Title: Anatomically-specific coupling between innate immune gene repertoire and microbiome structure during coral evolution

Short title: Coral innate immunity and the microbiome

Classification: Biological Sciences

Minor category: Microbiology

Authors:

Tanya Brown1,2

Dylan Sonett3

Ryan McMinds4

F. Joseph Pollock5

Mónica Medina6

Jesse R. Zaneveld1 #

Affiliations:

1 University of Washington, Bothell, School of Science, Technology, Engineering, and Mathematics, Division of Biological Sciences, UWBB-277, Bothell, WA 98011, USA

2 Department of Biology, University of Texas at Tyler, Tyler, TX 75799, USA

3 School of Pharmacy, University of Washington, Seattle, Washington, USA

4 Center for Global Health and Infectious Diseases Research, University of South Florida, 13201 Bruce B. Downs Blvd, MDC 56, Tampa, FL 33612, USA

5 Hawai՛i & Palmyra Program, The Nature Conservancy, Honolulu, HI, USA, 96817

6 Pennsylvania State University, Department of Biology, 208 Mueller Lab, University Park, PA 16802, USA

# Corresponding author

Correspondence to: Jesse Robert Zaneveld

University of Washington, Bothell, Department of Biology, UWBB-277, Bothell, WA 98011, USA

Telephone: 541-760-2411; Fax: 541-737-0496

Email: zaneveld@gmail.com

Keywords: coral, microbiome, coevolution

## Abstract

Background:

Tropical reef-building corals exist in intimate symbiosis with diverse microbes and viruses. Coral microbiomes vary greatly in biodiversity. It has previously been hypothesized that differences in coral innate immunity in general, and the copy number of innate immune genes containing Toll/Interleukin Repeat (TIR)-domains in particular, may drive interspecific differences in microbiome structure. Despite many existing studies of coral microbiomes, this hypothesis has previously been difficult to test due to a lack of consistently collected cross-species data on coral microbiomes.

Results:

In this manuscript, we reannotate TIR-domain containing genes across diverse coral genomes, and use phylogenetic comparative methods to compare these innate immune gene copy numbers against 16S rRNA marker gene data on coral mucus, tissue, and skeleton microbiomes from the Global Coral Microbiome Project (GCMP). The copy number of Toll-like receptor (TLRs) and Interleukin-1 receptor (IL-1Rs) gene families, as well as the total genomic count of their constituent domains (LRR and TIR domains; and Ig and TIR domains, respectively), explained most overall differences in microbiome richness and beta-diversity among coral species with sequenced genomes. We find that these correlations are also anatomically specific, with an especially strong correlation between IL-1R gene copy numbers and microbiome richness in the coral’s endolithic skeleton.

Conclusions:

During coral evolution, several gene families involved in innate immunity have undergone gene family expansions and contractions. Our results suggest that changes in the copy number of TLR and IL-1R genes may have played a key role in sculpting modern microbiome structure in corals.

## Background

The 1681 described species of scleractinian corals [1] are environmentally critical ecosystem engineers that underpin many tropical reef ecosystems. Microbiomes are important contributors to the health of these tropical corals, with competing and cooperating microbes influencing animal health [2]. Therefore, evaluating how corals regulate their microbiomes is of great importance. Numerous studies have uncovered important features of coral microbiomes, including the relative influence of differences across host anatomy [3,4], between species [3,5], among reefs [6–8], and along environmental gradients [8,9]. This literature has also extensively documented coral microbiome responses to various stressors, including heat [10,11], bleaching [12,13], sedimentation [14,15], nutrient pollution [10,15,16], predation [16], plastic pollution [17], turf or macro- algal competition [16,18,19], etc. Genetic studies within coral species have further found genotypic differences that correlate with microbiome composition [20,21]. Building on these ecological and population-genetic comparisons, specific coral microorganisms have been linked to important host health outcomes, such as protection against pathogens or susceptibility to them [22]. Recent microbiome manipulation experiments have even begun to establish the causal role of specific coral-associated bacteria in influencing key host traits like heat resistance [23]. Despite this thriving literature on coral microbiomes, the broader scale patterns of how modern coral microbiomes have evolved, and which host traits, if any, drive the large differences in microbiome structure and function seen between modern corals is not yet clear.

Increased attention to the question of how host traits have sculpted coral microbiomes over evolution is important. Comparative studies of coral microbiome evolution may identify host traits that have regulated coral microbiomes up to the present day. Comparative studies of coral microbiome evolution will also clarify a key part of the broader story of animal microbiome evolution. Vertebrate gut microbiomes are structured by both host phylogenetic relatedness and convergently evolved host traits like diet or flight [24–26]. One recent large-scale analysis of animal and plant microbiomes from the Earth Microbiome Project identified host factors [27], including the complexity of the adaptive immune system as key to shaping internal (digestive system) microbiomes, while external environmental and climatic factors, such as latitude, precipitation, and regional vegetation shaped external microbiomes such as plant surfaces [27]. At a finer-scale, genome-wide association studies have identified several mammalian and plant genes associated with microbiome structure and complexity [28]. Experimental work in gnotobiotic mice demonstrates that microbiome structure in turn can influence how diet modulates host immunity [29].

As diverse and early-diverging animals that lack traditional adaptive immunity, live in varied marine environments, can feed via predation and/or symbiosis and associate with complex microbiomes both internally and externally, the traits shaping the microbiome evolution of corals are less clear.

Among the many coral traits that could influence microbiome structure, differences in innate immunity are promising candidates, since they have the potential to directly regulate the microbiome by activating pathways that preferentially target particular groups of microbes. Results from studies of innate immunity in mammals and plants lend support to this idea. For example, using Toll-like receptor (TLR) 5 deficient compared to wild type mice, TLR genes have been correlated with membership of the mouse gut microbiome and host-microbial interactions with specific microbes, although the overall effect of TLR on gut microbiome β-diversity remains controversial (reviewed in [28]). Similarly, *Arabidopsis* with loss of function mutations in the pattern recognition receptor (PRR) gene FLS2, have altered rhizosphere microbiome β-diversity relative to wild type controls [30]. However, more research is needed to test if innate immunity drives changes in microbiome biodiversity over long periods of evolution.

In corals, the copy number of gene families containing Toll/Interleukin Repeat (TIR) domains in particular, has been proposed as a possible influence on microbiome structure [31]. TIR domains are a key intracellular signaling domain, found in multiple innate immune gene families, such as Toll-like Receptors (TLR, composed of TIR and leucine rich repeat (LRR) domains), Interleukin-1 Receptors (IL-1R, composed of TIR and immunoglobulin (Ig) domains), coral-specific TIR-only genes of unknown function, and Myeloid Differentiation Factor 88 (myD88). Collectively, these genes are known as TIR-domain containing genes [32]. In known invertebrate examples, the TIR domain serves as a key link in signaling cascades that trigger immune responses like melanization [33,34]. For example, in Toll-like receptors, after a microbe-associated molecular pattern (MAMP) has been detected by the TLR’s extracellular leucine rich repeats (LRR domains), it is the intracellular TIR domain that is required for transduction of that signal and the animal’s ultimate immune response [35].

While the molecular biology of TIR-domain-containing proteins in corals has not yet been independently confirmed, much research has been done on the mechanisms by which the homologues of TIR domain-containing proteins function in human innate immunity (reviewed in [36]). In us, TLRs dimerize, with different TLR hetero- or homo- dimers detecting different MAMPs (e.g. TLR1/TLR2 heterodimers detect triacylated lipoproteins from bacteria, TLR9 homodimers detect unmethylated CpG dinucleotides in bacterial or viral DNA, etc). Once bound to their target MAMP ligand, TLR dimers change conformation, bending in such a way as to bring their intracellular TIR domains nearby one another. This allows binding by five key adaptor proteins (myD88, MAL, TRIF, TRAM, & SARM), recruitment of protein kinases, and activation of transcription factors like NF-κB [36].

The genomic copy number of some gene families of TIR-domain containing genes is known to vary greatly between coral species[31,37], and on this basis these genes have been hypothesized to influence cross-species differences in coral microbiome structure [31]. However, a lack of consistently collected cross-species microbiome data has so far precluded empirical testing of this intriguing idea.

Given that only a tiny fraction of described sclearactinian species have sequenced genomes, analyses comparing innate immunity must be mindful to maintain statistical power. We reasoned that a targeted comparative analysis testing the previously published hypothesis that TIR-domain containing gene family expansions have altered coral microbiome richness would have much greater statistical power than a more general comparison involving all innate immune gene families, as it would involve relatively few comparisons.

In this study we tested whether changes in the copy number of coral TLR or IL-1R gene families have driven corresponding changes in microbiome richness, evenness, or composition during more than 250 million years of scleractinian coral evolution. Our analysis combined comparative genomics of all major lineages of scleractinian coral for which genomes are publicly available with data on coral mucus, tissue, and endolithic skeleton microbiomes from the Global Coral Microbiome Project (GCMP) dataset [4,38] — a collection of more than 1440 16S rRNA libraries from diverse coral species. We then used established phylogenetic comparative methods to test whether coral species’ TLR or IL-1R gene families correlated with their microbiome structure or composition.

The results identify potential drivers of microbiome structure among sequenced corals, and highlight the value of comparative analysis for supporting or refuting whether specific host traits underlie differences in microbiome structure between animal species. They further support the idea that different regions of coral anatomy differ not just in microbiome composition, but also in responsiveness to host traits. While we had hypothesized that coral tissue microbiomes would be most responsive to changes in host innate immune gene repertoire over evolution, the portion of coral anatomy that most strongly correlated with innate immunity proved to be a surprise.

## Methods

**Coral sampling and microbiome analysis**

The Global Coral Microbiome Project (GCMP) coral mucus, tissue and skeleton samples reanalyzed here were originally collected and processed following the methods outlined in Pollock *et al.,* 2018 [4]. Those methods are briefly restated here. All coral samples were collected by AAUS-certified scientific divers, in accordance with local regulations. These plus additional international samples were then resequenced with protocols standard for the Earth Microbiome Project [38]. Bacterial and archaeal DNA were extracted using the PowerSoil DNA Isolation Kit (MoBio Laboratories, Carlsbad, CA; now Qiagen, Venlo, Netherlands). To select for the 16S rRNA V4 gene region, polymerase chain reaction (PCR) was performed using the following primers with illumina adapter sequences (underlined) at the 5’ ends: 515F [39] 5′− TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG GTG YCA GCM GCC GCG GTA A −3′ and 806R [40] 5’− GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GGG ACT CAN VGG GTW TCT AAT −3′. PCR, library preparation and sequencing on an Illumina HiSeq (2x125bp) was performed by the EMP [38,41]. The resulting 16S rRNA amplicon sequences from the GCMP are available in the Qiita database Qiita (CRC32 id: 8817b8b8 and CRC32 id: ac925c85).

