealed *Psychromonas*,

Flavobacteriales and *Vibrio* (Figure 7). The effect size of *Propionigenium*, which was noticeably abundant in the gut tissue ( $5.7\% \pm 4.8\%$ ) and digesta ( $14.9\% \pm 9.6\%$ ), was not observed in LDA analysis, likely due to overlapping standard deviation values (as observed in Figure 3).



**Figure 6.** Heatmap of the top 53 taxa at the highest resolution, determined using the rarefied OTU table and generated using R (v3.3.2). The heatmap.2 function from the gplots (v3.0.1) package ([www.rdocumentation.org/packages/gplots](http://www.rdocumentation.org/packages/gplots)) was used. Sample dendrogram was generated using Vegan (v2.4.3), employing the Bray-Curtis metric of the grouped biological replicate count data. Color palette selected using the RColorBrewer package and set from “sky blue” for less abundant, to “blue” for more abundant (shown in color key). Relative abundance values of each taxon are also indicated through a trace line (black). The associated table includes the most resolvable taxonomic assignment according to the GreenGenes (v13.8) database, which is color-coded to the phylum level assignments (class for Proteobacteria) as indicated in the key and corresponding to the relative abundances in the Figure 2A relative abundance graphs. The figure has been generated using scalable graphics and, therefore, regions of interest can be viewed at a higher resolution digitally by increasing the magnification.



**Figure 7.** Linear discriminant analysis (LDA) effect size (LEfSe) performed on the microbial community relative abundance data at the of the gut tissue ( $n = 3$ ) and gut digesta ( $n = 3$ ). Grouped data were first analyzed using the Kruskal–Wallis test with a significance set to 0.05 to determine if the data was differentially distributed between groups, and those taxa that were differentially distributed were used for LDA model analysis to rank the relative abundance difference between groups. The LDA for significance was set to  $\pm 3$ , and the  $\log(10)$  transformed score is shown to demonstrate the effect size. Data were analyzed and prepared through Hutlab Galaxy provided through the Huttenhower lab. The gut tissue group is shown as green, and the gut digesta group as red.

### 3.6. Predicted Functional Capacity

The predicted functional capacity of the microbial communities of the digestive system was performed using PICRUSt (v1.1.2), showing an average NSTI value of 0.139 (range from 0.105 - 0.179). Scatter plot analysis using STAMP (v2.1.3) of the KEGG-Level-2 categories showed a preferential

abundance of energy metabolism in the gut tissue, as well as membrane transport, cell motility, and signal transduction (Figure 8A). For the gut digesta, amino acid metabolism, carbohydrate metabolism, metabolism of cofactors and vitamins, and replication and repair categories were observed. KEGG-Level-3 observances showed a preferential abundance of oxidative phosphorylation, carbon fixation, methane, and nitrogen metabolisms in the gut tissue, and categories related to the transporter and motility-related categories (2-component system, chemotaxis, bacterial motility proteins, and flagellar assembly) (Figure 8B). The gut digesta displayed categories related to pyrimidine metabolisms, as well as peptidases and amino acid enzymes, including arginine and proline metabolisms. Other categories that were enriched in the digesta included starch and sucrose metabolism, pentose phosphate pathway, glycolysis/gluconeogenesis, and ubiquinone and turpenoid-quinone biosynthesis necessary for electron transport. The KEGG-Level-2 and 3 categories identified through PICRUSt (v1.1.2) were listed in Table S5. Additionally, the LEfSe analysis of KO IDs (highest metabolic resolution) contributing most to the effect size difference between the gut tissue and gut digesta were determined and listed along with their associated metabolic definitions (Figure 9 and Table S6).

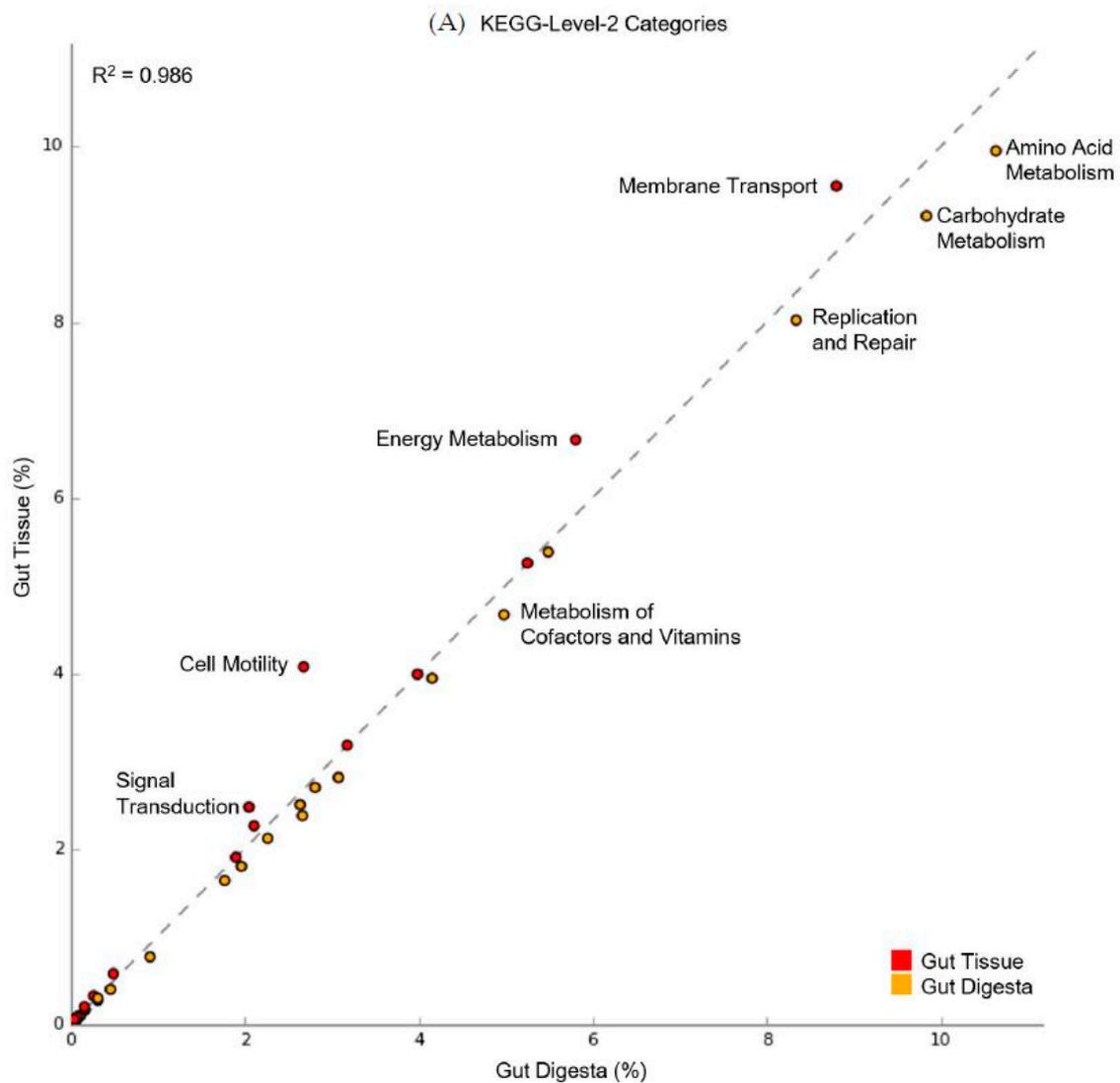
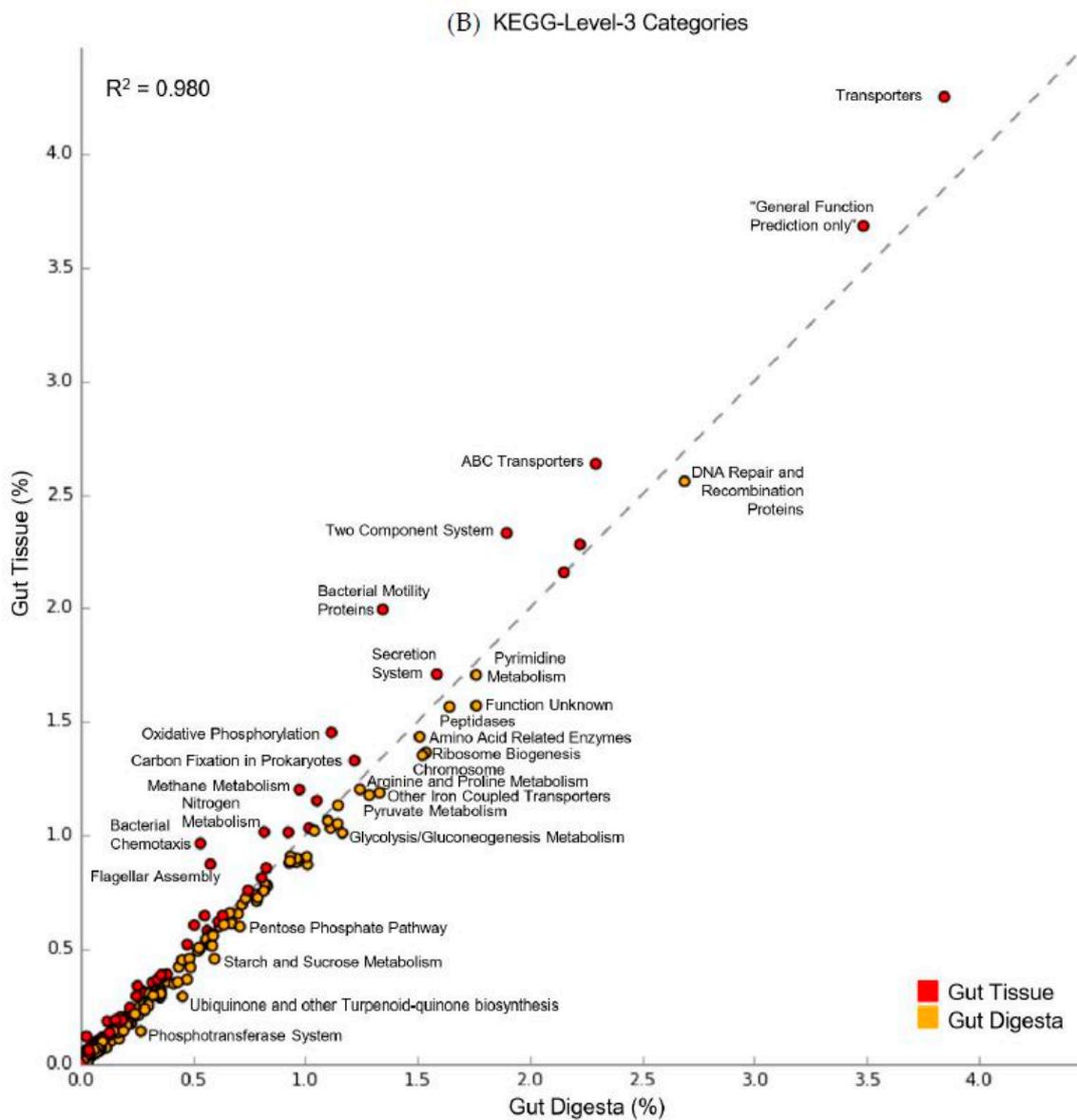
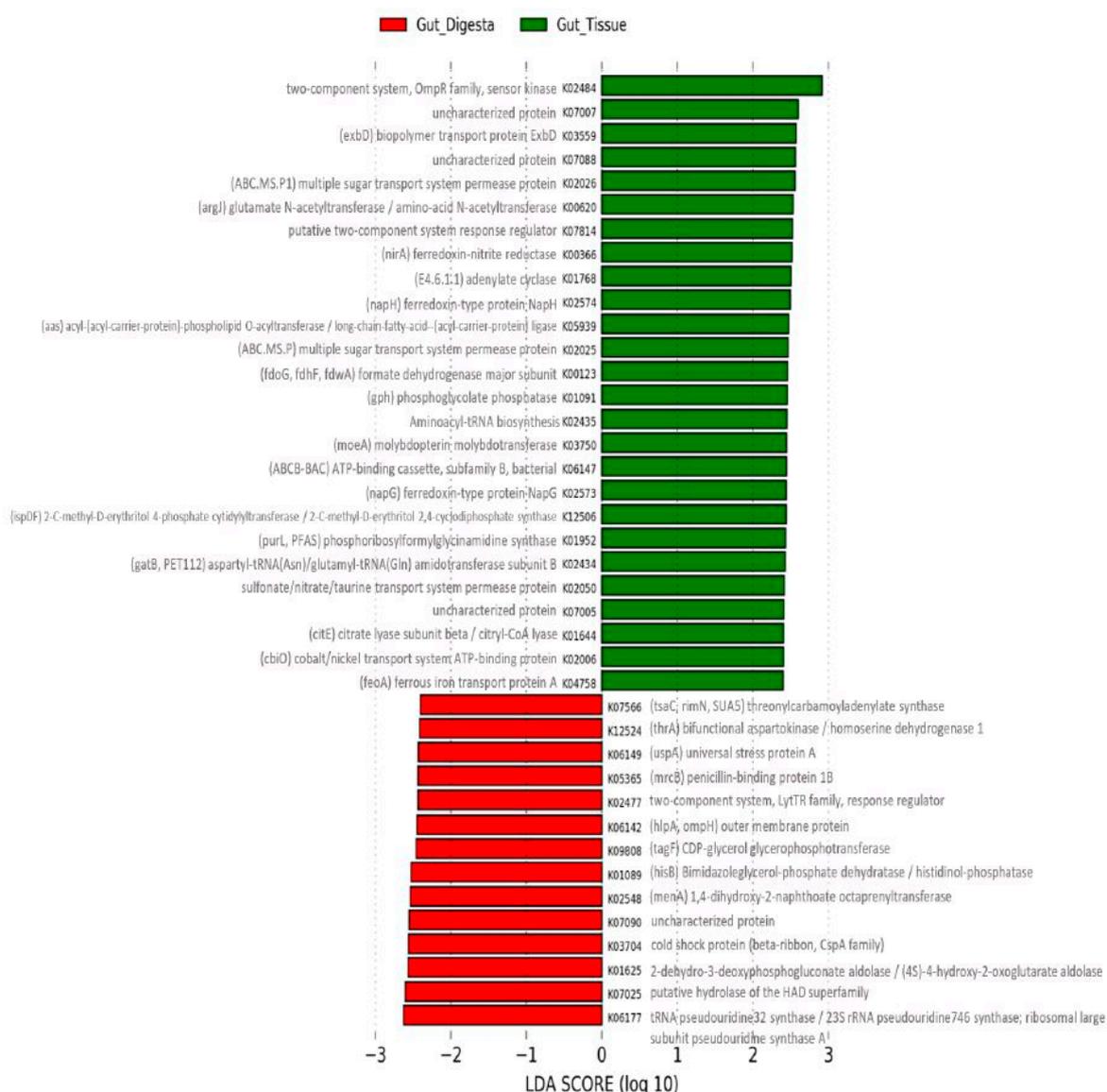


