Received: 5 May 2021 | Revised: 10 February 2022 | Accepted: 22 February 2022

DOI: 10.1111/mec.16414

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Adaptive variation in homologue number within transcript families promotes expression divergence in reef-building coral

Bradford Dimos[1] | Madison Emery[1] | Kelsey Beavers[1] | Nicholas MacKnight[1] | Marilyn Brandt[2] | Jeffery Demuth[1] | Laura Mydlarz[1]

[1]Department of Biology, University of Texas at Arlington, Arlington, Texas, USA

[2]Center for Marine and Environmental Studies, University of the Virgin Islands, St. Thomas, US Virgin Islands, USA

Correspondence
Bradford Dimos, Department of Biology, University of Texas at Arlington, Arlington, TX, USA.
Email: braddimos@gmail.com

Funding information
This work was supported by NSF Biological Oceanography Award No. 1712134 awarded to L.M.

Handling Editor: Pim Bongaerts

Abstract

Gene expression, especially in multispecies experiments, is used to gain insight into the genetic basis of how organisms adapt and respond to changing environments. However, evolutionary processes that can influence gene expression patterns between species such as the presence of paralogues which arise from gene duplication events are rarely accounted for. Paralogous transcripts can alter the transcriptional output of a gene, and thus exclusion of these transcripts can obscure important biological differences between species. To address this issue, we investigated how differences in transcript family size are associated with divergent gene expression patterns in five species of Caribbean reef-building corals. We demonstrate that transcript families that are rapidly evolving in terms of size have increased levels of expression divergence. Additionally, these rapidly evolving transcript families are enriched for multiple biological processes, with genes involved in the coral innate immune system demonstrating pronounced variation in homologue number between species. Overall, this investigation demonstrates the importance of incorporating paralogous transcripts when comparing gene expression across species by influencing both transcriptional output and the number of transcripts within biological processes. As this investigation was based on transcriptome assemblies, additional insights into the relationship between gene duplications and expression patterns will probably emergence once more genome assemblies are available for study.

KEYWORDS
CAFÉ, coral disease, coral immunity, homologous transcript families, multi-species gene expression, *Orbicella faveolata*

1 | INTRODUCTION

Quantifying gene expression patterns in order to understand how an organism interacts with its environment has become common place in ecological research. In particular, gene expression is well suited to understand the mechanisms organisms use to respond to the types of stress which are associated with climate change (Kelly, 2019; Logan & Cox, 2020; Rivera et al., 2021). It is also clear that investigating gene expression patterns across multiple species exposed to the same type of stress is important to predict how the community will be reshaped by that particular stressor (Avila-Magaña et al., 2021; Bernal et al., 2020; Dixon et al., 2020; Fuess et al., 2017; Traylor-Knowles et al., 2021). However, multispecies gene expression experiments present a suite of challenges, the first of which is obtaining a comparable list of homologous transcripts for use in downstream analysis (Rey et al., 2019). This barrier is commonly resolved through the identification of one-to-one orthologues across species (Bernal et al., 2020; Foissac et al., 2019; Hao et al., 2019; Huerlimann et al., 2020; Khan et al., 2013). However, this framework necessitates removing all transcripts that do not display this type of one-to-one

conservation, consequently excluding paralogues (Tekaia, 2016) and limiting the biological inferences that can be drawn. Additionally, as the number of species investigated or the phylogenetic distance between species increases, the number of one-to-one homologous transcripts declines. One solution to this problem is to group transcripts into homologous transcript families (HTFs) which include both orthologues and paralogues (Emms & Kelly, 2019), removing the requirement of the one-to-one relationship.

Confidently identifying orthologues and paralogues based on transcriptome data can be challenging given the technical limitations inherent in transcriptome assemblies. However, incorporating paralogues through the identification of HTFs may be particularly relevant to studies which investigate adaptation. Gene gains and losses, which dictate the number of paralogues within an HTF, can become fixed in populations due to both neutral and adaptive processes (Force et al., 1999; Lynch, 2000; Ohno, 1970). While expansion and contraction of HTF size may occur due to neutral processes, there is mounting evidence that some gene families gain and lose genes more rapidly than can be explained by a neutral model (Demuth & Hahn, 2009; Kondrashov, 2012). The most widely accepted explanations for gene family expansion are that it promotes neofunctionalization and subfunctionalization by reducing pleiotropic constraints (Des Marais & Rausher, 2008; Stearns, 2010), or simply increases the gene dose, thereby increasing the potential transcriptional output (Force et al., 1999; Ohno, 1970). Gene losses can also be adaptive through promoting genetic efficiency or removing genes whose costs are not outweighed by their benefits (Albalat & Cañestro, 2016). Numerous recent studies have described how changes in gene family size can promote adaptation and the diversification of organismal traits (Chain et al., 2014; Ronco et al., 2021; Tejada-Martinez et al., 2021). Thus, the exclusion of paralogues from comparative gene expression studies may obscure important biological differences.

In homologous gene families, adaptive changes in copy number vs. nonadaptive changes which arise over evolutionary time can be identified through describing the background rate of gene family turnover, followed by the identification of gene families which have significant deviations from this rate (De Bie et al., 2006; Hahn, 2005). As the background rate of gene turnover represents a genome-wide average, changes that have adaptive significance will demonstrate increased rates of change and can be described as rapidly evolving. This type of rapid gene evolution is particularly common in gene families involved in arms races such as immune genes (Machado & Ottolini, 2015; Malmstrøm et al., 2016). However, other types of ecologically relevant adaptation have been shown to follow this pattern of rapid changes in copy number such as expansions in hypoxia-related genes in high-altitude yaks (Qiu et al., 2012) and chemoreception and vision genes for predator responses in *Daphnia* (Zhang et al., 2021).

The rapid evolution of gene families is generally investigated from genome-derived gene models (Hahn, 2005). However, a similar approach can be applied to transcriptome assemblies, as the paralogues that result from gene duplication events will generate novel transcripts within HTFs (Chen et al., 2021). Using transcriptome assemblies to assess the evolution of gene families has several important limitations (Ungaro et al., 2017). Transcriptome assemblies generally have more missing genes than genomes, as the completeness of a transcriptome is dependent on the cell types and conditions under which a sample was taken (Moreton et al., 2016). Transcriptome completeness influences the reliability of estimating the rate of background turnover as genes that were not transcriptionally active would be incorrectly identified as gene losses. Additionally, for transcriptomes derived from short-read sequencing it can be challenging to disentangle genuine paralogues from isoforms or assembly artefacts. While computational approaches exist to remove isoforms (Haas et al., 2013) and limit the number of artefacts, these approaches increase the likelihood of collapsing recent paralogues (Huang et al., 2010), making rapidly evolving HTFs more difficult to identify. Despite these limitations to using transcriptomic data compared to genomic data, transcriptomes are substantially easier and more cost effective to produce, especially in organisms where high-quality genomes are difficult to generate and yet unavailable. Therefore, there is value in developing a framework for using transcriptomes for preliminary investigations into the evolution of HTFs.

The number of transcripts within HTFs is an important consideration when comparing expression across species, since in principle larger transcript family size results in more spatially or temporally specialized expression (Duarte et al., 2006; Guschanski et al., 2017). The increased transcriptional output associated with increased gene dose is often deleterious and will be quickly resolved through mechanisms such as epigenetic silencing (Huang & Chain, 2021) or functional subdivision of the two gene copies (Smet et al., 2017). The transcriptional activity of paralogues can be assessed through quantifying global gene expression (Song et al., 2020), which shows that expression levels are subject to selection (Veitia & Birchler, 2021), and can result in the generation of new expression patterns across species (Gillard et al., 2021). These types of species-level changes can be identified by comparing how much the expression of a gene varies between vs. within species (Rohlfs & Nielsen, 2015). This approach has proved successful in identifying divergent transcriptional responses to a heatwave in five species of coral reef fish (Bernal et al., 2020) and the unique expression of detoxification genes in *Poecilia Mexicana* which inhabit sulphide springs (Greenway et al., 2020). Importantly, this framework can be used to investigate if the rapid evolution of transcript families is associated with expression divergence.

Reef-building corals are an intriguing system to study how variation in the number of paralogues within HTFs influences expression divergence given the focus on using gene expression to investigate coral adaptation (Bay & Palumbi, 2015, 2017; Kenkel & Matz, 2017; Traylor-Knowles et al., 2021). Despite the importance and significance of gene expression data, those studies which investigate expression across multiple species either do so in a qualitative way (Fuess et al., 2017; Traylor-Knowles et al., 2021) or employ closely related species allowing reads to be mapped to a common reference assembly (Dixon et al., 2020). To date, only

one published study in coral has quantitatively compared expression across species from different genera (Avila-Magaña et al., 2021), but this study excluded paralogues from their analysis. Whole tissue homogenates provide a good sampling of coral cell types (Levy et al., 2021) due to their relatively simply body plan. Demographically, corals have historically possessed large effective population sizes (Matz et al., 2018; Prada et al., 2016) and high genetic connectivity (Davies et al., 2015), conditions which favour efficient selection. Further, adult corals are sessile, meaning that they cannot track favourable environmental conditions and instead must be able to dynamically adjust gene expression profiles in order to persist during unfavourable conditions. Adaptive changes in gene copy number have been demonstrated to accommodate this sessile life-history strategy in plants (Prunier et al., 2017) and thus there may be substantial variation in HTFs within corals. In support of this, some investigations have reported variation in multicopy gene families between coral species (Dimos et al., 2019; Hamada et al., 2013; Shinzato et al., 2020), but a systematic evaluation of this process between species from different genera and how this influences expression patterns has not been performed.

Anthropogenic impacts on the environment have caused widespread coral loss across the globe (Maynard et al., 2015; Stuart-Smith et al., 2018), prompting numerous investigations which seek to understand the stress responses and adaptive capacity of reef-building corals. In particular, Caribbean coral reefs have been severely depleted (Gardner, 2003), owing to the combined effects of thermally induced coral bleaching and the emergence of diseases affecting reef-building corals (Smith et al., 2013; van Woesik & Randall, 2017). The coral loss in the Caribbean has changed these communities by reducing coral cover and reshaping existing species assemblages (McWilliam et al., 2020). As an example, the once dominant branching corals in the genus *Acropora* have been functionally extirpated from most of their historical range due to white band disease (Aronson & Precht, 2001; Miller et al., 2009), and the emergence of white plague disease has led to substantial reductions in the abundance of corals in the genus *Orbicella* (Miller et al., 2009). The loss of these once dominant species has led to the increased relative abundance of species in other genera such as *Porites* and *Siderastrea* (Green et al., 2008; McWilliam et al., 2020). Studies focusing on comparable aspects of species biology, such as growth rate, immune activity and Symbiodiniaceae communities, have sought to draw links between species differences and the heterogenous responses these species show during environmental stress (Baumann et al., 2018; Bove et al., 2019; MacKnight et al., 2021; Pinzón et al., 2014). However, the genetic component and interspecies differences in expression have not been investigated in a quantitative way.

In this investigation we use gene expression data to investigate the evolution of HTFs, in five ecologically important species of Caribbean corals: *Orbicella faveolata*, *Montastraea cavernosa*, *Colpophyllia natans*, *Porites astreoides* and *Siderastrea siderea*. Our analysis demonstrates that HTFs which are rapidly evolving are enriched in numerous biological processes in each species and these transcript families have higher levels of expression divergence. These results highlight variation in transcript families which are linked to important immune mechanisms. Given the limitations of transcriptomic analysis noted in the introduction, it is possible that additional processes are experiencing rapid evolution that our analysis is not able to detect. Future work incorporating assembled genomes will improve the resolution of our findings and the robustness of the patterns observed. Overall, our study provides support for the association between rapid evolution of the number of paralogues within HTFs and its influence on gene expression patterns by providing a framework which allows quantitative comparison of gene expression patterns even among highly divergent species.