**Genomic Data Acquisition**

Reference coral genomes were selected based on completeness and overlap with the Global Coral Microbiome Project dataset. Twelve coral genomes were downloaded from NCBI (<https://www.ncbi.nlm.nih.gov/>) and Reef Genomics (<http://reefgenomics.org>).

**Genomic and Microbiome Metadata**

A list of urls for the genomes can be found in **Table S1A**. Metadata for all GCMP samples are available in **Table S1B,** while a data dictionary defining the meaning of column names is available in **Table** **S1C**.

**Domain Annotation**

We annotated TLR and IL-1R gene counts based on their domains (**Table S2A**). The downloaded coral genomes were used to locate TIR, leucine rich repeats (LRR), and immunoglobulin (Ig) containing genes using a custom pipeline available on github (https://github.com/zaneveld/GCMP\_genomics). Genome analysis occurred in two steps. First, genomes were analyzed using TransDecoder (v5.5.0) (<https://github.com/TransDecoder/TransDecoder/wiki>) to identify candidate coding regions within the genome files. During this process, TransDecoder converts the nucleotide sequences in the genome file to possible amino acid sequences using different open reading frames. In many cases more than one reading frame was possible. The predicted protein products of these different hypothesized reading frames are referred to as isoforms in transdecoder and also in this manuscript. The peptide file generated from TransDecoder was next searched with HMMER (hmmer.org, HMMER 3.3.2 (November 2020)) to locate putative TIR domains as well as known TIR-associated domains. The TIR, LRR, and Ig associated peptide alignment files used in this search were downloaded from the Pfam website (<http://pfam.xfam.org/>). The Pfam domain alignment files were used to build profiles in HMMER using hmmbuild by reading in the alignment file and creating a new Hidden Markov Model (HMM) profile. The HMM profiles were then used to search the TransDecoder peptide files, preserving hits with e-values < 1e-2. This liberal threshold was chosen since multiple domain annotations together (at minimum 2) were required to define TLR or IL-1R genes (see below). Because at least two independent false positive domain annotations would be needed to generate a false positive TLR or IL-1R annotation, we reasoned that this threshold gave a combined e-value of ~1e-4 for TLR or IL-1R annotations. The output file from hmmscan contained significant matches between the HMM profile and the transdecoder peptide file. The resultant hmmscan file was used to count the number of TIR-domain containing sequences found for each organism.

TIR-domain containing genes were further subdivided by scanning each for LRR or Ig domains using the same procedure. TIR-domain-containing genes were then subdivided into categories based on their domains: TIR only (defined by TIR domain and no LRR or Ig domains), TLR (TIR and one or more LRR domains), and IL-1R (TLR and one or more Ig domains) genes. Numbers of each type of gene, and average numbers of Ig or LRR domains within each type of gene were used in further analysis. The TIR domain was detected with the Pfam alignment PF01582. The presence of Pfam families across genes and ids for the 13 alignment files used to detect Ig or LRR domains, as well as the TIR domain alignment are found in **Table S2B, C**.

**Confirmation of uniqueness of annotated IL-1R and TLR copies.**

CD-hit clustering was used to check for redundancy among our annotations of IL-1R and TLR gene copies. First, the longest predicted peptide from each ORF was identified by TransDecoder. These were then clustered using CD-hit, with the default clustering threshold of 90%. Few predicted ORFs clustered together, but there were some examples where this analysis resulted in 3 or fewer ORFs that clustered together in certain genomes, indicating that a minority of predicted ORFs could originate from the same biological gene. To check if this could have impacted our assessment of TLR and IL-1R copy, CD-hit clusters were cross-referenced with gene ids for TLR and IL-1R, confirming that the few potentially redundant annotations were not TLR or IL-1R genes.

**16S library preparation, sequencing, and initial quality control**

16S rRNA sequencing data were processed in Qiita [42] using the standard EMP workflow. Briefly, sequences were demultiplexed based on 12bp Golay barcodes using “split\_libraries” with QIIME 1.9.1 default parameters [43] and trimmed to 100bp to remove low quality base pairs. Quality control (e.g., denoising, de-replication and chimera filtering) and identification amplicon sequence variants (ASVs) were performed using deblur 1.1.0 [44] with default parameters. The resulting biom and taxonomy tables were obtained from Qiita (CRC32 id: 8817b8b8 and CRC32 id: ac925c85) and processed using a customized QIIME2 v. 2020.8.0 [45] pipeline in python (github.com/zaneveld/GCMP\_global\_disease).

**Mitochondrial annotation and quality control**

Taxonomic assignment of ASVs was performed using vsearch [46] using a modified version of the SILVA v. 138 [47] taxonomic reference. The sole change to the SILVA138 reference was supplementation with diverse additional mitochondrial rRNA sequences obtained from METAXA2 [48,49]. In benchmarks, this change greatly improves annotation of coral mitochondrial rRNAs, without increasing false positive taxonomic assignments [49]. This expanded taxonomy is referred to as “silva\_metaxa2” in code. After taxonomic assignment, all mitochondrial and chloroplast reads were removed (**Table S3**).

The bacterial phylogenetic tree was built using the SATé-enabled phylogenetic placement (SEPP) insertion technique with the q2-fragment-insertion plugin [50], again using the SILVA v. 138 [47] database as reference taxonomy and phylogeny. The final output from this pipeline consisted of a taxonomy table, ASV feature table and phylogenetic tree that were used for downstream analyses.

Unrarefied feature tables were used in compositional data analyses (e.g. ANCOM-BC). In phylogenetic comparative analyses that are not compositional (e.g. phylogenetic independent contrasts) feature tables were rarefied to 1000 randomly-selected sequences per sample in order to avoid potential false-positives due to differences in library size across samples. These rarefied samples were then further pooled to form an average microbiome for the mucus, tissue, and skeleton of each species for phylogenetic analysis. Pooled sequences for each sample were again rarefied to 1000 sequences/sample in order to avoid species with more samples in the GCMP dataset from having artifactually higher microbiome biodiversity, which we reasoned could lead to false positive correlations with host traits. Rarefaction depth was chosen to maximizing taxonomic coverage (that is, to exclude as few samples as possible), and to allow direct comparison with prior analyses of GCMP microbiome samples (e.g. [4]). Examination of sequencing depths across the GCMP suggested that using a substantially higher rarefaction depth would result in loss of many samples in key species. For example, using a rarefaction depth of 10,000 sequences would reduce biological replication for *Acropora hyacinthus* to 2 (vs 10 at 1,000 sequences). However, such loss of biological replication might be worth it if it dramatically improved estimates of which species had the most vs. least diverse microbiomes. The effects of rarefaction depth on estimates of alpha diversity were therefore tested in a control analysis, which confirmed that rarefaction depth had little impact on the relative microbiome richness of coral species in this dataset.

**Tests of the effect of rarefaction depth**

While rarefaction will not induce false positive results, it can reduce statistical power. We therefore tested whether our use of a rarefaction depth of 1000 sequences per sample might bias microbiome results. We used samples with high sequencing depth (>100,000 sequences per sample) as an internal control to test the effect of sequencing depth on estimated relative alpha diversity. We emphasize relative alpha diversity across species rather than absolute alpha diversity because the focus of our analysis was correlating TIR containing gene repertoires with relative alpha diversity across coral species.

Coral skeletal samples were used for this analysis as skeleton is the compartment where TIR containing gene repertoires best correlated with microbiome diversity in our results. Sequence pools for each species were first rarefied to 100,000 reads and then those same samples were re-rarefied to 1,000 reads. If rarefaction depth biases our estimates of relative coral microbiome biodiversity, then the samples with the greatest apparent microbiome richness in the 100,000 read rarefaction should be only weakly correlated with apparent microbiome richness of the same samples at 1,000 reads. This was tested in both raw terms (correlation between observed features at 1k verses 100k reads) and by rank (correlation between rank observed features at 1k verses 100k reads).

**Phylogenetic comparison of innate immune gene repertoire and microbiome richness**

Genome TIR domains were compared to the microbiome data collected from the GCMP using phylogenetic independent contrasts analysis. This comparative method accounts for the non-independence of traits among species due to shared ancestry, and thereby avoids the high rates of false positive results that can result if phylogeny is ignored in cross-species comparisons. The α and ꞵ-diversity of the microbial community were compared to the TIR, TLR, and IL-R domain copy numbers using phylogenetically independent contrasts (PICs) with the phytools package in R [51]. In the PIC analysis, ordinary least squares (OLS) linear regression with the lm R function was used for regression of the contrasts. PIC regression results were manually checked for outliers and it was found that no results exceeded a Z-score threshold of +/- 3.

In α-diversity analysis of coral microbiomes, PICs were used to account for phylogeny while comparing corals genomic traits (e.g. the genomic copy number of IL-1R or TLR genes) against either richness (ln ASVs/1000 reads) or evenness (assessed with the Gini Index). For ꞵ-diversity analysis of coral microbiomes, Weighted or Uniweighted UniFrac ꞵ-diversity matrices were constructed, and then ordinated with Principal Coordinates Analysis (PCoA). The top 3 resulting PC axes (PC1, PC2, and PC3) were then extracted and regressed against genomic traits like IL-1R copy number or TLR copy number. Code used for this analysis is found on github (github.com/zaneveld/GCMP\_global\_disease).

**Microbial taxonomic analysis**

Microbial taxonomic analysis was conducted using the ANCOM-BC package in phyloseq in R [52]. Analyses were conducted at the class and family level to test for microbes whose abundance correlated with IL-1R and TLR copy numbers. To correct for multiple comparisons, we used an FDR cutoff of q < 0.05. Lists of significant microbes (those with a q < 0.05) were generated for mucus, tissue, skeleton, and all compartments. Because ANCOM-BC accounts for the compositionality of microbiome data, but not host phylogeny, we further confirmed the association of each putatively significant microbe identified by ANCOM-BC (at q < 0.05) using regression of phylogenetic independent contrasts for that microbe’s relative abundance against IL-1R or TLR gene copy number, and only retained those that also showed a significant assocation in PIC analysis.

