The *Pocillopora damicornis* genome and comparative genomic analysis highlights innate immune role in coral evolution

Ross Cunning, Rachael A. Bay, Philip Gillette, Andrew C. Baker, Nikki Traylor-Knowles

Abstract

Scleractinian corals face an uncertain future as they are threatened by environmental stress associated with global climate change. Studies of these organisms will benefit from expanding genomic resources, and comparative genomic analyses to reveal evolutionary adaptations in these organisms, with implications for their future responses. Here, we sequenced and annotated the genome of the widespread coral *Pocillopora damicornis*, and compare this genome with other publicly available genomes of scleractinians, cnidarians, and basal metazoans based on protein-coding gene orthology. We found that 45% of *P. damicornis* genes had orthologs in all other corals, defining a set of coral core genes, while 25% of genes were lineage-specific. These lineage-specific genes were primarily enriched in immune system functionality, including X, Y, and Z. Moreover, genes present in all corals but absent from other genomes were also enriched in immune functionality, indicating immune evolution is a prominent feature of coral clade and species evolution. Furthermore, we found numerous gene family expansions, especially in *A. digitifera* and *S. pistillata*, indicating lineage-specific adaptation that were also associated with immunity. Overall, these results suggest that immune evolution is a prominent feature of scleractinian evolution, and may underlie adaptive responses involving symbiosis, pathogen interactions, and environmental stress.

Introduction

Scleractinian corals serve the critical ecological role of building reefs that provide billions of dollars annually in goods and services[1] and sustain high levels of biodiversity[2]. Moreover, as basal metazoans, corals provide a model for studying the evolution of key traits such as symbiosis[3], immunity[4,5], and biomineralization[6]. However, corals are declining rapidly as ocean acidification impairs coral skeleton formation[7], and ocean warming disrupts their symbiosis with photosynthetic *Symbiodinium* spp.[8]. Understanding the genomic architecture of these traits is therefore key to understanding corals' success through evolutionary history[9] and under future climate scenarios. In particular, there is great interest in whether corals possess the genetic basis and genetic variation required to acclimatize and/or adapt to rapid climate change[10–12]. Addressing these questions requires expanding genomic resources for corals, and establishes a fundamental role for comparative genomic analysis in these organisms.

Genomic resources for corals have expanded rapidly in recent years, with genomic or transcriptomic information available for at least 20 coral species (see 9). Comparative analysis of these data has identified genes that may be important in biomineralization, symbiosis, and environmental stress responses[9]. However, complete genome sequences have only been analyzed and compared for two coral species, *Acropora digitifera* and *Stylophora pistillata*[13], which revealed extensive differences in genomic architecture and content. However, one of the main similarities noted was a highly diversified repertoire of immune-related genes, and the evolution of immune system functioning[13]. Therefore, additional complete coral genomes and more comprehensive comparative analysis may be transformative in our understanding of the genomic content and evolutionary history of reef-building corals, as well as the importance of specific gene repertoires and diversification within coral lineages.

Here, we present the genome of *Pocillopora damicornis*, one of the most abundant and widespread reef-building corals in the Indo-Pacific. This ecologically important coral is the subject of a large body of research on speciation[14–16], population genetics[17–20], symbiosis ecology[21–23], and reproduction[24–26], and is commonly used in experimental biology and physiology. Therefore, availability of the *P. damicornis* genome will advance

a number of fields in biology, ecology, and evolution, and provides a direct foundation for future studies in transcriptomics, population genomics, and functional genomics of corals.

Using this genome and other publicly available genomes of corals, cnidarians, and basal metazoans, we provide the most comprehensive comparative genomic analysis to date in Scleractinia. In particular, we address critical questions such as: 1) what genes are specific to or diversified within scleractinians, and 2) what genes are specific to or diversified within individual scleractinian species, and 3) what features distinguish the *P. damicornis* genome from other corals. We address these questions based on orthology of protein-coding genes, which generalizes the approaches taken by 9 and 13 to a larger set of complete genomes to describe both shared and unique adaptations in the Scleractinia. In comparing these genomes, we reveal prominent diversification and expansion of immune-related genes and pathways, demonstrating that the immune pathways are fundamental to evolutionary adaptation in corals.

Results and discussion

P. damicornis genome assembly and annotation

The estimated genome size of *P. damicornis* is 348 Mb, smaller than other scleractinian genomes analyzed to date. The size of the final assembly produced here was 234 Mb, and likely lacks high-identity repeat content that could not be assembled. The assembly comprises 96.3% contiguous sequence, and has the highest contig N50 (28.5 kb) of any cnidarian genome assembly (Table 1). We identified 26,077 gene models, with 21,389 (82%) of these being apparently complete with start and stop codons. This number of genes is consistent with other scleractinian and cnidarian genomes (Table 1). Among all gene models, 59.7% had identifiable homologs (E-value $\leq 10^{-5}$) in the UniProt-SwissProt database, and 83.7% contained protein domains annotated by InterProScan. In addition, 73% of genes contained identifiable homologs in at least one of the other 10 genomes analyzed here. Furthermore, a BUSCO search found that out of 978 metazoan single-copy orthologs, 865 (88.4%) were present and complete (5 of these were duplicated). An additional 28 orthologs were present but fragmented, and 85 (8.7%) were missing. Together, these statistics indicate the *P. damicornis* genome assembly is the highest quality and most complete scleractinian genome to date (Table 1).

	Pdam	Spis	Adig	Ofav*	Disc	Afen	Aipt	Nema	Hydr	Mlei	Aque
Genome size (Mb)	348†	434	420		428	350	260	329	1300		190
Assembly size (Mb)	234	400	419	486	445	370	258	356	852	156	167
Total contig size (Mb)	226	358	365	356	364	306	213	297	785	150	145
Contig / Assemly (%)	96.3	89.5	87	73.3	81.9	82.6	82.5	83.4	92.2	96.5	86.8
Contig N50 (kb)	25.9	14.9	10.9	7.4	18.7	20.1	14.9	19.8	9.7	11.9	11.2
Scaffold N50 (kb)	326	457	191	1162	770	510	440	472	92.5	187	120
No. gene models	26077	25769‡	23668	37660	23199	21372	29269	27273	31452	16554	29867
No. complete gene models	21389	25563‡	16434	29679	16082‡	15552‡	26658	13343			
BUSCO completeness (%)	88.4	72.2	34.3	71							
Mean exon length (bp)	245	262‡	230	240	226	218	354	208		314	
Mean intron length (bp)	667	917‡	952	1146	1119	1047	638	800		898	80
Protein length (mean aa)	455	615‡	424	413	450‡	475‡	517	331		154?	280(me

Scleractinian gene content

Analysis of ortholog groups revealed genes and gene families that were shared or unique to specific scleractinian species. Comparison of the four scleractinians (Fig. 1) identified 7,536 gene families shared by all corals, constituting putative coral 'core' genes. Each coral species also contained numerous gene families that were not shared by others, constituting lineage-specific genes. Since the number of shared ortholog groups between a given pair of species is related to the evolutionary distance that separates them, the highest number of

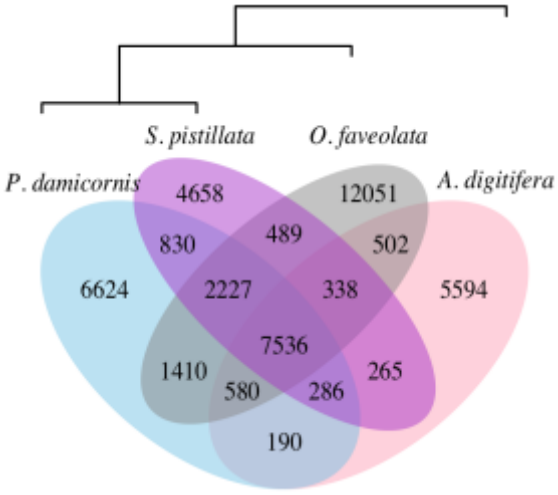

Figure 1: Figure 1. Species-specific and shared gene families across four scleractinian genomes. Numbers indicate the number of gene families, which include both single-copy genes and multi-copy gene families. Dendrogram is based on shared gene content, following 27.

shared ortholog groups was between *P. damicornis* and *S. pistillata*, the two most closely related corals in this analysis. Indeed, a dendrogram based on the number of shared gene families relative to the total number of gene families[27] reproduces the evolutionary relationships among these corals[28] (Fig. 1).