2 | METHODS

2.1 | Coral collection

Fragments of five colonies of each species, *Orbicella faveolata*, *Montastraea cavernosa*, *Colpophyllia natans*, *Porites astreoides* and *Siderastrea sidereal*, were collected via SCUBA from Brewer's Bay St. Thomas (18.34403, −64.98435) from ~35 ft of depth in June 2017 under the Indigenous Species Research and Export Permit number CZM17010T. After removal from the parent colony, coral fragments were placed in individual sealed bags and transported to the Center for Marine and Environmental Studies in coolers with seawater. Fragments were held in large flow-through seawater tables fed with filtered sea water from the bay which received natural sunlight under a neutral density shade cloth. After an acclimation period of 2 weeks, the corals were fragmented into 5 × 5-cm fragments with a sterilized chop saw or bandsaw, given identification numbers, photographed, then placed back into running seawater tables under shade to heal for a further 9 days. The temperature in the flow-through tanks approximates the temperature in the bay reaching a daily high of ~28.3°C and average of 27.71 ± 0.033°C, measured using an Onset HOBO U22-001 Water Temperature Data Logger. To further confirm the environmental conditions were comparable between each coral fragment from each species before collection for this analysis, each coral fragment was placed into individual mesocosms made from 5-L plastic multipurpose mixing buckets and each mesocosm was aerated with an airstone attached to the central air line. The containers were randomly placed in another large holding tank where the water line was below the top of the container to ensure that temperature remained constant among all samples, but that no water was exchanged between containers. All coral fragments from each species were randomized in the water table and held in these individual containers for 6 h. The environmental conditions in the water table were the same for each coral fragment both during the experiment and during the acclimation-healing period. After the experiment, the coral fragments were flash frozen in liquid nitrogen and transported back to the University of Texas at Arlington in a dry shipper. Samples were flash frozen in random order and the entire collection period lasted ~90 min.

2.2 | RNA extraction and sequencing

Coral tissue was removed from the frozen samples by excising ~1 cm^2 of tissue using bone cutters from the surface of the fragment while avoiding the underlying skeleton and homogenized in lysis buffer. Total RNA was extracted from coral fragments using the Ambion RNAeasy kit with DNase according to manufacturer's protocol and eluted in 100 μl of elution buffer. RNA integrity and concentration were checked with an Agilent Bioanalyzer 2100 using the RNA 6000 Nano kit, and all samples with a RIN >7.0 and >1 μg of total RNA were used for sequencing. Samples were sequenced through the Novagene company and mRNA libraries were prepared with the Illumina TruSeq RNA library prep kit, which uses Poly-A tail enrichment to purify mRNA. After library preparation, samples were sequenced on an Illumina HiSeqX at Novagene. Five colonies of each *M. cavernosa* and *C. natans* were sequenced, while four colonies of *O. faveolata*, *P. astreoides* and *S. siderea* were sequenced. Raw reads from this experiment can be found on the NCBI short read archive (PRJNA723585).

2.3 | Transcriptome assembly and cleaning

Transcriptome assemblies were generated with TRINITY version 2.5.1 (Grabherr et al., 2011; Haas et al., 2013) with parameters (--normalize_reads --seqType fq --SS_lib_type FR) after removing adapters and low-quality reads with TRIMMOMATIC version 0.32 (Bolger et al., 2014) using default parameters. Since transcriptome assemblies derived from adult corals contain transcripts originating from their symbiotic algae, we used a previously described *in silico* approach to remove symbiont transcripts (Davies et al., 2016). In short, multiple coral assemblies consisting of both genome-derived predicted gene models and transcriptomes covering a diversity of coral families (Davies et al., 2016; Kirk et al., 2018; Moya et al., 2012; van de Water et al., 2018) were combined to create a coral database. A symbiont database was constructed from the genome-derived predicted gene models from multiple genera of symbionts (Aranda et al., 2016; Liu et al., 2018; Parkinson et al., 2016; Shoguchi et al., 2021). Transcripts from each of our coral assemblies were then queried using BLASTN 2.2.27 against each database and a transcript was retained only if it had >80% identity over at least 100 bp to the coral database and that transcript did not have a lower e-value when blasted against the symbiont database.

To confidently identify HTFs, it is important to remove transcripts which are probably derived from alternative splice variants and isoforms to approximate a one-to-one correspondence between transcripts and genes. We took several steps to ensure these relationships held before conducting subsequent analyses. First, each assembly was filtered for the longest isoform to attempt to remove extraneous transcripts in our assemblies which derive from splice-variants of a gene using the get longest isoform script available through TRINITY version 2.5.1 (Grabherr et al., 2011). Next the program TRANSDECODER version 5.5.0 (Haas et al., 2013) was used to extract the longest open reading frame of each transcript and this reading frame was utilized to generate a predicted peptide sequence. The predicted proteomes for all species were then collapsed for sequence similarity using the program CD-HIT version 4.8.1 (Huang et al., 2010) at a similarity level of 0.9 to further exclude transcripts which are probably derived from alternative splice-variants of a gene. To ascertain the quality of these assemblies after transcript filtration, we employed several metrics of transcriptome completeness including Benchmarking Universal Single Copy Orthologs (BUSCO) version 5.2.2-0 (Simão et al., 2015) against the core metazoan database, DOGMA version 3.6 against the eukaryote protein domain database, N50 and total transcriptome assembly length. While we employed extensive efforts to remove splice isoforms and assembly artefacts inherent to transcriptome assemblies, it is likely that these efforts additionally removed recent paralogues and that future analysis would benefit from the inclusion of complete genome assemblies. To generate Gene Ontology (GO) annotations from our transcriptome assemblies we utilized the annotation software EGGNOG-MAPPER version 2 (Huerta-Cepas et al., 2017, 2019) through the online web portal against the eukaryote database.

2.4 | Orthogroup assignment

To identify homologous transcripts across our species and group them into HTFs we utilized the program ORTHOFINDER version 2.5.4 (Emms & Kelly, 2015, 2019). This software groups protein sequences predicted from the transcripts into groups containing both orthologues and paralogues, which we refer to as HTFs. A species tree was generated through the program Species Tree From All Genes (STAG) version 1.0.0 (Emms & Kelly, 2018) as implemented within ORTHOFINDER where branch lengths are measured in amino acid substitutions. We then converted the calculated branch lengths into time based on the estimated divergence times between these coral species (Pinzón et al., 2014).

2.5 | Transcript family evolution

To quantify the rate of gene family evolution we employed the program CAFÉ version 4.2.1 (Han et al., 2013). This program utilizes an ultrametric species tree with a set of gene families to determine the background rate of gene gain and loss per gene per unit time across the provided species tree. This background rate of gene turnover is then used as a null model to test for gene families which are experiencing significant deviations from this rate. For our analysis the species tree generated through STAG was converted into an ultrametric tree using PHYTOOLS version 0.7.70 (Revell, 2012) with the option "extend." We ran this analysis on the HTFs that contained at least one transcript in all species, which encompassed 4,904 HTFs. To account for the role that errors in our assemblies may have played in over-inflating our estimate of gene turnover, we utilized the error model

TABLE 1 Metrics of assembly completeness

Species	Source	Predicted peptides	No. of transcripts used in CAFÉ	Complete single copy	Fragmented	Complete and duplicated	Missing	N50	Protein domains
O. faveolata	*de novo*	47,678	7,986	698	39	84	157	11,934	1,091
M. cavernosa	*de novo*	57,740	9,978	757	34	78	109	13,783	1,186
C. natans	*de novo*	50,582	9,266	652	134	57	135	12,255	952
P. astreiodes	*de novo*	37,167	9,048	787	73	49	69	6,457	1,197
S. siderea	*de novo*	16,635	7,707	761	74	23	120	4,921	1,126

script included in the CAFÉ package (Han et al., 2013). This script iteratively searches a priori defined error distributions to identify which distribution maximizes the probability of observing the data. The error term with the highest probability of observation, 0.1 (which represents a predicted error in 10% of HTFs), was then used to run CAFÉ. To account for the possibility that assembly completeness may have influenced the estimates of transcript family evolution, we constructed a linear model to regress transcript number, number of complete and single-copy BUSCOs and number of identified protein domains against the number of rapidly evolving HTFs in each species. We also performed an additional regression using these same metrics against the number of transcripts within rapidly evolving HTFs in each species.

2.6 | Gene Ontology enrichment

To understand the processes which are undergoing rapid evolution in these species, we utilized the R script Gene Ontology with Mann–Whitney *U* test (GO_MWU) (Wright et al., 2015), to perform a Fisher's exact test where transcripts in rapidly evolving HTFs were assigned a value of 1 while all other contigs were assigned a value of 0. The enrichments were run against species-specific backgrounds. This method was employed separately for the transcripts in expanded vs. contracted HTFs. For each species identical parameters were used for biological processes with the smallest category set to 50, largest to 0.1 and clusterCutHeight to 0.25.

2.7 | Expression divergence

To determine which HTFs demonstrate species-specific expression shifts, we mapped reads from each coral colony to its respective species transcriptome with the read mapping program SALMON version 1.5.2 (Patro et al., 2017) using a k-mer index size of 31 in each species. Read counts were then extracted and summarized to transcript family level with the R package TXimport version 1.16.1 (Soneson et al., 2015) using the "salmon" option. These expression profiles were then combined across species and filtered for low abundance expression with less than an average of 10 counts. The HTFs above this cutoff were rlog normalized by species in the R package DESeq2 version 1.28.1 (Love et al., 2014). Species-level expression shifts were quantified using the R package evemodel version 0.0.0.9005 using a β shared test (Gillard et al., 2021; Rohlfs & Nielsen, 2015). The rlog-normalized counts of each HTF as well our generated phylogenetic tree were used as input files. This program compares expression variance within and across species to generate β values, where low β values represent HTFs with high levels of between-species variance and high β values represent HTFs with high within-species variance. The β values were $-\log_{10}$-transformed to generate our expression divergence metric. All statistical analysis were carried out in the R programing language (R Core Team (2020)).

2.8 | Comparison to genome-based approach

As both *O. faveolata* and *M. cavernosa* have genome resources available, we performed the previously described analysis again, incorporating the genome resources to determine if our results were robust to the increased completeness which genomes provide. Genome-derived predicted gene models were obtained from the *O. faveolata* assembly on NCBI (assembly ofav_dov_v1) (Prada et al., 2016), while for *M. cavernosa* a draft genome assembly is available (Rippe et al., 2021), and we chose to use this genome assembly to generate a genome-guided transcriptome assembly for *M. cavernosa*. Reads from our five *M. cavernosa* colonies were mapped to the draft genome with the read mapping program TOPHAT2 (Kim et al., 2013) using default parameters. The merged bam files were then used in the TRINITY genome-guided approach to generate our *M. cavernosa* transcriptome. We performed all the computational protocols described (isoform filtering, TRANSDECODER, CD-HIT, ORTHOFINDER, CAFÉ, SALMON, EVEMODEL) above with the exception of the *in silico* symbiont filtration as the genome assemblies are already free of symbiont sequences.

3 | RESULTS

3.1 | HTF assignment

Assembly metrics including transcript number, N50 values, BUSCO scores and DOGMA results can be found in Table 1. Collectively we identified 47,170 HTFs, of which 4,904 were represented by at least

one transcript in all species, and 1,629 were single copy in all species. The frequency of the size class distribution was similar across all species (Figure 1a); most transcripts are in HTFs with a single transcript and as the number of transcripts in an HTF increases the frequency of that size class decreases. HTFs with a single transcript are the most common size class in each species followed by HTFs with two transcripts, while HTFs with more than five transcripts are uncommon. The phylogenetic tree produced from these HTFs recapitulates the known phylogenetic relationship between these species of coral (Figure 1b).