Correlation heatmaps were constructed to visualize correlations between the relative abundance of particular microbes and IL-1R or TLR copy number. These took as their input the the coefficient data generated by ANCOM-BC, representing the correlation of each microbe with either TLR or IL-1R gene copy number. The correlation coefficient for each microbial taxon in each coral anatomical compartment (mucus, tissue or endolithic skeleton) was then compiled using custom python scripts and visualized using the matplotlib python library. Microbes are reported in the heatmap and manuscript text only if they were significantly associated with gene copy number by both ANCOM-BC and confirmatory phylogenetic regression as described above in at least one compartment. The color bar shows raw correlation coefficient values from ANCOM-BC.

Heatmaps to visualize microbial relative abundance in the Supplementary Material were created in the phyloseq R package. Colors depict the count of each microbial taxon in samples rarefied to 1000 sequences per sample, and assigned to colors using a log scale (we used plot\_heatmaps default log base 4 scaling). Samples were ordered by the IL-1R or TLR gene copy number of the species from which the sample was taken while microbes were ordered by ordination analysis using the phyloseq plot\_heatmap function’s “RDA” clustering method.

## **Tests controlling for effects of genome quality**

In order to control for potential effects of genome quality on regression results, Benchmarking Universal Single Copy Orthologs (BUSCO) scores were calculated for each genome. These scores reflect the proportion of single-copy orthologs present in all the vast majority of finished genomes that are also present in a coral genome, with low BUSCO scores suggesting low coverage or incomplete assembly. For phylogenetic correlations that appeared interesting based on initial analysis (e.g. skeleton microbiome richness vs. IL-1R gene copy number), the proportion of complete, single-copy orthologs (“BUSCOs”) in each genome was then added as an additional factor in multiple linear regression of phylogenetic independent contrasts for our dataset, in order to assess if observed correlations between innate immune gene repertoire and microbiome structure could be attributed to differences in genome quality.

## Results

Sequenced coral genomes vary greatly in TIR-domain containing gene copy number.

Previous analyses have reported significant variation in the copy number of TIR-domain containing genes among coral genomes. We annotated the genomic copy number of TIR-domain containing genes in 11 coral genomes that were also represented in microbiome data from the Global Coral Microbiome Project (GCMP) (**Table S1**). While a prior study analyzed both genomes and transcriptomes in order to maximize discovery of new TLR or IL-1R homologs [53], we chose to exclude transcriptomes from our analysis in order to prevent any potential confounding effects of some innate immune genes not being expressed in any given transcriptome. As a result of these annotations (Methods), we identified numerous TLR, IL-1R and TIR-only genes (**Table S2, 4**). Despite these methodological differences, our annotations mostly agree in trend and rank with prior studies of coral innate immune repertoires [37,53].

Many innate immune genes have modular structures based on the domains that they contain, and TIR domains commonly co-occur with several other domains within innate immune genes. We further subdivided TIR-domain containing genes based on the other domains present: genes with both TIR and immunoglobulin (Ig) domains were annotated as interleukin-1 receptors (IL-1R) while genes with both TIR domains and leucine rich repeat (LRR) domains were annotated as toll-like receptors (TLR) (**Fig. 1**). We also analyzed the total count of TIR domains, regardless of which gene they were part of and the presence or absence of other co-associated domains. Finally, clustering analysis was used to confirm that no two annotated TLR or IL-1R gene copies within a genome shared ≥ 90% sequence similarity, indicating that these annotated gene copies are non-redundant.

In keeping with past work [37,53] we find that both the copy number of TLR, IL-1R and TIR-only genes, and the total abundance of their component TIR, LRR and Ig domains varies greatly across coral genomes (**Table 1, Fig. 2, Fig. S1**).

**Coral IL-1R and TLR gene family copy numbers correlate with overall microbiome richness and evenness.**

In order to determine whether TLR or IL-1R might regulate microbiome biodiversity, we used phylogenetic independent contrasts (PIC) analysis to correlate changes in the genomic copy number of IL-1R and TLR against changes in microbiome richness (**Table S5A**) or evenness (**Table S5B**) during coral evolution.

We measured microbiome richness as the natural log of the observed number of amplicon sequence variants (ASVs) per 1000 sequences (**Table 1, Table S5A**) (see Methods). In control analyses using a subset of deeply sequenced samples, the relative microbiome richness of coral species was highly correlated across rarefaction depths of 1000 or 100,000 sequences per sample (Pearson R2 = 0.95, p = 3.13 x 10-25; Spearman R2 = 0.989, p=3.51 x 10-37). Thus, we chose to maximize inclusion of samples by including all Global Coral Microbiome Project samples with 1000 or more reads.

Microbiome richness was significantly reduced by increases in the copy number of IL-1R (PIC R2 = 0.835, qFDR = 0.00055; **Fig. 2B, D**; **Fig. S1C, D**; **Table S6A**), or TLR (PIC R2 = 0.707, qFDR = 0.0076; **Fig. 2 C, D**; **Fig. S1E, F**; **Table S6A**). Thus, corals that harbor more TLR or IL-1R gene copies tend to have less diverse microbiomes, and vice versa. In manual inspection of the data, there were several striking examples of this statistical trend. In the genus *Porites*, *P. lutea* had more than twice as many IL-1R and TLR gene copies as the closely related *P. rus* (14 IL-1R copies and eight TLR gene copies in *P. lutea* vs. six IL-1R copies and two TLR copies in *P. rus*; **Table 1**). Consistent with the idea that *P. lutea’*s expanded gene repertoire may play a role in microbiome filtering, *P. lutea’s* microbiome was roughly half as diverse as that of *P. rus* (56 ASVs/1000 reads vs. 114 ASVs/1000 reads, across all sample types). However, we caution that such pairwise species comparisons, although intriguing, are less reliable than the overall correlation (e.g. the 1000 reads from *P. lutea* were derived from only 2 microbiome samples per compartment, while the 1000 reads for *P. rus* were from 15 microbiome samples per compartment).

We also quantified the microbiome evenness of each coral species in the analysis using the Gini Index (**Table S5B, Fig. S2**), which takes on its highest value when ecological communities are least even. Gini index scores were strongly and significantly negatively correlated with the genomic copy number of IL-1R genes (PIC R2 = 0.876, qFDR = 0.000357; **Fig. S2C, D**; **Table S6A**) and TLR genes (PIC R2 = 0.805, qFDR = 0.0014; **Fig. S2E, F**; **Table S6A**), indicating that corals with more IL-1R or TLR gene copies have higher evenness.

These results support the prior hypothesis that IL-1R and TLR gene family expansions may influence coral microbiome structure [53]. The magnitude of the effect was, however, quite surprising, with IL-1R gene copy number explaining ~83% of the variation in microbiome richness, and ~88% of the variation in microbiome evenness, among coral species in the analysis. Although necessarily limited to coral species for which genomes are available, these correlations are much stronger than the effects of several biotic and abiotic factors previously tested in the same GCMP dataset, including depth, temperature, and turf-algal contact [4].

**Endolithic skeleton microbiomes drive correlations between TLR and IL-1R gene family expansion and microbiome structure.**

To this point, all our results were conducted by correlating genomic features of corals against the overall microbiome diversity of all available microbiome samples from each species in the GCMP dataset. However, microbiome richness has been shown to vary between coral compartments [4]. We expected that host tissues — where the immune system could most obviously act — would have microbiomes that most closely correlate with the innate immune gene repertoire of the host.

To test this idea, we separated coral microbiome samples into those that derive from mucus, tissue or endolithic skeleton, and repeated tests for correlations between coral innate immune repertoires and microbiome richness and evenness within each of those regions of anatomy.

Contrary to our hypothesis, coral endolithic skeleton — not tissue or mucus — was the sole driver of correlations between coral innate immune repertoire and microbiome richness (**Fig. 2E-G**; **Fig. S1, 3, 4, 5**; **Table S6 – 8**) and evenness (**Fig. S2, 6, 7, 8**; **Table S6 – 8**).

Microbiome richness in coral endolithic skeleton was significantly correlated with IL-1R (R2 = 0.944, qFDR = 2.05 x 10-5; **Fig. 2F**, **Fig. S5C, D**; **Table S6D**), and TLR (R2 = 0.90798, qFDR = 9.36 x 10-5; **Fig. 2G**; **Fig. S5E, F**; **Table S6D**) gene copies, whereas tissue (**Fig. S4**; **Table S6C**) and mucus (**Fig. S3**; **Table S6B**) microbiome richness was not. Microbiome evenness showed similar patterns (**Fig. S6-8**; **Table S6 – 8**). We confirmed these results using multiple linear models that modeled IL-1R or TLR copy number based on the additive effects of microbiome richness in coral mucus, tissue, and skeleton (e.g. IL-1R ~ richness\_mucus + richness\_tissue + richness\_skeleton). Skeleton microbiome richness alone was a significant predictor of both IL-1R and TLR copy number in post-hoc tests (IL-1R p = 3.32 x 10-05; TLR p = 6.66 x 10-05), while tissue or mucus microbiome richness were not.

Previous studies have reported that coral endolithic skeleton microbiomes are more species-rich than mucus or tissue, and show both species-specificity and signals of phylosymbiosis with their coral hosts [4]. Our results further suggest that endolithic skeleton microbiome diversity has tracked gene family expansions or contractions of coral TLR and IL-1R genes over evolution.

**Variation in genome quality does not explain immunity-microbiome correlations**

In principle, differences in the quality of genome assemblies used as inputs to this analysis could confound the apparent association between IL-1R gene copy number and microbiome richness in coral endolithic skeleton. To control for this, we calculated BUSCO scores for each genome, and regressed IL-1R copy number and BUSCO score together against skeleton microbiome richness. The resulting post-hoc tests show IL-1R copy number still significantly negatively correlated with skeleton microbiome richness (p = 0.0149), while BUSCO score (p=0.4691) or the interaction of BUSCO score and IL-1R gene copy number (p = 0.2733) did not.

**Microbiome composition varies with IL-R and TLR gene copy numbers.**

In addition to regulating microbiome richness and evenness, coral innate immune systems may also influence coral microbiome composition. If so, we might expect microbiome composition to correlate with the repertoire of innate immune proteins encoded in coral genomes. To test this, we compared differences in overall microbiome composition for each pair of coral species using two phylogenetic beta diversity metrics: Weighted UniFrac and Unweighted UniFrac. Using these beta-diversity distance metrics, we conducted principal coordinates analysis (PcoA), using the first three PC axes (PC1, PC2, and PC3) of the PcoA ordination (Methods). We correlated the microbiome PC coordinates against the number of predicted TLR or IL-1R gene copies (**Fig. 3**; **Table S9**).