Figure 8. Cont.



**Figure 8.** Scatter plot analysis of the predicted KEGG Orthology (KO) metabolic functions determined through Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt v1.1.2) performed on the gut tissue ( $n = 3$ ) and gut digesta ( $n = 3$ ). Biological replicates were grouped, and analysis was performed for the (A) KEGG-Level-2 and (B) KEGG-Level-3 hierarchical functional categories. The linear regression value calculated for the two groups is shown for each scatter plot graph. Preferentially enriched categories for the gut tissue are shown as red, and for the gut digesta as brown. Those categories with clearly preferentially abundant categories have been labeled. Data were analyzed and visualized using STAMP (v2.1.3) analytical software. The node labels have been generated using scalable graphics, and therefore regions of interest can be viewed at a higher resolution digitally by increasing the magnification.



**Figure 9.** Linear discriminant analysis (LDA) effect size (LEfSe) performed on the KEGG Orthology (KO) metabolic functions determined through PICRUSt (v1.1.2) for the gut tissue ( $n = 3$ ) and gut digesta ( $n = 3$ ). The KO IDs were determined through the “predict\_metagenomes.py” command in PICRUSt (v1.1.2). Grouped data were analyzed using the Kruskal–Wallis with a significance set to 0.05, and the significantly differentially distributed KO IDs were used for LDA model analysis ranking the relative abundance significance, at an LDA threshold showing entries ranking at  $\pm 2.4$ . The  $\log(10)$  transformed score is shown as the effect size. Data were analyzed and prepared through Hutlab Galaxy provided through the Huttenhower lab. The gut tissue group is shown as green, and the gut digesta group as red. The KO Id labels have been generated using scalable graphics, and therefore regions of interest can be viewed at a higher resolution digitally by increasing the magnification.

### 3.7. Co-Presence, Co-Exclusion, and Key Taxa in the Gut Environment

The resultant network generated through CoNet (v1.1.1) in Cytoscape (v3.6.0) yielded 71 nodes and 294 edges elucidating possible interactions occurring between taxa representing the distinct microbial communities of the sea urchin gut environment (Figure 10A). Analysis of network properties using NetworkAnalyzer (v2.7) showed an average network centralization of 0.128, the characteristic path length of 2.856, average number of neighbors at 8.282, with a network density of 0.118 and network heterogeneity of 0.452. Scatter plot analysis demonstrated the trends of closeness centrality

plotted against betweenness centrality, along with the nodes scaled to the degree (Figure 10B). The top 10 candidate key taxa based on the topological qualities of taxonomic nodes described in Berry and Widder et al. [30] were ranked by their closeness centrality, and showed *Propionigenium*, *Moritella*, SB-1 (Bacteroidetes), Desulfobacteraceae and *Desulfovibrio* in the gut digesta, and Rhodobacteraceae, Rhodophyta, Vibrionaceae, *Arcobacter* and *Bacilli* in the gut tissue (Table 3). Of these taxa, the gut digesta showed *Propionigenium* to have the highest degree of associations with the gut tissue taxa (17 total), with the majority of these associations shown as co-presence (14 total). The gut tissue showed Rhodobacteraceae to have the highest degree (9 total), and most of these associations were co-exclusion (8 total). The gut tissue also showed the highly abundant *Arcobacter* to be a likely key taxon, revealing a degree of 9, with the majority of these associations as co-presence (7 total).