3.2 | HTF evolution

By evaluating the evolution of the number of homologues within a transcript family we identified a background rate of turnover of 0.000854 homologue gains/losses per transcript family per million years (Myr) (Figure 1b) and 236 HTFs which are evolving significantly faster than this background rate of turnover. Given that these data are based on transcriptome assemblies, the background rate of HTF turnover is probably an overestimate as transcripts which were not sampled by our approach would be interpreted as a change in the

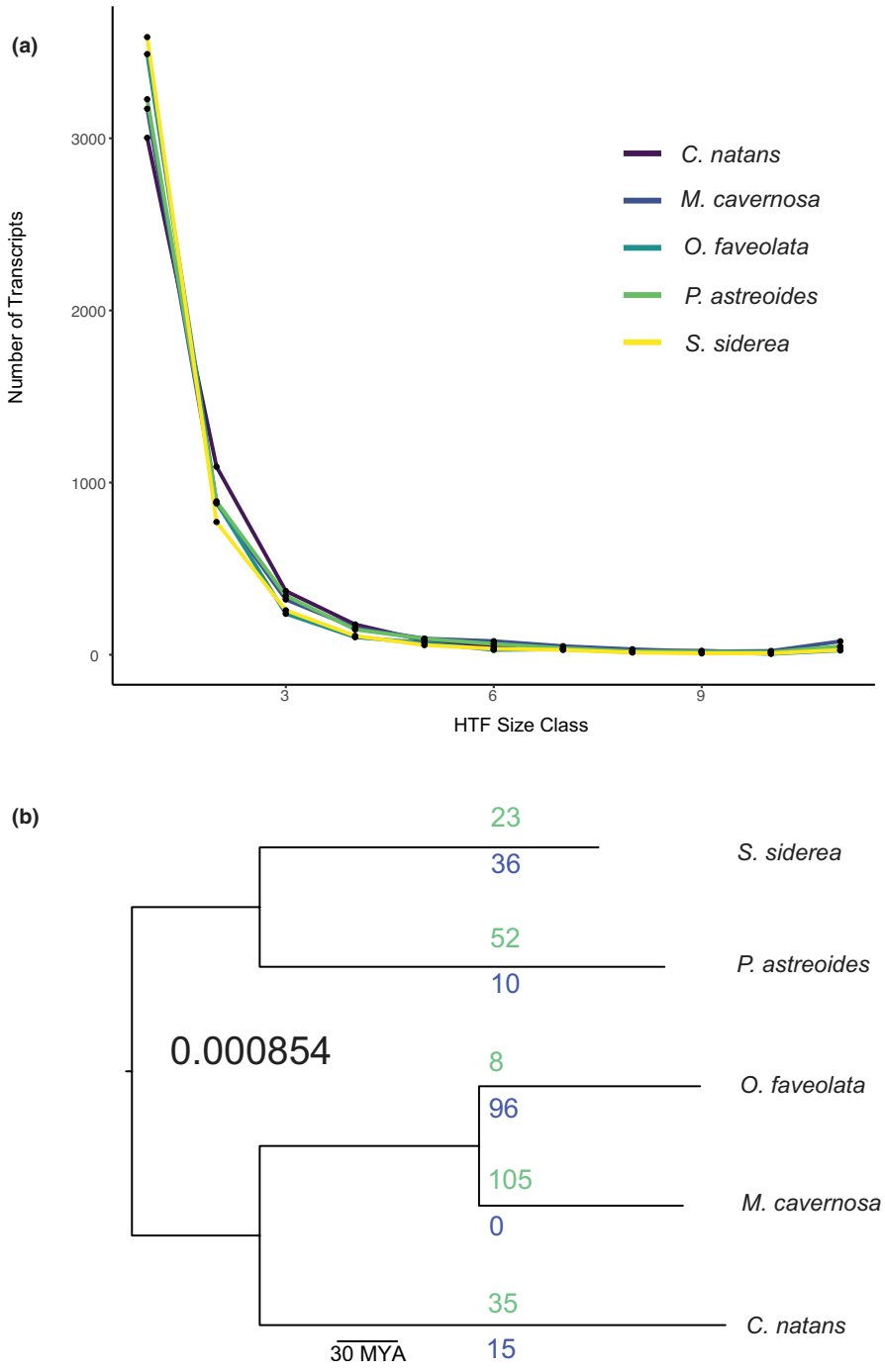

FIGURE 1 Quantifying HTF evolution. (a) Distribution of the number of transcripts contained in each transcript family size class separated by species. (b) Species tree generated from STAG for the five species of coral, with the number of expanded (green) and contracted (blue) transcript families as well as the background rate of transcript family turnover (8.54 e−4) superimposed on the tree. The number of expanded/contracted transcript families in a particular species is given on the branch leading to that species [Colour figure can be viewed at wileyonlinelibrary.com]

size of an HTF leading to a higher estimate of background turnover. The consequence of potentially overestimating the background rate of turnover is reduced power to detect rapidly evolving HTFs, which suggests that the 236 rapidly evolving HTFs we identified is likely to be an underestimate. The number of HTFs with adaptive variations in size differed by species: 104 in *Orbicella faveolata*, 105 in *Montastraea cavernosa*, 50 in *Colpophyllia natans*, 62 in *Porites astreoides* and 59 in *Siderastrea siderea* (Figure 1b). The number of rapidly evolving HTFs in a species was not significantly associated with standard metrics of transcriptome assembly completeness, including the percentage of complete single copy BUSCOs, number of transcripts in the assembly or the number of inferred protein domains in a linear regression ($p = .357$) (Figure S1). Additionally, the number of transcripts within these rapidly evolving HTFs was also not associated with these same metrics ($p = .379$) according to the linear regression (Figure S1). While neither of these relationships was significant, it should be noted that with only five species our power to detect such associations is low and therefore it is not possible to rule out assembly completeness as influencing our results. The number of transcripts within these rapidly evolving HTFs as well as the number of GO enrichments broken down by expansions vs. contractions of the number of homologues within a transcript family are given in Table 2. A subset of select GO term enrichments are visualized (Figure 2) based on terms commonly observed in coral gene expression studies including terms related to stress responses, metabolism and tissue formation. A full list of all enriched GO terms broken down by species can be found in Table S1.

3.3 | Expression divergence

To test if the rapidly evolving HTFs demonstrate species-specific expression shifts, we quantified gene expression patterns and tested for expression divergence across the 4,277 HTFs which were expressed across all species. The rapidly evolving HTFs had higher average expression divergence than the single-copy HTFs (0.830 vs. 0.447, $p = 2.2e{-}16$ two-sided *t*-test) (Figure 3). This elevated expression divergence among the rapidly evolving HTFs was also higher than a size-matched random down-sampling of genes (0.830 vs. 0.623, $p = 5.11e{-}06$ two-sided *t* test).

3.4 | Genome comparison

By performing the same analysis with the inclusion of the genomic data from *O. faveolata* and *M. cavernosa*, our results differed slightly, but our main conclusions were still supported. We identified more single-copy HTFs as well as more HTFs with at least one transcript present in all species with the genome-based assemblies, 2,009 and 6,324 respectively (Figure S2). This difference resulted in a slightly higher rate of background transcript family turnover (0.00088 transcript gains/losses per transcript family per Myr) and a greater number of rapidly evolving transcript families (354). The patterns of transcript family evolution were similar between the two analysis, as *O. faveolata* possessed the most rapidly contracting transcript families and *M. cavernosa* possessed the most rapidly expanding transcript families (Figure S2). Our observation of higher expression divergence within rapidly evolving transcript families compared to either single-copy transcript families (0.543 single copy vs. 0.718 rapidly evolving, $p = 7.687e{-}07$, *t* test) (Figure S2) or a downsampled number of transcript families was likewise supported with the inclusion of genomic data (0.599 downsampled vs. 0.718 rapidly evolving, $p = 7.77e{-}03$, *t* test). Overall, this additional analysis demonstrates that the primary findings of this study are robust even when different methods were used to generate the assemblies. The inclusion of genomic data from two coral species increased the number of identified HTFs that are rapidly evolving, indicating that similar analyses should continue to improve the resolution of rapidly evolving HTFs as genome assemblies become available for more species.

4 | DISCUSSION

Our investigation utilizes a framework to integrate paralogous transcripts into multispecies gene expression studies, and demonstrates that rapidly evolving HTFs affect numerous biological processes and are associated with divergent gene expression patterns in five species of Caribbean coral. These findings provide support that the evolution of the number of transcripts within an HTF, a proxy for gene family size, is associated with divergent expression patterns (Li et al., 2005; Taylor & Raes, 2004) and these processes are active in Caribbean corals. As these conclusions are based on our ability to estimate gene family size from transcriptome data, some caution in interpreting the data is warranted until genome assemblies become available for these species, allowing for a more thorough investigation of gene family expansion/contraction. However, these findings highlight both the evolution of HTFs and their resulting expression patterns as an understudied aspect of coral biology which has the potential to influence how these species responses to a changing climate.

4.1 | Feasibility

Our study focused on variation in the number of homologous transcripts within transcript families, which ideally should represent gene family size, and we employed extensive efforts to establish a one-to-one correspondence between transcripts and genes; however, some considerations need to be addressed. First, without genome assemblies available for all species it is not possible to confidently say that each unique transcript arises from a unique gene. However, using the genomes available for *Orbicella faveolata* and *Montastraea cavernosa* did not substantially change our main findings. Additionally, as the corals were only sampled under a single condition, there are probably genes which were not expressed and thus not sampled by our approach, leading to incomplete gene sampling, which has

TABLE 2 Transcript family evolution results

Species	Expanding HTFs	Transcripts in expanding HTFs	Biological Process enrichments expanded	Contracting HTFs	Transcripts in contracting HTFs	BP enrichments contracted
O. faveolata	8	265	16	95	239	33
M. cavernosa	110	1,550	229	0	0	0
C. natans	35	606	13	15	31	0
P. astreiodes	52	721	199	10	29	0
S. siderea	22	291	16	35	99	2

Abbreviation: HTF, homologous transcript family.

the potential to result in inflated estimates of gene family turnover. However, incomplete sampling of transcripts should only be problematic for our interpretations if it is biased among species, which we have no reason to suspect. Indeed, if genes are missed stochastically across species, it should only serve to inflate the background (i.e., neutral) rate of gene family turnover, making detection of rapidly (i.e., adaptively) expanding or contracting families more difficult. To this point, our calculated rate of gene family turnover approximates the observed rate among species of *Drosophila* (Han et al., 2013) and is lower than that of mammals (Matthew W. Hahn et al., 2007), which is interesting given our reduced ability to detect recent paralogues in transcriptomes. Overall, this demonstrates that by employing stringent filtering approaches and including an error term in the CAFÉ model, issues common to transcriptome assemblies can be reduced or accounted for to a level which allows investigation of the evolution of HTFs. Thus, by following the computational framework outlined here, other investigations can incorporate the evolution of HTFs in species currently lacking genome assemblies, as long as whole organism homogenates can be acquired.

4.2 | The benefit of incorporating paralogues

By incorporating paralogues into our analysis we were able to increase the number of HTFs used for multispecies comparisons by more than three-fold compared to only using only single-copy HTFs. This includes the additional 254 rapidly evolving HTFs, indicating an adaptive role of this variation (De Bie et al., 2006; Hahn, 2005). While it is not possible to ascertain from these data the selective benefit of this rapid evolution in particular HTFs, these transcript families were enriched for GO terms in all five species, consistent with a role in species-level adaptation. Population-level data coupled to fitness effects would be required to conclusively determine the adaptive significance of this variation and associated GO enrichments. Despite this limitation some of the enriched terms do qualitatively correspond to host biology. For example, *O. faveolata*, the species which can attain the largest colony diameter (Madin et al., 2016), is enriched for "regulation of anatomical structure size" among the rapidly contracting HTFs, which may represent a loss of negative regulators of growth and could potentially underly the large colony size this species can attain. Similarly, *Siderastrea siderea* is enriched for "response to pH" among the rapidly expanding HTFs, which may be linked to the exceptional ability of this species to handle highly acidified water (Castillo et al., 2014). While such connections between species biology and the rapid evolution of HTFs are speculative, these examples underline how variation of HTFs could influence each species biology. These results also highlight potential new avenues of study such as the rapid evolution of HTFs involved in carbohydrate and lipid metabolism. These changes in metabolic processes may have important consequences for a coral's ability to efficiently assimilate symbiont-derived carbohydrates or potentially store nutrients, both of which have been associated with corals' ability to withstand periods of temperature stress (Rädecker et al., 2021; Roach et al., 2021). Thus, the observed differences in the number of homologues within these transcript families may play an important role in each species biology.

4.3 | Rapidly evolving HTFs have increased expression divergence

Rapidly evolving HTFs have an overall higher level of expression divergence, demonstrating an association between the processes of gene duplication and gene expression. This finding is in line with previous work showing that gene copy number influences transcriptional output, although this may be modified by dosage compensation mechanisms (Jiang & Assis, 2019). While our findings are based on transcripts, rather than genome-derived gene models, the association between the rapid change in HTF size and expression divergence provides support for the idea that gene expression is often modified among paralogues in response to gene duplication or loss (Des Marais & Rausher, 2008; Stearns, 2010). Note that gene expression divergence can arise from other processes which are not dependent on paralogues. However, the substantially higher levels of expression divergence within the rapidly evolving HTFs suggest that variation in the number of paralogues may be an important contributor to the development of divergent expression patterns. To our knowledge this is the first investigation to link the rate of transcript family evolution with divergent expression patterns.