Innate immune gene copy numbers strongly and significantly correlated with microbiome composition. When all samples were considered together (irrespective of anatomical compartment) Weighted UniFrac PC1 correlated strongly with the genomic copy number of IL-1R genes (PIC Weighted UniFrac PC1 R2 = 0.917, qFDR = 0.00012; **Fig. 3A, C, D**; **Table S9A**) and TLR genes (PIC Weighted UniFrac PC1 R2 = 0.77, qFDR = 0.0028; **Fig. 3B, E, F**; **Table S9A**). Unweighted UniFrac PC1 showed similar trends (**Table S9A**). These results indicate that gene family expansions or contractions of the TLR and IL-1R gene families over evolution corresponded to changes in overall microbiome composition.

We repeated the above protocol separately on the mucus, tissue, and skeleton compartments separately (**Table S9B-D**). In coral mucus, tissue, and skeleton, IL-1R gene copy number significantly correlated with Weighted UniFrac PC1 of the microbiome (PIC R2 = 0.72-0.81, FDR q < 0.01; **Table S9A-C**). TLR gene copy number significantly correlated with Weighted UniFrac PC1 in coral mucus and tissue (PIC R2 = 0.73, qFDR = 0.004; **Table S9B, C**) but not endolithic skeleton (**Table S9D**). In qualitative (presence/absence) analysis of microbiome β-diversity, IL-1R gene copy number correlated with Unweighted UniFrac PC1 in coral tissue and skeleton compartments, but not mucus (**Table S9B-D**), while TLR gene copy number correlated with Unweighted UniFrac PC1 in coral tissue but not mucus or skeleton (**Table S9B-D**).

These β-diversity results were more complex than the straightforward pattern of associations between innate immune gene copy number and microbiome richness. In coral tissue both IL-1R gene copy number and TLR gene copy number correlated with key aspects of quantitative (Weighted UniFrac PC1) and qualitative (Unweighted UniFrac PC1) microbiome β-diversity, in keeping with our original expectation that immunity should act strongly on the tissue-associated microbiome. In mucus, both TLR and IL-1R gene copy numbers appeared correlated with microbiome β-diversity when using quantitative but not qualitative metrics, suggesting immunity might influence microbial relative abundance more strongly than microbial presence/absence. In skeleton, IL-1R appeared correlated with IL-1R gene copy number, but not TLR gene copy number regardless of β-diversity metric. Together, these results suggest anatomically specific correlations between gene family expansion of some key innate immune genes and microbiome β-diversity.

**IL-1R and TLR copy number is associated with differential abundance of key microbes.**

We sought to identify microbial taxa that may be influenced by IL-1R or TLR gene copy number. This analysis could be confounded by the compositional nature of coral microbiome data. Therefore ANCOM-BC [52], which accounts for compositionality, was used for the analysis. However, because ANCOM-BC does not account for host-phylogeny, which is vital for cross-species comparisons, every microbe that was found to be significantly associated with a host trait in ANCOM-BC was also tested using phylogenetic independent contrasts against that same trait (Methods), and only reported if significant in both tests – a conservative approach aimed at addressing the confounding effects of both microbiome compositionality and the non-independence of comparative data.

Using this approach, we first tested whether IL-1R copy number correlated with microbial differential abundance in all samples regardless of tissue compartment (i.e. ‘all’). In this overall analysis, IL-1R copy number correlated with the differential abundance of 102 families of bacteria and archaea, while TLR copy number correlated with 88 families (**Table S10A**).

Correlations between microbiome composition and IL-1R or TLR copy number varied with anatomy. The abundance of 38 bacterial families were significantly and consistently correlated with TLR copy number (ANCOM-BC qFDR < 0.05 in mucus, tissue, and skeleton; **Table S10B-D**; **Fig. S9-S11**) and 38 bacterial families were also significantly and consistently correlated with IL-1R copy number (ANCOM-BC qFDR < 0.05 in mucus, tissue, and skeleton; **Table S10B-D**; raw heatmaps in **Fig. S12-S14**), both regardless of compartment. In contrast to the bacterial families with consistent interrelationships, some were specific to one or more compartments. One example are Rickettsiaceae, which live intracellularly in host cells. Rickettsiaceae relative abundance in tissue decreased with IL-1R gene copy number (ANCOM-BC coef = -0.011, W = 0.33, qFDR = 0; **Fig. 4**, **Fig. S13**; **Table S10C**) and TLR gene copy number (ANCOM-BC coef = -0.028, W =1.36, qFDR = 0; **Fig. 4**, **Fig S10**, **Table S10C**), but this taxon was not present in mucus (**Fig. S9**, **Fig S12**) or skeleton (**Fig. S13**, **Fig. S14**). The association of Rickettsiaceae with TLR copy number was confirmed by PIC, while it’s association with IL-1R was not, and is therefore omitted from the heatmap in **Fig. 4**. *Nitrosopumiliaceae* archaea were notable for correlating with IL-1R and TLR gene copy number across all compartments in ANCOM-BC analysis (all ANCOM-BC qFDR < 0.05; **Fig S9-S14**), though only their association with IL-1R was confirmed by PIC (**Fig. 4)**. These observations suggest that genomic changes in coral innate immune repertoires may have differential impact on lineages of coral-associated microbes.

**The domain architecture of TLR but not IL-1R genes is associated with microbiome richness.**

IL-1R and TLR genes are known to vary in the domain architecture of their extracellular regions, with variable numbers of Ig or LRR sensing domains, respectively. We reasoned that if higher copy numbers of TLR or IL-1R genes are associated with microbiome richness because they act as filters on the microbiome, the same selective pressures (for greater specificity in microbial associates) might also influence the domain architecture of extracellular sensing components within TLR or IL-1R genes. To test this, we analyzed the average number of LRR and Ig domain copy numbers associated with each TLR or IL-1R gene in each coral species.

Interestingly, we find that there is a significant positive correlation between LRR domain copy number per TLR genes and TLR copies (R2 = 0.437, p = 0.0159; **Fig. 5A, B**; **Fig. S15A**) but not Ig copy number with IL-1R (R2 = 0.00067, p = 0.345; **Fig. 5C, D**; **Fig. S15B**). In other words, corals with more TLR genes also have TLR genes with more LRR domains each.

Furthermore, we find that microbiome richness is correlated negatively with the average number of LRR domain copies (PIC R2 = 0.748, p = 0.00059; **Fig 5B**) but not Ig domain copies (PIC R2 = 0.341, p = 0.059; **Fig 5D**). That is, corals whose TLR genes typically have more elaborate domain architectures also tend to have simpler (or more specific) microbial associations. Importantly, the genomic copy number of LRR or Ig domains in total (rather than per gene) did not correlate with microbiome richness and beta diversity (**Table S8, S11**), so this finding does not simply recapitulate our previous findings regarding gene copy number.

**Responses to IL-1R and TLR are correlated between bacterial families.**

The microbe associated molecular patterns (MAMPs) detected by species-specific members of the TLR and IL-1R gene families could in principle be the same or independent. If independent, the correlation between any given microorganism and TLR gene family expansions would be independent of IL-1R gene family expansions. To test this, we analyze whether the correlation coefficients between each microbial family's relative abundance with TLR or IL-1R copy number were themselves correlated with one another. Microbial families’ responses to IL-1R and TLR gene family copy numbers were themselves very strongly correlated (R2 = 0.9062; **Fig S16**). Thus, microbes that are in high abundance in corals with many IL-1R gene copies also tend to be in high abundance in corals with many TLR gene copies. This suggests that TLR and IL-1R gene family expansions sculpt coral microbiomes in similar, rather than contrasting ways.

## Discussion

By combining coral microbiome and genomic data in a comparative framework, we demonstrate that TIR-domain containing innate immune gene repertoires strongly correlate with microbiome structure and diversity. These results suggest that gene family expansions of innate immune genes may have contributed to differences in the structure of coral microbiomes across millions of years of evolution. Those differences in immunity and microbiome structure, in turn, may influence the ability of modern scleractinian corals to survive escalating challenges from disease and climate change.

**Coral innate immune systems may be filtering microbiome membership.**

A key finding of the Earth Microbiome Project was that the diverse communities of microbes associated with animals and plants are nonetheless much less diverse than most environmental microbiomes [54]. Indeed, these animal-associated microbiomes were found to show high nestedness relative to the environment, indicating that what lives on animals is mostly — though certainly not entirely — a subset of environmental microbial lineages [54]. This suggests that filtering of environmental microbes is one key process shaping animal and plant microbiomes. That filtering is likely to be genomically encoded, since more closely related animals tend to have more similar microbiomes (‘phylosymbiosis’ [4,55–57]), excepting important deviations from this trend such as those driven by the evolution of specialized diets [58,59] or flight [24]. Could evolutionary changes in host immunity drive this trend, by altering which environmental microbes are excluded?

Our results in corals suggest that, for this group at least, immunity plays an important role in overall microbiome richness and evenness. Both coral microbiome composition [4], and coral innate immune responses [60,61] have long been known to vary between species. Indeed, gene expression of TLR and other innate immune genes is among the fastest evolving in coral transcriptomes [62]. Within coral populations, there is evidence that variation in immune activity between coral fragments correlates with their microbiome structure. For example, in *Montipora capitata*, increases in phenoloxidase activity correlated with decreased microbiome richness [63]. If such correlations extended across species, it could begin to explain why coral microbiomes differ so greatly in bacterial and archaeal biodiversity.

TIR-domain containing genes have undergone large gene-family expansions in some lineages of corals, which has been proposed as a driver of differences in coral microbiome richness [53]. Our findings support this hypothesis, and suggest that expansions of key innate immune genes reduce microbiome richness by allowing for sensing and exclusion of more diverse groups of microbes.

**Microbiome composition changes with TLR and IL-1R copy number**

We find a relationship between microbiome β-diversity and the copy number of TIR-containing innate immune genes. Our findings suggest that differences in innate immunity may explain substantial portions of the known differences in microbiome structure between coral species.

In our results, the relative abundances of diverse microbial taxa correlated with TLR or IL-1R gene copy number. Interestingly, these effects were not independent: microbes that correlated with TLR copy number also tended to be correlated with IL-1R copy number. This might reflect either that TLR and IL-1R gene family expansions are driven by similar selective pressures and therefore tend to co-occur, or that both types of gene family expansion influence the microbiome in similar ways.