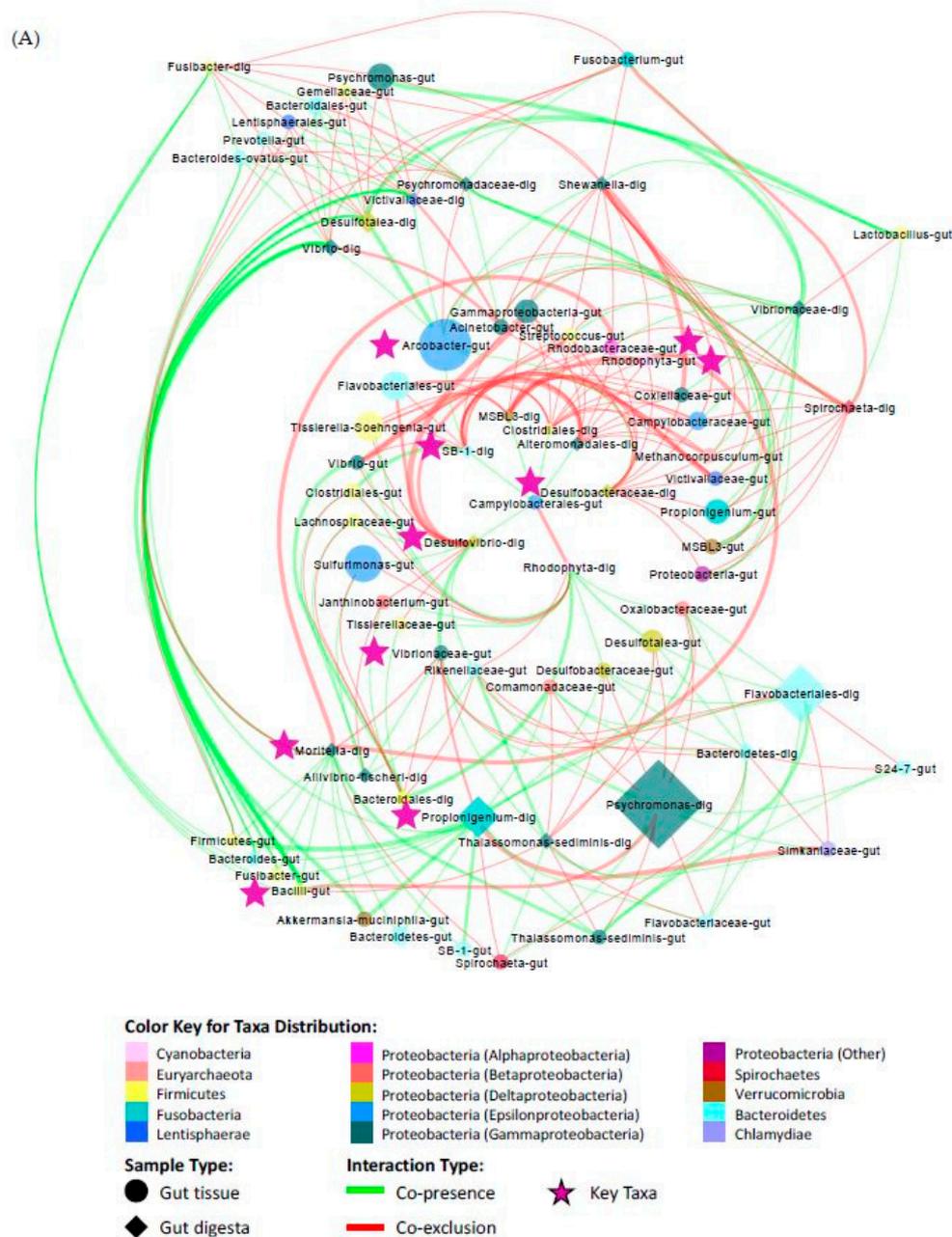
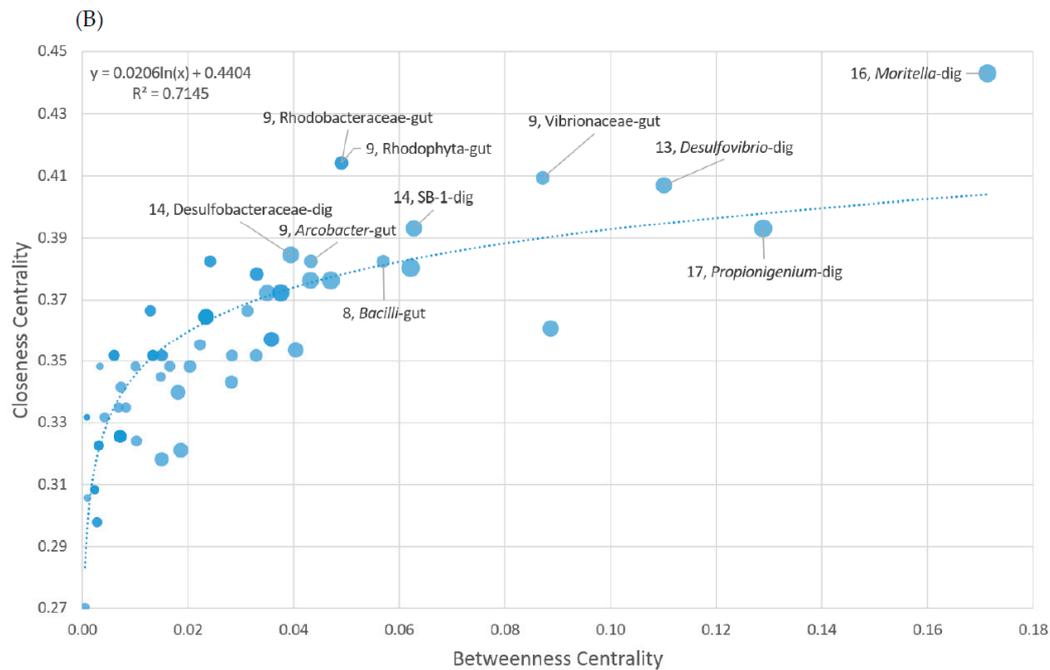


Figure 10. Cont.



**Figure 10.** Co-occurrence patterns between taxonomic entries of the gut tissue and gut digesta, determined through Co-occurrence Network inference (CoNet v1.1.1), and analyzed through Cytoscape (v3.6.0). Taxonomic entries with a cumulative row sum of 200 or above with 2/3 of samples showing non-zero value entries were used through an ensemble approach that incorporated the Pearson, Spearman, Bray–Curtis, Kullback–Leibler, and mutual information metrics. The top and bottom 200 edges were selected and merged by the union method. (A) The network analysis shows the edges represented by the q-value (merged with the Brown method at  $p < 0.05$  for each metric) and are shown as green (co-presence) and red (co-exclusion), with the nodes representing taxa were scaled according to relative abundance and colored according to the phyla (class for Proteobacteria) assignments. The final network was arranged using the yFiles (v1.0) Cytoscape (v3.6.0) add-on radial layout, and taxonomic entries shown at the highest resolution are denoted with the sample type (circle for gut tissue, “-gut”; diamond for gut digesta; “-dig”). The network has been generated using scalable graphics, and therefore nodes of interest can be viewed at a higher resolution digitally by increasing the magnification. (B) Scatter plot analysis was performed using topological metrics determined by NetworkAnalyzer (v2.7), to demonstrate patterns of key (keystone) species between taxonomic entries of the gut tissue and gut digesta based on closeness and betweenness centrality, as well as the degree (number of co-presence and co-exclusion edges). Linear regression between closeness and betweenness centrality was shown as logarithmic ( $R^2$  value = 0.7145), and the top 10 entries ranked by closeness centrality are depicted. Note, the taxa *Rhodophyta* and *Rhodobacteraceae* had the same closeness and betweenness centrality measurements, and their corresponding plot is indistinguishable. Linear regression determined through Microsoft Excel Software (Seattle, WA).

**Table 3.** Candidate key taxa resulting from CoNet (v1.1.1) analysis between the gut tissue and gut digesta microbial communities. The group assignment is shown (gut tissue,  $n = 3$ ; gut digesta,  $n = 3$ ), along with the phylum and taxon at the highest resolution. The average abundance (Av. Ab.) was determined for each taxon based on the respective group (gut tissue or digesta as indicated). The entries were ranked according to closeness centrality (Clos. Cent.), highest to lowest (a feature of keystone-ness), and then by degree (edges; Deg.), highest to lowest. Also shown is the betweenness centrality (Bet. Cent.), and the total, positive (co-presence; Pos. Deg.) and negative (co-exclusion; Neg. Deg.) degree values.

| Sample Type | Phylum                                    | Taxon                 | Av. Ab. | Clos. Cent. | Bet. Cent. | Deg. | Pos. Deg. | Neg. Deg. |
|-------------|---|-----------------------|---------|-------------|------------|------|-----------|-----------|
| Gut Digesta | Fusobacteria                              | <i>Propionigenium</i> | 14.89%  | 0.39        | 0.13       | 17   | 14        | 4         |
| Gut Digesta | Proteobacteria<br>(Gammaproteobacteria)   | <i>Moritella</i>      | 0.15%   | 0.44        | 0.17       | 16   | 10        | 6         |
| Gut Digesta | Bacteroidetes                             | SB-1                  | 0.38%   | 0.39        | 0.06       | 14   | 3         | 11        |
| Gut Digesta | Proteobacteria<br>(Deltaproteobacteria)   | Desulfobacteraceae    | 0.24%   | 0.38        | 0.04       | 14   | 2         | 12        |
| Gut Digesta | Proteobacteria<br>(Deltaproteobacteria)   | <i>Desulfovibrio</i>  | 0.24%   | 0.41        | 0.11       | 13   | 6         | 7         |
| Gut Tissue  | Proteobacteria<br>(Alphaproteobacteria)   | Rhodobacteraceae      | 0.34%   | 0.41        | 0.05       | 9    | 1         | 8         |
| Gut Tissue  | Cyanobacteria                             | Rhodophyta            | 0.22%   | 0.41        | 0.05       | 9    | 1         | 8         |
| Gut Tissue  | Proteobacteria<br>(Gammaproteobacteria)   | Vibrionaceae          | 0.36%   | 0.41        | 0.09       | 9    | 5         | 5         |
| Gut Tissue  | Proteobacteria<br>(Epsilonproteobacteria) | <i>Arcobacter</i>     | 20.59%  | 0.38        | 0.04       | 9    | 7         | 2         |
| Gut Tissue  | Firmicutes                                | <i>Bacilli</i>        | 0.16%   | 0.38        | 0.06       | 8    | 6         | 2         |

#### 4. Discussion

Overall, the *S. purpuratus* gut ecosystem exhibited high abundances of *Arcobacter* and *Sulfurimonas* in the gut tissues. These taxa belong to phylum Epsilonproteobacteria, and specifically within the order Campylobacterales, members of which are known to be chemolithoautotrophic [88]. Both *Arcobacter* and *Sulfurimonas* have been implicated as marine sulfur-oxidizing bacteria, with members of *Arcobacter* forming oxidized sulfur filaments in response to geothermally produced sulfide in hydrothermal vents [89,90], and *Sulfurimonas* has been observed to oxidize sulfur in anoxic deep-sea hydrothermal sediments [91,92]. Additionally, Epsilonproteobacteria have been reported in the gut systems of marine organisms, with *Arcobacter* previously observed in the deep-sea vent-dwelling shrimp *Rimicaris exoculata* [93] and the marine Chilean oyster *Tiostrea chilensis* [94], and *Sulfurimonas* in the stomach of the hydrothermal vent crab *Xenograpsus testudinatus* [95] and as symbionts in marine gastropod mollusks *Alviniconcha* [96,97]. These bacteria have been implicated as potential symbionts, assisting in the oxidation of sulfur in the highly sulfidic environments [98]. Previous studies using Illumina MiSeq HTS of the V4 region of the collective 16S rRNA genes conducted on the gut environment of the sea urchin *L. variegatus* from the Gulf of Mexico similarly revealed Campylobacterales to be near-exclusively abundant in the gut tissue [15,16], of which further analysis of the highly-represented Campylobacterales sequence showed a ~90-91% similarity match to *Arcobacter/Sulfuricum* through NCBI BLAST [15].

