Supplemental File 3 (S03) for

Integration of microscopy and multi-omic data reveals pathology of a coral disease outbreak

Carly E. Karrick[1,2]*, Hannah L. Swain[3]*, Ashley M. Rossin[4,5], Mallory Kastner[6,7], Gillian Coleman[5,8], Morgan Coleman[5, 9], Lauren I. Howe-Kerr[2,10], Michelle A. Johnston[11], Matthew D. Meyer[12], Marissa F. Nuttall[11], Mia K. Olson[2], Amy Apprill[6], Marilyn Brandt[8], Daniel M. Holstein[5,13], Laura Mydlarz[3], Adrienne M.S. Correa[1,2]

[1]Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA, USA
[2]Biosciences Department, Rice University, Houston, TX, USA
[3]Department of Biology, The University of Texas at Arlington, TX, USA
[4]Cooperative Institute for Marine & Atmospheric Studies, University of Miami, FL, USA
[5]Department of Oceanography & Coastal Sciences, Louisiana State University, Baton Rouge, LA, USA
[6]Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, MA, USA
[7]MIT-WHOI Joint Program in Oceanography/Applied Ocean Science and Engineering, Earth, Atmospheric and Planetary Sciences, Cambridge and Woods Hole, MA, USA
[8]Center for Marine and Environmental Studies, University of the Virgin Islands, USVI, USA
[9]Harte Research Institute, Texas A&M University-Corpus Christi, Corpus Christi, TX, USA
[10]Minderoo Foundation, Perth, WA, Australia
[11]National Oceanic and Atmospheric Administration (NOAA) Office of National Marine Sanctuaries, Flower Garden Banks National Marine Sanctuary, Galveston, TX, United States,
[12]Shared Equipment Authority, Rice University, Houston, TX, USA
[13]School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, NY, USA

*Co-first authors

Materials and methods

Sample collection

Following the disease outbreak, fifteen samples of *Pseudodiploria strigosa* corals were collected in the Flower Garden Banks National Marine Sanctuary at the West Flower Garden Bank (Buoy 2: 27° 52.526 N, -93° 48.836 W and Buoy 5: 27° 52.519 N, -93° 48.889 W) and East Flower Garden Bank (Buoy 2: 27° 54.516 N, 93° 35.831 W) (Figure 1). Samples were collected on SCUBA using a hammer and chisel at a depth range of 18 – 24 m. Corals (N = 10 colonies) were sampled *in situ* as sets of diseased and apparently healthy *P. strigosa*, where two samples were

collected per diseased coral (DD: adjacent to the tissue loss lesion, n = 5; HD: apparently healthy tissue on the same coral, n = 5) and one sample was collected from a neighboring apparently healthy conspecific coral (HH; n = 5). Each sample was divided into three fragments that were preserved for histopathology, transmission electron microscopy (TEM), and DNA/RNA sequencing. Histopathology tissue samples were fixed in zinc-buffered formalin (Z-fix concentrate from Anatech Ltd. diluted with filtered seawater) for 24 hours, washed with seawater for 24 hours, then stored in 70% ethanol for transport to Louisiana State University, Baton Rouge, LA, USA. TEM samples were preserved in 750 µL of 2% PFA in 3X PBS at 4ºC for transport to Rice University, Houston, TX, USA. For DNA/RNA extraction, samples were collected, flash frozen in liquid nitrogen, and stored at -150ºC for transport to the University of Texas at Arlington, Arlington, TX, USA.

Histopathology imaging and analysis

Histopathology samples were decalcified with a 1% HCl EDTA solution and stored in 70% ethanol until further processing. Samples were processed (serially dehydrated and flushed with xylenes) using a Leica ASP6025, embedded in paraffin wax blocks on a Leica EG1150H embedding machine, and sectioned at 5µm thickness on a Leica RM2125RTS microtome. Five serial sections were made ~500 µm apart for each sample. Histological slides were stained with hematoxylin and eosin stain on a Leica ST5020, viewed on an Olympus BX41 microscope with an Olympus SC180 camera attachment, and analyzed using ImageJ software (*ImageJ for Microscopy*, 2007). A quantitative histopathological diagnostic tool was applied to these slides for this analysis. The diagnostic tool used followed the methodology in Rossin et al. (2026), where images from the slides were measured for vacuolization, exocytosis, gastrodermal separation, degraded Symbiodiniaceae, and Symbiodiniaceae and vacuole sizes. We used a one-way analysis of variance (ANOVA) to test for differences in each of the five disease-related measurements among health states.