**The relationship between innate immune gene repertoire and the microbiome is anatomically-specific**

While there are clear reasons to expect coral innate immune gene repertoire to affect tissue-associated coral microbiomes, our results suggest that these gene family expansions have even more clear-cut effects on coral’s endolithic skeleton microbiomes. Corals show compartmentalized differences in molecular function [64] and microbial richness and composition across anatomy [4]. Perhaps surprisingly, the microbiome of coral endolithic skeleton has been shown to be far more diverse than coral tissues or mucus [4]. In our results, endolithic skeleton microbiomes were more strongly correlated with host innate immunity than either mucus or tissue.

These results suggest that coral immunity strongly influences endolithic microbiomes. They also underline the importance of efforts to reevaluate how we conceptualize coral immunity and calcification. While coral calcification and immunity are typically thought of separately, recent work synthesizes these two fields [65], noting that skeleton is a key barrier against pathogens, and in many species contains diverse defensive chemical compounds and enzymes. These notably include melanin, which is a key aspect of immune defense in many non-vertebrate lineages [66].

**Disease-susceptible corals have more, not fewer, TIR-domain containing genes.**

One counter-intuitive aspect of our results is that the coral taxa with the largest numbers of TIR-containing innate immune genes tend to be those, such as *Acropora,* regarded as more susceptible to disease [67]. This is surprising, because we might expect more disease-susceptible coral species to have less diverse innate immune gene repertoires. How gene-family expansion, the commensal microbiome, and coral disease interact is a rich topic for future investigation. Our results here suggest several hypotheses that could be explored. It could be that more diverse microbial ecosystems are less susceptible to invasion by pathogens, as per Elton’s 1958 biotic resistance hypothesis [68,69]. Alternatively, less rich microbiomes may be less functionally redundant, increasing the risk that loss of any particular beneficial microbe may degrade host health [70,71]. Finally, the damage threshold hypothesis [72] proposes that corals differ in immune gene expression strategy, with many slow-growing corals constitutively expressing a variety of innate immune genes, while other fast-growing corals adopt a reactive strategy, which avoids expending resources on baseline expression of many innate immune genes in order to maximize growth or fecundity. If so, then reactive, fast-growing corals may experience stronger selective pressures to diversify sensors that can detect pathogens or cellular damage, since their baseline level of protection is low. TIR-domain containing gene family expansions in these reactive corals may enable more sensitive and/or specific responses to pathogen exposure.

**Conclusions**

We find that among sequenced coral genomes, gene family expansions of TLR and IL-1R genes are correlated with alterations in microbiome structure, and reductions in microbiome richness — with this apparent interplay between innate immunity and the microbiome most noticeable in coral endolithic skeleton. These findings are consistent with the idea that animal immunity sculpts microbiome structure and composition in part by sensing and filtering out many environmental microbes. This interpretation of the correlations we found between gene copy number and the microbiome is reinforced by the correlation between microbiome richness and the domain architecture within TLR genes, wherein there are more sensing (i.e. LRR) domains per TLR gene in corals that have lower microbiome richness. Our results further underscore the importance of distinguishing coral microbiomes across anatomy, and of exploring how coral innate immunity regulates corals’ diverse endolithic microbiomes. Finally, these results emphasize that integrating expanding coral genome and microbiome datasets in comparative frameworks is a promising approach that will help to uncover the interactions between immunity, the microbiome, and reef health.

**Declarations**

**Ethics approval and consent to participate.** Not applicable.

**Consent for publication.** Not applicable.

**Availability of Data and Material.** Global Coral Micorbiome Project sequence data, feature tables, and taxonomy tables are available on Qiita (<https://qiita.ucsd.edu/>; Study ID 10895). Analysis code is available on GitHub: <https://github.com/zaneveld/GCMP_Genomics/>. Additional intermediate data products are available in Supplementary Data Tables.

**Competing Interests.** The authors declare no competing interests.

**Funding.** The GCMP dataset was supported by a National Science Foundation Dimensions of Biodiversity grant (#1442306) to Rebecca Vega Thurber and MM, with sequencing provided by an in-kind ​​UC San Diego Seed Grant in Microbiome Science grant to JZ. Analysis in this manuscript was supported by NSF IoS CAREER grant (# 1942647) and an in-kind ​​UC San Diego Seed Grant in Microbiome Science grant to JZ. GCMP data collection was supported by a National Science Foundation Dimensions of Biodiversity grant (#1442306) to MM

**Author Contributions.** TB, DS, JZ analyzed the data; TB, JZ, DS wrote the manuscript; RM, JFP, MM, and JZ collected GCMP data and metadata; all authors edited the manuscript.

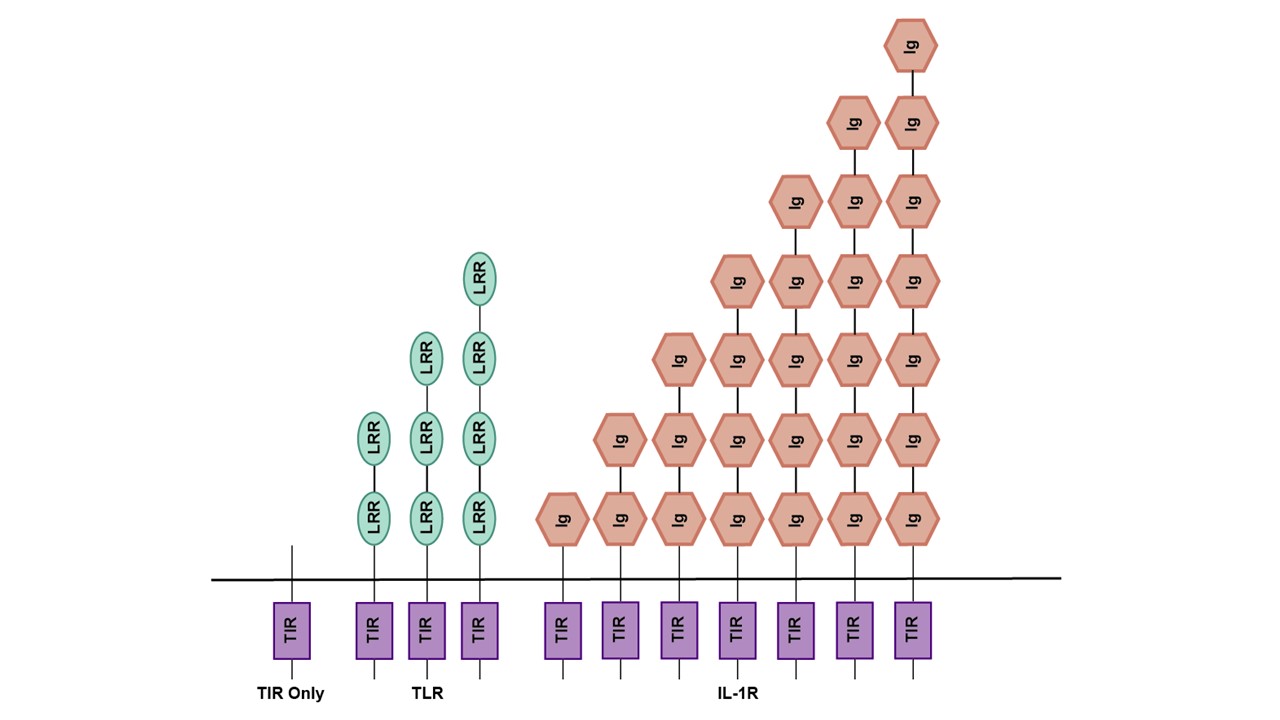
**Acknowledgements.**  The authors would like to acknowledge many contributors for their field assistance during past collections of Global Coral Microbiome Project data reanalyzed in this manuscript, including: Rebecca Vega Thurber, a leader of the original GCMP project; Tasman Douglass, Margaux Hein, Frazer McGregor, Kathy Morrow, Katia Nicolet, Cathie Page, and Gergely Torda for their field assistance in Australia; Valeria Pizarro, Mateo López-Victoria, Styles Smith, Alaina Weinheimer, Claudia Tatiana Galindo for Assistance in Colombia; Chris Voolstra, Maren Ziegler, Anna Roik, and many others at KAUST for field assistance in Saudi Arabia; Mark Vermeij, Kristen Marhaver, Pedro Frade, Ben Mueller, and others at CARMABI for field assistance during sampling in Curaçao; Danwei Huang for field assistance in Singapore; Ruth Gates, Katie Barrott, Courtney Couch, Keoki Stender for field assistance in Hawai’i; Le Club de Plongee Suwan Macha and Jean-Pascal Quod for field assistance, and Amelia Foster and Jerome Payet for sampling in Isle de la Réunion; the Burkepile lab for field assistance in Mo’orea; Yossi Loya, Raz Tamir for sampling assistance in Israel; and Lyndsy Gazda, Jamie Lee Proffitt, Gabriele Swain, and Alaina Weinheimer for their assistance in the laboratory. The authors also acknowledge the staff of the Coral Bay Research Station, Lizard Island Research Station, Lord Howe Island Marine Park, Lord Howe Island Research Station, and RV Cape Ferguson for their logistical support.

