Environmental Epigenetics, 2026, **12**(1), 1–17
https://doi.org/10.1093/eep/dvag021
Published: 19 June 2026
Research article

Characterization of long noncoding RNAs, microRNAs, and piwiRNAs in reef-building corals

Jill Ashey[1,*], Javier A. Rodriguez-Casariego[2,3], Kathleen M. Durkin[4], Zachary Bengtsson[4], Samuel J. White[4], Ariana S. Huffmyer[1,4], Danielle M. Becker[1], Jose M. Eirin-Lopez[2], Hollie M. Putnam[1], Steven B. Roberts[4]

[1]Department of Biological Sciences, University of Rhode Island, Kingston, RI 02881, United States
[2]Environmental Epigenetics Laboratory, Institute of Environment, Florida International University, Miami, FL 33199, United States
[3]Rosenstiel School of Marine, Atmospheric and Earth Sciences, University of Miami, Miami, FL 33149, United States
[4]School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA 98195, United States

*Corresponding author. Department of Biology, Univerisity of Pennsylvania, Philadelphia, PA 19104, United States. E-mail: jillashey12@gmail.com

Abstract

Noncoding RNAs (ncRNAs) play critical regulatory roles in gene expression regulation that influence diverse biological processes. Yet, their characterization in nonmodel organisms, particularly sessile, benthic ecosystem engineers such as reef-building corals, remains limited. This study provides the first comprehensive analysis of expressed ncRNAs from three ecologically important coral genera from Mo'orea, French Polynesia: *Acropora*, *Pocillopora*, and *Porites*. We identified homologs for ncRNA biogenesis and functional machinery, characterized long ncRNAs (lncRNAs), microRNAs (miRNAs), and piwiRNAs (piRNAs), and assessed their genomic context and potential targets. Our findings reveal the presence of conserved ncRNA machinery across these coral species, indicating their capability to generate and utilize ncRNAs for gene expression regulation. We identified a single miRNA conserved with corals and Eumetazoans (miR-100), four miRNAs shared across all three species and previously identified in other cnidarian taxa (miR-100, miR-2023, miR-2025, and miR-2036), as well as several species-restricted miRNAs. Predicted gene targets of characterized miRNAs included immune response regulation in *A. pulchra* and *P. tuahiniensis* and signal transduction pathways in *P. evermanni* and *P. tuahiniensis*. Proximity analysis indicated ∼71%–99% of piRNAs overlapped with genes, with genomic maintenance and stability identified as the primary functional enrichment of those genes. lncRNA characterization found little sequence overlap across species (<1%), although lncRNAs in all species were often in proximity to immune-related genes. This study characterizes the genomic features of diverse ncRNAs in reef-building corals and identifies putative regulatory associations between ncRNAs and genes involved in core biological processes.

Keywords **noncoding RNAs, long ncRNAs, microRNAs, piwiRNAs, corals**

Introduction

Long thought to be the molecular "junk" of the cell, noncoding RNAs (ncRNAs) have emerged as pivotal regulators of cellular structure and function through gene expression regulation across diverse cell types and tissues [1]. Unlike protein-coding RNAs (i.e. messenger RNAs; mRNAs), ncRNAs do not encode for proteins, but instead perform crucial structural and regulatory functions [2]. ncRNAs can be categorized into two broad classes: structural RNAs, such as ribosomal RNAs (rRNAs) and transfer RNAs (tRNAs) that primarily facilitate translation through their secondary structures, and regulatory ncRNAs, including long and small noncoding RNAs [2] that interact with mRNA transcripts to regulate their translation or degradation. Regulatory ncRNAs can be further subdivided into different families based on their functions in the cell and the length of the transcript, typically being categorized as long ncRNA (lncRNA; >200 nucleotides) and small ncRNA (sncRNA; <200 nucleotides). sncRNAs can be further divided into microRNAs (miRNAs, ∼21–24 nucleotides) that

Received: 31 December 2025. **Revised:** 7 May 2026. **Accepted:** 1 June 2026

regulate mRNA translation and degradation, piwi interacting RNAs (piRNAs, ~24–32 nucleotides) with a mRNA silencing role, and other nuclear ncRNAs, such as small nuclear RNAs (snRNAs, ~150 nucleotides), which aid in spliceosome assembly and function in the nucleus [3,4].