Transmission electron microscopy (TEM) imaging and analysis

Coral fragments ~3 mm^2 long were decalcified in 1000 µL of 1.5% ascorbic acid for 7-14 days, replacing the ascorbic acid solution daily until the entire skeleton was dissolved. The decalcified sample was fixed in Karnovsky's fixative for at least 24 hours before being post-fixed and processed according to Howe-Kerr et al. (2023). Electron micrographs were imaged using a 120 kV JEOL JEM-1400Flash Transmission Electron Microscope equipped with a high-contrast pole piece and a 15 megapixel AMT NanoSprint15 sCMOS camera. Electron micrographs showed 438 total Symbiodiniaceae cells: 123 from HH tissue, 142 from HD tissue, and 173 from DD tissue. Each cell was scored for the presence/absence of degradation characteristics, abnormalities, and putative viral characteristics. A cell was considered degraded if it contained ≥2 of the following: (i) indistinct organelles (<2 identifiable organelles), (ii) an abnormal cell membrane (oblong or discontinuous), (iii) indistinct thylakoid membranes (blurred together or unevenly spaced), and (iv) large cavities. Other features/abnormalities included the absence of

the pyrenoid body and starch granules. Putative viral characteristics included filamentous virus-like particles (VLPs), masses of filamentous VLPs, spherical/icosahedral VLPs, crystalline arrays of spherical/icosahedral VLPs, crystal structures, viroplasms (both whorled and non-whorled), and multi-membrane vesicles (MMVs). ImageJ (v1.54g) was used to measure the diameters of spherical/icosahedral VLPs and MMVs. Representative images of all features identified can be found in Supplementary Figure 1.

Cells considered too degraded for accurate evaluation (n = 16) were removed from downstream analyses (representative images in Supplementary Figure 2). All statistical analyses were performed in RStudio (v2023.06.2+561). Binomial logistic regressions were used to identify associations between health state and viral and/or degradation characteristics/abnormalities using the glm() function of the stats package (v4.4.1). p-values were calculated using the Anova() function in the car package (v3.1-3; Fox & Weisberg, 2019). For binomial logistic regressions with statistically significant results ($p < 0.05$), post-hoc tests were performed as pairwise comparisons (pairs() function of the graphics package v4.4.1) of estimated marginal means (emmeans() function of the emmeans package v1.11.1; Lenth & Piaskowski, 2026). Pearson correlation tests were run with the cor.test() function in the stats package to assess correlations between viral/degradation characteristics/abnormalities.

Nucleic acid extraction and sequencing

Subsets of the fifteen *P. strigosa* samples were first extracted for total RNA using a pea-sized amount of coral tissue was harvested from the frozen samples using stainless steel sterilized bone cutters. For DD samples, the tissue was taken ~1 – 2cm from the visible lesion line. Extractions were conducted using the Invitrogen RNAqueous Kit (AM1914). Potential DNA and chromatin contamination was removed from each sample using the Ambion DNase I (RNase-free) kit from Invitrogen (Life Technologies AM2222). Total RNA aliquots were sent to Novogene Co., LTD for library preparation and paired-end sequencing. All 15 *P. strigosa* samples had a RNA integrity number (RIN) ≥ 4 and were pre-processed with poly-A tail enrichment to capture all eukaryotic reads. Library prep was completed using the NEBNext Ultra II RNA Library Prep kit by Illumina and samples were processed for 150bp, paired-end sequencing on the Illumina Novaseq 6000 platform with raw read file sizes at a minimum of 5 million.

Upon confirmation of successful RNA extraction, DNA was extracted from the remainder of each coral. Coral samples were airbrushed with a PBS solution to remove coral tissue from the skeleton. The resulting tissue slurry was spun in a centrifuge at maximum speed at 4ºC for 40 minutes or until there was a well-formed pellet with near-transparent supernatant. DNA was extracted from ~0.040 g of the pelleted coral holobiont biomass with the Qiagen DNeasy PowerBioFilm Kit (Qiagen, Inc., Germantown, MD, 24000-50) according to the manufacturer's protocol and stored at -80ºC until further processing. One 50 µL aliquot of extracted DNA was transferred to Rice University and another was transferred to Woods Hole Oceanographic

Institution for sequencing of Symbiodiniaceae ITS-2 rDNA and 16S rRNA genes for bacteria/archaea, respectively.