In contrast to the gut tissue, the microbial composition of gut digesta was dominated by *Psychromonas* (phylum Gammaproteobacteria). This halophilic genus is known to occur in cold marine environments (~4 °C) [99], and members are capable of hydrolyzing starch and other insoluble sugars [100–103], as well as producing  $\omega$ -3 polyunsaturated fatty acids [99]. It has recently been shown that the total  $\omega$ -3 fatty acids of the gut contents ( $23.26 \pm 6.88$ ; mean  $\pm$  SD) of experimentally fed *S. purpuratus* are higher than the total  $\omega$ -3 fatty acids of their algal diets ( $13.41 \pm 6.32$ ); the dominant driver of this pattern was from the  $\omega$ -3 eicosapentaenoic acid (EPA; 20:5 $\omega$ -3) [20], but the mechanism for this enrichment was unknown. The psychrophilic capabilities of *Psychromonas*

are consistent with the temperature of the Oregon Pacific Coast seawater temperature from which *S. purpuratus* was sampled, which was recorded at 13.1 °C, within the range of optimum growth *Psychromonas* spp. [100]. Also observed in the gut digesta were *Propionigenium* of phylum Fusobacteria, members of which are anaerobic, are capable of fermenting succinate to propionate [104] and have been implicated in carbohydrate metabolisms that include cellulose, producing short-chain fatty acids [105]. *Propionigenium* is involved in a myriad of host health benefits, including its association in the modulation of the lifespan of the Turquoise Killifish (*Nothobranchius furzeri*) [105]. Order Flavobacteriales of phylum Bacteroidetes were also observed, which have been identified in the intestines of shrimp *Litopenaeus stylirostris* raised in aquaculture and clear waters [106]. Importantly, members of Flavobacteriales have been investigated with consideration of their metabolic capability to break-down alginate from brown seaweeds common in both temperate and polar coastal environments [107]. Lastly, the genus *Desulfotalea* of Deltaproteobacteria was observed at approximately equal relative abundances in the gut digesta and gut tissue (~5% in each). Certain members of *Desulfotalea* are adapted to extremely cold environments, and these bacteria are capable of sulfur reduction [108,109].

Shannon and Simpson's diversity measured for all samples showed the microbial communities of the gut environment to have low diversity, whereas the pharynx and environmental samples had a high diversity value. Such high microbial diversity in the marine environment indicates a rich microbial community and is expected due to the fluctuations of abiotic conditions and nutrients, spatial dispersion of microbes through the marine environment, and dynamic host-species populations that shape the microbiota to the environment [21]. The reduction of species richness from the environment to the gut tissue and digesta indicated a more preferential microbial community comprised of specific taxa in the gut system.

Compartmentalization of microbial community profiles in the gut environment of *S. purpuratus* was demonstrated through Bray-Curtis based 2D MDS plot and dendrogram cluster analysis and was supported through heatmap and ANOSIM/Adonis analysis. Grouping based on sample type was found to be significant through ANOSIM/Adonis statistics, indicating the consistency of microbial taxa across biological replicates representing each group (pharynx, gut tissue, gut digesta, water, and algae). 2D MDS plot analysis supported this low intra-sample variation within groups, and dendrogram analysis showed the gut tissues to cluster nearer the gut digesta, and the pharynx nearer the water samples. 2D MDS plot and dendrogram analysis of the top 100 OTUs strengthened the trends observed when using the rarefied OTU table. Additionally, an intra-sample variation of the rare taxa was not markedly transformed when determined through 2D MDS plot and dendrogram analysis. However, analysis using rare taxa slightly altered the cluster patterns of groups, showing the gut tissue to cluster nearer the water and pharynx, suggesting shared rare microbial taxa between these groups. Interestingly, although the weighted and unweighted Unifrac approaches did support the significant within-group similarities observed through Bray-Curtis based 2D MDS and dendrogram analysis, the between-group cluster patterns were slightly transformed. For weighted Unifrac, which measures the distance between samples using the relative abundances and phylogenetic relationships of the presented microorganisms [70], the gut tissue and digesta samples clustered closer together. However, through the unweighted method, which considers the presence/absence of taxa and their phylogeny [70], the gut tissue clustered with the water and pharynx samples. Such differences suggest an underlying phylogenetic relatedness between those bacterial microbiota colonizing the gut system based on their abundance. For all ordination analyses performed in this study, the microbial communities of the algae remained the least similar to all samples of the study.

Heatmap analyses supported the 2D MDS plot and dendrogram analysis, showing the contribution of Epsilonproteobacteria (*Arcobacter* and *Sulfurimonas*) to the unique microbial profile of the gut tissue, and *Psychromonas* in the gut digesta. LEfSe analysis demonstrated the significant effect size of certain key bacterial microbiota contributing to microbial community dissimilarity of the gut tissue and digesta, showing *Arcobacter*, *Sulfurimonas*, and *Tissierella\_Soehngenii* in the gut tissue, and

*Psychromonas*, Flavobacteriales, and *Vibrio* in the gut digesta. Although *Propionigenium* was found to be more heightened in the gut digesta ( $14.9\% \pm 9.6\%$ ) than the gut tissue ( $5.7\% \pm 4.8\%$ ), this taxon did not contribute to a high LDA score. This is likely due to the overlapping standard deviation of the relative abundance values of this taxon (Figure 3). Thus, these beta diversity analyses demonstrate the contribution of all taxonomic groups (heightened and rare taxa) in shaping both intra-sample and group-wise microbial profile similarity in the sea urchin *S. purpuratus*, as well as the specific taxa contributing to the unique cluster patterns and effect size of the compartmentalized gut tissue and digesta bacteria.

Public repositories of taxonomic information, such as the GreenGenes (v13.8) database used in this study, often possess repetitive entries, resulting in OTU tables with many redundant taxonomic assignments [25]. By using PhyloToAST (v1.4.0), such bias in the representation of 16S rRNA gene and associated taxonomic information can be improved, by binning highly similar taxonomic OTU sequences and assigning taxonomy to a higher resolution (genus/species when possible) through BLAST. Other strategies to improving the diversity measures of OTU tables generated from a targeted HTS approach have been proposed, such as the LULU algorithm that utilizes a post-clustering strategy to curate an OTU table based on co-occurrence patterns to help eliminate spuriously generated OTUs from the final dataset [110]. In this study, the initial OTU table generated following quality checks resulted in 44,664 unique assignments. However, condensing of redundant taxonomic IDs through PhyloToAST (v1.4.0) reduced the number of unique observations to 776, without any loss of read-count data when compared to the filtered OTU table. Additionally, subsampling of the condensed OTU table to the minimum read count (49,641) increased the beta diversity significance between biological replicate groups. Thus, the combined use of PhyloToAST (v1.4.0) and subsampling to the minimum value, enhanced the resolution of taxa, as well as increased the robustness of the beta analyses measures.

The popularity of targeted high throughput sequencing of microbial 16S rRNA genes has spurred the development of multiple bioinformatics tools, techniques, and approaches, designed to ensure the quality and reliability of the generated results. Traditionally, following the quality assessment and filtering of raw sequence read data, OTUs are chosen based on clustering highly similar sequences at a designated threshold, such as the 97% pairwise similarity described in our analysis, and commonly used as a proxy for species based on the 16S rRNA gene [111,112]. This OTU picking approach will often generate an initially high amount of unique observations, spuriously inflate the alpha diversity, and skew beta diversity measures, due to PCR and/or sequencing errors generating a high number of very rare OTUs [49]. Therefore, an additional filtration step is often necessary, such as the removal of OTUs occurring at  $<0.0005\%$  average abundance across all samples as applied in our study and others [50,51], and described for Illumina MiSeq generated sequence read data [49,52]. Alternative approaches to this strategy include a current practice of selecting ASVs as the operational units, which are generated by coupling a denoising step on the raw data that employs a parametric error model to define the frequency of observed sequence variants as it relates to real biological sequence data [43,44]. One such ASV selection tool, DADA2 (v1.10) as implemented in the recently developed QIIME2 package (v2018.11), utilizes a Poisson model to determine these repeated sequence variants [43,44], and was used as an alternative approach to our OTU method. For our OTU-based approach, the initial unfiltered OTU table generated a high number of unique observations, which were further filtered to eliminate potential spurious sequence reads, and then condensed through a BLAST approach with PhyloToAST (v1.4.0) and rarefied to the minimum sequence value. As previously stated, this produced 776 unique OTUs with 49,641 reads across all samples. After using DADA2 (v1.10) through QIIME2 (v2018.11), a total of 1134 unique ASVs with an average read count of  $\sim 31,191$  were generated across all samples, producing a quantitatively comparable sequence read count to the rarefied OTU table used in our downstream analyses. Additionally, following taxonomic assignment of the representative sequences generated for both the OTUs and ASVs using GreenGenes (v13.8), a similar relative abundance distribution was determined at comparable levels of phylogenetic resolution. Some exceptions included the additional resolution of one OTU assigned as phylum

Gammaproteobacteria in the pharynx tissue, which was further resolved to order Legionellales through the ASV method. However, the similarities of relative abundance are consistent with the most current literature comparing the two strategies [52]. Lastly, it should be noted that although DADA2 denoising may be advantageous in determining real biological sequences associated with rare taxa, this may come at the cost of determining “false positives” as described elsewhere [52].

Predicted functional profiles generated for the gut tissue and digesta at KEGG-Level-3 showed enrichment of metabolic qualities related to energy metabolisms, including oxidative phosphorylation, carbon fixation, methane, and nitrogen metabolisms in the gut tissue. The gut tissue also included heightened motility-related categories, suggesting a potential role of these functional attributes in the colonization of the gut tissue lumen [113], potentially through abiotic factors driving microbial colonization into the more oxygenated sea urchin gut tissue surface [114]. In the gut digesta, KEGG-Level-3 categories related to protein metabolisms (e.g., peptidases and amino acid enzymes, arginine and proline metabolisms) and carbohydrate metabolisms (e.g., starch and sucrose metabolism, pentose phosphate pathway, glycolysis/gluconeogenesis, and ubiquinone and other terpenoid-quinone biosynthesis) suggest a potential role of these bacteria in the digestive physiology of the purple sea urchin. These results are consistent with previous culture-dependent studies of sea urchin gut bacteria in digestion [9] and reflect potential anaerobic metabolisms likely to be performed in this mucous-sequestered niche [115]. LEfSe analysis at the KO Id level indicated the categories that contributed most to the differences between the distinct microbial communities of the gut system and supported the observed KEGG-Level-2 and 3 categories. The gut tissue presented seven categories related to membrane transporters, such as multiple sugar transport system permease proteins, cobalt/nickel transport system ATP-binding proteins and ferrous iron transport protein A, including ABC transport proteins such as the sulfonate/nitrate/taurine transport system permease protein category. Categories related to nitrogen metabolism, and specifically nitrate reduction, were also observed as heightened. Such nitrate reductive activity occurring in the sea urchin gut has been previously suggested to occur by the gut microbiota of sea urchins [14,116,117]. The gut digesta represented categories related to amino acid metabolism related to threonine and histidine biosynthesis, vitamin metabolisms such as Menaquinone (vitamin K<sub>2</sub>) production, and categories related to carbohydrate metabolism in the pentose phosphate pathway. These categories support the trends observed in the scatter plot analysis and offer insight into the functional qualities that accompany the life strategies of the gut microbes that may drive their distribution in the compartmentalized gut system. Lastly, although the calculated NSTI values (avg. = 0.139) for the data in this study indicated adequate confidence in the functional predictions of non-human-associated microbial communities based on taxonomic inference alone [75], a shotgun metagenomics approach targeting the metacommunity DNA of a microbiome sample would offer more reliable insight into those genes likely involved in host health and digestion.