Regulatory ncRNAs interact with mRNA via direct and indirect mechanisms. miRNAs interact with Argonaute proteins (i.e. family of proteins involved in gene expression regulation) to form the RNA-induced silencing complex (RISC), which facilitates translational repression or degradation of mRNA targets [5]. Similarly, piRNAs associate with Piwi-proteins (Argonaute/PIWI family) to form piRISC complexes that silence transposable elements and other complementary RNA targets (e.g. pseudogenes, *nanos* mRNA, transcripts) in the germline and the soma [6,7]. lncRNA function is broader in scope than other ncRNAs, with evidence that they play key regulatory roles in gene expression, chromatin organization, and cellular architecture, through RNA–RNA, RNA–DNA, and RNA–protein interactions [8].

In plants and animals, ncRNAs are widely implicated in gene expression regulation in numerous biological contexts, including cellular homeostasis and response to environmental perturbations [3,9,10]. Experimental studies in model systems have demonstrated that disruption of ncRNAs can influence organismal responses. For example, depletion of a miRNA (miR-85) in *Caenorhabditis elegans* during heat stress resulted in overexpression of heat shock proteins and reduced survival [11], while the knockdown of the lncRNA *COLD INDUCED lncRNA 1* in *Arabidopsis* increased susceptibility to cold stress, as evidenced by increased endogenous reactive oxygen species and reduced osmoregulatory substances [12]. Further, *C. elegans* can also modulate piRNA expression in response to increased temperatures and bacterial infection, with these expression profiles persisting across multiple generations [13]. Collectively, these studies demonstrate that ncRNAs play important roles in regulatory processes across a range of phyla and kingdoms, motivating investigation of their diversity and organization in nonmodel organisms.

Despite extensive characterization of ncRNAs in model organisms, comparatively little is known about ncRNAs in nonmodel taxa, including aquatic and marine organisms [14,15]. Reef-building corals (phylum Cnidaria) are basal metazoans [16], sessile for most of their life [17,18], and foundational ecosystem engineers on coral reefs worldwide [19]. As sessile organisms, corals must dynamically regulate gene expression to accommodate predictable sources of natural variability, such as seasonal [20–22] and diel environmental cycles [23–26], as well as to stochastic environmental changes due to both natural disturbances and increasing anthropogenic stressors [27–30]. While much attention has been dedicated to coral responses to increasing sea surface temperatures and heatwaves [31,32], which causes coral bleaching or a breakdown in the coral and endosymbiotic dinoflagellate symbiosis [33], comparatively little is known about the baseline ncRNA regulatory architecture that underlies coral gene expression more generally.

Given the central role of ncRNAs in gene expression regulation in other systems, establishing a foundational understanding of ncRNA diversity and putative mechanisms represents a critical step toward resolving RNA-mediated regulatory mechanisms in corals. To date, only a limited number of studies have examined coral ncRNAs, primarily focusing on miRNAs in a single species context, with suggested associations to calcification and symbiosis [34]. Additional work has documented changes in miRNA and lncRNA expression under thermal stress [35–38] and disease challenges [39], but the broader landscape of ncRNAs across coral lineages remains largely unexplored. A comparative, descriptive framework is therefore needed to contextualize these findings and to assess how ncRNA complements vary across coral taxa.

Here, we provide the first in-depth characterization of ncRNAs in three important reef-building coral genera: *Acropora*, *Pocillopora*, and *Porites*. We focus on three species (*Acropora pulchra*, *Pocillopora tuahiniensis*, and *Porites evermanni*) that are representative of genera commonly found on reefs worldwide and are evolutionary distant [40,41]. These species display contrasting symbiotic and life history strategies and thermal sensitivities, presenting a compelling system to examine variation in ncRNA in corals [42]. For instance, branching *Pocillopora* spp. (robust clade; [40]) and *Acropora* spp. (complex clade; [40]) are often characterized by faster growth and higher sensitivity to environmental changes, whereas massive *Porites* spp. (complex clade; [40]) tend to exhibit slower growth and greater environmental tolerance [43]. These genera also differ in their association with Symbiodiniaceae; *Pocillopora* and *Acropora* spp. host a mixture of *Cladocopium*, *Symbiodinium*, and/or *Durusdinium*, while *Porites* exhibits high fidelity primarily with specific *Cladocopium* symbionts (e.g. C15; [42,44–46]). Comparative analysis of ncRNA repertoires across these taxa provides an unique opportunity to explore how regulatory RNA landscapes vary across coral genera and to establish a critical baseline for future functional and experimental studies.