To characterize Symbiodiniaceae composition, the ITS-2 region of Symbiodiniaceae rDNA was PCR amplified with the Sym_VAR primer set (Sym_VAR_5.8SII and Sym_VAR_REV; Hume et al., 2018). Each 16 µL PCR reaction contained: 1 µL DNA template, 9 µL EconoTaq PLUS GREEN 2X Master Mix, 0.42 µL each of forward and reverse primer at 100 µM, and 5.16 µL DNAse/RNAse-free water. PCR reactions were cycled as: initial denaturation at 94°C for 3 minutes; followed by 35 cycles of 94°C for 75 seconds, 58°C for 75 seconds, and 74°C for 90 seconds; and a final extension at 74°C for 7 minutes, before holding at 4°C. Library preparation and ITS-2 sequencing were performed at Oregon State University's Center for Quantitative Life Sciences (OSU CQLS, Corvallis, OR, USA) with 2x300 bp paired-end reads on an Illumina MiSeq platform. Blanks (nuclease-free PCR-grade water) were used as negative controls during sequencing.

To characterize bacteria and archaea, single barcoded PCR reactions were performed to amplify the V4 region of the small subunit rRNA gene using primers 515FY and 806RB (Apprill et al., 2015; Parada et al., 2015). As a sequencing control, we used genomic DNA from Microbial Mock Community B (even, low concentration), v5.1L, HM-782D, and nuclease-free PCR-grade water for the negative control. Each 50 µL PCR reaction contained: 2 µL DNA template, 0.5 µL GoTaq DNA Polymerase (Promega), 1 µL of each forward and reverse primer at 10 µM, 1 µL of 10 mM deoxynucleoside triphosphate (dNTP) mix (Promega), 5 µL $MgCl_2$, 10 µL GoTaq 5X colourless flexi buffer (Promega), and 29.5 µL nuclease-free water. The PCR cycling conditions were as follows: initial denaturation at 95°C for 2 minutes; followed by 34 or 37 cycles of 95°C for 20 seconds, 55°C for 15 seconds, and 72°C for 5 minutes; and a final extension at 72°C for 10 minutes, before holding at 4°C. We used 34 or 37 cycles for samples to obtain comparable amounts of amplified DNA for subsequent purification and sequencing. Barcoded PCR products were purified using the MinElute PCR Purification Kit (Qiagen, Inc., Germantown, MD), and concentrations were measured using a Qubit 2.0 fluorometer. Purified samples were diluted to 1 ng/µL and pooled for sequencing. Sequencing was performed with 2x250 bp paired-end reads on an Illumina MiSeq platform at the University of Illinois at Urbana-Champaign (Kozich et al., 2013). During analysis of the resulting fastq files, two samples (HD and DD tissues from coral 007) were excluded from analysis due to a mislabeling error during laboratory processing.

Analysis of Symbiodiniaceae composition

The libraries of Symbiodiniaceae ITS-2 rDNA reads were run against the remote SymPortal database (SymPortal.org; Hume et al., 2019) to identify Symbiodiniaceae taxonomy as ITS-2 type profiles. The SymPortal output was imported into RStudio (v2023.06.2+561) with the phyloseq package (v1.48.0). Because multiple Symbiodiniaceae genera were present in these samples, between-genus ITS-2 copy number variation was accounted for by dividing the

absolute abundances by the average ITS-2 copy number estimates reported by Saad et al. (2020), as performed by Howe-Kerr et al. (2023) and Karrick et al. (in review). Per-sample ITS-2 type profile absolute abundance was converted to relative abundance to normalize data prior to visualizing Symbiodiniaceae composition as stacked relative abundance plots with the package ggplot2 (v3.5.2). There were no significant differences in dispersion between tissue health states according to ANOVA tests of PERMDISP2 tests (Df = 2, Sum Sq = 0.08569, Mean Sq = 0.042846, F = 0.2111, p = 0.8127; anova function of the stats package v4.4.1 and betadisper function of the vegan package v2.7-1, respectively). So, to assess whether Symbiodiniaceae ITS-2 type profile composition varied between tissue health states, a global PERMANOVA test (adonis2 function in the vegan package; v2.7-1) was used.