Coral core genes

The coral 'core' gene families (i.e., those found in all four corals) comprise 17.3% of all gene families identified among the four genomes. In *P. damicornis*, core gene family members totaled 12,147, or 46.6% of all genes. Functional profiling of this core genome in *P. damicornis* revealed significant enrichment of 44 GO terms associated with basic cellular and metabolic functions, including nucleic acid synthesis and processing, cellular signaling and transport, and lipid, carbohydrate, and protein metabolism (Fig. 2, SUPP DATA). This basic functionality explains why over 30% of these gene families are also found in all other cnidarians, and 96.3% have orthologs in at least one non-coral. This is consistent with the identification of basic housekeeping functions in the coral core protein sets characterized by 9 and 13.

Coral-specific genes are highly enriched for immunity and copper-ion related genes

Among the coral core gene families, 278 (3.7%) had no orthologs outside Scleractinia, indicating these may be coral-specific gene families. These coral-specific genes in *P. damicornis* (n=349) were significantly enriched for GO terms related to immune function, such as viral defense, signal transduction, and NF-κB pathway regulation (Table 2). NF-κB signaling plays a central role in innate immunity[29,30], and was recently demonstrated to be conserved and responsive to immune challenge in *O. faveolata*[31]. The 32 genes in this set associated with signal transduction may also represent coral-specific immune pathways; SwissProt annotations of these genes included dopamine receptors, neuropeptide receptors, G-protein coupled receptors, and tumor necrosis factor (TNF) receptor-associated factors (TRAFs) (Supplementary Data S1). The TNF receptor superfamily in *A. digitifera* is comprised of 40 proteins, and is more diverse in corals than any organism described thus far other than choanoflagellates[5,32]. That similar genes were present in each coral genome but absent from other cnidarians suggests that diversification of these immune signaling pathways is a common feature of corals. Indeed, *P. damicornis* contained 39 proteins with TNFR cysteine-rich domains. Finally, caveola assembly, or the formation of structures in cell membranes that anchor transmembrane proteins, may

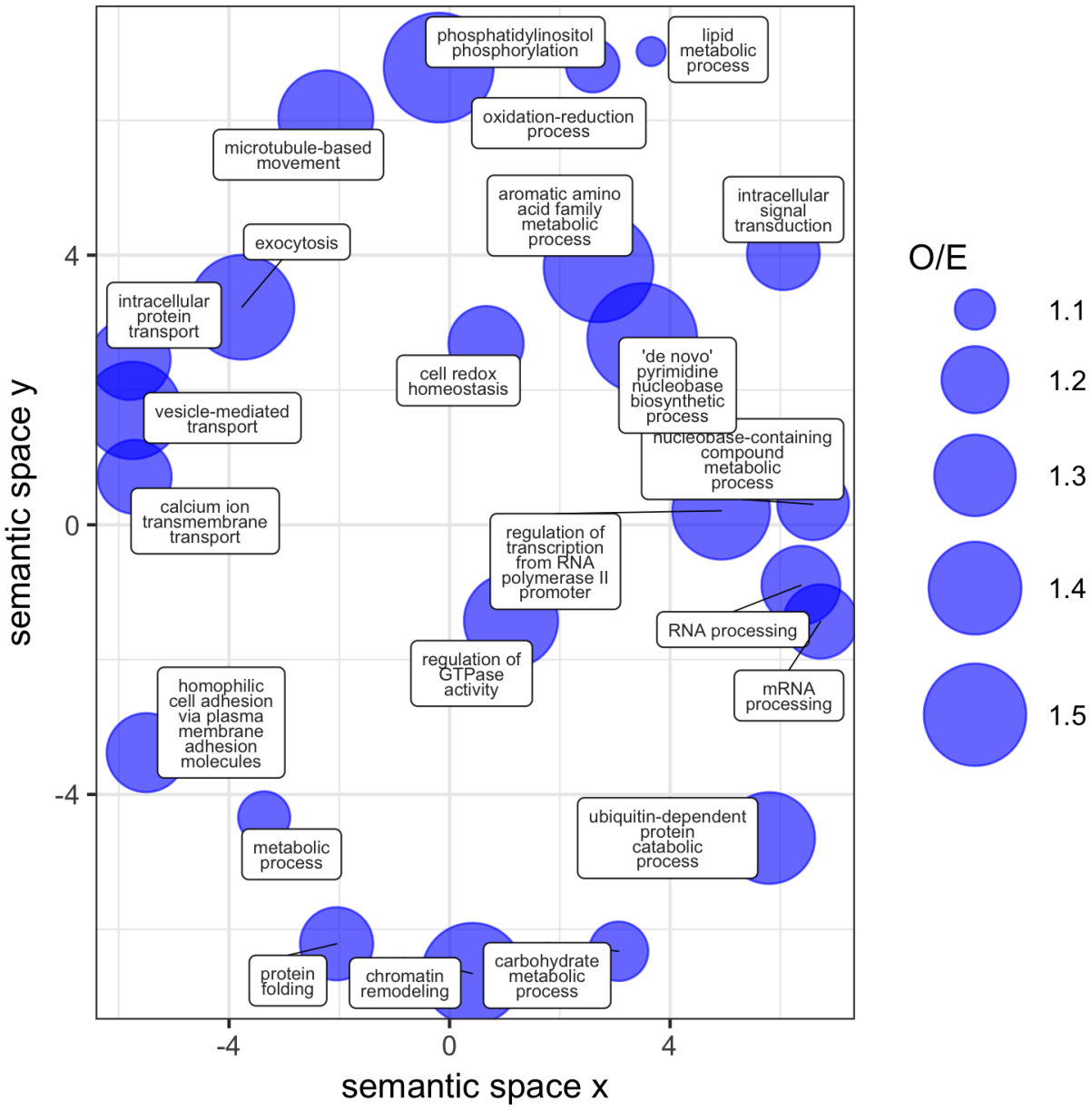

Figure 2: Figure 2. Functional profile of the core coral genome in *P. damicornis*. Significantly enriched GO terms ($p < 0.05$) within this gene set were reduced (allowed similarity = 0.4) and visualized in semantic space by REVIGO. Point size is scaled to the ratio of observed to expected (O/E) occurrences of each GO term in the coral core gene set.

also serve an important role in signal transduction and immunity[33]. Together, these results suggest that corals as a group have evolved a diverse set of immune signaling genes for interacting with and responding to pathogens and the environment.