We used CoNet (v1.1.1) analysis to identify theoretical modeling of relationships occurring between taxa in the distinctly compartmentalized gut tissue and gut digesta microbial communities [27–29]. Although taxa that were noticeably abundant in this study were represented (such as *Propionigenium* in the gut digesta and *Arcobacter* in the gut tissue), there were many low-abundance taxa representing likely key taxa. For example, *Moritella* (Gammaproteobacteria, <1% in the gut digesta) were observed to have a high closeness centrality relative to betweenness centrality and a high degree of edges (16 total) with 10 positive associations to taxa in the gut digesta. Interestingly, two taxa identified as Desulfobacteraceae and *Desulfovibrio* to the highest resolution were included in the top 10 likely key taxa in the gut digesta, were sulfur-reducing bacteria (<1% in the gut digesta) and represented a total degree of 14 and 13 respectively, with the majority of these indicated as co-exclusion relationships. These taxa belong to the phylum Deltaproteobacteria and represent species known to utilize sulfate as electron acceptors [118]. In the gut tissue, Rhodobacteraceae (Alphaproteobacteria) and order Rhodophyta (classified in phylum Cyanobacteria according to the GreenGenes v13.8 database) revealed the same pattern of degree (nine each), closeness and betweenness

centrality measures. Additionally, *Arcobacter*, which was highly abundant in the gut tissue (~ 20%), was identified as a key taxon with a degree of nine (with seven positive associations). This genus is comprised of species capable of utilizing elemental sulfur, hydrogen sulfide, and thiosulfate as terminal electron acceptors [119], alluding to a biogeochemical basis for the observed ecological relationships as suggested previously [120]. However, the relationship between microbial taxa of the gut tissue and gut digesta, including the role or key status of particular microbial taxa will require further verification.

In summary, the results of this study provide insight into the gut bacterial microbiota of *S. purpuratus* grown *in situ*, specifically elaborating (1) the taxonomic distribution, (2) the predicted functional categories assigned to the gut-associated bacterial communities, and (3) key taxa likely involved in maintaining the distribution patterns of gut microbiota through co-occurrence relationships in this evolutionarily and ecologically significant deuterostome. The gut environment demonstrated an allocation of microbial communities in the gut tissue, with a heightened abundance of Epsilonproteobacteria (namely *Arcobacter* and *Sulfurimonas*), and in the gut digesta, a higher abundance of Gammaproteobacteria (such as the psychrophilic genus *Psychromonas*). This trend of microbial compartmentalization has been previously observed in the laboratory [15,24] and naturally occurring [16] *L. variegatus* sea urchin bacteriome from the Gulf of Mexico. In those studies, certain microbial taxa were found to be consistent between the gut systems of both laboratory-raised and naturally occurring organisms, with the naturally occurring organisms showing slightly higher species diversity and richness in the gut system. More specifically, a near-exclusive abundance of *Arcobacter* (Epsilonproteobacteria) was observed in the gut tissue, whereas *Vibrio* (Gammaproteobacteria), a genus common to halophilic temperate marine environments, was most dominant in the gut digesta. Other bacterial microbiota found consistent in the gut digesta of *L. variegatus* included *Propionigenium* and taxa assigned as family Rhodobacteraceae. For both the laboratory-raised and naturally occurring sea urchins, beta diversity analysis of the pharynx showed cluster patterns that diverged from the gut tissue and digesta and more closely resembled that of their environment. Interestingly, many of the same trends of microbial integration in *L. variegatus* were also observed in the naturally occurring sea urchin *S. purpuratus* gut system of this study. This included an abundance of Epsilonproteobacteria such as *Arcobacter* and *Sulfurimonas* in the gut tissue, Gammaproteobacteria such as *Psychromonas* and *Vibrio*, as well as other common species such as *Propionigenium* in the gut digesta, and a pharynx microbial community resembling the environment. Such selective enrichment and compartmentalization of bacteria in both *S. purpuratus* and *L. variegatus* despite geographical separation support an essential role of specific bacterial taxa to their hosts' health and digestion, and could be further supported by future studies establishing the gut microbiome of laboratory-raised counterparts and/or sea urchins collected at different time-points and locations. Lastly, whether this trend of bacterial enrichment into the sea urchin gut system is the result of (1) naturally occurring microbes in adjacent seawater finding a suitable habitat in the urchin gut environment to flourish, or (2) a host-mediated selective integration of key bacterial microbiota, remains to be verified. However, the similarities identified in taxa across geographical scales suggests that this could be an interesting avenue for future study.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2076-2607/7/2/35/s1>: Figure S1: 2D multidimensional scale (MDS) plot and dendrogram analysis performed using the Bray–Curtis metrics calculated for the top 100 and rare taxa across all samples of the study; Figure S2: 3D principle coordinates analysis (PCoA) plot and dendrogram analysis performed using the weighted Unifrac metrics calculated for the rarefied OTU table data across all samples of the study; Figure S3: 3D principle coordinates analysis (PCoA) plot and dendrogram analysis performed using the unweighted Unifrac metrics calculated for the rarefied OTU table data across all samples of the study; Table S1: The OTU table data to most resolvable level of the merged biological replicate samples of this study (pharynx,  $n = 3$ ; gut tissue,  $n = 3$ ; gut digesta,  $n = 3$ ; water,  $n = 3$ ; and algae,  $n = 3$ ) with taxonomy assigned using the GreenGenes (v13.8) database; Table S2: The unfiltered OTU table data of all samples of this study with taxonomy assigned using the GreenGenes (v13.8) database; Table S3: The QIIME2 (v2018.11) ASV table data generated through DADA2 (v1.10) for all samples in this study with taxonomy assigned using the GreenGenes (v13.8) database; Table S4: Top 100 representative sequences aligned to multiple databases; Table S5: Results of the predicted functional profiles determined through PICRUSt (v1.1.2) and analyzed through STAMP (v2.1.3) of the Kyoto Encyclopedia of Genes and Genomes (KEGG) Level-2 and KEGG-Level-3 categories (gut tissue,  $n = 3$ ; gut digesta,  $n = 3$ ); Table S6: Description of each functional category determined through LEfSe

analysis (LDA threshold at  $\pm 2.4$ ) of the KEGG Ortholog (KO) Ids determined through PICRUSt (v1.1.2) analysis of the gut tissue and gut digesta microbial communities.

**Author Contributions:** J.A.H. and A.K.B. conceived the study, analysis, and interpretation of the data, and wrote the first draft of the manuscript. J.B.S. collected all field data and sea urchin samples, processed the samples for HTS, contributed to the manuscript preparation, reviewed and provided valuable edits/comments. A.W.E.G., S.A.W., and C.D.M. critically reviewed the manuscript, edited and helped interpret the data. C.D.M. and M.R.C. conducted the HTS, helped in data processing, and provided critical comments on the manuscript.

**Funding:** The following are acknowledged for their support of the Microbiome Resource at the University of Alabama at Birmingham, School of Medicine: Comprehensive Cancer Center (P30AR050948), Center for Clinical Translational Science (UL1TR000165), the Heflin Center for Genomic Sciences and the UAB Microbiome Center. Research supported in part by NIH NORC Nutrition Obesity Research Center (P30DK056336) (S.A.Watts). Animal care and use for the experiments conducted in this study was approved by the Oregon Institute of Marine Biology, University of Oregon, and the collection of sea urchins was conducted under the collection permit issued by the Oregon Department of Fish and Wildlife (is: #20366), and the Institutional Animal Care and Use Committee (IACUC-10043; 17 Jan, 2014–current).

**Acknowledgments:** We would like to thank Peter Eipers of the Department of Cell, Developmental and Integrative Biology at the University of Alabama at Birmingham (UAB), for his assistance in high-throughput sequencing for this study. We would also like to thank the Biology Department at UAB for logistics and graduate tuition and stipend support.