**Tables**

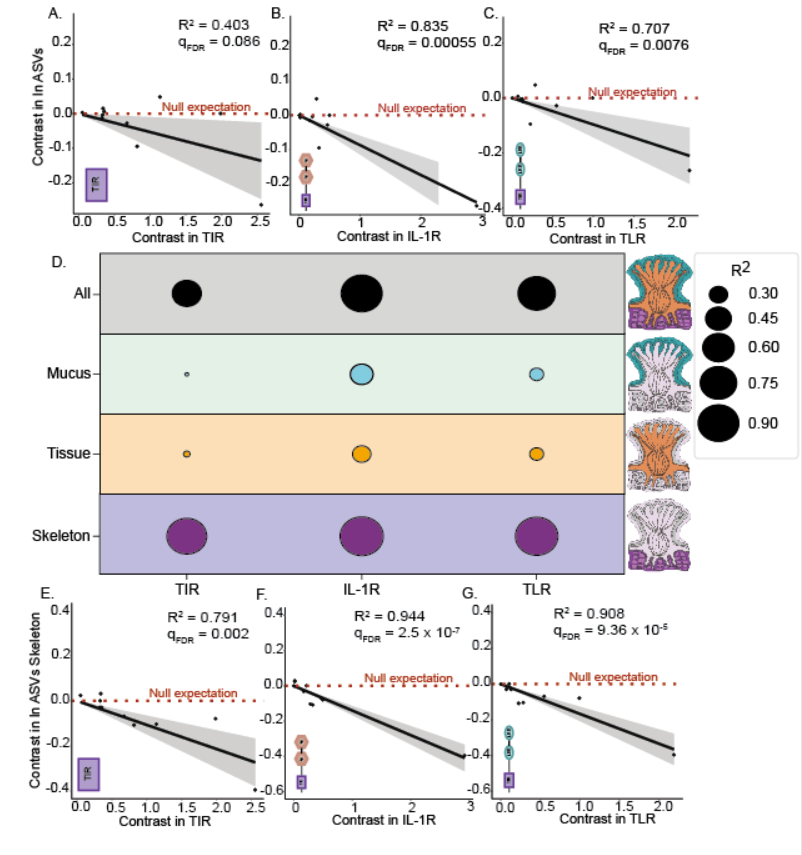
**Table 1.** Toll/Interleukin Repeat (TIR) domain containing elements and their associated numbers per species.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Coral Species** | **TIR Isoforms** | **IL-1R Copies** | **TLR Copies** | **Richness (ln ASVs /1000 reads)** | **Genome Sequence Reference** |
| *Acropora hyacinthus* | 37 | 11 | 8 | 3.788 | Sinzato *et al* 2020 |
| *Acropora cytherea* | 33 | 10 | 7 | 3.619 | Sinzato *et al* 2020 |
| *Pocillopora damicornis* | 24 | 6 | 5 | 4.368 | Cunning *et al* 2018 |
| *Pocillopora verrucosa* | 28 | 7 | 7 | 4.369 | Buitrago-Lopez *et al* 2020 |
| *Orbicella faveolata* | 27 | 7 | 6 | 4.484 | Prada *et al* 2016 |
| *Stylophora* | 23 | 8 | 6 | 4.236 | Chen *et al* 2008 |
| *Galaxea fascicularis* | 19 | 3 | 1 | 5.266 | Ying et al 2018 |
| *Porites lutea* | 36 | 14 | 8 | 4.016 | Robbins et al 2019 |
| *Fungia fungites* | 32 | 7 | 8 | 4.3 | Ying et al 2018 |
| *Montipora capitata* | 27 | 5 | 1 | 4.034 | Helmkampf *et al* 2019 |
| *Porites rus* | 29 | 6 | 2 | 4.739 | Celis et al 2018 |

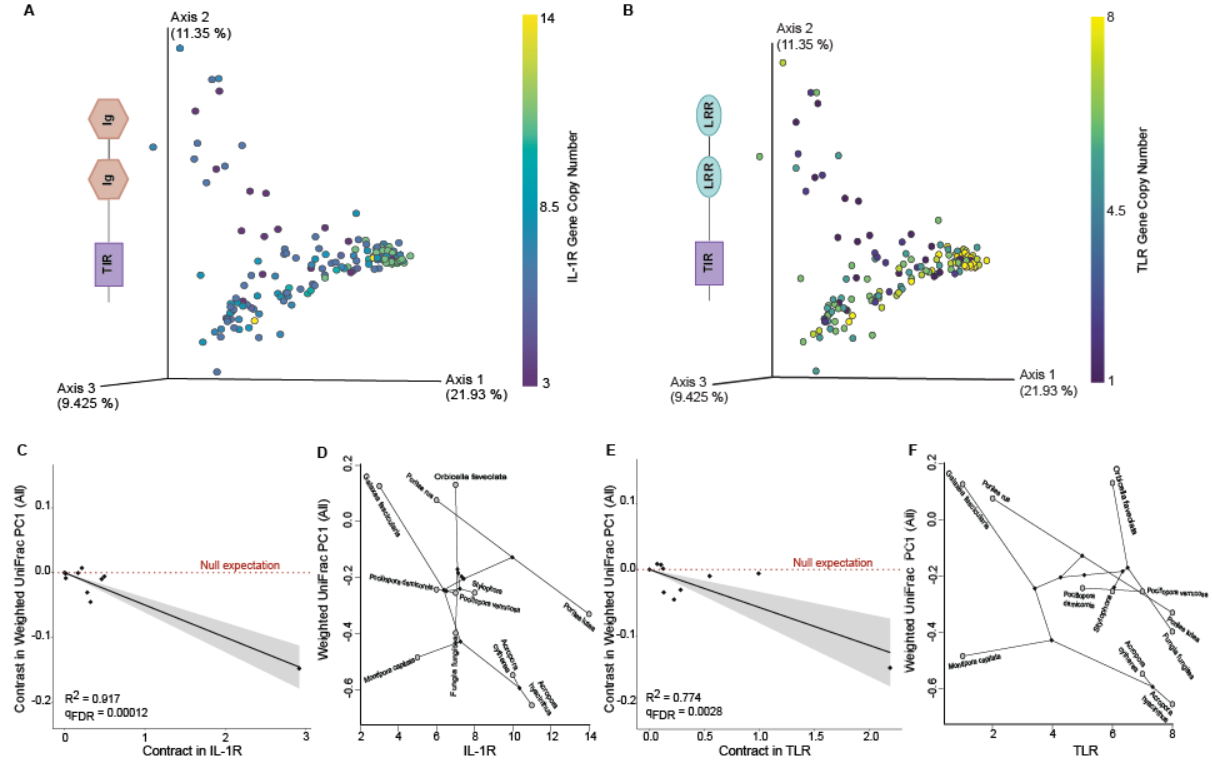
**Figures and Figure Legends**

****

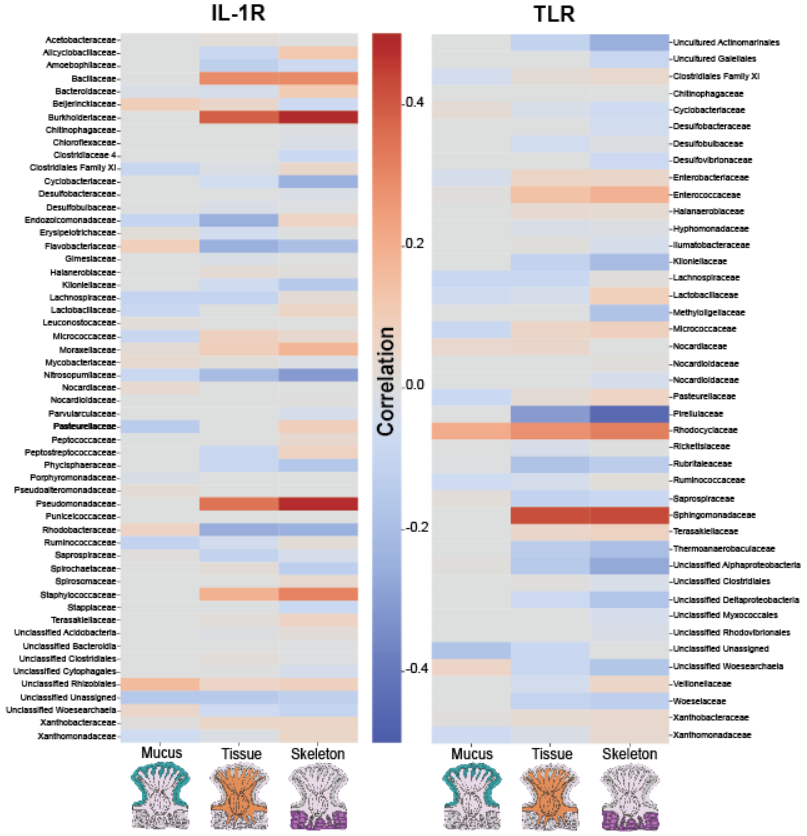
**Fig. 1.** TIR containing gene architecture in coral genomes. Diagram depicts all domains present in TIR only, TLR, or IL-1R genes identified in this study. TIR: Toll-Interleukin Repeat; Ig: Immunoglobulin; LRR: Leucine Rich Repeat; TLR: Toll/Interleukin-like Receptor; IL-1R: Interleukin-1 Receptor. Counts and domain structure for each of these genes from each coral species are found in Table S2.



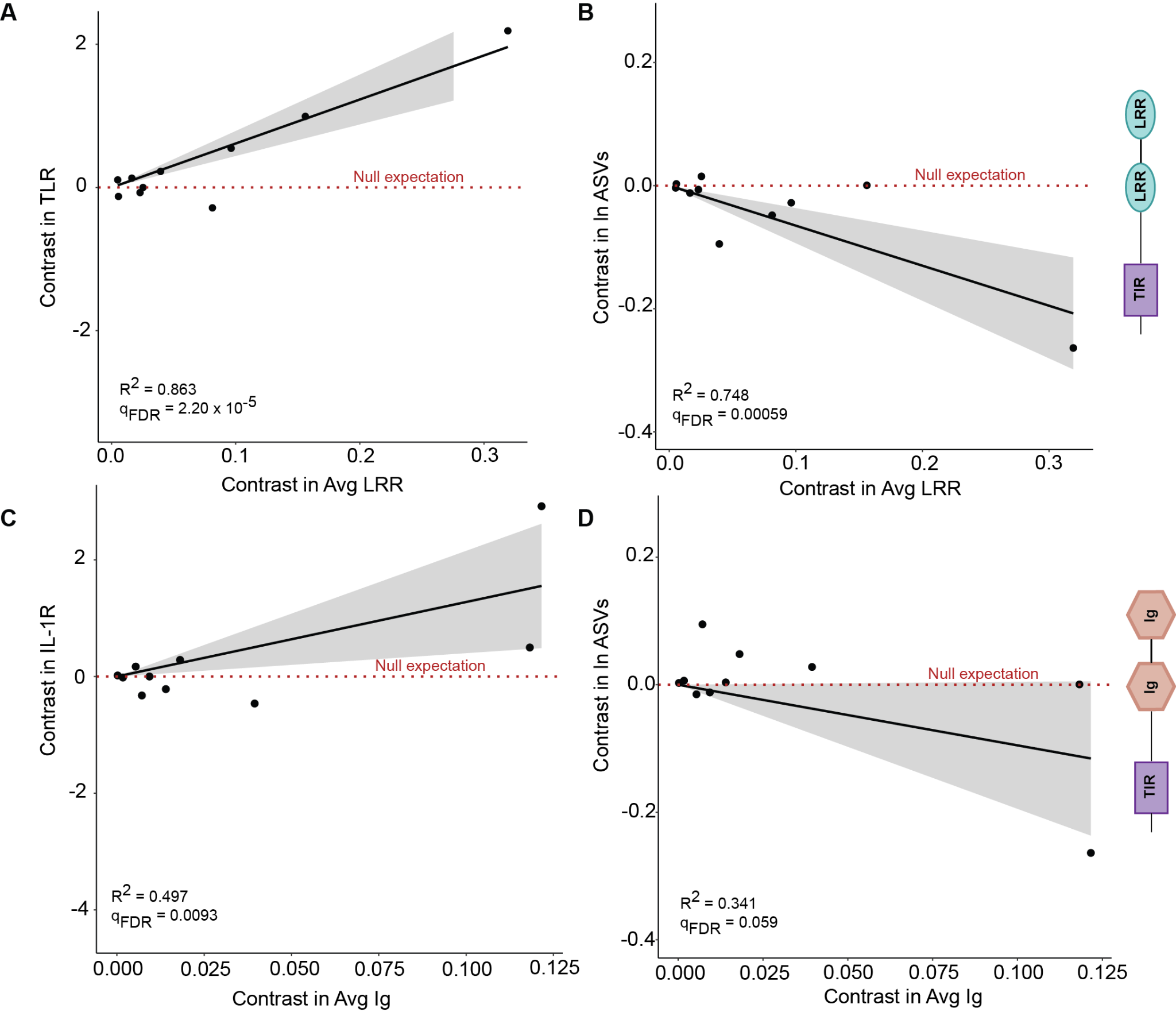
**Fig. 2. Coral innate immune gene repertoire correlates with microbiome richness. The top** rows shows phylogenetic independent contrast regressions of genomic copy number of innate immune components against log microbiome richness (ln ASVs) across all regions of coral anatomy, for **A.** TIR domains (R2 =0.403, qFDR = 0.086, p = 0.021); **B.** IL-1R genes (R2 = 0.835, qFDR = 0.00055, p = 5.13 x 10-15); **C.**TLR genes (R2 = 0.707, qFDR = 0.0076, p = 0.0012). Shading in phylogenetic independent contrasts analysis indicates the 95% confidence interval of the mean. **D.** R2 values from correlations of the genomic copy number of innate immune components against microbiome richness, organized by compartment. The bottom row uses microbiome data from coral skeleton only to show phylogenetic independent contrasts of innate immune components vs. log microbiome richness for: **E.** TIR-only genes (R2 =0.791, qFDR=0.002, p = 0.000252); **F.** IL-1R genes (R2 = 0.944, qFDR=2.15 x 10-5, p = 6.39 x 10-7); **G.** TLR genes (R2 = 0.90798, qFDR= 9.36 x 10-5, p = 5.85 x 10-6). No correlations had outliers, as defined by studentized residuals of >3 in absolute value.