Importantly, the association between the variation in the number of HTFs and expression divergence may be an important mechanism of adaptation. If the new paralogues arise from a gene

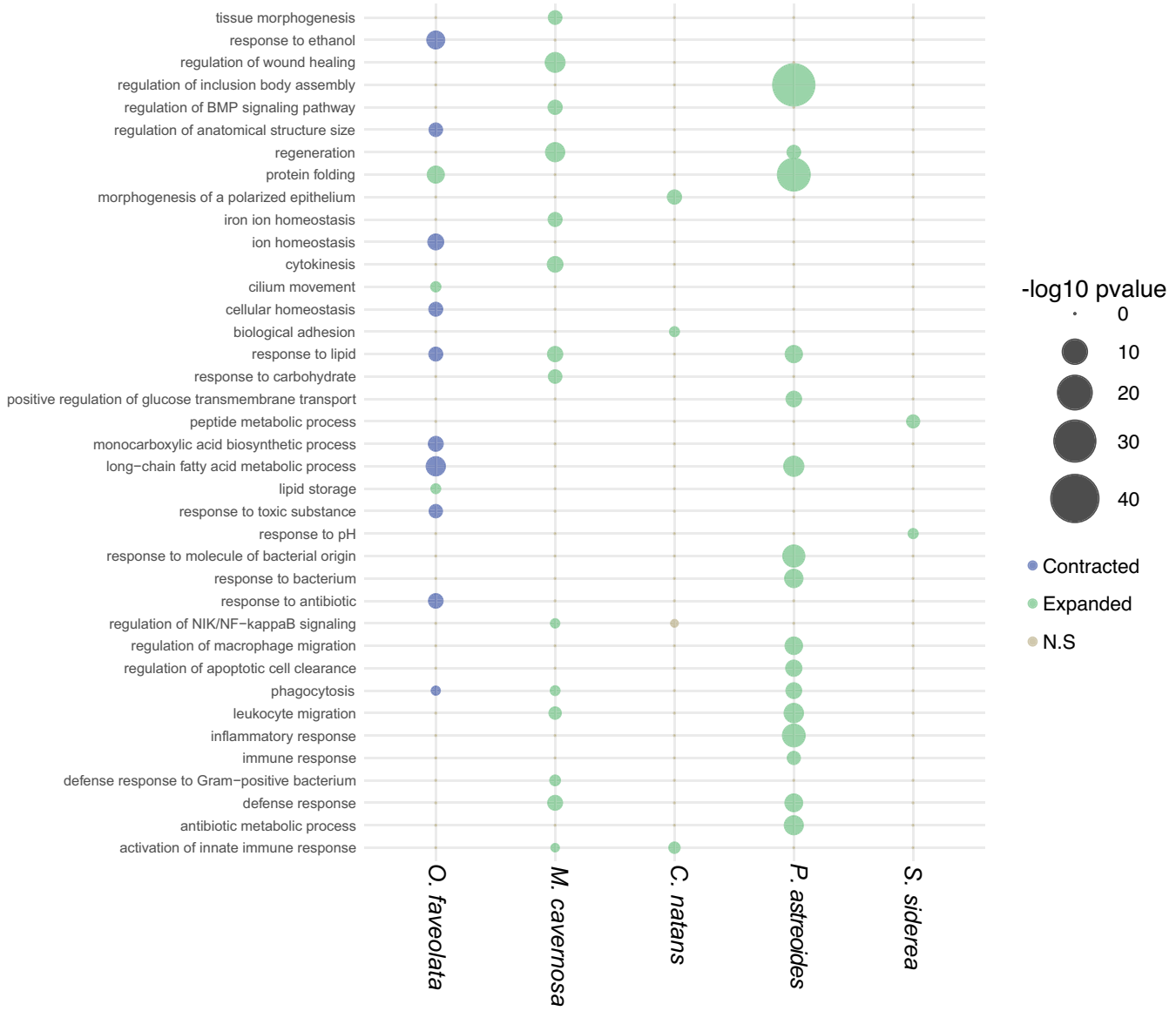

FIGURE 2 Selected subset of GO enrichments among the expanded and contracted transcript families. Bubble plot visualizing selected GO term enrichments in each species. Size of the circle corresponds to the significance ($-\log_{10}$-transformed *p*-value) after false discovery rate correction, and colour denotes if the term enrichment is in expanded (green) or contracted (blue) transcript families, or grey if the term is nonsignificant after false discovery rate correction. GO enrichments were selected for visualization based on their potential involvement in either stress responses, metabolism or tissue formation. A table containing all of the GO enrichments can be found in Table S1 [Colour figure can be viewed at wileyonlinelibrary.com]

duplication event, this can relax pleiotropic constraints, promoting the processes of neofunctionalization and subfunctionalization (J. Zhang, 2003). Gene neofunctionalization in newly arisen paralogues has been demonstrated to promote evolutionary innovation and adaptation in several systems (Deng et al., 2010; Zimmer et al., 2018). Gene subfunctionalization can also function as an adaptive mechanism (Abascal et al., 2013; Spady et al., 2006) though there are fewer examples of this processes. While our data do not explicitly demonstrate either of these processes *per se*, they do highlight that rapidly evolving HTFs whose paralogues are candidates for neofunctionalization and subfunctionalization have acquired new expression patterns consistent with changing function. These findings thus highlight that as more coral genome assemblies are produced and sequencing technology capable of detecting gene copy number variation becomes more accessible, investigating the processes of gene neofunctionalization and subfunctionalization may yield important insights.

4.4 | Rapid evolution of immunity

We find that HTFs which have immune-related annotations are common among the rapidly evolving transcript families leading to immune-related enrichments in several species (Figure 4). This reflects findings from other systems which have characterized the rapid evolution and subsequent change in copy number of immune

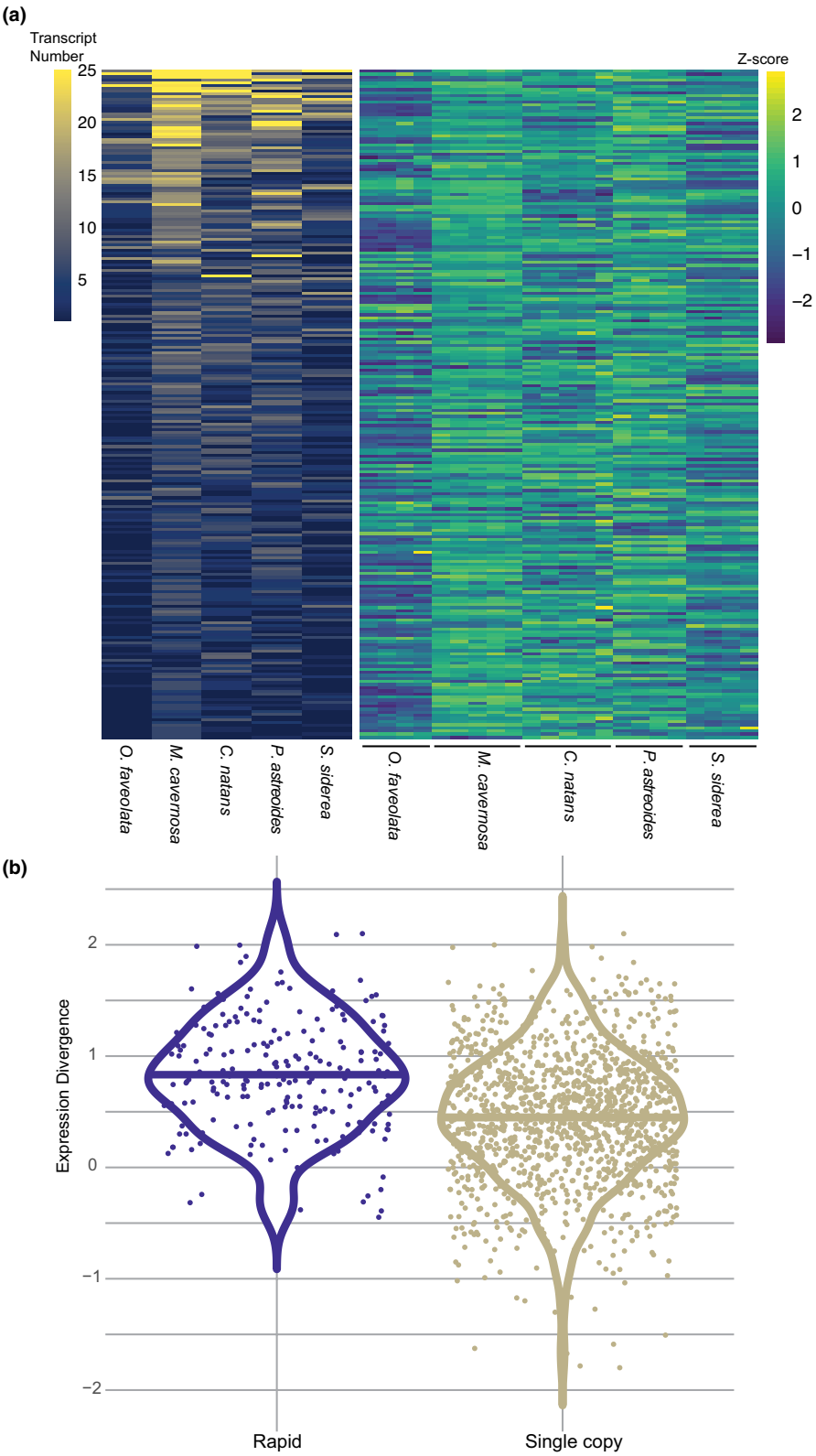

FIGURE 3 Rapidly evolving transcript families have increased expression divergence. (a) The number of transcripts contained within the rapidly evolving transcript families as well as the expression of those transcript families. The left heatmap visualizes transcript number where transcript families with more than 25 transcripts were reduced to 25 for visualization purposes. The right heatmap shows the corresponding z-score normalized expression value of each transcript family in the left heatmap. (b) Violin plot showing the expression divergence of the rapidly evolving transcript families compared to the single-copy transcript families. The thick line shows the mean of each group. The difference in expression divergence is significant between the two groups (0.830 vs. 0.447, $p = 2.2e{-}16$ two-sided *t* test) [Colour figure can be viewed at wileyonlinelibrary.com]

genes (Evans et al., 2006; Sackton et al., 2007, 2017; Waterhouse et al., 2007). This common pattern is thought to be due to the propensity of immune genes to engage in evolutionary arms races through their ability to mediate interactions with other organisms (Lazzaro & Clark, 2012). This process is probably important in reef-building corals that live in a microbe-rich environment (van Oppen & Blackall, 2019). The immune system of corals functions to both regulate commensal microbes (beneficial bacteria and the algal Symbiodiniaceae) as well as defend against pathogenic ones (Kvennefors et al., 2010; Mansfield & Gilmore, 2019; Mydlarz et al., 2006; Wu et al., 2019). Thus, the rapid evolution of immune transcript families indicates that these interactions may have played an important role in shaping

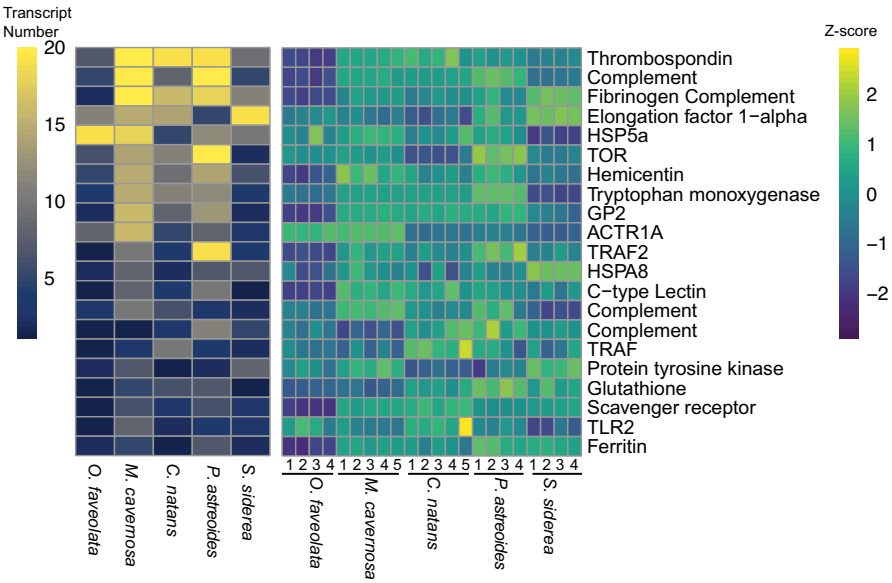

FIGURE 4 Rapid evolution of immune genes. The number of transcripts contained within the immune-related rapidly evolving transcript families as well as the expression of those transcript families. Immune transcript families were selected based on the transcripts within that family possessing either the GO annotation "Immune system process, GO: 0002376" or "Innate immune response GO: 0002226." The left heatmap visualizes transcript number while the right heatmap shows the corresponding z-score normalized expression value of each transcript family. The transcript family names were assigned based on the lowest seed orthologue evalue from EGGNOG-MAPPER of all transcripts contained within the given transcript family [Colour figure can be viewed at wileyonlinelibrary.com]

the evolutionary history of these coral species. Other investigations have found similar trends, including phylosymbiosis between coral species and bacterial communities on the Great Barrier Reef (Pollock et al., 2018), as well as highly stable host–microbe associations in the Red Sea (Ziegler et al., 2019). Overall, this indicates that interactions with microbes may broadly play a role in shaping the evolution of corals.