In this study, we aimed to:

1. Identify homologs for ncRNA-related protein machinery across coral species to assess the presence and conservation of pathways involved in ncRNA biogenesis and processing.
2. Characterize the features of ncRNAs, focusing on the lncRNAs, miRNAs, and piRNAs, and compare patterns of sequence similarity and divergence among species.
3. Examine the genomic context of lncRNAs and piRNAs by identifying nearby protein-coding genes and assessing functional enrichment of associated annotations.
4. Predict mRNA–miRNA interactions and evaluate functional enrichment among putative targets to describe potential regulatory associations.

Results

Species ID

mtORF haplotype ID identified all Pocillopora samples as *Pocillopora tuahiniensis* (Haplotype_10; [47]; Fig. S1). Coral nuclear histone region spanning H2A–H4 (H2; [48]) indicated all *Porites* samples as *Porites evermanni* (Fig. S2). Sequencing of the *Pax*-C 46/47 intron [49,50] and the mitochondrial putative control region (933+ bp) plus 83 bp of cytochrome oxidase III as described by [51] confirmed the *Acropora* species as *Acropora pulchra*.

NcRNA protein machinery

At least one protein homolog of all ncRNA proteins of interest (Table 1) was identified in each species.

Long ncRNA

A total of 15 libraries were sequencing, yielding an average of 49M raw reads per sample (Table S3). Quality filtering and trimming retained >98.8% of reads across all samples (Table S3). Alignment rates

Table 1 Presence of ncRNA protein machinery in each species. The number indicates the number of unique proteins of that type that were found.

Protein	Function	*A. pulchra*	*P. evermanni*	*P. tuahiniensis*
AGO2	Part of miRNA/siRNA silencing complex	10	2	3
DGCR8	Forms microprocessor complex with Drosha that cleaves the pri-miRNA into a pre-miRNA (stem-loop structure)	3	1	1
Dicer	Cleaves loop from pre-miRNA to form mature miRNA	4	4	5
Drosha	Forms microprocessor complex with DGCR8 that cleaves the pri-miRNA into a pre-miRNA (stem-loop structure)	2	2	2
Pip5k1a	Blocks pre-miRNA transport into cytoplasm	3	3	2
Piwi	Interacts with piRNAs, facilitates "ping-pong" piRNA biogenesis	3	2	2
Rnase P	Cleaves lncRNA to produce mature 3' ends	9	8	7
Xpo5	Exports pre-miRNA from nucleus to cytoplasm	8	1	1

varied across species, likely reflecting the use of congeneric references for species without published genomes. *Porites evermanni* had the highest overall alignment rate (mean 86.54% with 43 M mapped reads), while *A. pulchra* and *P. tuahiniensis* exhibited lower alignment rates (mean 53.51% with 24 M mapped reads and 59.65% with 30M mapped reads, respectively; Table S3). Despite the lower percentages, the number of mapped reads for these species remained above 20M (Table S3), above the threshold required for robust transcriptomic analyses (typically 10–20 M). *Porites evermanni* samples also displayed a higher proportion of multi-mapped reads (mean 21.26%) compared to *A. pulchra* (4.47%) and *P. tuahiniensis* (6.46%; Table S3), potentially reflecting differences in genome repetitiveness or duplication. Importantly, the final mapping depth (mean >24 M reads per sample) ensures that the biological resolution of downstream analyses was not compromised by the use of cross-species mapping.

Following expression and size filtering, as well as excluding known coding transcripts and predicted coding sequences, the number of putative lncRNA transcripts identified was 6 798 in *A. pulchra*, 2812 in *P. evermanni*, and 4924 in *P. tuahiniensis.* This estimate represents primarily polyadenylated intergenic lncRNAs, due to library prep methods and coding potential filtering. The mean lncRNA lengths for *A. pulchra*, *P. evermanni*, and *P. tuahiniensis* were 696, 1004, and 748 bp, respectively, while the median lengths were 504, 646, and 526 bp, respectively (Fig. 1A). Comparison of shared lncRNAs between the three species indicated 8 (0.06%) shared transcripts, whereas 6565 (45.2%), 2647 (18.2%), and 4818 (33.1%) lncRNAs were unique to *A. pulchra*, *P. evermanni*, and *P. tuahiniensis*, respectively. *Acropora pulchra* shared more lncRNAs with *P. evermanni* (142; 1.1%) and fewer lncRNAs with *P. tuahiniensis* (83; 0.6%). *Porites evermanni* and *P. tuahiniensis* shared the fewest lncRNAs (15; 0.1%). We acknowledge that lncRNAs may be conserved at other levels other than primary sequence, such as syntenic, motif, and/or secondary structure conservation; however, we have restricted our results to the empirically supported lncRNAs.