**Conflicts of Interest:** The authors declare that this research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

1. Ebert, T.A.; Schroeter, S.C.; Dixon, J.D.; Kalvass, P. Settlement patterns of red and purple sea urchins (*Strongylocentrotus franciscanus* and *S. purpuratus*) in California, USA. *Mar. Ecol. Prog. Ser.* **1994**, *111*, 41–52. [[CrossRef](#)]
2. Dethier, M.N. Disturbance and recovery in intertidal pools: Maintenance of mosaic patterns. *Ecol. Monogr.* **1984**, *54*, 99–118. [[CrossRef](#)]
3. Lafferty, K.D.; Kushner, D. Population regulation of the purple sea urchin (*Strongylocentrotus purpuratus*) at the California Channel Islands. In *Proceedings of the 5th California Islands Symposium*; Browne, D.R., Mitchell, K.L., Chaney, H.W., Eds.; Publication 99-0038; Minerals Management Service: Santa Barbara, CA, USA, 2000; pp. 379–381.
4. Watanabe, J.M.; Harrold, C. Destructive grazing by sea urchins *Strongylocentrotus* spp. in a central California kelp forest: Potential roles of recruitment, depth, and predation. *Mar. Ecol. Prog. Ser.* **1991**, 125–141. [[CrossRef](#)]
5. Metaxas, A.; Scheibling, R.E. Community structure and organization of tidepools. *Mar. Ecol. Prog. Ser.* **1993**, 187–198. [[CrossRef](#)]
6. Tegner, M.J.; Dayton, P.K. Ecosystem effects of fishing in kelp forest communities. *ICES J. Mar. Sci.* **2000**, *57*, 579–589. [[CrossRef](#)]
7. Davidson, T.M.; Grupe, B.M. Habitat modification in tidepools by bioeroding sea urchins and implications for fine-scale community structure. *Mar. Ecol.* **2015**, *36*, 185–194. [[CrossRef](#)]
8. Sea Urchin Genome Sequencing Consortium. The genome of the sea urchin *Strongylocentrotus purpuratus*. *Science* **2006**, *314*, 941–952. [[CrossRef](#)]
9. Lasker, R.; Giese, A.C. Nutrition of the sea urchin, *Strongylocentrotus purpuratus*. *Biol. Bull.* **1954**, *106*, 328–340. [[CrossRef](#)]
10. Holland, N.D.; Ghiselin, M.T. A comparative study of gut mucous cells in thirty-seven species of the class Echinoidea (Echinodermata). *Biol. Bull.* **1970**, *138*, 286–305. [[CrossRef](#)]
11. De Ridder, C.; Jangoux, M. Digestive system: Echinoidea. In *Echinoderm Nutrition*; Jangoux, M., Lawrence, J.M., Eds.; A.A. Balkema Publ.: Rotterdam, The Netherlands, 1982; pp. 213–234. ISBN 90-6191-080-3.
12. Ziegler, A.; Mooi, R.; Rolet, G.; De Ridder, C. Origin and evolutionary plasticity of the gastric caecum in sea urchins (Echinodermata: Echinoidea). *BMC Evol. Biol.* **2010**, *10*, 313. [[CrossRef](#)]
13. Sawabe, T.; Oda, Y.; Shiomi, Y.; Ezura, Y. Alginate degradation by bacteria isolated from the gut of sea urchins and abalones. *Microb. Ecol.* **1995**, *30*, 193–202. [[CrossRef](#)] [[PubMed](#)]

14. Fong, W.; Mann, K.H. Role of gut flora in the transfer of amino acids through a marine food chain. *Can. J. Fish. Aquat. Sci.* **1980**, *37*, 88–96. [[CrossRef](#)]
15. Hakim, J.A.; Koo, H.; Dennis, L.N.; Kumar, R.; Ptacek, T.; Morrow, C.D.; Lefkowitz, E.J.; Powell, M.L.; Bej, A.K.; Watts, S.A. An abundance of Epsilonproteobacteria revealed in the gut microbiome of the laboratory cultured sea urchin, *Lytechinus variegatus*. *Front. Microbiol.* **2015**, *6*, 1047. [[CrossRef](#)] [[PubMed](#)]
16. Hakim, J.A.; Koo, H.; Kumar, R.; Lefkowitz, E.J.; Morrow, C.D.; Powell, M.L.; Watts, S.A.; Bej, A.K. The gut microbiome of the sea urchin, *Lytechinus variegatus*, from its natural habitat demonstrates selective attributes of microbial taxa and predictive metabolic profiles. *FEMS Microbiol. Ecol.* **2016**, *92*, fiw146. [[CrossRef](#)] [[PubMed](#)]
17. Sauchyn, L.K.; Lauzon-Guay, J.S.; Scheibling, R.E. Sea urchin fecal production and accumulation in a rocky subtidal ecosystem. *Aquat. Biol.* **2011**, *13*, 215–223. [[CrossRef](#)]
18. Sauchyn, L.K.; Scheibling, R.E. Degradation of sea urchin feces in a rocky subtidal ecosystem: Implications for nutrient cycling and energy flow. *Aquat. Biol.* **2009**, *6*, 99–108. [[CrossRef](#)]
19. Sauchyn, L.K.; Scheibling, R.E. Fecal production by sea urchins in native and invaded algal beds. *Mar. Ecol. Prog. Ser.* **2009**, *396*, 35–48. [[CrossRef](#)]
20. Schram, J.B.; Kobelt, J.N.; Dethier, M.N.; Galloway, A.W.E. Trophic transfer of macroalgal fatty acids in two urchin species: Digestion, egestion, and tissue building. *Front. Ecol. Evol.* **2018**, *6*, 83. [[CrossRef](#)]
21. Troussellier, M.; Escalas, A.; Bouvier, T.; Mouillot, D. Sustaining rare marine microorganisms: Macroorganisms as repositories and dispersal agents of microbial diversity. *Front. Microbiol.* **2017**, *8*, 947. [[CrossRef](#)]
22. Ley, R.E.; Turnbaugh, P.J.; Klein, S.; Gordon, J.I. Microbial ecology: Human gut microbes associated with obesity. *Nature* **2006**, *444*, 1022–1023. [[CrossRef](#)]
23. Turnbaugh, P.J.; Ley, R.E.; Mahowald, M.A.; Magrini, V.; Mardis, E.R.; Gordon, J.I. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **2006**, *444*, 1027–1031. [[CrossRef](#)] [[PubMed](#)]
24. Brothers, C.J.; Van Der Pol, W.J.; Morrow, C.D.; Hakim, J.A.; Koo, H.; McClintock, J.B. Ocean warming alters predicted microbiome functionality in a common sea urchin. *Proc. R. Soc. B* **2018**, *285*, 20180340. [[CrossRef](#)] [[PubMed](#)]
25. Dabdoub, S.M.; Fellows, M.L.; Paropkari, A.D.; Mason, M.R.; Huja, S.S.; Tsigarida, A.A.; Kumar, P.S. PhyloToAST: Bioinformatics tools for species-level analysis and visualization of complex microbial datasets. *Sci. Rep.* **2016**, *6*, 29123. [[CrossRef](#)] [[PubMed](#)]
26. Caporaso, J.G.; Kuczynski, J.; Stombaugh, J.; Bittinger, K.; Bushman, F.D.; Costello, E.K.; Fierer, N.; Gonzalez Peña, A.; Goodrich, J.K.; Gordon, J.G.; et al. QIIME allows analysis of high-throughput community sequencing data. *Nat. Meth.* **2010**, *7*, 335–336. [[CrossRef](#)] [[PubMed](#)]
27. Faust, K.; Sathirapongsasuti, J.F.; Izard, J.; Segata, N.; Gevers, D.; Raes, J.; Huttenhower, C. Microbial co-occurrence relationships in the human microbiome. *PLOS Comput. Biol.* **2012**, *8*, e1002606. [[CrossRef](#)] [[PubMed](#)]
28. Faust, K.; Raes, J. Microbial interactions: From networks to models. *Nat. Rev. Microbiol.* **2012**, *10*, 538–550. [[CrossRef](#)] [[PubMed](#)]
29. Faust, K.; Raes, J. CoNet app: Inference of biological association networks using Cytoscape. *F1000Research* **2016**, *5*, 1519. [[CrossRef](#)]
30. Berry, D.; Widder, S. Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Front. Microbiol.* **2014**, *5*, 219. [[CrossRef](#)]
31. Banerjee, S.; Schlaeppi, K.; Heijden, M.G. Keystone taxa as drivers of microbiome structure and functioning. *Nat. Rev. Microbiol.* **2018**, *16*, 567–576. [[CrossRef](#)]
32. Song, S.J.; Amir, A.; Metcalf, J.L.; Amato, K.R.; Xu, Z.Z.; Humphrey, G.; Knight, R. Preservation methods differ in fecal microbiome stability, affecting suitability for field studies. *mSystems* **2016**, *1*, e00021-16. [[CrossRef](#)]
33. Caporaso, J.G.; Lauber, C.L.; Walters, W.A.; Berg-Lyons, D.; Lozupone, C.A.; Turnbaugh, P.J.; Fierer, N.; Knight, R. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci. USA* **2011**, *108* (Suppl. 1), 4516–4522. [[CrossRef](#)] [[PubMed](#)]
34. Caporaso, J.G.; Lauber, C.L.; Walters, W.A.; Berg-Lyons, D.; Huntley, J.; Fierer, N.; Owens, S.M.; Betley, J.; Fraser, L.; Bauer, M.; et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* **2012**, *6*, 1621–1624. [[CrossRef](#)] [[PubMed](#)]