Of particular interest is the interaction formed with Symbiodiniaceae as this relationship requires the modulation of host immune processes (Mansfield & Gilmore, 2019; Matthews et al., 2017; Neubauer et al., 2017) in order to receive a supply of translocated carbohydrates from their symbionts. Different species of coral are known to preferentially associate with one or a few species of Symbiodiniaceae (Thornhill et al., 2006), although the mechanism by which hosts distinguish between these different symbionts is unknown. The rapid evolution of immune transcripts by the host may be one mechanism to facilitate this relationship through neofunctionalization of immune genes, but this hypothesis would need to be functionally confirmed. One interesting trend in our data is that *O. faveolata*, which demonstrates flexibility in its symbiont associations (Edmunds et al., 2014), has reduced numbers of homologues in rapidly evolving transcript families which are involved in immunity while *Porites astreoides*, which transmits its symbionts vertically and demonstrates tight control of its symbiotic association (Chornesky & Peters, 1987; Kenkel et al., 2013), has an expanded number of homologues in rapidly evolving immune transcript families. This pattern raises the possibility that symbiont fidelity could be related to the size of immune transcript families as species with larger immune transcript repertoires may have an improved ability to select for preferred symbiotic partners. The cellular mechanisms that corals use to detect different strains of symbionts is still an ongoing area of research (Jacobovitz et al., 2021), though immune specificity is thought to play a critical role. In particular, scavenger receptors and thrombospondin (Neubauer et al., 2016; Neubauer et al., 2017) have been demonstrated to influence symbiont recognition and density, and we observe rapid evolution of transcripts matching these annotations. Thus, the changes observed between species within these and other immune-related rapidly evolving HTFs could potentially be an evolved mechanism to distinguish between different strains of Symbiodiniaceae.

In addition to modulating the relationship with beneficial microorganisms, corals must also defend themselves against disease-causing microbes. This is particularly pressing in the Caribbean as disease outbreaks affecting reef-building corals have reshaped communities (Aronson & Precht, 2001; Gardner, 2003; Miller et al., 2009; Mydlarz et al., 2006) and continue to be among the most pressing selective forces these organisms face (Vega Thurber et al., 2020). In response to disease, corals utilize a complex innate immune system involving pattern recognition receptors including Toll-like receptors (TLRs) (Poole & Weis, 2014; Williams et al., 2018) and C-type lectins (Emery et al., 2021; Kvennefors et al., 2008) to sense threats and initiate immune responses including the complement cascade (Poole et al., 2016). Interestingly, we observe that *M. cavernosa*, *Colpophyllia natans* and *P. astreoides* have an expanded number of homologues within rapidly evolving transcript families matching these annotations, while *O. faveolata* has a contracted number of homologues within these transcript families. This variation may influence disease-susceptibility as *O. faveolata* is highly susceptible to several coral

diseases whereas *M. cavernosa* and *P. astreoides* are considered more resilient to disease (Aeby et al., 2019; Smith et al., 2013; Williams et al., 2020). In support of this, species-level variation in immune activity is widespread in Caribbean corals (Palmer et al., 2011; Pinzón et al., 2014; Rosales et al., 2020), which may be the result of the variation of immune transcript repertoires we observe. However, *C. natans* is also considered to be susceptible to disease (Aeby et al., 2019; MacKnight et al., 2021; Sutherland et al., 2004) and demonstrates expansions in immune-related transcript families. Thus, the relationship between an expanded immune transcript repertoire may not always lead to resistance to a particular disease *in situ*.

Together, these data indicate that the immune system is a major target of evolution in Caribbean coral. As coral diseases are currently among the strongest selective forces acting on Caribbean reefs (Vega Thurber et al., 2020), understanding the differing immune repertoires each species possesses and how this relates to disease resilience may be an important determinant of which corals will persist on the reef. Indeed, in the Caribbean, coral species that have historically dominated stable forereef environments such as *O. faveolata* have been declining in abundance (Gardner, 2003) while corals such as *P. astreoides* that have traditionally been found in the variable environment of the reef edge are rising in abundance (Green et al., 2008; McWilliam et al., 2020) with disease playing a major role in these shifts.

5 | CONCLUSION

We found that by expanding our analysis of multispecies gene expression through the identification of HTFs, we were able to identify adaptive variation in homologue number within transcript families, which is associated with divergent expression patterns. However, given the inherent limitations in using transcriptome assemblies to assess gene family evolution, these conclusions should be considered preliminary and that follow-up work based on genome assemblies will probably yield additional insights not captured by this study. Despite this, we identified numerous biological processes influenced by this variation, which notably include several families of immune genes previously linked to both symbiont recognition and pathogen defence. These findings highlight the role that interactions with microbes have played in the evolution of these coral species and how this may influence species' ability to persist as coral diseases become an increasingly pressing threat in the Caribbean. Therefore, further work towards linking variation within HTFs that arise through changes in gene copy number to adaptive capacity may help predict which coral species will be able to adapt and which are likely to face extirpation.

ACKNOWLEDGEMENTS

We thank Marilyn Brandt and the support dive team at the Center for Marine and Environmental Studies, University of the Virgin Islands, for their help with sample collecting as well as the high-performance computing centre at the University of Texas at Arlington.

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

AUTHOR CONTRIBUTIONS

B.D., M.E. and J.D. designed and preformed research. L.M. and M.E. assisted with writing. B.D., K.B. and N.M. performed laboratory work and constructed assemblies. M.B. assisted with sample collection.

DATA AVAILABILITY AND BENEFIT-SHARING

Data availability: All raw reads and appropriate sample metadata from this project are available on the NCBI short read archive under Bioproject no.: PRJNA723585. All code and data needed to recreate the major findings from this paper are publicly available through both the data dryad repository: https://doi.org/10.5061/dryad.tht76hf1d as well as on github: https://github.com/braddimo/Gene-expansion. Benefit-Sharing: Benefits generated: there are no benefits outlined in the Nagoya protocol associated with this study to report.

ORCID

Bradford Dimos https://orcid.org/0000-0003-0899-6378
Laura Mydlarz https://orcid.org/0000-0002-8371-0766

REFERENCES

Abascal, F., Corpet, A., Gurard-Levin, Z. A., Juan, D., Ochsenbein, F., Rico, D., Valencia, A., & Almouzni, G. (2013). Subfunctionalization via adaptive evolution influenced by genomic context: the case of histone chaperones ASF1a and ASF1b. *Molecular Biology and Evolution*, *30*(8), 1853–1866. https://doi.org/10.1093/molbev/mst086

Aeby, G. S., Ushijima, B., Campbell, J. E., Jones, S., Williams, G. J., Meyer, J. L., Häse, C., & Paul, V. J. (2019). Pathogenesis of a tissue loss disease affecting multiple species of corals along the Florida reef tract. *Frontiers in Marine Science*, *6*, 678. https://doi.org/10.3389/fmars.2019.00678

Albalat, R., & Cañestro, C. (2016). Evolution by gene loss. *Nature Reviews Genetics*, *17*(7), 379–391. https://doi.org/10.1038/nrg.2016.39

Aranda, M., Li, Y., Liew, Y. J., Baumgarten, S., Simakov, O., Wilson, M. C., Piel, J., Ashoor, H., Bougouffa, S., Bajic, V. B., Ryu, T., Ravasi, T., Bayer, T., Micklem, G., Kim, H., Bhak, J., LaJeunesse, T. C., & Voolstra, C. R. (2016). Genomes of coral dinoflagellate symbionts highlight evolutionary adaptations conducive to a symbiotic lifestyle. *Scientific Reports*, *6*(1), 39734. https://doi.org/10.1038/srep39734

Aronson, R. B., & Precht, W. F. (2001). White-band disease and the changing face of Caribbean coral reefs. In J. W. Porter (Ed.), *The ecology and etiology of newly emerging marine diseases* (pp. 25–38). Springer Netherlands. https://doi.org/10.1007/978-94-017-3284-0_2

Avila-Magaña, V., Kamel, B., DeSalvo, M., Gómez-Campo, K., Enríquez, S., Kitano, H., Rohlfs, R. V., Iglesias-Prieto, R., & Medina, M. (2021). Elucidating gene expression adaptation of phylogenetically divergent coral holobionts under heat stress. *Nature Communications*, *12*(1), 5731. https://doi.org/10.1038/s41467-021-25950-4

Baumann, J. H., Davies, S. W., Aichelman, H. E., & Castillo, K. D. (2018). Coral symbiodinium community composition across the Belize Mesoamerican barrier reef system is influenced by host species and thermal variability. *Microbial Ecology*, *75*(4), 903–915. https://doi.org/10.1007/s00248-017-1096-6

Bay, R. A., & Palumbi, S. R. (2015). Rapid acclimation ability mediated by transcriptome changes in reef-building corals. *Genome Biology and Evolution*, *7*(6), 1602–1612. https://doi.org/10.1093/gbe/evv085

Bay, R. A., & Palumbi, S. R. (2017). Transcriptome predictors of coral survival and growth in a highly variable environment. *Ecology and Evolution*, *7*(13), 4794–4803. https://doi.org/10.1002/ece3.2685

Bernal, M. A., Schunter, C., Lehmann, R., Lightfoot, D. J., Allan, B. J. M., Veilleux, H. D., Rummer, J. L., Munday, P. L., & Ravasi, T. (2020). Species-specific molecular responses of wild coral reef fishes during a marine heatwave. *Science Advances*, *6*(12), eaay3423. https://doi.org/10.1126/sciadv.aay3423

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bove, C. B., Ries, J. B., Davies, S. W., Westfield, I. T., Umbanhowar, J., & Castillo, K. D. (2019). Common Caribbean corals exhibit highly variable responses to future acidification and warming. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1900), 20182840. https://doi.org/10.1098/rspb.2018.2840

Castillo, K. D., Ries, J. B., Bruno, J. F., & Westfield, I. T. (2014). The reef-building coral *Siderastrea siderea* exhibits parabolic responses to ocean acidification and warming. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1797), 20141856. https://doi.org/10.1098/rspb.2014.1856

Chain, F. J. J., Feulner, P. G. D., Panchal, M., Eizaguirre, C., Samonte, I. E., Kalbe, M., Lenz, T. L., Stoll, M., Bornberg-Bauer, E., Milinski, M., & Reusch, T. B. H. (2014). Extensive copy-number variation of young genes across stickleback populations. *PLoS Genetics*, *10*(12), e1004830. https://doi.org/10.1371/journal.pgen.1004830

Chen, C.-L., Zhang, L., Li, J.-L., Mao, X.-X., Zhang, L.-S., Hu, Q.-J., Liu, J.-Q., & Xi, Z.-X. (2021). Phylotranscriptomics reveals extensive gene duplication in the subtribe Gentianinae (Gentianaceae). *Journal of Systematics and Evolution*, *59*(6), 1198–1208. https://doi.org/10.1111/jse.12651

Chornesky, E. A., & Peters, E. C. (1987). Sexual reproduction and colony growth in the scleractinian coral *Porites astreoides*. *The Biological Bulletin*, *172*(2), 161–177. https://doi.org/10.2307/1541790