Another enriched function in the coral-specific gene set was copper ion transmembrane transport (Table 2), which may reflect an important role for delivery of copper to endosymbionts, where it is a critical component of proteins involved in photosynthesis (plastocyanin) and antioxidant activity (superoxide dismutase)[34]. Indeed, in mycorrhizal symbioses, fungi are known to deliver copper to the photosynthetic plant partner[35]. Other enriched functions in the coral-specific genes include regulation of transcription and chromatin silencing, potentially indicating coral-specific pathways of controlling gene expression.

Table 2: Enrichment of GO terms in the coral-specific and coral-diversified genes in *P. damicornis*.

Gene set	GO Accession	GO Term Name	Observed	Expected	p
Coral-specific	GO:0035434	copper ion transmembrane transport	2	0.10	0.00
	GO:0007165	signal transduction	32	23.78	0.00
	GO:0046654	tetrahydrofolate biosynthetic process	1	0.01	0.01
	GO:0051092	NF-kappaB activation	1	0.02	0.02
	GO:0070836	caveola assembly	1	0.04	0.04
	GO:0051607	defense response to virus	2	0.32	0.04
	GO:0006355	regulation of gene-specific transcription	12	6.06	0.05
	GO:0000183	chromatin silencing at rDNA	1	0.05	0.05
	GO:0045132	meiotic chromosome segregation	1	0.05	0.05
Coral-diversified	GO:0006857	oligopeptide transport	3	0.05	0.00
	GO:0007264	small GTPase mediated signal transduction	6	1.05	0.00
	GO:0001822	kidney development	2	0.05	0.00
	GO:0006468	protein phosphorylation	11	3.82	0.00
	GO:0007219	Notch signaling pathway	2	0.11	0.00
	GO:0033151	V(D)J recombination	1	0.01	0.01
	GO:0045737	positive regulation of CDK activity	1	0.01	0.01
	GO:0008152	metabolic process	31	30.87	0.01
	GO:0016255	attachment of GPI anchor to protein	1	0.04	0.04

Gene families diversified in the scleractinian clade

In addition to gene families that were only found in scleractinians, there were also gene families (n=21) that were significantly larger in scleractinians compared to corallimorphs and anemones ($p < 0.01$; Fig. 3), which may indicate diversification within the scleractinian lineage. Members of these gene families in *P. damicornis* (n=339) were significantly enriched for 9 GO terms suggesting roles in cellular signaling and immunity (Table 2). To further investigate these 21 coral-diversified gene families, we compared representative proteins to the Swissprot database and found significant similarities to proteins with known immune functions (Fig. 3, Supplementary Data S2).

The top SwissProt hits of the coral-diversified gene families included receptors for pathogen recognition, such as a C-type lectin, a G-protein-coupled receptors (GPCRs), and both Notch and Wnt-signaling receptors (lipoprotein receptor-related protein). Notch and Wnt signaling are critical developmental gene pathways that may also have a role in coral innate immunity[36], particularly in wound-healing processes[30,37]. Other coral-diversified genes were similar to Ras-related proteins with leucine-rich repeats, and a tetratricopeptide repeat-containing protein, which may also play roles in signal transduction[38]. Many of these tetratricopeptide repeat proteins also contained a CHAT domain characteristic of caspases, indicating a potential role in apoptotic signaling[39]. Another coral-diversified gene group was similar to Poly (ADP-ribose) polymerase, which may act as an anti-apoptotic signal transducer[40]. Lastly, lactadherin-like genes in the coral-diversified gene set may be involved in phagocytosis and clearance of apoptotic cells[41]. Several of the other coral-diversified gene families were similar to genes that were differentially expressed in corals in response to stress or immune challenges, including the HSP70 co-chaperone sacsin[42], the oligopeptide transporter solute carrier

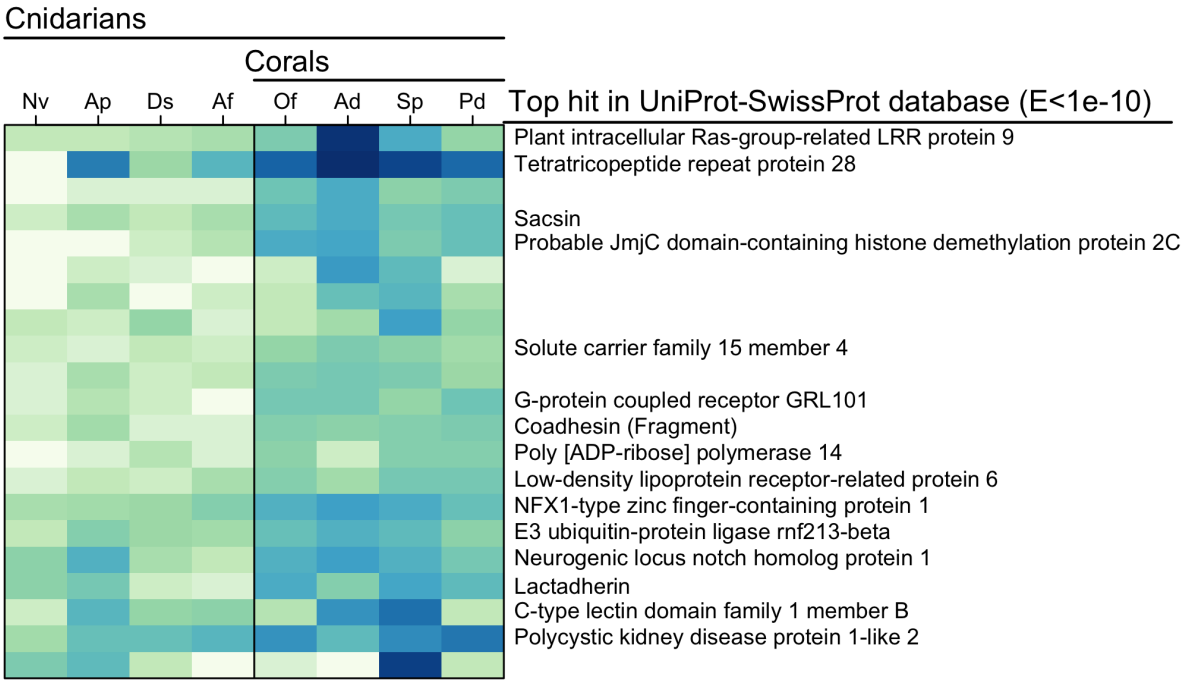

Figure 3: Figure 3. Heatmap showing gene ortholog groups that were larger in scleractinians compared to other cnidarians (Nv=*N. vectensis*, Ap=*A. pallida*, Ds=*Discosoma sp.*, Af=*A. fenestrafer*, Of=*O. faveolata*, Ad=*A. digitifera*, Sp=*S. pistillata*, Pd=*P. damicornis*). For each ortholog group, the longest protein sequence from *P. damicornis* was compared to the UniProt-SwissProt database using blastp, and the top hit was selected based on the lowest E-value (if < 1e-10).

family 15[44], and NFX1-type zinc finger protein[45], indicating that these gene families may also have important roles in the innate immune response.