35. Kumar, R.; Eipers, P.; Little, R.B.; Crowley, M.; Crossman, D.K.; Lefkowitz, E.J.; Morrow, C.D. Getting started with microbiome analysis: Sample acquisition to bioinformatics. *Curr. Protoc. Hum. Genet.* **2014**, *82*, 18.8.1–18.8.29. [[CrossRef](#)] [[PubMed](#)]
36. Kozich, J.J.; Westcott, S.L.; Baxter, N.T.; Highlander, S.K.; Schloss, P.D. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Environ. Microbiol.* **2013**, *79*, 5112–5120. [[CrossRef](#)] [[PubMed](#)]
37. Ausubel, F.M. *Current Protocols in Molecular Biology*; Ausubel, F.M., Brent, R., Kingston, R.F., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K., Eds.; Publishing Associates and Wiley-Interscience: New York, NY, USA, 1987; ISBN 978-047-162-594-0.
38. Cock, P.J.A.; Fields, C.J.; Goto, N.; Heuer, M.L.; Rice, P.M. The Sanger FASTQ File Format for Sequences with Quality Scores, and the Solexa/Illumina FASTQ Variants. *Nucleic Acids Res.* **2009**, *38*, 1767–1771. [[CrossRef](#)] [[PubMed](#)]
39. Andrews, S. FastQC: A quality Control Tool for High Throughput Sequence Data. 2010. Available online: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc> (accessed on 12 January 2019).
40. Gordon, A.; Hannon, G.J. “FASTX-Toolkit,” FASTQ/A Short-Reads pre-Processing Tools. 2012. Available online: [http://hannonlab.cshl.edu/fastx\\_toolkit](http://hannonlab.cshl.edu/fastx_toolkit) (accessed on 12 January 2019).
41. Edgar, R.C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **2010**, *26*, 2460–2461. [[CrossRef](#)] [[PubMed](#)]
42. Bolyen, E.; Rideout, J.R.; Dillon, M.R.; Bokulich, N.A.; Abnet, C.; Al-Ghalith, G.A.; Alexander, H.; Alm, E.J.; Arumugam, M.; Asnicar, F.; et al. QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science. *PeerJ* **2018**. [[CrossRef](#)]
43. Callahan, B.J.; McMurdie, P.J.; Holmes, S.P. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* **2017**, *11*, 2639–2643. [[CrossRef](#)]
44. Callahan, B.J.; McMurdie, P.J.; Rosen, M.J.; Han, A.W.; Johnson, A.J.A.; Holmes, S.P. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **2016**, *13*, 581. [[CrossRef](#)]
45. Blaxter, M.; Mann, J.; Chapman, T.; Thomas, F.; Whitton, C.; Floyd, R.; Abebe, E. Defining operational taxonomic units using DNA barcode data. *Philos. Trans. Royal Soc. B* **2005**, *1462*, 1935–1943. [[CrossRef](#)]
46. Wang, Q.; Garrity, G.M.; Tiedje, J.M.; Cole, J.R. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **2007**, *73*, 5261–5267. [[CrossRef](#)] [[PubMed](#)]
47. Desantis, T.Z.; Hugenholtz, P.; Larsen, N.; Rojas, M.; Brodie, E.L.; Keller, K.; Huber, T.; Dalevi, D.; Hu, P.; Andersen, G.L. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **2006**, *72*, 5069–5072. [[CrossRef](#)] [[PubMed](#)]
48. McDonald, D.; Price, M.N.; Goodrich, J.; Nawrocki, E.P.; DeSantis, T.Z.; Probst, A.; Anderson, G.L.; Knight, R.; Hugenholtz, P. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J.* **2011**, *6*, 610–618. [[CrossRef](#)]
49. Bokulich, N.A.; Subramanian, S.; Faith, J.J.; Gevers, D.; Gordon, J.I.; Knight, R.; Mills, D.A.; Caporaso, J.G. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nat. Methods* **2013**, *10*, 57–59. [[CrossRef](#)] [[PubMed](#)]
50. De Souza, R.S.C.; Okura, V.K.; Armanhi, J.S.L.; Jorrín, B.; Lozano, N.; Da Silva, M.J.; Imperial, J. Unlocking the bacterial and fungal communities assemblages of sugarcane microbiome. *Sci. Rep.* **2016**, *6*, 28774. [[CrossRef](#)] [[PubMed](#)]
51. Pechal, J.L.; Schmidt, C.J.; Jordan, H.R.; Benbow, M.E. Frozen: Thawing and its effect on the postmortem microbiome in two pediatric cases. *J. Forensic Sci.* **2017**, *62*, 1399–1405. [[CrossRef](#)] [[PubMed](#)]
52. Nearing, J.T.; Douglas, G.M.; Comeau, A.M.; Langille, M.G. Denoising the Denoisers: An independent evaluation of microbiome sequence error-correction approaches. *PeerJ* **2018**, *6*, e5364. [[CrossRef](#)]
53. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [[CrossRef](#)]
54. De Cárcer, D.A.; Denman, S.E.; McSweeney, C.; Morrison, M. Evaluation of subsampling-based normalization strategies for tagged high-throughput sequencing datasets from gut microbiomes. *Appl. Environ. Microbiol.* **2011**, *77*, 8795–8798. [[CrossRef](#)]
55. Pruesse, E.; Peplies, J.; Glöckner, F.O. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* **2012**, *28*, 1823–1829. [[CrossRef](#)]

56. Cole, J.R.; Wang, Q.; Cardenas, E.; Fish, J.; Chai, B.; Farris, R.J.; Kulam-Syed-Mohideen, A.S.; McGarrell, D.M.; Marsh, T.; Garrity, G.M.; et al. The Ribosomal Database Project: Improved alignments and new tools for rRNA analysis. *Nucleic Acids Res.* **2008**, *37* (Suppl. 1), D141–D145. [[CrossRef](#)] [[PubMed](#)]
57. Pruesse, E.; Quast, C.; Knittel, K.; Fuchs, B.M.; Ludwig, W.; Peplies, J.; Glöckner, F.O. SILVA: A comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res.* **2007**, *35*, 7188–7196. [[CrossRef](#)] [[PubMed](#)]
58. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
59. Parks, D.H.; Tyson, G.W.; Hugenholtz, P.; Beiko, R.G. STAMP: Statistical analysis of taxonomic and functional profiles. *Bioinformatics* **2014**, *30*, 3123–3124. [[CrossRef](#)] [[PubMed](#)]
60. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [[CrossRef](#)]
61. Hill, T.C.; Walsh, K.A.; Harris, J.A.; Moffett, B.F. Using ecological diversity measures with bacterial communities. *FEMS Microbiol. Ecol.* **2003**, *43*, 1–11. [[CrossRef](#)] [[PubMed](#)]
62. Marcon, E.; Scotti, I.; Hérault, B.; Rossi, V.; Lang, G. Generalization of the partitioning of Shannon diversity. *PLoS ONE* **2014**, *9*, e90289. [[CrossRef](#)] [[PubMed](#)]
63. Simpson, E.H. Measurement of diversity. *Nature* **1949**, *163*, 688. [[CrossRef](#)]
64. Kruskal, W.H.; Wallis, W.A. Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc.* **1952**, *47*, 583–621. [[CrossRef](#)]
65. Bray, J.R.; Curtis, J.T. An ordination of the upland forest communities of southern Wisconsin. *Ecol. Monograph.* **1957**, *27*, 325–349. [[CrossRef](#)]
66. Anderson, M.J. A new method for non-parametric multivariate analysis of variance. *Austral Ecol.* **2001**, *26*, 32–46. [[CrossRef](#)]
67. Clarke, K.R. Non-parametric multivariate analyses of changes in community structure. *Austral Ecol.* **1993**, *18*, 117–143. [[CrossRef](#)]
68. Oksanen, J.; Kindt, R.; Legendre, P.; O'Hara, B.; Stevens, M.H.H.; Oksanen, M.J.; Suggests, M. The vegan package. *Community Ecol. Package* **2007**, *10*, 631–637.
69. Clarke, K.R.; Gorley, R.N. *PRIMER v6: User Manual/Tutorial (Plymouth Routines in Multivariate Ecological Research)*; PRIMER-E Ltd.: Plymouth, UK, 2006.
70. Lozupone, C.; Lladser, M.E.; Knights, D.; Stombaugh, J.; Knight, R. UniFrac: An effective distance metric for microbial community comparison. *ISME J.* **2011**, *5*, 169–172. [[CrossRef](#)] [[PubMed](#)]
71. gplots: Various R Programming Tools for Plotting Data. Available online: [cran.r-project.org/web/packages/gplots](http://cran.r-project.org/web/packages/gplots) (accessed on 22 October 2018).
72. RColorBrewer: ColorBrewer Palettes. Available online: [cran.r-project.org/web/packages/RColorBrewer](http://cran.r-project.org/web/packages/RColorBrewer) (accessed on 22 October 2018).
73. Segata, N.; Izard, J.; Waldron, L.; Gevers, D.; Miropolsky, L.; Garrett, W.S.; Huttenhower, C. Metagenomic biomarker discovery and explanation. *Genome Biol.* **2011**, *12*, R60. [[CrossRef](#)] [[PubMed](#)]
74. Fisher, R.A. The use of multiple measurements in taxonomic problems. *Ann. Hum. Genet.* **1936**, *7*, 179–188. [[CrossRef](#)]
75. Langille, M.G.; Zaneveld, J.; Caporaso, J.G.; McDonald, D.; Knights, D.; Reyes, J.A.; Clemente, J.C.; Burkepille, D.E.; Thurber, R.L.V.; Knight, R.; et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* **2013**, *31*, 814. [[CrossRef](#)]
76. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498–2504. [[CrossRef](#)]
77. Choma, M.; Bárta, J.; Šantrůčková, H.; Urich, T. Low abundance of Archaeorhizomycetes among fungi in soil metatranscriptomes. *Sci. Rep.* **2016**, *6*, 38455. [[CrossRef](#)]
78. Galton, F. Regression towards mediocrity in hereditary stature. *J. R. Anthropol. Inst.* **1886**, *15*, 246–263. [[CrossRef](#)]
79. Pearson, K. Note on regression and inheritance in the case of two parents. *Proc. R. Soc. Lond.* **1895**, *58*, 240–242. [[CrossRef](#)]
80. Spearman, C. The proof and measurement of association between two things. *Am. J. Psychol.* **1904**, *15*, 72–101. [[CrossRef](#)]