Short ncRNA: miRNAs

13 small RNA-seq libraries yielded an average of 22.6 M raw reads per sample (Table S4). Two libraries were omitted due to low sequencing depth. Quality filtering and trimming retained >80% of reads across all samples (Table S4). Alignment rates for miRNAs via ShortStack was consistent across species (*A. pulchra* mean 70% with 13 M mapped reads, *P. tuahiniensis* 67.95% with 12 M mapped reads, *P. evermanni* 78.1% with 13 M mapped reads). A portion of aligned reads were placed using a guided probability model in ShortStack for multimapping reads (*A. pulchra* mean 35.64%, *P. tuahiniensis* 31.52%, *P. evermanni* 24.93%). The relatively high percentage of multimapped guided placements across all species ensures that multimapped miRNAs were accurately quantified.

A total of 38, 46, and 37 putative miRNAs were identified from *A. pulchra*, *P. evermanni*, and *P. tuahiniensis*, respectively. In all species, these miRNAs were 21–24 nucleotides long, with an average length of 22 nucleotides (Fig. 1B). Of these miRNAs, 24 from *A. pulchra*, 9 from *P. evermanni*, and 9 from *P. tuahiniensis*, respectively, had matches to the curated database of previously described cnidarian miRNAs. The miRNAs that matched known miRNA sequences from the reference database were named after the known miRNA [i.e. if sequence matched nve-mir-100 and/or adi-mir-100, that miRNA sequence was named (coral species)-mir-100)] If a sequence did not match any known sequence, it was named [coral species]-mir-novel-# in sequential order (Table S5). Sequences were identified that had identical mature sequences, but unique locations in the genome in *A. pulchra* and *P. evermanni*; these sequences were given the same number and differentiated with an "a" or "b." *Pocillopora tuahiniensis* did not have any identical sequences in their putative miRNAs.

Four miRNA sequences were conserved across the three species. These include the eumetazoan-conserved mir-100 and three miRNAs previously described in the cnidarian *Nematostella vectensis*: mir-2023, mir-2025, and mir-2036. Additionally, *A. pulchra* and *P. evermanni* shared 1 miRNA (mir-2030); *P. evermanni* and *P. tuahiniensis* shared 1 miRNA (mir-novel-20); and *A. pulchra* and *P. tuahiniensis* shared 1 miRNA (mir-novel-7).

Short ncRNA: piRNAs

piRNA sequence analyses identified 38.2, 10.7, and 5.4 million sequences with lengths between 24 and 31 bp in *A. pulchra*, *P. tuahiniensis*, and *P. evermanni*, respectively. These sequences did not match mRNA nor any other class of ncRNAs and were thus considered putative piRNA for further analyses. Putative piRNAs were between 24 and 31 bp in length, with most sequences having between 28 and 30 bp (Fig. 1C). Evidence of ping-pong amplification signatures, which are indicative of secondary piRNA biogenesis, was observed in the putative piRNA transcripts through a sequence bias towards uracil on the 5′end and an adenine in the 10th position (Fig. S3; [52]).

piRNA cluster prediction was applied for reliable characterization of putative piRNAs primary transcripts. A total of 97, 107, and 47 piRNA clusters were identified in *A. pulchra*, *P. evermanni*, and *P. tuahiniensis*, respectively. The proportion of aligned putative piRNA