**Fig. 3.** Comparison of innate immune gene repertoire and microbiome composition. **A.** Principle Coordinates Analysis (PCoA) of Weighted UniFrac Beta diversity distances for samples from all coral compartments, colored by genomic IL-1R copy number in the coral host. **B.** PCoA ordination of Weighted UniFrac distances for all compartments colored by genomic copy number for TLR genes. **C.** Phylogenetic independent contrasts regression of IL-1R copy number vs. Weighted UniFrac PC1 (R2 = 0.917, qFDR = 0.00012). **D.** Phylomorphospace of IL-1R copy number vs. Weighted UniFrac PC1. **E**. Phylogenetic independent contrasts regression of TLR copy number vs. Weighted UniFrac PC1 (R2 = 0.774, qFDR = 0.0028). **F.** Phylomorphospace of IL-1R copy number vs. Weighted UniFrac PC1. No correlations had outliers, as defined by studentized residuals of >3 in absolute value.



**Fig. 4.** Heatmap showing microbial families significantly correlated with IL-1R and TLR copy numbers. Columns show correlation coefficients from ANCOM-BC analysis comparing the relative abundance of microbial families against IL-1R and TLR copy numbers in the mucus, tissue, and skeleton, respectively. All associations shown were confirmed by phylogenetic independent contrasts analysis. Red colors indicate a positive correlation between IL-1R and TLR copy number and the relative abundance of that microbial family, while blue colors indicates a negative correlation. Taxa present in only some compartments are marked in grey (i.e. similar to a 0.0 correlation) in compartments where they are absent.



**Fig. 5.** Domain copy numbers influence microbiome richness.Phylogenetic contrasts comparing **A.** average LRR domain copies per TLR vs TLR domain copy number (R2 = 0.863, qFDR = 2.20 x 10-5, **B.** average LRR domain copies per TLR vs ln ASVs (R2 = 0.748, qFDR = 0.00059), **C.** average Ig domain copies per IL-1R vs IL-1R copy number (R2 = 0.497, qFDR = 0.0093), and **D.** average Ig domain copies per IL-1R vs ln ASVs per genome (R2 = 0.341, qFDR = 0.059).

**Supplementary Data Tables**

**Supplementary Table 1.** Sample and Genomic metadata. (**A**) Genomes used in the analysis (B) Per sample metadata for the GCMP data (**C**) Mapping file for the GCMP data.

**Supplementary Table 2.** Annotations of TIR-domain containing gene families in coral genomes for (**A**) Total number of IL-1R and TLR genes,(**B**) Domain make up of IL-1R genes, (**C**) Domain make up of TLR genes.

**Supplementary Table 3.** Quality control information for the sequencing depth of microbiome samples in the dataset.

**Supplementary Table 4.** Distribution of TIR, LRR, and Ig domains across unique predicted protein isoforms in sequenced coral genomes.

**Supplementary Table 5.** Microbiome (**A**) richness and (**B**) evenness across samples in the analysis.

**Supplementary Table 6**. Phylogenetic comparison of IL-1R and TLR gene family copy number vs. microbiome alpha diversity in (**A**) All Samples (**B**) Mucus (**C**) Tissue, (**D**) Skeleton.

**Supplementary Table 7.** Trait table of the genomic and microbiome traits of coral species in the analysis.

**Supplementary Table 8.** Phylogenetic comparison of TIR LRR and Ig domains vs. microbiome alpha diversity in A) All Samples B) Mucus c) Tissue, d) Skeleton.

**Supplementary Table 9**. Phylogenetic comparison of IL-1R and TLR gene family copy number vs. PC axes from beta-diversity PCoA analysis in (**A**) All Samples (**B**) Mucus (**C**) Tissue, (**D**) Skeleton.

**Supplementary Table 10.** ANCOMBC comparison of relative abundance of microbial taxa vs. IL-1R and TLR gene family copy number in (**A**) All Samples (**B**) Mucus (**C**) Tissue, (**D**) Skeleton.

**Supplementary Table 11.** Phylogenetic comparison of TIR LRR and Ig domains vs. microbiome PC axes from beta-diversity PCoA analysis in (**A**) All Samples (**B**) Mucus (**C**) Tissue, (**D**) Skeleton.

**NEW References**

1. Hoeksema, BW, Cairns S. World List of Scleractinia. [Internet]. Available from: https://www.marinespecies.org/scleractinia

2. Rohwer F, Seguritan V, Azam F, Knowlton N. Diversity and distribution of coral-associated bacteria. Mar Ecol Prog Ser. 2002;243:1–10.

3. Ricci F, Tandon K, Black JR, Lê Cao K-A, Blackall LL, Verbruggen H. Host Traits and Phylogeny Contribute to Shaping Coral-Bacterial Symbioses. Msystems. 2022;7:e00044-22.

4. Pollock FJ, McMinds R, Smith S, Bourne DG, Willis BL, Medina M, et al. Coral-associated bacteria demonstrate phylosymbiosis and cophylogeny. Nat Commun. 2018;9:4921.

5. Hernandez-Agreda A, Leggat W, Bongaerts P, Herrera C, Ainsworth TD. Rethinking the coral microbiome: simplicity exists within a diverse microbial biosphere. MBio. 2018;9:e00812-18.

6. Littman RA, Willis BL, Pfeffer C, Bourne DG. Diversities of coral-associated bacteria differ with location, but not species, for three acroporid corals on the Great Barrier Reef. FEMS Microbiol Ecol. 2009;68:152–63.

7. Rodriguez‐Lanetty M, Granados‐Cifuentes C, Barberan A, Bellantuono AJ, Bastidas C. Ecological Inferences from a deep screening of the C omplex B acterial C onsortia associated with the coral, P orites astreoides. Mol Ecol. 2013;22:4349–62.

8. Pantos O, Bongaerts P, Dennis PG, Tyson GW, Hoegh-Guldberg O. Habitat-specific environmental conditions primarily control the microbiomes of the coral Seriatopora hystrix. ISME J. 2015;9:1916–27.

9. Dunphy CM, Gouhier TC, Chu ND, Vollmer SV. Structure and stability of the coral microbiome in space and time. Sci Rep. 2019;9:1–13.

10. Wang L, Shantz AA, Payet JP, Sharpton TJ, Foster A, Burkepile DE, et al. Corals and their microbiomes are differentially affected by exposure to elevated nutrients and a natural thermal anomaly. Front Mar Sci. 2018;5:101.

11. Pootakham W, Mhuantong W, Yoocha T, Putchim L, Jomchai N, Sonthirod C, et al. Heat‐induced shift in coral microbiome reveals several members of the Rhodobacteraceae family as indicator species for thermal stress in Porites lutea. MicrobiologyOpen. 2019;8:e935.

12. Ricci F, Rossetto Marcelino V, Blackall LL, Kühl M, Medina M, Verbruggen H. Beneath the surface: community assembly and functions of the coral skeleton microbiome. Microbiome. 2019;7:1–10.

13. Kusdianto H, Kullapanich C, Palasuk M, Jandang S, Pattaragulwanit K, Ouazzani J, et al. Microbiomes of Healthy and Bleached Corals During a 2016 Thermal Bleaching Event in the Upper Gulf of Thailand. Front Mar Sci. 2021;8:643962.

14. Fifer JE, Bui V, Berg JT, Kriefall N, Klepac C, Bentlage B, et al. Microbiome structuring within a coral colony and along a sedimentation gradient. Cell Stress Response Physiol Adapt Corals Subj Environ Stress Pollut. 2022;

15. Marchioro GM, Glasl B, Engelen AH, Serrão EA, Bourne DG, Webster NS, et al. Microbiome dynamics in the tissue and mucus of acroporid corals differ in relation to host and environmental parameters. PeerJ. 2020;8:e9644.

16. Zaneveld JR, Burkepile DE, Shantz AA, Pritchard CE, McMinds R, Payet JP, et al. Overfishing and nutrient pollution interact with temperature to disrupt coral reefs down to microbial scales. Nat Commun. 2016;7:1–12.

17. Corinaldesi C, Canensi S, Dell’Anno A, Tangherlini M, Di Capua I, Varrella S, et al. Multiple impacts of microplastics can threaten marine habitat-forming species. Commun Biol. 2021;4:1–13.

18. Pratte ZA, Longo GO, Burns AS, Hay ME, Stewart FJ. Contact with turf algae alters the coral microbiome: contact versus systemic impacts. Coral Reefs. 2018;37:1–13.