Besides immunity, coral-diversified gene families may also play roles in other unique functional traits of corals. One defining trait of scleractinians is their production of calcium carbonate skeletons. Coral-diversified gene families that may be involved in this process include genes similar to calcium ion channels (e.g., polycystins) and cell adhesion proteins (e.g., coadhesin, Fig. 3) that have previously been identified as components of the skeletal organic matrix with potential roles in biomineralization[6,46]. Finally, we also found significant similarities between coral-diversified genes and a histone demethylation protein, potentially indicating an important role of histone modification in regulating gene expression, which could drive phenotypic plasticity and epigenetic inheritance in corals[47].

Coral species-specific gene family expansions

The number of gene families in all corals decreased exponentially as gene family size increased, consistent with patterns of gene family size in other organisms[48] (Fig. 5). *P. damicornis* had smaller gene families overall, and the fewest large gene families (n=3 with size>32, max size=75). The most large gene families were observed in *A. digitifera* (n=25 with size>32, max size=255), consistent with pervasive gene duplication in this species suggested by 13. However, statistical comparison of gene family sizes across the four coral species, which accounts for differences in total genome size, indicated that *S. pistillata* had the most significantly expanded gene families relative to other corals (n=15), followed by *A. digitifera* (n=10). *O. faveolata* only had one gene family that was significantly expanded, while *P. damicornis* had none. This finding confirms that uneven gene family size[13] is characteristic of scleractinians, suggesting that lineage-specific gene family

expansion is a major driver of evolution in the Scleractinia.

Lineage-specific gene family expansion may represent an important mechanism of molecular evolution driving adaptation, or may reflect the presence of 'genetic parasites' that propagate across the genome[49]. Among the gene families with lineage-specific expansions in corals, several were similar to reverse transcriptases and transposable elements (see SwissProt annotations in Supplementary Data S3 and Supplementary Table S1), suggesting they may indeed be genetic parasites. However, others may be important in interactions with the environment, cellular signaling, and immunity (Supplementary Data S3), which are the functional categories most likely to undergo lineage-specific expansions[50]. While gene families with different sizes across the four scleractinian genomes could also reflect variation in assembly completeness and quality[51], their functional annotations are consistent with expectations of true gene family expansion.

One expanded gene family in *A. digitifera* was similar to NOD-like receptors (NLRs), which are cytoplasmic pattern recognition receptors that play a key role in pathogen detection and immune activation[52]. Characterized by the presence of NACHT domains, NLR genes have been found to be highly diversified, yet highly variable in number, in the genomes of different cnidarians[53] and other species[49]. The expansion in *A. digitifera* is consistent with these observations, and may represent adaptation to a new pathogen environment[54], or to species-specific symbiotic interactions with microbial eukaryotes and prokaryotes[30]. Another expanded gene family in *A. digitifera* was similar to ephrin-like receptors, which may mediate signaling cascades and cell-cell communication[55]. In *S. pistillata*, one expanded gene family was similar to tachylectin-2, a pattern recognition receptor that has been identified in many cnidarians[30]. A tachylectin-2 homolog was found to be under selection in the coral *Oculina*[56], providing more evidence that such genes are involved in adaptive evolution in corals. The one significantly expanded gene family in *O. faveolata* did not have a strong hit in the SwissProt database, but did contain a caspase-like domain, suggesting a potential role in apoptotic signaling. Overall, differential expansion of genes related to the immune system is consistent with the findings of 13, and suggests that this phenomenon is generally applicable to corals.

In addition to putative immune-related function, genes that have undergone lineage-specific expansions in corals may also play roles in biomineralization. For example, one significantly expanded gene family in *A. digitifera* was similar to a CUB and peptidase domain-containing protein that was found to be secreted in the skeletal organic matrix[46], and another in *S. pistillata* was similar to fibrillar collagen with roles in biomineralization[57].

P. damicornis-specific genes

Although the *P. damicornis* genome did not contain any gene families that were significantly expanded relative to the other corals, it did contain many genes (n=6966) with no orthologs in other genomes. These *P. damicornis*-specific genes were significantly enriched for 11 GO terms. The most significantly enriched GO terms in the Pdam-specific genes were G-protein coupled receptor (GPCR) signaling pathway, bioluminescence, activation of NF-kappaB-inducing kinase, and positive regulation of JNK cascade. The mitogen-activated protein kinase JNK plays a role in responses to stress stimuli, inflammation, and apoptosis[58]. JNK prevents the accumulation of reactive oxygen species (ROS) in corals in response to thermal and UV stress, and, inhibition of JNK leads to coral bleaching and cell death[59]. The NF-kappaB transcription factor may also link oxidative stress and apoptosis involved in coral bleaching[60], in addition to its central role in innate immunity[30]. The occurrence of a lineage-specifc genes that may function in these pathways indicates that *P. damicornis* may have evolved unique strategies for coping with environmental stress based on these immune pathways.

An expanded role of immunity in *P. damicornis* may explain how *Pocillopora* has achieved such a widespread distribution[14,61]. Indeed, *Pocillopora* corals are a fast-growing and weedy pioneer species in Hawaii[62], GBR[63], and Panama[64]. In fact, throughout the Eastern Pacific, *Pocillopora* thrives in marginal habitats, often dealing with elevated turbidity and reduced salinity after heavy rainfall events, subaerial exposure during extreme low tides, and warm- and cold-water stress due to ENSO events and periodic upwelling[65].