Davies, S. W., Marchetti, A., Ries, J. B., & Castillo, K. D. (2016). Thermal and pCO2 stress elicit divergent transcriptomic responses in a resilient coral. *Frontiers in Marine Science*, *3*, https://doi.org/10.3389/fmars.2016.00112

Davies, S. W., Treml, E. A., Kenkel, C. D., & Matz, M. V. (2015). Exploring the role of Micronesian islands in the maintenance of coral genetic diversity in the Pacific Ocean. *Molecular Ecology*, *24*(1), 70–82. https://doi.org/10.1111/mec.13005

De Bie, T., Cristianini, N., Demuth, J. P., & Hahn, M. W. (2006). CAFE: A computational tool for the study of gene family evolution. *Bioinformatics*, *22*(10), 1269–1271. https://doi.org/10.1093/bioinformatics/btl097

De Smet, R., Sabaghian, E., Li, Z., Saeys, Y., & Van de Peer, Y. (2017). Coordinated functional divergence of genes after genome duplication in *Arabidopsis thaliana*. *The Plant Cell*, *29*(11), 2786–2800. https://doi.org/10.1105/tpc.17.00531

Demuth, J. P., & Hahn, M. W. (2009). The life and death of gene families. *BioEssays*, *31*(1), 29–39. https://doi.org/10.1002/bies.080085

Deng, C., Cheng, C.-H.-C., Ye, H., He, X., & Chen, L. (2010). Evolution of an antifreeze protein by neofunctionalization under escape from adaptive conflict. *Proceedings of the National Academy of Sciences*, *107*(50), 21593–21598. https://doi.org/10.1073/pnas.1007883107

Des Marais, D. L., & Rausher, M. D. (2008). Escape from adaptive conflict after duplication in an anthocyanin pathway gene. *Nature*, *454*(7205), 762–765. https://doi.org/10.1038/nature07092

Dimos, B. A., Butler, C. C., Ricci, C. A., MacKnight, N. J., & Mydlarz, L. D. (2019). Responding to threats both foreign and domestic: NOD-like receptors in corals. *Integrative and Comparative Biology*, *59*(4), 819–829. https://doi.org/10.1093/icb/icz111

Dixon, G., Abbott, E., & Matz, M. (2020). Meta-analysis of the coral environmental stress response: *Acropora* corals show opposing responses depending on stress intensity. *Molecular Ecology*, *29*(15), 2855–2870. https://doi.org/10.1111/mec.15535

Duarte, J. M., Cui, L., Wall, P. K., Zhang, Q., Zhang, X., Leebens-Mack, J., Ma, H., Altman, N., & dePamphilis, C. W. (2006). Expression pattern shifts following duplication indicative of subfunctionalization and neofunctionalization in regulatory genes of arabidopsis. *Molecular Biology and Evolution*, *23*(2), 469–478. https://doi.org/10.1093/molbev/msj051

Edmunds, P., Pochon, X., Levitan, D., Yost, D., Belcaid, M., Putnam, H., & Gates, R. (2014). Long-term changes in Symbiodinium communities in *Orbicella annularis* in St. John, US Virgin Islands. *Marine Ecology Progress Series*, *506*, 129–144. https://doi.org/10.3354/meps10808

Emery, M. A., Dimos, B. A., & Mydlarz, L. D. (2021). Cnidarian pattern recognition receptor repertoires reflect both phylogeny and life history traits. *Frontiers in Immunology*, *12*, 689463. https://doi.org/10.3389/fimmu.2021.689463

Emms, D. M., & Kelly, S. (2015). OrthoFinder: Solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology*, *16*(1), 157. https://doi.org/10.1186/s13059-015-0721-2

Emms, D. M., & Kelly, S. (2018). STAG: Species tree inference from all genes [Preprint]. *Evolutionary Biology*. https://doi.org/10.1101/267914

Emms, D. M., & Kelly, S. (2019). OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biology*, *20*(1), 238. https://doi.org/10.1186/s13059-019-1832-y

Evans, J. D., Aronstein, K., Chen, Y. P., Hetru, C., Imler, J.-L., Jiang, H., Kanost, M., Thompson, G. J., Zou, Z., & Hultmark, D. (2006). Immune pathways and defence mechanisms in honey bees *Apis mellifera*. *Insect Molecular Biology*, *15*(5), 645–656. https://doi.org/10.1111/j.1365-2583.2006.00682.x

Foissac, S., Djebali, S., Munyard, K., Vialaneix, N., Rau, A., Muret, K., Esquerré, D., Zytnicki, M., Derrien, T., Bardou, P., Blanc, F., Cabau, C., Crisci, E., Dhorne-Pollet, S., Drouet, F., Faraut, T., Gonzalez, I., Goubil, A., Lacroix-Lamandé, S., … Giuffra, E. (2019). Multi-species annotation of transcriptome and chromatin structure in domesticated animals. *BMC Biology*, *17*(1), 108. https://doi.org/10.1186/s12915-019-0726-5

Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y. L., & Postlethwait, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. *Genetics*, *151*(4), 1531–1545. https://doi.org/10.1093/genetics/151.4.1531

Fuess, L. E., Pinzón C, J. H., Weil, E., Grinshpon, R. D., & Mydlarz, L. D. (2017). Life or death: Disease-tolerant coral species activate autophagy following immune challenge. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1856), 20170771. https://doi.org/10.1098/rspb.2017.0771

Gardner, T. A., Côté, I. M., Gill, J. A., Grant, A., & Watkinson, A. R. (2003). Long-term region-wide declines in Caribbean corals. *Science*, *301*(5635), 958–960. https://doi.org/10.1126/science.1086050

Gillard, G. B., Grønvold, L., Røsæg, L. L., Holen, M. M., Monsen, Ø., Koop, B. F., Rondeau, E. B., Gundappa, M. K., Mendoza, J., Macqueen, D. J., Rohlfs, R. V., Sandve, S. R., & Hvidsten, T. R. (2021). Comparative regulomics supports pervasive selection on gene dosage following whole genome duplication. *Genome Biology*, *22*(1), 103. https://doi.org/10.1186/s13059-021-02323-0

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644–652. https://doi.org/10.1038/nbt.1883

Green, D., Edmunds, P., & Carpenter, R. (2008). Increasing relative abundance of Porites astreoides on Caribbean reefs mediated by an overall decline in coral cover. *Marine Ecology Progress Series*, *359*, 1–10. https://doi.org/10.3354/meps07454

Greenway, R., Barts, N., Henpita, C., Brown, A. P., Arias Rodriguez, L., Rodríguez Peña, C. M., Arndt, S., Lau, G. Y., Murphy, M. P., Wu, L., Lin, D., Tobler, M., Kelley, J. L., & Shaw, J. H. (2020). Convergent evolution of conserved mitochondrial pathways underlies repeated adaptation to extreme environments. *Proceedings of the National Academy of Sciences*, *117*(28), 16424–16430. https://doi.org/10.1073/pnas.2004223117

Guschanski, K., Warnefors, M., & Kaessmann, H. (2017). The evolution of duplicate gene expression in mammalian organs. *Genome Research*, *27*(9), 1461–1474. https://doi.org/10.1101/gr.215566.116

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B. O., Lieber, M., MacManes, M. D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C. N., … Regev, A. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, *8*(8), 1494–1512. https://doi.org/10.1038/nprot.2013.084

Hahn, M. W., De Bie, T., Stajich, J. E., Nguyen, C., & Cristianini, N. (2005). Estimating the tempo and mode of gene family evolution from comparative genomic data. *Genome Research*, *15*(8), 1153–1160. https://doi.org/10.1101/gr.3567505

Hahn, M. W., Demuth, J. P., & Han, S.-G. (2007). Accelerated rate of gene gain and loss in primates. *Genetics*, *177*(3), 1941–1949. https://doi.org/10.1534/genetics.107.080077

Hamada, M., Shoguchi, E., Shinzato, C., Kawashima, T., Miller, D. J., & Satoh, N. (2013). The complex NOD-like receptor repertoire of the coral *Acropora digitifera* includes novel domain combinations. *Molecular Biology and Evolution*, *30*(1), 167–176. https://doi.org/10.1093/molbev/mss213

Han, M. V., Thomas, G. W. C., Lugo-Martinez, J., & Hahn, M. W. (2013). Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. *Molecular Biology and Evolution*, *30*(8), 1987–1997. https://doi.org/10.1093/molbev/mst100

Hao, Y., Xiong, Y., Cheng, Y., Song, G., Jia, C., Qu, Y., & Lei, F. (2019). Comparative transcriptomics of 3 high-altitude passerine birds and their low-altitude relatives. *Proceedings of the National Academy of Sciences*, *116*(24), 11851–11856. https://doi.org/10.1073/pnas.1819657116

Huang, K. M., & Chain, F. J. J. (2021). Copy number variations and young duplicate genes have high methylation levels in sticklebacks. *Evolution*, *75*(3), 706–718. https://doi.org/10.1111/evo.14184

Huang, Y., Niu, B., Gao, Y., Fu, L., & Li, W. (2010). CD-HIT suite: A web server for clustering and comparing biological sequences. *Bioinformatics*, *26*(5), 680–682. https://doi.org/10.1093/bioinformatics/btq003

Huerlimann, R., Maes, G. E., Maxwell, M. J., Mobli, M., Launikonis, B. S., Jerry, D. R., & Wade, N. M. (2020). Multispecies transcriptomics reveals evolutionary diversity in the mechanisms regulating shrimp tail muscle excitation-contraction coupling. *Gene*, *752*, 144765. https://doi.org/10.1016/j.gene.2020.144765

Huerta-Cepas, J., Forslund, K., Coelho, L. P., Szklarczyk, D., Jensen, L. J., von Mering, C., & Bork, P. (2017). Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Molecular Biology and Evolution*, *34*(8), 2115–2122. https://doi.org/10.1093/molbev/msx148

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S. K., Cook, H., Mende, D. R., Letunic, I., Rattei, T., Jensen, L. J., von Mering, C., & Bork, P. (2019). eggNOG 5.0: A hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Research*, *47*(D1), D309–D314. https://doi.org/10.1093/nar/gky1085

Jacobovitz, M. R., Rupp, S., Voss, P. A., Maegele, I., Gornik, S. G., & Guse, A. (2021). Dinoflagellate symbionts escape vomocytosis by host cell immune suppression. *Nature Microbiology*, *6*(6), 769–782. https://doi.org/10.1038/s41564-021-00897-w

Jiang, X., & Assis, R. (2019). Rapid functional divergence after small-scale gene duplication in grasses. *BMC Evolutionary Biology*, *19*(1), 97. https://doi.org/10.1186/s12862-019-1415-2

Kelly, M. (2019). Adaptation to climate change through genetic accommodation and assimilation of plastic phenotypes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*(1768), 20180176. https://doi.org/10.1098/rstb.2018.0176

Kenkel, C. D., Goodbody-Gringley, G., Caillaud, D., Davies, S. W., Bartels, E., & Matz, M. V. (2013). Evidence for a host role in thermotolerance divergence between populations of the mustard hill coral (*Porites astreoides*) from different reef environments. *Molecular Ecology*, *22*(16), 4335–4348. https://doi.org/10.1111/mec.12391

Kenkel, C. D., & Matz, M. V. (2017). Gene expression plasticity as a mechanism of coral adaptation to a variable environment. *Nature Ecology & Evolution*, *1*(1), 0014. https://doi.org/10.1038/s41559-016-0014

Khan, Z., Ford, M. J., Cusanovich, D. A., Mitrano, A., Pritchard, J. K., & Gilad, Y. (2013). Primate transcript and protein expression levels evolve under compensatory selection pressures. *Science*, *342*(6162), 1100–1104. https://doi.org/10.1126/science.1242379

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., & Salzberg, S. L. (2013). TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology*, *14*(4), R36. https://doi.org/10.1186/gb-2013-14-4-r36

Kirk, N. L., Howells, E. J., Abrego, D., Burt, J. A., & Meyer, E. (2018). Genomic and transcriptomic signals of thermal tolerance in heat-tolerant corals (*Platygyra daedalea*) of the Arabian/Persian Gulf. *Molecular Ecology*, *27*(24), 5180–5194. https://doi.org/10.1111/mec.14934

Kondrashov, F. A. (2012). Gene duplication as a mechanism of genomic adaptation to a changing environment. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1749), 5048–5057. https://doi.org/10.1098/rspb.2012.1108