Analysis of bacteria and archaea composition

16S rRNA gene amplicon sequencing data were processed in RStudio (v. 4.4.0) using the DADA2 pipeline (dada2 v. 1.16; Callahan et al., 2016). Demultiplexed fastq files were quality filtered, merged, and used to infer amplicon sequence variants (ASVs). Parameters for filterAndTrim were set as follows: truncLen = 240, 220; maxN = 0; maxEE = 2,2; truncQ = 2; rm.phix = TRUE; compress = TRUE; multithread = TRUE. After filtering and trimming, error rates were learned and the DADA2 sample inference algorithm was applied. Paired reads were merged, chimeric sequences were removed, and taxonomy was assigned with the Silva reference database (v. 138.1, Quast et al., 2012). The resulting ASV count, taxonomy, and metadata tables were imported into phyloseq (v. 1.48.0). Sequences classified as eukaryotic, mitochondrial, or chloroplast were removed.

Variation in beta diversity by health state was assessed with a principal component analysis, using Aitchison distance to account for the compositional nature of sequence data (Gloor et al., 2017). PERMANOVA analysis using adonis2 from the *vegan* package (v. 2.6.0, Oksanen et al., 2026) examined the impact of the health state covariate on community composition. Differences in similarity by health state were assessed with average similarity percentage (SIMPER) which was calculated using PRIMER software (PRIMER-E Ltd, Plymouth, UK). Significant changes in relative abundance of taxa based on health state were tested with the R package *corncob* (v. 0.4.1), which uses a BH global False Discovery Rate (FDR) control approach (Martin et al., 2025). To determine whether these *P. strigosa* corals exhibited a core microbiome, taxa present in >80% of samples within a tissue state and greater than 0.05% relative abundance were examined.

Coral de novo *transcriptome assembly and annotation*

Raw Illumina Novaseq reads generated from the *P. strigosa* RNA were trimmed and quality-filtered using the program FastP (v0.23). Of the fifteen sequenced samples, the six with the highest Q30 values (> 90%) were used to assemble a de novo *P. strigosa* transcriptome. First, a *de novo* metatranscriptome was assembled using Trinity (v. 2.15.1) on the University of Texas

– Arlington's High Performance Computing Cluster (HPC). From this *de novo* metatranscriptome, coral transcripts were identified by extracting the longest isoforms from the Trinity output and performing BLASTn (v. 2.15.0) searches against a diverse master coral database containing genome-derived gene models and transcriptomes. Of the resulting coral-only transcriptome file, reads with less than 90% identity and shorter than 150 bp were removed. BUSCO (v. 5.5.0) was used to assess the quality of the de novo transcriptome. The final coral transcriptome was annotated using BLASTx (v. 2.15.0) against the NCBI Uniprot Database. Transdecoder (v5.5.0) was used to isolate the predicted protein sequences of likely coding regions from the coral-only transcripts. Similar sequences were removed from the predicted protein sequence file using cdhit (v. 4.8.1) to be used in functional enrichment analysis.

Isolation and quantification of target reads

BBsplit (v38.79) was used to isolate coral host and Symbiodinaceae reads using the *Pseudodiploria strigosa de novo* transcriptome and Symbiodinaceae reference transcriptomes from three genera: *Symbiodinium* (Bayer et al., 2012), *Breviolum* (Parkinson et al., 2016)*, Cladocopium* (Davies et al., 2018)*, Durusdinium* (Bellantuono et al., 2019). Reads that did not map to reference transcriptomes were filtered out as "unmatched" and were used for viral community metatranscriptomics. The dominant Symbiodinaceae genus was identified based on the quantity of transcripts mapped to the reference transcriptomes. Coral host and Symbiodiniaceae reads were reformatted into paired end reads to be used for quantification.

Salmon (v1.40) was used for transcript quantification. First, an index was created using the reference transcriptome of *P. strigosa* and *Breviolum minutum* (Parkinson et al., 2016). The quantification step used a quasi-mapping program to map the samples to the created indices. The resulting mapping rates (%) indicated how well the samples mapped to the indices that were created using the reference transcriptomes.