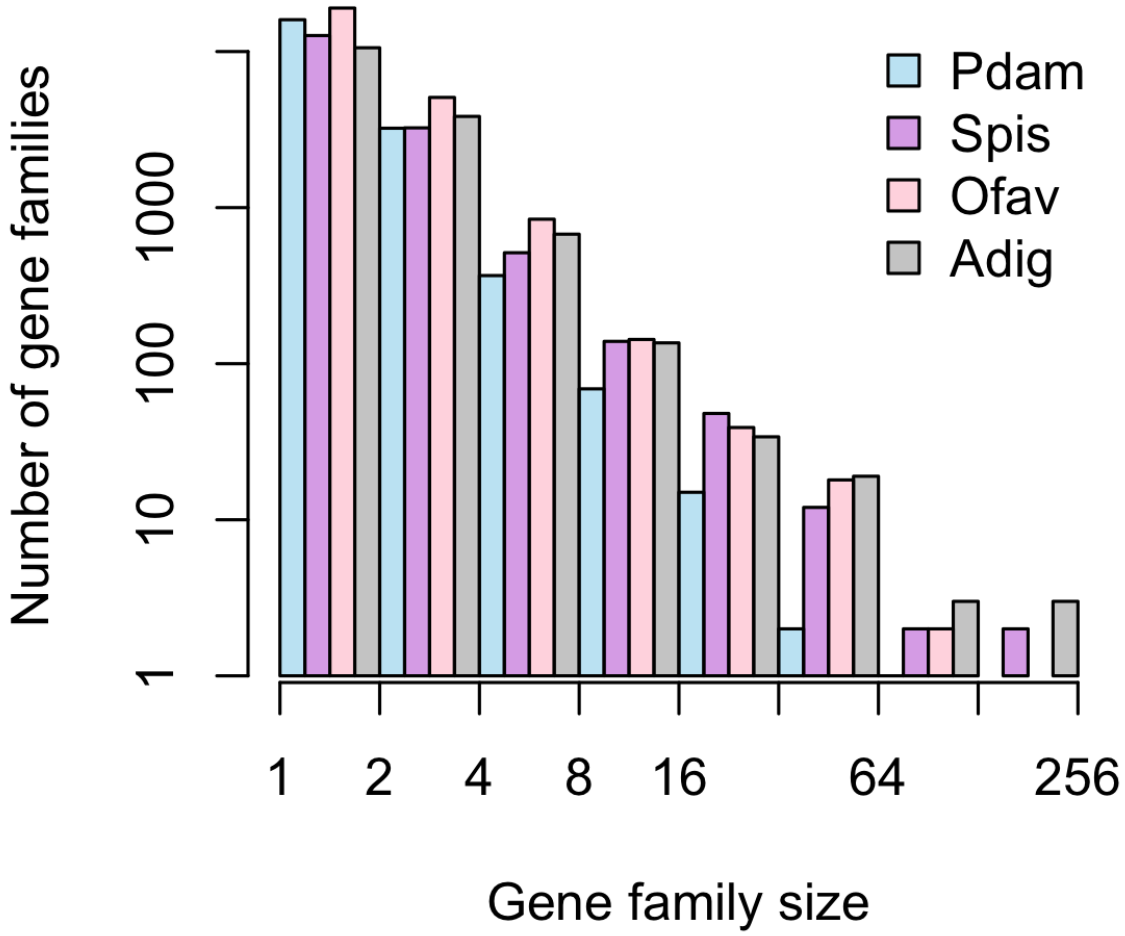

Figure 4: Figure 4. Gene family size distribution in four coral genomes. Pdam=*P. damicornis*, Spis=*S. pistillata*, Ofav=*O. faveolata*, Adig=*A. digitifera*. Bars represent the total number of gene families in a given size class using exponential binning, with each interval open on the left (i.e., the first interval contains gene families of size 1, the second interval contains gene families of size 2 and 3, etc.).

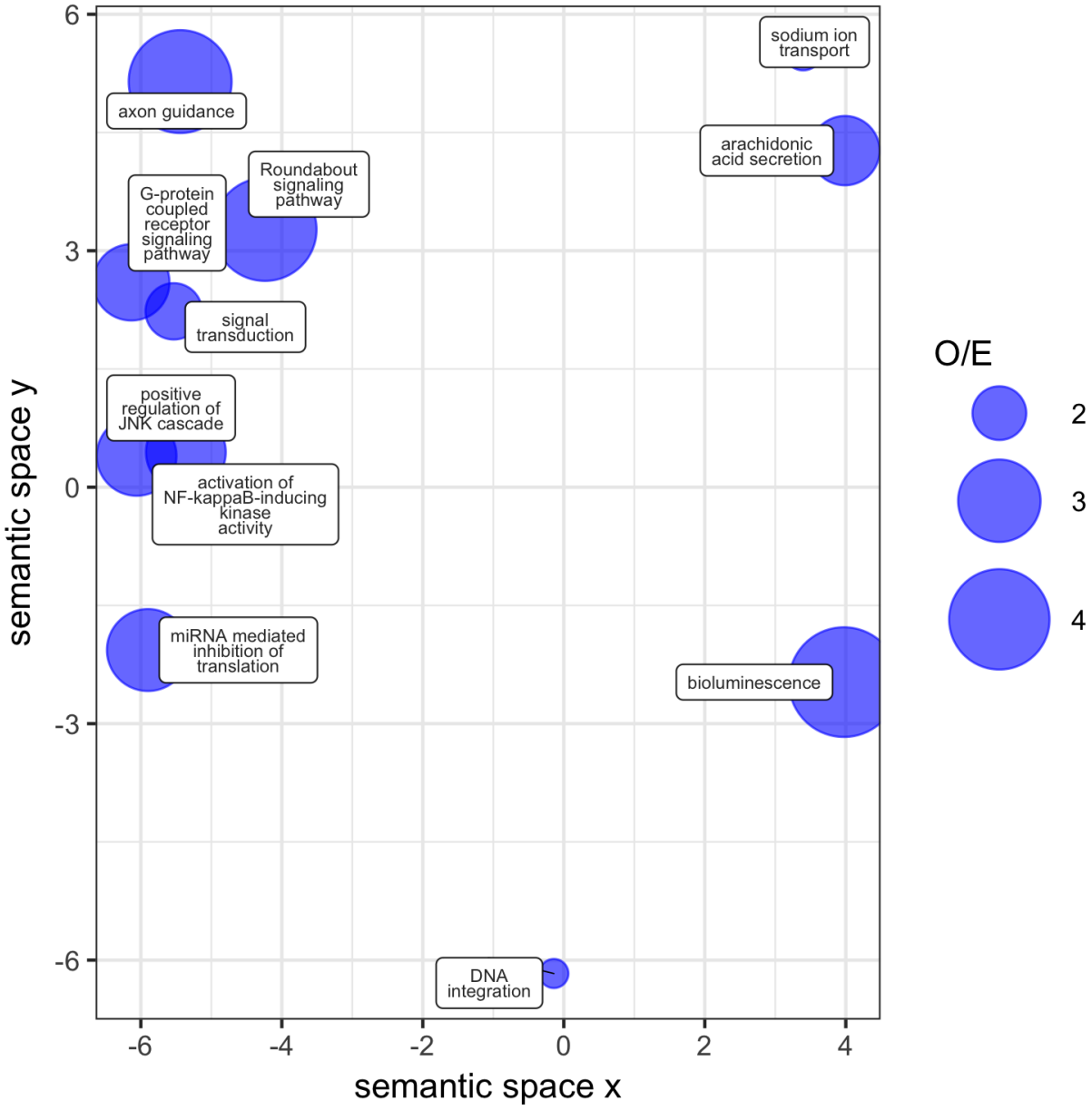

Figure 5: Figure 5. Functional profile of the gene families only found in *P. damicornis*. Significantly enriched GO terms ($p < 0.05$) within this gene set were visualized in semantic space by REVIGO. Point size is scaled to the ratio of observed to expected (O/E) occurrences of each GO term in the coral core gene set.

Conclusions

This comparative analysis revealed significant expansion of immune-related pathways within the scleractinian clade, and further lineage-specific diversification in each scleractinian species. Different immune genes were diversified in each species (e.g., Nod-like and tachylectin-like receptors in *A. digitifera* and *S. pistillata*, and caspase-like and JNK signaling genes in *O. faveolata* and *P. damicornis*), suggesting diverse adaptive roles. Indeed, immune pathways govern the interactions between corals and their algal endosymbionts[66,67], the susceptibility of corals to disease[68], and their responses to environmental stress[59]. Therefore, prominent diversification of immune-related functionality across the Scleractinia is not surprising, and may underlie responses to selection involving symbiosis, self-defense, and stress-susceptibility.

The specific roles of the scleractinian and species-specific immune repertoires should be elucidated with further study to better understand the genomic underpinnings of coral responses to stress and their future under global climate change. Indeed, factors placing high selection pressure on corals, such as bleaching and disease, both which involve dysfunction of the immune system. Shared adaptations potentially evolved in a common ancestor, while species-specific adpatations indicate corals continue to evolve novel immune-related functionality in response to niche-specific selection pressures. These results suggest that evolution of the innate immune system has been a defining feature of the success of scleractinian corals, and likewise may mediate their continued success under scenarios of global climate change. Importantly, the complete sequences of more coral genomes, such as *P. damicornis*, will help elucidate coral futures, and provide the genomic resources to enable future study of these systems moving forward.