Kvennefors, E. C. E., Leggat, W., Hoegh-Guldberg, O., Degnan, B. M., & Barnes, A. C. (2008). An ancient and variable mannose-binding lectin from the coral *Acropora millepora* binds both pathogens and symbionts. *Developmental & Comparative Immunology*, *32*(12), 1582–1592. https://doi.org/10.1016/j.dci.2008.05.010

Kvennefors, E. C. E., Leggat, W., Kerr, C. C., Ainsworth, T. D., Hoegh-Guldberg, O., & Barnes, A. C. (2010). Analysis of evolutionarily conserved innate immune components in coral links immunity and symbiosis. *Developmental & Comparative Immunology*, *34*(11), 1219–1229. https://doi.org/10.1016/j.dci.2010.06.016

Lazzaro, B. P., & Clark, A. G. (2012). Rapid evolution of innate immune response genes. In R. S. Singh, J. Xu, & R. J. Kulathinal (Eds.), *Rapidly evolving genes and genetic systems* (pp. 203–210). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199642274.003.0020

Levy, S., Elek, A., Grau-Bové, X., Menéndez-Bravo, S., Iglesias, M., Tanay, A., Mass, T., & Sebé-Pedrós, A. (2021). A stony coral cell atlas illuminates the molecular and cellular basis of coral symbiosis, calcification, and immunity. *Cell*, *184*(11), 2973–2987.e18. https://doi.org/10.1016/j.cell.2021.04.005

Li, W.-H., Yang, J., & Gu, X. (2005). Expression divergence between duplicate genes. *Trends in Genetics*, *21*(11), 602–607. https://doi.org/10.1016/j.tig.2005.08.006

Liu, H., Stephens, T. G., González-Pech, R. A., Beltran, V. H., Lapeyre, B., Bongaerts, P., Cooke, I., Aranda, M., Bourne, D. G., Forêt, S., Miller, D. J., van Oppen, M. J. H., Voolstra, C. R., Ragan, M. A., & Chan, C. X. (2018). Symbiodinium genomes reveal adaptive evolution of functions related to coral-dinoflagellate symbiosis. *Communications Biology*, *1*(1), 95. https://doi.org/10.1038/s42003-018-0098-3

Logan, M. L., & Cox, C. L. (2020). Genetic constraints, transcriptome plasticity, and the evolutionary response to climate change. *Frontiers in Genetics*, *11*, 538226. https://doi.org/10.3389/fgene.2020.538226

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12). https://doi.org/10.1186/s13059-014-0550-8

Lynch, M., & Conery, J. S. (2000). The evolutionary fate and consequences of duplicate genes. *Science*, *290*(5494), 1151–1155. https://doi.org/10.1126/science.290.5494.1151

Machado, L. R., & Ottolini, B. (2015). An evolutionary history of defensins: A role for copy number variation in maximizing host innate and adaptive immune responses. *Frontiers in Immunology*, *6*. http://dx.doi.org/10.3389/fimmu.2015.00115

MacKnight, N. J., Cobleigh, K., Lasseigne, D., Chaves-Fonnegra, A., Gutting, A., Dimos, B., Antoine, J., Fuess, L., Ricci, C., Butler, C., Muller, E. M., Mydlarz, L. D., & Brandt, M. (2021). Microbial dysbiosis reflects disease resistance in diverse coral species. *Communications Biology*, *4*(1), 679. https://doi.org/10.1038/s42003-021-02163-5

Madin, J. S., Anderson, K. D., Andreasen, M. H., Bridge, T. C. L., Cairns, S. D., Connolly, S. R., Darling, E. S., Diaz, M., Falster, D. S., Franklin, E. C., Gates, R. D., Harmer, A. M. T., Hoogenboom, M. O., Huang, D., Keith, S. A., Kosnik, M. A., Kuo, C.-Y., Lough, J. M., Lovelock, C. E., … Baird, A. H. (2016). The Coral Trait Database, a curated database of trait information for coral species from the global oceans. *Scientific Data*, *3*(1), 160017. https://doi.org/10.1038/sdata.2016.17

Malmstrøm, M., Matschiner, M., Tørresen, O. K., Star, B., Snipen, L. G., Hansen, T. F., Baalsrud, H. T., Nederbragt, A. J., Hanel, R., Salzburger, W., Stenseth, N. C., Jakobsen, K. S., & Jentoft, S. (2016). Evolution of the immune system influences speciation rates in teleost fishes. *Nature Genetics*, *48*(10), 1204–1210. https://doi.org/10.1038/ng.3645

Mansfield, K. M., & Gilmore, T. D. (2019). Innate immunity and cnidarian-Symbiodiniaceae mutualism. *Developmental & Comparative Immunology*, *90*, 199–209. https://doi.org/10.1016/j.dci.2018.09.020

Matthews, J. L., Crowder, C. M., Oakley, C. A., Lutz, A., Roessner, U., Meyer, E., Grossman, A. R., Weis, V. M., & Davy, S. K. (2017). Optimal nutrient exchange and immune responses operate in partner specificity in the cnidarian-dinoflagellate symbiosis. *Proceedings of the National Academy of Sciences*, *114*(50), 13194–13199. https://doi.org/10.1073/pnas.1710733114

Matz, M. V., Treml, E. A., Aglyamova, G. V., & Bay, L. K. (2018). Potential and limits for rapid genetic adaptation to warming in a Great Barrier Reef coral. *PLOS Genetics*, *14*(4), e1007220. https://doi.org/10.1371/journal.pgen.1007220

Maynard, J., van Hooidonk, R., Eakin, C. M., Puotinen, M., Garren, M., Williams, G., Heron, S. F., Lamb, J., Weil, E., Willis, B., & Harvell, C. D. (2015). Projections of climate conditions that increase coral disease susceptibility and pathogen abundance and virulence. *Nature Climate Change*, *5*(7), 688–694. https://doi.org/10.1038/nclimate2625

McWilliam, M., Pratchett, M. S., Hoogenboom, M. O., & Hughes, T. P. (2020). Deficits in functional trait diversity following recovery on coral reefs. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1918), 20192628. https://doi.org/10.1098/rspb.2019.2628

Miller, J., Muller, E., Rogers, C., Waara, R., Atkinson, A., Whelan, K. R. T., Patterson, M., & Witcher, B. (2009). Coral disease following massive bleaching in 2005 causes 60% decline in coral cover on reefs in the US Virgin Islands. *Coral Reefs*, *28*(4), 925–937. https://doi.org/10.1007/s00338-009-0531-7

Moreton, J., Izquierdo, A., & Emes, R. D. (2016). Assembly, assessment, and availability of De novo generated eukaryotic transcriptomes. *Frontiers in Genetics*, *6*. https://doi.org/10.3389/fgene.2015.00361

Moya, A., Huisman, L., Ball, E. E., Hayward, D. C., Grasso, L. C., Chua, C. M., & Miller, D. J. (2012). Whole transcriptome analysis of the coral *Acropora millepora* reveals complex responses to CO_2-driven acidification during the initiation of calcification: Transcriptomic response of corals to acidification. *Molecular Ecology*, *21*(10), 2440–2454. https://doi.org/10.1111/j.1365-294X.2012.05554.x

Mydlarz, L. D., Jones, L. E., & Harvell, C. D. (2006). Innate immunity, environmental drivers, and disease ecology of marine and freshwater invertebrates. *Annual Review of Ecology, Evolution, and Systematics*, *37*(1), 251–288. https://doi.org/10.1146/annurev.ecolsys.37.091305.110103

Neubauer, E.-F., Poole, A. Z., Neubauer, P., Detournay, O., Tan, K., Davy, S. K., & Weis, V. M. (2017). A diverse host thrombospondin-type-1 repeat protein repertoire promotes symbiont colonization during establishment of cnidarian-dinoflagellate symbiosis. *eLife*, *6*, e24494. https://doi.org/10.7554/eLife.24494

Neubauer, E. F., Poole, A. Z., Weis, V. M., & Davy, S. K. (2016). The scavenger receptor repertoire in six cnidarian species and its putative role in cnidarian-dinoflagellate symbiosis. *PeerJ*, *4*, e2692. https://doi.org/10.7717/peerj.2692

Ohno, S. (1970). *Evolution by gene duplication*. Springer.

Palmer, C. V., McGinty, E. S., Cummings, D. J., Smith, S. M., Bartels, E., & Mydlarz, L. D. (2011). Patterns of coral ecological immunology: Variation in the responses of Caribbean corals to elevated temperature and a pathogen elicitor. *Journal of Experimental Biology*, *214*(24), 4240–4249. https://doi.org/10.1242/jeb.061267

Parkinson, J. E., Baumgarten, S., Michell, C. T., Baums, I. B., LaJeunesse, T. C., & Voolstra, C. R. (2016). Gene expression variation resolves species and individual strains among coral-associated dinoflagellates within the genus *Symbiodinium*. *Genome Biology and Evolution*, *8*(3), 665–680. https://doi.org/10.1093/gbe/evw019

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*, *14*(4), 417–419. https://doi.org/10.1038/nmeth.4197

Pinzón C., J. H., Beach-Letendre, J., Weil, E., & Mydlarz, L. D. (2014). Relationship between phylogeny and immunity suggests older caribbean coral lineages are more resistant to disease. *PLoS One*, *9*(8), e104787. https://doi.org/10.1371/journal.pone.0104787

Pollock, F. J., McMinds, R., Smith, S., Bourne, D. G., Willis, B. L., Medina, M., Thurber, R. V., & Zaneveld, J. R. (2018). Coral-associated bacteria demonstrate phylosymbiosis and cophylogeny. *Nature Communications*, *9*(1), 4921. https://doi.org/10.1038/s41467-018-07275-x

Poole, A. Z., Kitchen, S. A., & Weis, V. M. (2016). The role of complement in cnidarian-dinoflagellate symbiosis and immune challenge in the sea anemone *Aiptasia pallida*. *Frontiers in Microbiology*, *7*. https://doi.org/10.3389/fmicb.2016.00519

Poole, A. Z., & Weis, V. M. (2014). TIR-domain-containing protein repertoire of nine anthozoan species reveals coral–specific expansions and uncharacterized proteins. *Developmental & Comparative Immunology*, *46*(2), 480–488. https://doi.org/10.1016/j.dci.2014.06.002

Prada, C., Hanna, B., Budd, A. F., Woodley, C. M., Schmutz, J., Grimwood, J., Iglesias-Prieto, R., Pandolfi, J. M., Levitan, D., Johnson, K. G., Knowlton, N., Kitano, H., DeGiorgio, M., & Medina, M. (2016). Empty niches after extinctions increase population sizes of modern corals. *Current Biology: CB*, *26*(23), 3190–3194. https://doi.org/10.1016/j.cub.2016.09.039

Prunier, J., Caron, S., & MacKay, J. (2017). CNVs into the wild: Screening the genomes of conifer trees (Picea spp.) reveals fewer gene copy number variations in hybrids and links to adaptation. *BMC Genomics*, *18*(1), 97. https://doi.org/10.1186/s12864-016-3458-8

Qiu, Q., Zhang, G., Ma, T., Qian, W., Wang, J., Ye, Z., Cao, C., Hu, Q., Kim, J., Larkin, D. M., Auvil, L., Capitanu, B., Ma, J., Lewin, H. A., Qian, X., Lang, Y., Zhou, R., Wang, L., Wang, K., … Liu, J. (2012). The yak genome and adaptation to life at high altitude. *Nature Genetics*, *44*(8), 946–949. https://doi.org/10.1038/ng.2343

R Core Team (2020). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*. https://www.R-project.org/

Rädecker, N., Pogoreutz, C., Gegner, H. M., Cárdenas, A., Roth, F., Bougoure, J., Guagliardo, P., Wild, C., Pernice, M., Raina, J.-B., Meibom, A., & Voolstra, C. R. (2021). Heat stress destabilizes symbiotic nutrient cycling in corals. *Proceedings of the National Academy of Sciences*, *118*(5), e2022653118. https://doi.org/10.1073/pnas.2022653118

Revell, L. J. (2012). phytools: An R package for phylogenetic comparative biology (and other things): phytools: R package. *Methods in Ecology and Evolution*, *3*(2), 217–223. https://doi.org/10.1111/j.2041-210X.2011.00169.x

Rey, C., Veber, P., Boussau, B., & Sémon, M. (2019). CAARS: Comparative assembly and annotation of RNA-Seq data. *Bioinformatics*, *35*(13), 2199–2207. https://doi.org/10.1093/bioinformatics/bty903