19. Briggs AA, Brown AL, Osenberg CW. Local versus site-level effects of algae on coral microbial communities. R Soc Open Sci. 2021;8:210035.

20. Miller N, Maneval P, Manfrino C, Frazer TK, Meyer JL. Spatial distribution of microbial communities among colonies and genotypes in nursery-reared Acropora cervicornis. PeerJ. 2020;8:e9635.

21. Palacio-Castro AM, Rosales SM, Dennison CE, Baker AC. Microbiome signatures in Acropora cervicornis are associated with genotypic resistance to elevated nutrients and heat stress. Coral Reefs. 2022;41:1389–403.

22. Rosales SM, Miller MW, Williams DE, Traylor-Knowles N, Young B, Serrano XM. Microbiome differences in disease-resistant vs. susceptible Acropora corals subjected to disease challenge assays. Sci Rep. 2019;9:1–11.

23. Rosado PM, Leite DCA, Duarte GAS, Chaloub RM, Jospin G, Nunes da Rocha U, et al. Marine probiotics: increasing coral resistance to bleaching through microbiome manipulation. ISME J. 2019;13:921–36.

24. Song SJ, Sanders JG, Delsuc F, Metcalf J, Amato K, Taylor MW, et al. Comparative Analyses of Vertebrate Gut Microbiomes Reveal Convergence between Birds and Bats. mBio. 2020;11:e02901-19.

25. Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, et al. Evolution of mammals and their gut microbes. science. 2008;320:1647–51.

26. Bodawatta KH, Koane B, Maiah G, Sam K, Poulsen M, Jønsson KA. Species-specific but not phylosymbiotic gut microbiomes of New Guinean passerine birds are shaped by diet and flight-associated gut modifications. Proc R Soc B. 2021;288:20210446.

27. Woodhams DC, Bletz MC, Becker CG, Bender HA, Buitrago-Rosas D, Diebboll H, et al. Host-associated microbiomes are predicted by immune system complexity and climate. Genome Biol. 2020;21:1–20.

28. Thaiss CA, Levy M, Suez J, Elinav E. The interplay between the innate immune system and the microbiota. Curr Opin Immunol. 2014;26:41–8.

29. Alexander M, Turnbaugh PJ. Deconstructing Mechanisms of Diet-Microbiome-Immune Interactions. Immunity. 2020;53:264–76.

30. Fonseca JP, Lakshmanan V, Boschiero C, Mysore KS. The Pattern Recognition Receptor FLS2 Can Shape the Arabidopsis Rhizosphere Microbiome β-Diversity but Not EFR1 and CERK1. Plants. 2022;11:1323.

31. Poole AZ, Weis VM. TIR-domain-containing protein repertoire of nine anthozoan species reveals coral–specific expansions and uncharacterized proteins. Dev Comp Immunol. 2014;46:480–8.

32. Medzhitov R, Janeway CA. Innate immunity: the virtues of a nonclonal system of recognition. Cell. 1997;91:295–8.

33. Ligoxygakis P, Pelte N, Ji C, Leclerc V, Duvic B, Belvin M, et al. A serpin mutant links Toll activation to melanization in the host defence of Drosophila. EMBO J. 2002;21:6330–7.

34. van de Water JA, Ainsworth TD, Leggat W, Bourne DG, Willis BL, Van Oppen MJ. The coral immune response facilitates protection against microbes during tissue regeneration. Mol Ecol. 2015;24:3390–404.

35. Nie L, Cai S-Y, Shao J-Z, Chen J. Toll-like receptors, associated biological roles, and signaling networks in non-mammals. Front Immunol. 2018;9:1523.

36. O’Neill LA, Bowie AG. The family of five: TIR-domain-containing adaptors in Toll-like receptor signalling. Nat Rev Immunol. 2007;7:353–64.

37. Emery MA, Dimos BA, Mydlarz LD. Cnidarian pattern recognition receptor repertoires reflect both phylogeny and life history traits. Front Immunol. 2021;12:689463.

38. Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, Locey KJ, et al. A communal catalogue reveals Earth’s multiscale microbial diversity. Nature. 2017;551:457–63.

39. Parada AE, Needham DM, Fuhrman JA. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. Environ Microbiol. 2016;18:1403–14.

40. Apprill A, McNally S, Parsons R, Weber L. Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. Aquat Microb Ecol. 2015;75:129–37.

41. Caporaso JG. EMP 16S Illumina Amplicon Protocol. 2018 [cited 2020 Jul 6]; Available from: https://www.protocols.io/view/emp-16s-illumina-amplicon-protocol-nuudeww

42. Gonzalez A, Navas-Molina JA, Kosciolek T, McDonald D, Vázquez-Baeza Y, Ackermann G, et al. Qiita: rapid, web-enabled microbiome meta-analysis. Nat Methods. 2018;15:796–8.

43. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 2010;7:335–6.

44. Amir A, McDonald D, Navas-Molina JA, Kopylova E, Morton JT, Zech Xu Z, et al. Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. mSystems. 2017;2:e00191-16.

45. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019;37:852–7.

46. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH: a versatile open source tool for metagenomics. PeerJ. 2016;4:e2584.

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

48. Bengtsson-Palme J, Hartmann M, Eriksson KM, Pal C, Thorell K, Larsson DGJ, et al. METAXA2: improved identification and taxonomic classification of small and large subunit rRNA in metagenomic data. Mol Ecol Resour. 2015;15:1403–14.

49. Sonett D, Brown T, Bengtsson-Palme J, Padilla-Gamiño JL, Zaneveld JR. The Organelle in the Ointment: cryptic mitochondria account for many unknown sequences in cross-species microbiome comparisons [Internet]. bioRxiv; 2023 [cited 2023 Sep 19]. p. 2021.02.23.431501. Available from: https://www.biorxiv.org/content/10.1101/2021.02.23.431501v2

50. Janssen S, McDonald D, Gonzalez A, Navas-Molina JA, Jiang L, Xu ZZ, et al. Phylogenetic Placement of Exact Amplicon Sequences Improves Associations with Clinical Information. mSystems [Internet]. 2018 [cited 2021 Feb 19];3. Available from: https://msystems.asm.org/content/3/3/e00021-18

51. Revell LJ. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol Evol. 2012;217–23.

52. Lin H, Peddada SD. Analysis of compositions of microbiomes with bias correction. Nat Commun. 2020;11:1–11.

53. Poole AZ, Weis VM. TIR-domain-containing protein repertoire of nine anthozoan species reveals coral–specific expansions and uncharacterized proteins. Dev Comp Immunol. 2014;46:480–8.

54. Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, Locey KJ, et al. A communal catalogue reveals Earth’s multiscale microbial diversity. Nature. 2017;551:457–63.

55. Mazel F, Davis KM, Loudon A, Kwong WK, Groussin M, Parfrey LW. Is host filtering the main driver of phylosymbiosis across the tree of life? Msystems. 2018;3:e00097-18.

56. Brooks AW, Kohl KD, Brucker RM, Opstal EJ van, Bordenstein SR. Phylosymbiosis: Relationships and Functional Effects of Microbial Communities across Host Evolutionary History. PLOS Biol. 2016;14:e2000225.

57. O’Brien PA, Tan S, Yang C, Frade PR, Andreakis N, Smith HA, et al. Diverse coral reef invertebrates exhibit patterns of phylosymbiosis. ISME J. 2020;14:2211–22.

58. Delsuc F, Metcalf JL, Wegener Parfrey L, Song SJ, González A, Knight R. Convergence of gut microbiomes in myrmecophagous mammals. Mol Ecol. 2014;23:1301–17.

59. Song SJ, Sanders JG, Baldassarre DT, Chaves JA, Johnson NS, Piaggio AJ, et al. Is there convergence of gut microbes in blood-feeding vertebrates? Philos Trans R Soc B Biol Sci. 2019;374:20180249.

60. Palmer CV, McGinty ES, Cummings DJ, Smith SM, Bartels E, Mydlarz LD. Patterns of coral ecological immunology: variation in the responses of Caribbean corals to elevated temperature and a pathogen elicitor. J Exp Biol. 2011;214:4240–9.

61. Palmer C, Bythell J, Willis B. Enzyme activity demonstrates multiple pathways of innate immunity in Indo-Pacific anthozoans. Proc R Soc B Biol Sci. 2012;279:3879–87.

62. Dimos B, Emery M, Beavers K, MacKnight N, Brandt M, Demuth J, et al. Adaptive variation in homologue number within transcript families promotes expression divergence in reef‐building coral. Mol Ecol. 2022;31:2594–610.

63. Brown T, Sonett D, Zaneveld JR, Padilla-Gamiño JL. Characterization of the microbiome and immune response in corals with chronic Montipora white syndrome. Mol Ecol. 2021;30:2591–606.

64. Axworthy JB, Timmins-Schiffman E, Brown T, Rodrigues LJ, Nunn BL, Padilla-Gamiño JL. Shotgun Proteomics Identifies Active Metabolic Pathways in Bleached Coral Tissue and Intraskeletal Compartments. Front Mar Sci. 2022;

65. Levy S, Mass T. The Skeleton and Biomineralization Mechanism as Part of the Innate Immune System of Stony Corals. Front Immunol [Internet]. 2022 [cited 2023 Feb 17];13. Available from: https://www.frontiersin.org/articles/10.3389/fimmu.2022.850338

66. Cerenius L, Söderhäll K. Immune properties of invertebrate phenoloxidases. Dev Comp Immunol. 2021;122:104098.

67. Vega Thurber R, Mydlarz LD, Brandt M, Harvell D, Weil E, Raymundo L, et al. Deciphering coral disease dynamics: integrating host, microbiome, and the changing environment. Front Ecol Evol. 2020;8:575927.

68. Elton CS. The Ecology of Invasions by Animals and Plants. Springer Nature; 2020.

69. Marraffini ML, Geller JB. Species richness and interacting factors control invasibility of a marine community. Proc R Soc B Biol Sci. 2015;282:20150439.

70. Sommer F, Rühlemann MC, Bang C, Höppner M, Rehman A, Kaleta C, et al. Microbiomarkers in inflammatory bowel diseases: caveats come with caviar. Gut. 2017;66:1734–8.

71. Sommer F, Anderson JM, Bharti R, Raes J, Rosenstiel P. The resilience of the intestinal microbiota influences health and disease. Nat Rev Microbiol. 2017;15:630–8.

72. Palmer CV. Immunity and the coral crisis. Commun Biol. 2018;1:91.