Differential gene expression analysis of coral host and dominant Symbiodinaceae genus

The unannotated *P. strigosa* transcriptome and unannotated *B. minutum* transcriptome were uploaded to RStudio (v2023.12.1+402) and formatted with TxImport (v1.30.0) into a respective gene-level matrices. Raw gene counts from Salmon were used to create a gene count matrix. DESeq2 (v1.42.0) was used to normalize for sequencing depth and differential gene expression analysis, with the design *~Site + Health State.* Gene counts were regularized log (rlog) transformed. Expression profiles for *P. strigosa* and *B. minutum* were filtered, removing transcripts with average expression < 10. Differentially expressed genes (DEGs) were selected based on significance (log2FoldChange > +/- 1 and padj ≤ 0.05) between health states: HH vs HD, HH vs DD, HD vs DD. A principal component analysis (PCA) was performed to identify sample outliers, as well as to visualize the dispersion of the samples across the two principal components using the R-package PCAtools (v2.14.0). DEGs were annotated using the NCBI Uniprot Database.

Analysis of viral composition

The BBsplit-generated "unmatched" reads were normalized with bbnorm.sh within BBMap (v39.09; Bushnell, 2014). *De novo* transcriptome assembly was performed with rnaviralSPAdes using the option --rnaviral in SPAdes (v4.1.0; kmers: 21, 33, 55, 77; Bankevich et al., 2012). MVP (v1.1.4; Coclet et al., 2024)) was used to annotate and map reads to the assembled scaffolds. MVP uses a modular approach, and all modules except Module 07 (optional binning) were used. With the exclusion of Module 07, MVP uses geNomad (v1.7.6; Camargo et al., 2024) for virus taxonomic identification; CheckV (v1.0.1; Nayfach et al., 2021) for both quality filtering based on scaffold length and completeness and average nucleotide identity clustering; Bowtie2 (v2.5.3; Langmead & Salzberg, 2012) for short read mapping and CoverM (v0.7.0; Aroney et al., 2025) for abundance/coverage estimations normalized to reads per kilobase per million reads (RPKM); MMseqs2 (v14.7e284; Steinegger & Söding, 2017) "search" workflow against the PHROGs (v.14; Terzian et al., 2021) and PFAM (v37.0; Paysan-Lafosse et al., 2025) databases and blastp (v2.14.1; Camacho et al., 2009) against RdRp HMM profiles for functional annotation of geNomad-predicted protein sequences predicted with pyrodigal-gv. All MVP modules except modules 01 and 05 were run with default parameters; modules 01 and 05 were run with options --genomad-relaxed and --filtration relaxed, respectively.

The final MVP-generated abundance tables were imported into RStudio (v2023.06.2+561) with the phyloseq package (v1.48.0; McMurdie & Holmes, 2013). Per-sample vOTU composition was visualized at the vOTU, Order, Class, and Realm levels with relative abundance plots generated with ggplot2 (v3.5.0; Wickham, 2016). Per-sample vOTU alpha diversity was measured using Shannon diversity indices calculated with the estimate_richness() function in the phyloseq package. These data were not normally distributed and variances were not homogeneous, as determined by a Shapiro-Wilk test ($W = 0.84166$, $p = 0.01326$; shapiro.test function of the stats package v4.4.1) and Barlett test ($K^2 = 92.117$, $df = 14$, $p = 1.509 \times 10^{-13}$), bartlett.test function of the stats package), respectively. Therefore, a Kruskal-Wallis rank sum test was used to assess significant differences in vOTU alpha diversity between categories. Per-sample beta diversity was measured using Bray-Curtis distance matrices with the vegdist function of the vegan package (v2.7-1; Oksanen et al., 2026), and these measurements were plotted with NMDS plots. Heterogeneous dispersion was not detected in these data based on ANOVA analysis ($df = 2$, Sum Sq = 0.000301, Mean Sq = 0.0001503, $F = 0.0326$, $p = 0.968$; anova function of the stats package v4.4.1) of PERMDISP2 tests performed with the betadisper() function of the vegan package. Using the adonis2() function in the vegan package, a global PERMANOVA test was used to assess differences in vOTU beta diversity between health states. DESeq2 (v1.44.0; Love et al., 2014) was used with the design *~Site + Health State* to assess the presence of significantly differentially abundant vOTUs between tissue health states.