Rippe, J. P., Dixon, G., Fuller, Z. L., Liao, Y., & Matz, M. (2021). Environmental specialization and cryptic genetic divergence in two massive coral species from the Florida Keys Reef Tract. *Molecular Ecology*, *mec.15931*, https://doi.org/10.1111/mec.15931

Rivera, H. E., Aichelman, H. E., Fifer, J. E., Kriefall, N. G., Wuitchik, D. M., Wuitchik, S. J. S., & Davies, S. W. (2021). A framework for understanding gene expression plasticity and its influence on stress tolerance. *Molecular Ecology*, *30*(6), 1381–1397. https://doi.org/10.1111/mec.15820

Roach, T. N. F., Dilworth, J., H., C. M., Jones, A. D., Quinn, R. A., & Drury, C. (2021). Metabolomic signatures of coral bleaching history. *Nature Ecology & Evolution*, *5*(4), 495–503. https://doi.org/10.1038/s41559-020-01388-7

Rohlfs, R. V., & Nielsen, R. (2015). Phylogenetic ANOVA: The expression variance and evolution model for quantitative trait evolution. *Systematic Biology*, *64*(5), 695–708. https://doi.org/10.1093/sysbio/syv042

Ronco, F., Matschiner, M., Böhne, A., Boila, A., Büscher, H. H., El Taher, A., Indermaur, A., Malinsky, M., Ricci, V., Kahmen, A., Jentoft, S., & Salzburger, W. (2021). Drivers and dynamics of a massive adaptive radiation in cichlid fishes. *Nature*, *589*(7840), 76–81. https://doi.org/10.1038/s41586-020-2930-4

Rosales, S. M., Clark, A. S., Huebner, L. K., Ruzicka, R. R., & Muller, E. M. (2020). Rhodobacterales and Rhizobiales are associated with stony coral tissue loss disease and its suspected sources of transmission. *Frontiers in Microbiology*, *11*, 681. https://doi.org/10.3389/fmicb.2020.00681

Sackton, T. B., Lazzaro, B. P., & Clark, A. G. (2017). Rapid expansion of immune-related gene families in the house fly, *Musca domestica*. *Molecular Biology and Evolution*, *34*(4), 857–872. https://doi.org/10.1093/molbev/msw285

Sackton, T. B., Lazzaro, B. P., Schlenke, T. A., Evans, J. D., Hultmark, D., & Clark, A. G. (2007). Dynamic evolution of the innate immune system in Drosophila. *Nature Genetics*, *39*(12), 1461–1468. https://doi.org/10.1038/ng.2007.60

Shinzato, C., Khalturin, K., Inoue, J., Zayasu, Y., Kanda, M., Kawamitsu, M., & Satoh, N. (2020). Eighteen coral genomes reveal the evolutionary origin of acropora strategies to accommodate environmental changes. *Molecular Biology and Evolution*, *38*(1), 16–30. https://doi.org/10.1093/molbev/msaa216

Shoguchi, E., Beedessee, G., Hisata, K., Tada, I., Narisoko, H., Satoh, N., Kawachi, M., & Shinzato, C. (2021). A new dinoflagellate genome illuminates a conserved gene cluster involved in sunscreen biosynthesis. *Genome Biology and Evolution*, *13*(2), evaa235. https://doi.org/10.1093/gbe/evaa235

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*(19), 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Smith, T. B., Brandt, M. E., Calnan, J. M., Nemeth, R. S., Blondeau, J., Kadison, E., Taylor, M., & Rothenberger, P. (2013). Convergent mortality responses of Caribbean coral species to seawater warming. *Ecosphere*, *4*(7), art87. https://doi.org/10.1890/ES13-00107.1

Soneson, C., Love, M. I., & Robinson, M. D. (2015). Differential analyses for RNA-seq: Transcript-level estimates improve gene-level inferences. *F1000Research*, *4*, 1521. https://doi.org/10.12688/f1000research.7563.1

Song, M. J., Potter, B. I., Doyle, J. J., & Coate, J. E. (2020). Gene balance predicts transcriptional responses immediately following ploidy change in *Arabidopsis thaliana*. *The Plant Cell*, *32*(5), 1434–1448. https://doi.org/10.1105/tpc.19.00832

Spady, T. C., Parry, J. W. L., Robinson, P. R., Hunt, D. M., Bowmaker, J. K., & Carleton, K. L. (2006). Evolution of the cichlid visual palette through ontogenetic subfunctionalization of the opsin gene arrays. *Molecular Biology and Evolution*, *23*(8), 1538–1547. https://doi.org/10.1093/molbev/msl014

Stearns, F. W. (2010). One hundred years of pleiotropy: A retrospective. *Genetics*, *186*(3), 767–773. https://doi.org/10.1534/genetics.110.122549

Stuart-Smith, R. D., Brown, C. J., Ceccarelli, D. M., & Edgar, G. J. (2018). Ecosystem restructuring along the Great Barrier Reef following mass coral bleaching. *Nature*, *560*(7716), 92–96. https://doi.org/10.1038/s41586-018-0359-9

Sutherland, K., Porter, J., & Torres, C. (2004). Disease and immunity in Caribbean and Indo-Pacific zooxanthellate corals. *Marine Ecology Progress Series*, *266*, 273–302. https://doi.org/10.3354/meps266273

Taylor, J. S., & Raes, J. (2004). Duplication and divergence: The evolution of new genes and old ideas. *Annual Review of Genetics*, *38*(1), 615–643. https://doi.org/10.1146/annurev.genet.38.072902.092831

Tejada-Martinez, D., de Magalhães, J. P., & Opazo, J. C. (2021). Positive selection and gene duplications in tumour suppressor genes reveal clues about how cetaceans resist cancer. *Proceedings of the Royal Society B: Biological Sciences*, *288*(1945), 20202592. https://doi.org/10.1098/rspb.2020.2592

Tekaia, F. (2016). Inferring orthologs: Open questions and perspectives. *Genomics Insights*, *9*, GEI.S37925. https://doi.org/10.4137/GEI.S37925

Thornhill, D. J., Fitt, W. K., & Schmidt, G. W. (2006). Highly stable symbioses among western Atlantic brooding corals. *Coral Reefs*, *25*(4), 515–519. https://doi.org/10.1007/s00338-006-0157-y

Traylor-Knowles, N., Connelly, M. T., Young, B. D., Eaton, K., Muller, E. M., Paul, V. J., Ushijima, B., DeMerlis, A., Drown, M. K., Goncalves, A., Kron, N., Snyder, G. A., Martin, C., & Rodriguez, K. (2021). Gene expression response to stony coral tissue loss disease transmission in *M. cavernosa* and *O. faveolata* from Florida. *Frontiers in Marine Science*, *8*, 681563. https://doi.org/10.3389/fmars.2021.681563

Ungaro, A., Pech, N., Martin, J.-F., McCairns, R. J. S., Mévy, J.-P., Chappaz, R., & Gilles, A. (2017). Challenges and advances for transcriptome assembly in nonmodel species. *PLoS One*, *12*(9), e0185020. https://doi.org/10.1371/journal.pone.0185020

van de Water, J. A. J. M., Chaib De Mares, M., Dixon, G. B., Raina, J.-B., Willis, B. L., Bourne, D. G., & van Oppen, M. J. H. (2018). Antimicrobial and stress responses to increased temperature and bacterial pathogen challenge in the holobiont of a reef-building coral. *Molecular Ecology*, *27*(4), 1065–1080. https://doi.org/10.1111/mec.14489

van Oppen, M. J. H., & Blackall, L. L. (2019). Coral microbiome dynamics, functions and design in a changing world. *Nature Reviews Microbiology*, *17*(9), 557–567. https://doi.org/10.1038/s41579-019-0223-4

van Woesik, R., & Randall, C. J. (2017). Coral disease hotspots in the Caribbean. *Ecosphere*, *8*(5), e01814. https://doi.org/10.1002/ecs2.1814

Vega Thurber, R., Mydlarz, L. D., Brandt, M., Harvell, D., Weil, E., Raymundo, L., Willis, B. L., Langevin, S., Tracy, A. M., Littman, R., Kemp, K. M., Dawkins, P., Prager, K. C., Garren, M., & Lamb, J. (2020). Deciphering coral disease dynamics: Integrating host,

microbiome, and the changing environment. *Frontiers in Ecology and Evolution*, *8*, 575927. https://doi.org/10.3389/fevo.2020.575927

Veitia, R. A., & Birchler, J. A. (2021). Gene-dosage issues: A recurrent theme in whole genome duplication events. *Trends in Genetics*, *38*(1), 1–3. https://doi.org/10.1016/j.tig.2021.06.006

Waterhouse, R. M., Kriventseva, E. V., Meister, S., Xi, Z., Alvarez, K. S., Bartholomay, L. C., Barillas-Mury, C., Bian, G., Blandin, S., Christensen, B. M., Dong, Y., Jiang, H., Kanost, M. R., Koutsos, A. C., Levashina, E. A., Li, J., Ligoxygakis, P., MacCallum, R. M., Mayhew, G. F., … Christophides, G. K. (2007). Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. *Science*, *316*(5832), 1738–1743. https://doi.org/10.1126/science.1139862

Williams, L. M., Fuess, L. E., Brennan, J. J., Mansfield, K. M., Salas-Rodriguez, E., Welsh, J., Awtry, J., Banic, S., Chacko, C., Chezian, A., Dowers, D., Estrada, F., Hsieh, Y.-H., Kang, J., Li, W., Malchiodi, Z., Malinowski, J., Matuszak, S., McTigue, T., … Gilmore, T. D. (2018). A conserved Toll-like receptor-to-NF-κB signaling pathway in the endangered coral *Orbicella faveolata*. *Developmental & Comparative Immunology*, *79*, 128–136. https://doi.org/10.1016/j.dci.2017.10.016

Williams, L., Smith, T. B., Burge, C. A., & Brandt, M. E. (2020). Species-specific susceptibility to white plague disease in three common Caribbean corals. *Coral Reefs*, *39*(1), 27–31. https://doi.org/10.1007/s00338-019-01867-9

Wright, R. M., Aglyamova, G. V., Meyer, E., & Matz, M. V. (2015). Gene expression associated with white syndromes in a reef building coral, *Acropora hyacinthus*. *BMC Genomics*, *16*(1). https://doi.org/10.1186/s12864-015-1540-2

Wu, Y., Zhou, Z., Wang, J., Luo, J., Wang, L., & Zhang, Y. (2019). Temperature regulates the recognition activities of a galectin to pathogen and symbiont in the scleractinian coral *Pocillopora damicornis*. *Developmental & Comparative Immunology*, *96*, 103–110. https://doi.org/10.1016/j.dci.2019.03.003

Zhang, J. (2003). Evolution by gene duplication: An update. *Trends in Ecology & Evolution*, *18*(6), 292–298. https://doi.org/10.1016/S0169-5347(03)00033-8

Zhang, X., Blair, D., Wolinska, J., Ma, X., Yang, W., Hu, W., & Yin, M. (2021). Genomic regions associated with adaptation to predation in *Daphnia* often include members of expanded gene families. *Proceedings of the Royal Society B: Biological Sciences*, *288*(1955), 20210803. https://doi.org/10.1098/rspb.2021.0803

Ziegler, M., Grupstra, C. G. B., Barreto, M. M., Eaton, M., BaOmar, J., Zubier, K., Al-Sofyani, A., Turki, A. J., Ormond, R., & Voolstra, C. R. (2019). Coral bacterial community structure responds to environmental change in a host-specific manner. *Nature Communications*, *10*(1), 3092. https://doi.org/10.1038/s41467-019-10969-5

Zimmer, C. T., Garrood, W. T., Singh, K. S., Randall, E., Lueke, B., Gutbrod, O., Matthiesen, S., Kohler, M., Nauen, R., Davies, T. G. E., & Bass, C. (2018). Neofunctionalization of duplicated P450 genes drives the evolution of insecticide resistance in the brown Planthopper. *Current Biology*, *28*(2), 268–274.e5. https://doi.org/10.1016/j.cub.2017.11.060

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Dimos, B., Emery, M., Beavers, K., MacKnight, N., Brandt, M., Demuth, J., & Mydlarz, L. (2022). Adaptive variation in homologue number within transcript families promotes expression divergence in reef-building coral. *Molecular Ecology*, *31*, 2594–2610. https://doi.org/10.1111/mec.16414