81. Kullback, S.; Leibler, R.A. On information and sufficiency. *Ann. Math. Stat.* **1951**, *22*, 79–86. [[CrossRef](#)]
82. Cover, T.M.; Thomas, J.A. Information Theory and Statistics. In *Elements of Information Theory*; John Wiley and Sons, Inc.: New York, NY, USA, 1991; Volume 1, pp. 279–335. ISBN 0-471-06259-6.
83. Brown, M.B. 400: A method for combining non-independent, one-sided tests of significance. *Biometrics* **1975**, *31*, 987–992. [[CrossRef](#)]
84. Wiese, R.; Eiglsperger, M.; Kaufmann, M. yFiles—Visualization and Automatic Layout of Graphs. In *Graph Drawing Software (Mathematics and Visualization)*; Jünger, M., Mutzel, P., Eds.; Springer: Berlin/Heidelberg, Germany, 2004; pp. 173–191. ISBN 978-3-642-18638-7.
85. Assenov, Y.; Ramírez, F.; Schelhorn, S.E.; Lengauer, T.; Albrecht, M. Computing topological parameters of biological networks. *Bioinformatics* **2007**, *24*, 282–284. [[CrossRef](#)] [[PubMed](#)]
86. Ma, B.; Wang, H.; Dsouza, M.; Lou, J.; He, Y.; Dai, Z.; Brookes, P.C.; Xu, J.; Gilbert, J.A. Geographic patterns of co-occurrence network topological features for soil microbiota at continental scale in eastern China. *ISME J.* **2016**, *10*, 1891. [[CrossRef](#)]
87. Wang, Z.; Bafadhel, M.; Haldar, K.; Spivak, A.; Mayhew, D.; Miller, B.E.; Tal-Singer, R.; Johnston, S.L.; Ramsheh, M.Y.; Barer, M.R. Lung microbiome dynamics in chronic obstructive pulmonary disease exacerbations. *Eur. Respir. J.* **2016**, *47*, 1082–1092. [[CrossRef](#)] [[PubMed](#)]
88. Wang, Y.; Sheng, H.F.; He, Y.; Wu, J.Y.; Jiang, Y.X.; Tam, N.F.Y.; Zhou, H.W. Comparison of the levels of bacterial diversity in freshwater, intertidal wetland, and marine sediments by using millions of Illumina tags. *Appl. Environ. Microbiol.* **2012**, *78*, 8264–8271. [[CrossRef](#)]
89. Wirsén, C.O.; Sievert, S.M.; Cavanaugh, C.M.; Molyneaux, S.J.; Ahmad, A.; Taylor, L.; Delong, E.F.; Taylor, C.D. Characterization of an autotrophic sulfide-oxidizing marine *Arcobacter* sp. that produces filamentous sulfur. *Appl. Environ. Microbiol.* **2002**, *68*, 316–325. [[CrossRef](#)]
90. Sievert, S.M.; Hüglér, M.; Taylor, C.D.; Wirsén, C.O. Sulfur Oxidation at Deep-Sea Hydrothermal Vents. In *Microbial Sulfur Metabolism*; Dahl, C., Friedrich, C.G., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 238–258. ISBN 978-3-540-72682-1.
91. Inagaki, F.; Takai, K.; Kobayashi, H.; Nealson, K.H.; Horikoshi, K. *Sulfurimonas autotrophica* gen. nov.; sp. nov.; a novel sulfur-oxidizing  $\epsilon$ -proteobacterium isolated from hydrothermal sediments in the Mid-Okinawa Trough. *Int. J. Syst. Evol. Microbiol.* **2003**, *53*, 1801–1805. [[CrossRef](#)]
92. Campbell, B.J.; Engel, A.S.; Porter, M.L.; Takai, K. The versatile  $\epsilon$ -proteobacteria: Key players in sulphidic habitats. *Nat. Rev. Microbiol.* **2006**, *4*, 458–468. [[CrossRef](#)] [[PubMed](#)]
93. Durand, L.; Zbinden, M.; Cuffe-Gauchard, V.; Duperron, S.; Roussel, E.G.; Shillito, B.; Cambon-Bonavita, M.A. Microbial diversity associated with the hydrothermal shrimp *Rimicaris exoculata* gut and occurrence of a resident microbial community. *FEMS Microbiol. Ecol.* **2009**, *71*, 291–303. [[CrossRef](#)] [[PubMed](#)]
94. Romero, J.; Garcia-Varela, M.; Laclette, J.; Espejo, R. Bacterial 16S rRNA gene analysis revealed that bacteria related to *Arcobacter* spp. constitute an abundant and common component of the oyster microbiota (*Tiostrea chilensis*). *Microb. Ecol.* **2002**, *44*, 365–371. [[CrossRef](#)] [[PubMed](#)]
95. Ho, T.W.; Hwang, J.S.; Cheung, M.K.; Kwan, H.S.; Wong, C.K. Dietary analysis on the shallow-water hydrothermal vent crab *Xenograpsus testudinatus* using Illumina sequencing. *Mar. Biol.* **2015**, *162*, 1787–1798. [[CrossRef](#)]
96. Urakawa, H.; Dubilier, N.; Fujiwara, Y.; Cunningham, D.E.; Kojima, S.; Stahl, D.A. Hydrothermal vent gastropods from the same family (Provannidae) harbour  $\epsilon$ - and  $\gamma$ -proteobacterial endosymbionts. *Environ. Microbiol.* **2005**, *7*, 750–754. [[CrossRef](#)] [[PubMed](#)]
97. Beinart, R.A.; Sanders, J.G.; Faure, B.; Sylva, S.P.; Lee, R.W.; Becker, E.L.; Gartman, A.; Luther, G.W.; Seewald, J.S.; Fisher, C.R. Evidence for the role of endosymbionts in regional-scale habitat partitioning by hydrothermal vent symbioses. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E3241–E3250. [[CrossRef](#)] [[PubMed](#)]
98. Beinart, R.; Gartman, A.; Sanders, J.; Luther, G.; Girguis, P. The uptake and excretion of partially oxidized sulfur expands the repertoire of energy resources metabolized by hydrothermal vent symbioses. *Proc. R. Soc. B* **2015**, *282*, 20142811. [[CrossRef](#)] [[PubMed](#)]
99. Nichols, D.S. Prokaryotes and the input of polyunsaturated fatty acids to the marine food web. *FEMS Microbiol. Lett.* **2003**, *219*, 1–7. [[CrossRef](#)]
100. Kawasaki, K.; Nogi, Y.; Hishinuma, M.; Nodasaka, Y.; Matsuyama, H.; Yumoto, I. *Psychromonas marina* sp. nov.; a novel halophilic, facultatively psychrophilic bacterium isolated from the coast of the Okhotsk Sea. *Int. J. Syst. Evol. Microbiol.* **2002**, *52*, 1455–1459. [[CrossRef](#)]

101. Groudieva, T.; Grote, R.; Antranikian, G. *Psychromonas arctica* sp. nov.; a novel psychrotolerant, biofilm-forming bacterium isolated from Spitzbergen. *Int. J. Syst. Evol. Microbiol.* **2003**, *53*, 539–545. [[CrossRef](#)]
102. Breeze, J.; Cady, N.; Staley, J. Subfreezing growth of the sea ice bacterium “*Psychromonas ingrahamii*”. *Microb. Ecol.* **2004**, *47*, 300–304. [[CrossRef](#)] [[PubMed](#)]
103. Hosoya, S.; Jang, J.-H.; Yasumoto-Hirose, M.; Matsuda, S.; Kasai, H. *Psychromonas agarivorans* sp. nov.; a novel agarolytic bacterium. *Int. J. Syst. Evol. Microbiol.* **2009**, *59*, 1262–1266. [[CrossRef](#)] [[PubMed](#)]
104. Schink, B. The genus *Propionigenium*. In *The Prokaryotes*; Springer: New York, NY, USA, 2006; pp. 955–959. ISBN 978-0387-25494-4.
105. Smith, P.; Willemsen, D.; Popkes, M.; Metge, F.; Gandiwa, E.; Reichard, M.; Valenzano, D.R. Regulation of life span by the gut microbiota in the short-lived African turquoise killifish. *eLife* **2017**, *66*, e27014. [[CrossRef](#)] [[PubMed](#)]
106. Cardona, E.; Gueguen, Y.; Magré, K.; Lorgeoux, B.; Piquemal, D.; Pierrat, F.; Noguier, F.; Saulnier, D. Bacterial community characterization of water and intestine of the shrimp *Litopenaeus stylirostris* in a biofloc system. *BMC Microbiol.* **2016**, *16*, 157. [[CrossRef](#)] [[PubMed](#)]
107. Thomas, F.; Barbeyron, T.; Tonon, T.; Génicot, S.; Czjzek, M.; Michel, G. Characterization of the first alginate operons in a marine bacterium: From their emergence in marine Flavobacteriia to their independent transfers to marine Proteobacteria and human gut Bacteroides. *Environ. Microbiol.* **2012**, *14*, 2379–2394. [[CrossRef](#)] [[PubMed](#)]
108. Rabus, R.; Ruepp, A.; Frickey, T.; Rattei, T.; Fartmann, B.; Stark, M.; Bauer, M.; Zibat, A.; Lombardot, T.; Becker, I.; Amann, J.; et al. The genome of *Desulfotalea psychrophila*, a sulfate-reducing bacterium from permanently cold Arctic sediments. *Environ. Microbiol.* **2004**, *6*, 887–902. [[CrossRef](#)]
109. Lockhart, C.; Fraga, D. Characterization of the arginine kinase from *Desulfotalea psychrophila* LSv54: The effects of environmental conditions and catalytic domain sequence variations on enzymatic turnover. *FASEB J.* **2007**, *21*, A299.
110. Frøslev, T.G.; Kjølner, R.; Bruun, H.H.; Ejrnæs, R.; Brunbjerg, A.K.; Pietroni, C.; Hansen, A.J. Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. *Nat. Commun.* **2017**, *8*, 1188. [[CrossRef](#)]
111. Konstantinidis, K.T.; Tiedje, J.M. Genomic insights that advance the species definition for prokaryotes. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 2567–2572. [[CrossRef](#)]
112. Stackebrandt, E.; Goebel, B.M. Taxonomic note: A place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int. J. Syst. Evol. Microbiol.* **1994**, *44*, 846–849. [[CrossRef](#)]
113. Tamar, E.; Koler, M.; Vaknin, A. The role of motility and chemotaxis in the bacterial colonization of protected surfaces. *Sci. Rep.* **2016**, *6*, 19616. [[CrossRef](#)] [[PubMed](#)]
114. Thorsen, M.S. Microbial activity, oxygen status and fermentation in the gut of the irregular sea urchin *Echinocardium cordatum* (Spatangoida: Echinodermata). *Mar. Biol.* **1998**, *132*, 423–433. [[CrossRef](#)]
115. Meziti, A.; Kormas, K.A.; Pancucci-Papadopoulou, M.A.; Thessalou-Legaki, M. Bacterial phylotypes associated with the digestive tract of the sea urchin *Paracentrotus lividus* and the ascidian *Microcosmus* sp. *Russ. J. Mar. Biol.* **2007**, *33*, 84–91. [[CrossRef](#)]
116. Guerinot, M.L.; Fong, W.; Patriquin, D.G. Nitrogen fixation (acetylene reduction) associated with sea urchins (*Strongylocentrotus droebachiensis*) feeding on seaweeds and eelgrass. *J. Fish. Res. Board Can.* **1977**, *34*, 416–420. [[CrossRef](#)]
117. Guerinot, M.L.; Patriquin, D.G. N<sub>2</sub>-fixing vibrios isolated from the gastrointestinal tract of sea urchins. *Can. J. Microbiol.* **1981**, *27*, 311–317. [[CrossRef](#)] [[PubMed](#)]
118. Achá, D.; Hintelmann, H.; Yee, J. Importance of sulfate reducing bacteria in mercury methylation and demethylation in periphyton from Bolivian Amazon region. *Chemosphere* **2011**, *82*, 911–916. [[CrossRef](#)] [[PubMed](#)]

119. Roalkvam, I.; Drønen, K.; Stokke, R.; Daae, F.L.; Dahle, H.; Steen, I.H. Physiological and genomic characterization of *Arcobacter anaerophilus* IR-1 reveals new metabolic features in Epsilonproteobacteria. *Front. Microbiol.* **2015**, *6*, 987. [[CrossRef](#)]
120. Becker, P.T.; Samadi, S.; Zbinden, M.; Hoyoux, C.; Compère, P.; De Ridder, C. First insights into the gut microflora associated with an echinoid from wood falls environments. *Cah. Biol. Mar.* **2009**, *50*, 343–352.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).