Integration of microbial and viral communities

To perform integrative analyses of bacterial/archaeal (ASV), Symbiodiniaceae (ITS-2 type profile), and viral (vOTU) communities, raw counts were converted to relative abundances and filtered taxa not present at ≥1% abundance in ≥1 sample. After calculating within-dataset alpha (Shannon diversity indices: estimate_richness() function of R package phyloseq v1.48.0) and beta diversity (Bray-Curtis dissimilarity: vegdist() function of R package vegan v2.7-2) metrics, correlation analyses were used to test whether these metrics were correlated across datasets in both global and health-state-specific subsets of the data containing either HH-, HD- or DD-only samples. Spearman correlations comparing within-dataset Shannon indices with the cor.test() function of the R package stats (v4.4.1) were used to test whether richness and evenness were correlated across datasets. The R package ggplot2 (v4.0.0; Wickham, 2016) was used to create scatterplots comparing vOTU and ASV or ITS-2 sequence ~~gene~~ Shannon diversity. Mantel tests comparing within-dataset Bray-Curtis dissimilarity matrices with the mantel() function (permutations = 999, method = "spearman") in the R package vegan (v2.7-2; Oksanen et al., 2025) were used to test whether community dissimilarity was correlated between datasets. Two pairwise Mantel tests were conducted: vOTU-ASV and vOTU-ITS2. Bray-Curtis dissimilarity matrices were ordinated with the cmdscale() function (k = 3, eig = TRUE) in the R package stats (v4.4.1), and these ordinations were plotted with the R package ggplot2 (v4.0.0; Wickham, 2016). Procrustes analyses were used to test whether these ordinations could be superimposed and were performed with the protest() function (permutations = 999) in the R package vegan (v2.7-2; Oksanen et al., 2025) between vOTU-ASV and vOTU-ITS2 ordinations. Procrustes results were plotted with ggplot2 (v4.0.0; Wickham, 2016).

Pairwise correlations and network analyses on a concatenated table of ASV, ITS-2 type profile, and vOTU relative abundance in all samples and HH-, HD-, and DD-only samples were used to assess the significance of putative interactions between microbial taxa. To test whether individual sequences covaried across samples (i.e., to identify proportional relationships between sequences), the propr() function (metric = "rho", symmetrize = TRUE, p = 999) in the R package propr (v5.1.8; Quinn et al., 2017) was used to calculate proportionality metrics between all pairwise combinations of taxa. 999 permutations were used to calculate the false discovery rate (FDR) with the UpdateCutoffs() function (number_of_cutoffs = 100) in propr (v5.1.8; Quinn et al., 2017). The FDR was used to extract top associations where the rho thresholds yielded FDR < 0.05. Heatmaps of features with rho > 0.7 were plotted using the pheatmap() function in the R package pheatmap (v1.0.13; Kolde, 2019). To further explore relationships between microbial taxa, co-occurrence networks were created with the learn_network() function (sensitive = true, heterogeneous = false, n_obs_min = 10 [for global analyses] or 3 [for health subsets]) in the FlashWeave package (v0.19.2; Tackmann et al., 2019) in Julia (v1.10.2; Bezanson et al., 2017). Using the Julia packages DataFrames (v1.6.1; Bouchet-Valat & Kamiński, 2023) and CSV (v.0.10.14), CSV files of network edges and weights were extracted from the .edgelist file created with the save_network() function in FlashWeave for further analysis in R. In R, edges between metadata nodes (n = 20) were filtered out for all downstream analyses to isolate putative

microbe-microbe relationships. A tidygraph object from the network was created using the tbl_graph() function (directed = FALSE) in the package tidygraph (v1.3.1; Pedersen, 2025b) with Louvain community detection using the group_louvain() function of tidygraph to identify clusters of co-occurring microbial consortia or functional modules. The networks were visualized with the ggraph() function (layout = "fr") of package ggraph (v2.2.2; Pedersen, 2025a). Node-level metrics were calculated with the centrality_degree(with and without option weights = absolute value of weights; posweights), centrality_betweenness(weights = posweights), and centrality_closeness(weights = posweights) functions in the package tidygraph v.1.3.1 (Pedersen, 2025b). Network-level statistics were calculated with the edge_density(), modularity(), and transitivity() functions of the package igraph (v.2.2.1; Antonov et al., 2023; Csárdi et al., 2026; Csárdi & Nepusz, 2006) and differences in these metrics were visualized with barplots created using the package ggplot2 (v.4.4.0; Wickham, 2016). Due to low sample sizes, bootstrapping could not be performed to statistically compare network statistics between networks.