Methods

P. damicornis genome sequencing and assembly

The *P. damicornis* genotype used for sequencing was collected at Isla de Saboga, Panama in March 2005, and cultured indoors at the University of Miami Coral Resource Facility until the time of sampling. Genomic DNA was extracted from two healthy fragments and two bleached fragments of this genotype in September 2016 using a Qiagen DNAeasy Midi kit. DNA was shipped overnight on dry ice to Dovetail Genomics (Santa Cruz, CA, USA), where Chicago libraries were prepared and sequenced on an Illumina XXX platform, and genome scaffolds were assembled *de novo* using the HiRise software pipeline[69]. The Dovetail HiRise scaffolds were then filtered to remove those of potential non-coral origin using BLAST[70] searches against three databases: 1) *Symbiodinium*, containing the genomes of *S. minutum*[71] and *S. microadriaticum*[72], 2) bacteria, containing 6954 complete bacterial genomes from NCBI, and 3) viruses, containing 2996 viral genomes from the phantome database (phantome.org; accessed 2017-03-01). Scaffolds with a BLAST hit to any of these databases with an e-value $< 10^{-20}$ and a bitscore > 1000 were considered to be non-coral in origin and removed from the assembly[73].

P. damicornis genome annotation

The filtered assembly was analyzed for completeness using BUSCO[74] to search for 978 universal metazoan single-copy orthologs. The `--long` option was passed to BUSCO in order to train the *ab initio* gene prediction software Augustus[75]. Augustus gene prediction parameters were then used in the MAKER pipeline[76] to annotate gene models, using as supporting evidence two RNA-seq datasets from *P. damicornis* [Mayfield?; Traylor-Knowles/Bhattacharya?], one from closely-related *S. pistillata*[13], and protein sequences from 20 coral species[9]. Results from this initial MAKER run were used to train a second gene predictor (SNAP[77]) prior to an iterative MAKER run to refine gene models. Predicted protein sequences were then extracted from the assembly and putative functional annotations were added by searching for homologous proteins in the Uniprot-Swissprot database[78] using BLAST ($E<10^{-5}$), and protein domains using InterProScan[79]. Genome annotation summary statistics were generated using the Genome Annotation Generator software[80].

Comparative genomic analyses

We identified ortholog groups (gene families) among the predicted proteins of four scleractinians, two corallimorpharians, two anemones, one hydrozoan, one sponge, and one ctenophore (Table 1) using the software fastOrtho (http://enews.patricbrc.org/fastortho/) based on the MCL algorithm with a blastp E-value cutoff of 10^{-5}. Based on these orthologous gene families, we defined and extracted several gene sets of interest: 1) gene families that were shared by all four scleractinians (i.e., coral 'core' genes), 2) gene families that were present in all four scleractinians but absent from other organisms (i.e., coral-specific genes), 3) gene families that were significantly larger in scleractinians relative to other anthozoans (Binomial generalized linear model, FDR-adjusted $p<0.01$; i.e., coral-diversified genes), 4) gene families that were significantly larger in each scleractinian species relative to other scleractinians (pairwise comparisons using Fisher's exact test, FDR-adjusted $p<0.01$; i.e., coral species-specific gene family expansions), and 5) genes present in *P. damicornis* with no orthologs in any other genome (i.e., *P. damicornis*-specific genes).

Functional characterization

Putative gene functionality was characterized using Gene Ontology (GO) analysis. GO terms were assigned to predicted *P. damicornis* protein sequences using InterProScan[81]. Significantly enriched GO terms in gene sets of interest relative to the whole genome were identified using the R package topGO[82], and significantly enriched GO terms were clustered and visualized using REVIGO[83]. These analyses were implemented using custom scripts in R, Python, and Unix shell, which are available in the accompanying data repository (http://www.github.com/jrcunning/pdam-genome).

References

Acknowledgements

RC was supported by NSF OCEXXXXXXX. We thank Peter Glynn for the original field collection of the *P. damicornis* genotype used for genome sequencing.

Author contributions

NTK conceived and directed sequencing of the genome of *P. damicornis*, which was cultured by PG. RC, RAB, and NTK designed the comparative genomic study. RC and RAB analyzed the data, and RC prepared figures, tables, and the first draft of this manuscript. All authors interpreted the results, and reviewed, commented on and approved the final manuscript.

Additional information

Data availability: All data and code to reproduce the analyses and figures described herein can be found at github.com/jrcunning/pdam-genome.
Competing interests: The authors declare that they have no competing interests.

1. Costanza, R., Groot, R. de & Sutton, P. Changes in the global value of ecosystem services. *Glob. Environ. Chang.* **26,** 152–158 (2014).

2. Knowlton, N. *et al.* Coral Reef Biodiversity. *Life World's Ocean. Divers. Distrib. Abundance* 65–78

(2010). doi:10.1002/9781444325508.ch4

3. Neubauer, E. F., Poole, A. Z., Weis, V. M. & Davy, S. K. The scavenger receptor repertoire in six cnidarian species and its putative role in cnidarian-dinoflagellate symbiosis. *PeerJ* **4,** e2692 (2016).

4. Palmer, C. V. & Traylor-Knowles, N. Towards an integrated network of coral immune mechanisms. *Proc. R. Soc. B Biol. Sci.* **279,** 4106–4114 (2012).

5. Quistad, S. D. *et al.* Evolution of TNF-induced apoptosis reveals 550 My of functional conservation. *Proc. Natl. Acad. Sci.* **111,** 9567–9572 (2014).

6. Takeuchi, T., Yamada, L., Shinzato, C., Sawada, H. & Satoh, N. Stepwise evolution of coral biomineralization revealed with genome-wide proteomics and transcriptomics. *PLoS One* **11,** (2016).

7. Kleypas, J. A. *et al.* Geochemical consequences of increased atmospheric carbon dioxide on coral reefs. *Science (80-.).* **284,** 118–120 (1999).

8. Jokiel, P. L. & Coles, S. L. Effects of temperature on the mortality and growth of Hawaiian reef corals. *Mar. Biol.* **43,** 201–208 (1977).

9. Bhattacharya, D. *et al.* Comparative genomics explains the evolutionary success of reef-forming corals. *Elife* **5,** e13288 (2016).

10. Oppen, M. J. H. van, Oliver, J. K., Putnam, H. M. & Gates, R. D. Building coral reef resilience through assisted evolution. *Proc. Natl. Acad. Sci.* **112,** 2307–2313 (2015).

11. Dixon, G. B. *et al.* Genomic determinants of coral heat tolerance across latitudes. *Science (80-.).* **348,** 1460–1462 (2015).

12. Bay, R. A., Rose, N. H., Logan, C. A. & Palumbi, S. R. Genomic models predict successful coral adaptation if future ocean warming rates are reduced. 1–10 (2017). doi:10.1126/sciadv.1701413

13. Voolstra, C. R. *et al.* Comparative analysis of the genomes of Stylophora pistillata and Acropora digitifera provides evidence for extensive differences between species of corals. *Sci. Rep.* **7,** 1–14 (2017).

14. Pinzón, J. H. *et al.* Blind to morphology: genetics identifies several widespread ecologically common species and few endemics among Indo-Pacific cauliflower corals. *J. Biogeogr.* **40,** 1–20 (2013).

15. Schmidt-Roach, S., Miller, K. J., Lundgren, P. & Andreakis, N. With eyes wide open: a revision of species within and closely related to the Pocillopora damicornis species complex (Scleractinia; Pocilloporidae) using morphology .... *Zool. J. Linn. Soc.* (2014).

16. Johnston, E. C. *et al.* A genomic glance through the fog of plasticity and diversification in Pocillopora. *Sci. Rep.* **7,** 5991 (2017).

17. Stoddart, J. a. Genetical structure within populations of the corals Pocillopora damicornis. *Mar. Biol.* **81,** 19–30 (1984).

18. Pinzón, J. H. & LaJeunesse, T. C. Species delimitation of common reef corals in the genus Pocillopora using nucleotide sequence phylogenies, population genetics and symbiosis ecology. *Mol. Ecol.* **20,** 311–325 (2011).

19. Combosch, D. J. & Vollmer, S. V. Population Genetics of an Ecosystem-Defining Reef Coral Pocillopora damicornis in the Tropical Eastern Pacific. *PLoS One* **6,** e21200 (2011).

20. Thomas, L., Kennington, W. J., Evans, R. D., Kendrick, G. A. & Stat, M. Restricted gene flow and local adaptation highlight the vulnerability of high-latitude reefs to rapid environmental change. *Glob. Chang. Biol.* **23,** 2197–2205 (2017).

21. Glynn, P. W., Maté, J. L., Baker, A. C. & Calderón, M. O. Coral bleaching and mortality in Panama and Ecuador during the 1997-1998 El Ni{ñ}o-Southern Oscillation event: Spatial/temporal patterns and

comparisons with the 1982-1983 event. *Bull. Mar. Sci.* **69,** 79–109 (2001).

22. McGinley, M. P. *et al.* Symbiodinium spp. in colonies of eastern Pacific Pocillopora spp. are highly stable despite the prevalence of low-abundance background populations. *Mar. Ecol. Prog. Ser.* **462,** 1–7 (2012).

23. Cunning, R. & Baker, A. C. Excess algal symbionts increase the susceptibility of reef corals to bleaching. *Nat. Clim. Chang.* **3,** 259–262 (2013).

24. Stoddart, J. A. Asexual production of planulae in the coral Pocillopora damicornis. *Mar. Biol.* **76,** 279–284 (1983).

25. Ward, S. Evidence for broadcast spawning as well as brooding in the scleractinian coral Pocillopora damicornis. *Mar. Biol.* **112,** 641–646 (1992).

26. Schmidt-Roach, S., Miller, K. J., Woolsey, E., Gerlach, G. & Baird, A. H. Broadcast Spawning by Pocillopora Species on the Great Barrier Reef. *PLoS One* **7,** e50847 (2012).

27. Snel, B., Bork, P. & Huynen, M. a. Genome phylogeny based on gene content. *Nat. Genet.* **21,** 108–110 (1999).

28. Fukami, H. *et al.* Mitochondrial and nuclear genes suggest that stony corals are monophyletic but most families of stony corals are not (Order Scleractinia, Class anthozoa, phylum cnidaria). *PLoS One* **3,** (2008).

29. Hatada, E. N., Krappmann, D. & Scheidereit, C. NF-kB and the innate immune response. *Curr. Opin. Immunol.* **12,** 52–58 (2000).

30. Mydlarz, L. D., Fuess, L., Mann, W., Pinzón, J. H. & Gochfeld, D. J. Cnidarian Immunity: From Genomes to Phenomes. in *Cnidaria, past, present futur.* (eds. Goffredo, S. & Dubinsky, Z.) 441–466 (Springer International Publishing, 2016). doi:10.1007/978-3-319-31305-4_28

31. Williams, L. M. *et al.* A conserved toll-like receptor-to-NF-κB signaling pathway in the endangered coral Orbicella faveolata. *Dev. Comp. Immunol.* (2017). doi:10.1016/j.dci.2017.10.016

32. Quistad, S. D. & Traylor-Knowles, N. Precambrian origins of the TNFR superfamily. *Cell Death Discov.* **2,** 16058 (2016).

33. Patel, H. H. & Insel, P. A. Lipid rafts and caveolae and their role in compartmentation of redox signaling. *Antioxid. Redox Signal.* **11,** 1357–1372 (2009).

34. Festa, R. A. & Thiele, D. J. Copper: an Essential Metal in Biology. *Curr. Biol.* **21,** R877–R883 (2011).

35. González-Guerrero, M., Escudero, V., Saéz, Á. & Tejada-Jiménez, M. Transition Metal Transport in Plants and Associated Endosymbionts: Arbuscular Mycorrhizal Fungi and Rhizobia. *Front. Plant Sci.* **7,** 1–21 (2016).

36. Anderson, D. A., Walz, M. E., Weil, E., Tonellato, P. & Smith, M. C. RNA-Seq of the Caribbean reef-building coral Orbicella faveolata (Scleractinia-Merulinidae) under bleaching and disease stress expands models of coral innate immunity. *PeerJ* **4,** e1616 (2016).

37. DuBuc, T. Q. *et al.* Initiating a regenerative response; cellular and molecular features of wound healing in the cnidarian Nematostella vectensis. *BMC Biol.* **12,** 24 (2014).

38. Schapire, A. L., Valpuesta, V. & Botella, M. A. TPR proteins in plant hormone signaling. *Plant Signal. Behav.* **1,** 229–230 (2006).

39. Aravind, L. & Koonin, E. V. Classification of the caspase-hemoglobinase fold: Detection of new families and implications for the origin of the eukaryotic separins. *Proteins Struct. Funct. Genet.* **46,** 355–367 (2002).

40. Iwata, H. *et al.* PARP9 and PARP14 cross-regulate macrophage activation via STAT1 ADP-ribosylation. *Nat. Commun.* **7,** (2016).

41. Ait-Oufella, H. *et al.* Lactadherin deficiency leads to apoptotic cell accumulation and accelerated

atherosclerosis in mice. *Circulation* **115,** 2168–2177 (2007).

42. Weiss, Y. *et al.* The acute transcriptional response of the coral Acropora millepora to immune challenge: Expression of GiMAP/IAN genes links the innate immune responses of corals with those of mammals and plants. *BMC Genomics* **14,** (2013).

43. Mayfield, A. B., Chen, Y.-J., Lu, C.-Y. & Chen, C.-S. The proteomic response of the reef coral Pocillopora acuta to experimentally elevated temperatures. *PeerJPreprints* (2017). doi:10.7287/peerj.preprints.3252v1

44. Moya, A. *et al.* Near-future pH conditions severely impact calcification, metabolism and the nervous system in the pteropod Heliconoides inflatus. *Glob. Chang. Biol.* **22,** 3888–3900 (2016).

45. DeSalvo, M. K. *et al.* Differential gene expression during thermal stress and bleaching in the Caribbean coral Montastraea faveolata. *Mol. Ecol.* **17,** 3952–3971 (2008).

46. Ramos-Silva, P. *et al.* The skeletal proteome of the coral acropora millepora: The evolution of calcification by co-option and domain shuffling. *Mol. Biol. Evol.* **30,** 2099–2112 (2013).

47. Putnam, H. M., Davidson, J. M. & Gates, R. D. Ocean acidification influences host DNA methylation and phenotypic plasticity in environmentally susceptible corals. *Evol. Appl.* **9,** 1165–1178 (2016).