Integration of microscopy and metatranscriptomic measures

Regularized log expression (rlog) of *P. strigosa* was integrated with statistically significant disease phenotypes as continuous variables from histopathology (average vacuole area, average Symbiodiniaceae area) and TEM (proportion of spherical/icosahedral VLPs, proportion of multi-membrane vesicles (MMV) and proportion of cells with starch granules) using a Weighted Gene Co-Expression Network Analysis (WGCNA; v1.71; Langfelder & Horvath, 2008). For the coral host, a minimum module size of 50 and soft-thresholding power of 12 was used to identify networks of genes correlated to the above disease phenotypes. Top modules were identified using the following criteria: eigengene > 0.40 and p-values ≤ 0.05. Top modules were combined for each disease phenotype, and used in the STRING Database (v12.0). The *P. strigosa* predicted protein sequence output file from Trandecoder was uploaded to STRING and used in place of a proteome. WGCNA modules were separated by disease phenotype, into positive and negative correlation gene sets. Gene sets were uploaded to STRING with the analysis criteria of protein-protein interactions of high confidence (0.7) and an addition of protein interactions of no more than 20 per gene set. Significant functional enrichments were identified and filtered for enrichments in Biological Process, Molecular Function, KEGG Pathways and Reactome Pathways.

Regularized log expression (rlog) of *Breviolum* was integrated with statistically significant disease phenotypes as continuous variables from histopathology (average vacuole area, average Symbiodiniaceae area) and TEM (proportion of spherical/icosahedral VLPs, proportion of multi-membrane vesicles (MMV) and proportion of cells with starch granules) using a Weighted Gene Co-Expression Network Analysis (WGCNA; v1.71; Langfelder & Horvath, 2008). A minimum module size of 30 and soft-thresholding power of 13 was used to identify networks of genes correlated to the above disease phenotypes. Top modules were identified using the following criteria: eigengene > 0.40 and p-values ≤ 0.05. Top modules were combined for each

disease phenotype, and used in the STRING Database (v12.0). The *Breviolum* predicted protein sequence output file from Trandecoder was uploaded to STRING and used in place of a proteome. WGCNA modules were separated by disease phenotype, into positive and negative correlation gene sets. Gene sets were uploaded to STRING with the analysis criteria of protein-protein interactions of high confidence (0.7) and an addition of protein interactions of no more than 20 per geneset. Significant functional enrichments were identified and filtered for enrichments in Biological Process, Molecular Function, KEGG Pathways and Reactome Pathways.

Histopathology measurements and the proportion of cells with TEM characteristics were integrated into vOTU differential abundance analyses as continuous variables in DESeq2. For histopathology measurements, the average Symbiodiniaceae cell and vacuole areas had means > 5, so these variables were centered and scaled with the scale() function before running DESeq2. The Log2 Fold Change ± standard error for vOTUs in these analyses were visualized using the ggplot2 package (v3.5.2; Wickham, 2016).

Integration of bacteria, archaea and Symbiodinaceae sequencing and RNA-Seq data
An integrative NMDS analysis in RStudio (v2023.06.2+561) was used to assess whether adding omics data layers increased our ability to explain variation in these data. Using a custom R function, tables of per-sample ASV relative abundance, ITS-2 type profile relative abundance, vOTU relative abundance, rlog-transformed host gene expression, and rlog-transformed Symbiodiniaceae gene expression were range-standardized before concatenating all tables into one. Then, Bray-Curtis dissimilarity matrices of all 31 possible combinations of omics layers were created using the combined data and the vegdist() function in the R package vegan (v2.7-2; Oksanen et al., 2025). NMDS ordinations of these dissimilarity matrices were calculated with the metaMDS function (k = 2, trymax = 100) in vegan and plotted using ggplot2 (v.4.0.0; Wickham, 2016). For each ordination/data combination, PERMANOVA tests comparing health states with 999 permutations were run using the adonis2() function in vegan (v2.7-2). Based on ANOVA comparisons (anova() function of package stats v.4.4.1) of PERMDISP2 results (betadisper() function of package vegan v.2.7-2), dispersions were homogeneous, so random subsampling was not performed prior to running PERMANOVAs. To test whether dissimilarity matrices generated from data including multiple omics layers explained a higher proportion of the variance in community dissimilarity between health states, a Kruskal-Wallis test was run to compare the mean R^2 between combinations grouped by the number of omics layers included.