48. Huynen, M. A. & Nimwegen, E. van. The frequency distribution of gene family sizes in complete genomes. *Mol. Biol. Evol.* **15,** 583–589 (1998).

49. Schiffer, P., Gravemeyer, J., Rauscher, M. & Wiehe, T. Ultra Large Gene Families: A Matter of Adaptation or Genomic Parasites? *Life* **6,** 32 (2016).

50. Lespinet, O., Wolf, Y. I., Koonin, E. V. & Aravind, L. The role of lineage-specific gene family expansion in the evolution of eukaryotes. *Genome Res.* **12,** 1048–1059 (2002).

51. Hahn, M. W., Han, M. V. & Han, S. G. Gene family evolution across 12 Drosophila genomes. *PLoS Genet.* **3,** 2135–2146 (2007).

52. Kanneganti, T. D., Lamkanfi, M. & Núñez, G. Intracellular NOD-like Receptors in Host Defense and Disease. *Immunity* **27,** 549–559 (2007).

53. Hamada, M. *et al.* The complex NOD-like receptor repertoire of the coral acropora digitifera includes novel domain combinations. *Mol. Biol. Evol.* **30,** 167–176 (2013).

54. Stein, C., Caccamo, M., Laird, G. & Leptin, M. Conservation and divergence of gene families encoding components of innate immune response systems in zebrafish. *Genome Biol.* **8,** R251 (2007).

55. Kullander, K. & Klein, R. Mechanisms and functions of Eph and ephrin signalling. *Nat. Rev. Mol. Cell Biol.* **3,** 475–486 (2002).

56. Hayes, M. L., Eytan, R. I. & Hellberg, M. E. High amino acid diversity and positive selection at a putative coral immunity gene (tachylectin-2). *BMC Evol. Biol.* **10,** (2010).

57. Christiansen, H. E., Lang, M. R., Pace, J. M. & Parichy, D. M. Critical early roles for col27a1a and col27a1b in zebrafish notochord morphogenesis, vertebral mineralization and post-embryonic axial growth. *PLoS One* **4,** 1–10 (2009).

58. Newton, K. & Dixit, V. M. Signaling in innate immunity and inflammation. *Cold Spring Harb. Perspect. Biol.* **4,** (2012).

59. Courtial, L. *et al.* The c-Jun N-Terminal kinase prevents oxidative stress induced by UV and thermal stresses in corals and human cells. *Sci. Rep.* **7,** 1–10 (2017).

60. Weis, V. M. Cellular mechanisms of Cnidarian bleaching: stress causes the collapse of symbiosis. *J. Exp. Biol.* **211,** 3059–3066 (2008).

61. Veron, J. E. N. & Stafford-Smith, M. G. *Corals of the World.* (Australian Institute of Marine Science,

2000).

62. Grigg, R. W. & Maragos, J. E. Recolonization of hermatypic corals on submerged lava flows in Hawaii. *Ecology* **55,** 387–395 (1974).

63. Whitney, S. M., Baldet, P. & Hudson, G. S. Form I Rubiscos from non-green algae are expressed abundantly but not assembled in tobacco chloroplasts. *Plant ...* (2001).

64. Glynn, P. W., Stewart, R. H. & McCosker, J. E. Pacific coral reefs of Panama: structure, distribution, and predators. *Geol. Rundschau* **61,** 483–519 (1972).

65. Glynn, P. W. *et al. Coral Reefs of the Eastern Tropical Pacific.* **8,** (2017).

66. Davy, S. K., Allemand, D. & Weis, V. M. Cell biology of cnidarian-dinoflagellate symbiosis. *Microbiol. Mol. Biol. Rev.* **76,** 229–261 (2012).

67. Poole, A. Z., Kitchen, S. A. & Weis, V. M. The role of complement in cnidarian-dinoflagellate symbiosis and immune challenge in the sea anemone Aiptasia pallida. *Front. Microbiol.* **7,** (2016).

68. Libro, S., Kaluziak, S. T. & Vollmer, S. V. RNA-seq profiles of immune related genes in the staghorn coral Acropora cervicornis Infected with white band disease. *PLoS One* **8,** 1–11 (2013).

69. Putnam, N. H. *et al.* Chromosome-scale shotgun assembly using an in vitro method for long-range linkage. *Genome Res.* **26,** 342–350 (2016).

70. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. - PubMed - NCBI. *J. Mol. Biol.* **215,** 403–410 (1990).

71. Shoguchi, E. *et al.* Draft Assembly of the Symbiodinium minutum Nuclear Genome Reveals Dinoflagellate Gene Structure. *Curr. Biol.* **23,** 1399–1408 (2013).

72. Aranda, M. *et al.* Genomes of coral dinoflagellate symbionts highlight evolutionary adaptations conducive to a symbiotic lifestyle. *Sci. Rep.* **6,** 39734 (2016).

73. Baumgarten, S. *et al.* The genome of Aiptasia, a sea anemone model for coral symbiosis. *PNAS* 201513318–201513362 (2015).

74. Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31,** 3210–3212 (2015).

75. Stanke, M., Steinkamp, R., Waack, S. & Morgenstern, B. AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Res.* **32,** W309–W312 (2004).

76. Campbell, M. S., Holt, C., Moore, B. & Yandell, M. Genome Annotation and Curation Using MAKER and MAKER-P. *Curr. Protoc. Bioinforma.* **48,** 4.11.1–39 (2014).

77. Korf, I. Gene finding in novel genomes. *BMC Bioinformatics* **5,** 59 (2004).

78. The UniProt Consortium. UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* **45,** D158–D169 (2017).

79. Finn, R. D. *et al.* InterPro in 2017—beyond protein family and domain annotations. *Nucleic Acids Res.* **45,** D190–D199 (2017).

80. Hall, N. The Kardashian index: a measure of discrepant social media profile for scientists. *Genome Biol.* **15,** 424 (2014).

81. Jones, P. *et al.* InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **30,** 1236–1240 (2014).

82. Alexa, A. & Rahnenfuhrer, J. topGO: Enrichment Analysis for Gene Ontology. R package version 2.24.0. (2016).

83. Supek, F., Bošnjak, M., Škunca, N. & Šmuc, T. Revigo summarizes and visualizes long lists of gene

ontology terms. *PLoS One* **6,** (2011).

84. Srivastava, M. *et al.* The Amphimedon queenslandica genome and the evolution of animal complexity. *Nature* **466,** 720–726 (2010).

85. Wang, X. *et al.* Draft genomes of the corallimorpharians Amplexidiscus fenestrafer and Discosoma sp. *Mol. Ecol. Resour.* **2017,** doi: 10.1111/1755–0998.12680 (2017).

86. Shinzato, C. *et al.* Using the Acropora digitifera genome to understand coral responses to environmental change. *Nature* **476,** 320–323 (2011).

87. Prada, C. *et al.* Empty Niches after Extinctions Increase Population Sizes of Modern Corals. *Curr. Biol.* **26,** 3190–3194 (2016).

88. Putnam, N. H. *et al.* Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science* **317,** 86–94 (2007).

89. Chapman, J. a *et al.* The dynamic genome of Hydra. *Nature* **464,** 592–6 (2010).

90. Ryan, J. F. *et al.* The genome of the ctenophore Mnemiopsis leidyi and its implications for cell type evolution. *Science (80-.).* **342,** 1242592 (2013).