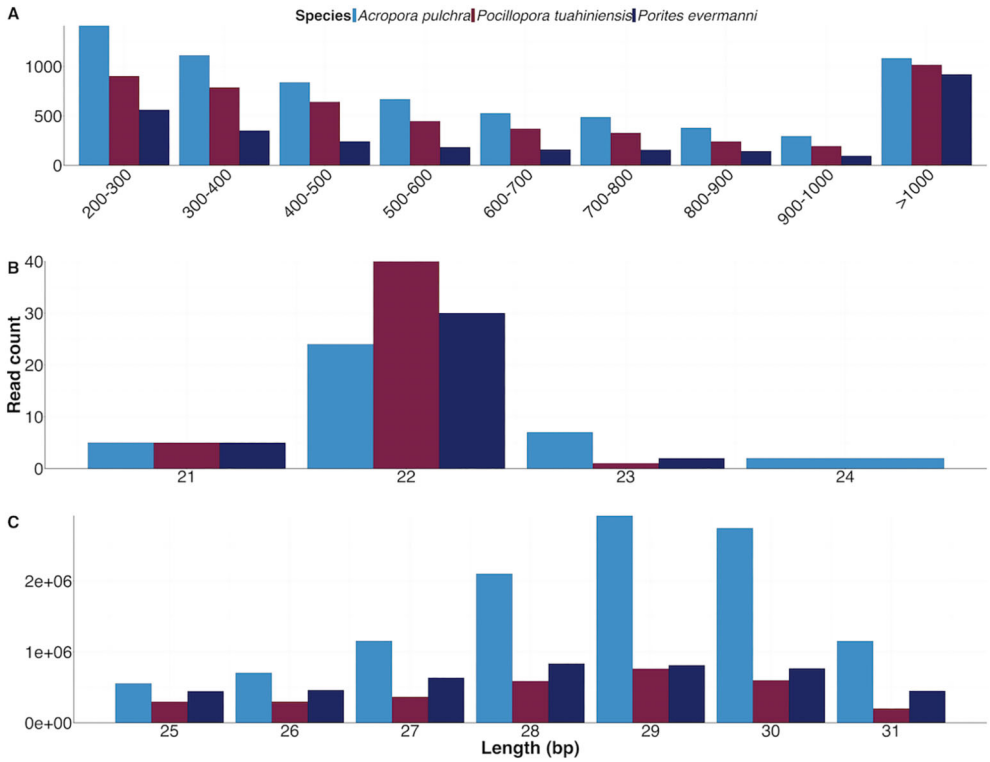

Figure 1 Length distribution of ncRNAs across species. (A) Histogram of lncRNA lengths by species. (B) Histogram of miRNA lengths by species. (C) Histogram of piRNA lengths by species. Alt text: Three bar charts labeled A, B, and C display RNA read length distribution for three coral species: *Acropora pulchra, Pocillopora tuahiniensis,* and *Porites evermanni*. The x-axis represents the sequence length in base pairs, and the y-axis indicates read counts.

reads included in clusters was higher for *A. pulchra* (mean 30.27%) and *P. evermanni* (31.53%) than for *P. tuahiniensis* (19.02%). In all species, the clusters were highly enriched for piRNA signatures, specifically a strong 5′ uracil bias, exceeding 90% for all clusters (Fig. 2A; Table S6). Clusters in all species also demonstrated the ping-pong amplification, with high z-scores across all species, indicative of secondary piRNA biogenesis via the ping-pong cycle (Fig. S3C; Table S6).

Overlaps of clusters with genomic regions of interest followed a similar pattern to all putative piRNAs (Fig. 2B, Fig. S4) with genes and intergenic elements being highly represented. Overlaps with TEs were less abundant in the piRNA clusters than in the putative-piRNA pool, except for *P. tuahiniensis* where overlaps with TEs were 11% more abundant in the clusters, compared to the putative-piRNA pool. Only a small portion of transposon families were enriched in clusters respective to their presence in the genome (Fig. 2C). Long terminal repeats and DNA transposons were significantly enriched in the clusters of all three species, while long interspersed nuclear element, rolling-circle transposons, and short interspersed nuclear elements transposons were differentially enriched between species.

LncRNA and piRNA genomic proximity to genes

Across all species, the vast majority of identified lncRNAs were intergenic, showing no overlap with their nearest protein coding genes (Table 2). However, lncRNAs were often found in relatively close proximity to genes, with mean distances of 8.1 kb for *A. pulchra*, 8.4 kb for *P. tuahiniensis*, and 5.9 kb kb for *P. evermanni*. While no single GO term was significantly enriched across all three species, lncRNA-neighboring genes were broadly associated with innate immunity and cellular proliferation (Fig. 3). Both *A. pulchra* and *P. tuahiniensis* exhibited enrichment for terms related to regulation of interleukin production, cell adhesion, and transcriptional control (Fig. 3; Table S7). Similarly, *A. pulchra* and *P. evermanni* shared enrichment in pathways such as NF-kappaB signaling, wound healing and cell division (Fig. 3; Table S7). Finally, *P. tuahiniensis* and *P. evermanni* showed shared enrichment for blastocyst and megakaryocyte development (Fig. 3; Table S7).

Table 2 Percentages of ncRNAs that overlap (≥1 bp overlap) with a gene.

	A. pulchra (%)	*P. evermanni* (%)	*P. tuahiniensis* (%)
lncRNA	0.2	0.3	0.3
piRNA	71.3	84.7	99.2

In all species, a majority of piRNAs clusters overlapped with coding genes (Table 2). Across all species, overlaps were substantial, with mean lengths of 2558 bp, 1856 bp, and 1417 bp for *A. pulchra*, *P. tuahiniensis*, and *P. evermanni*, respectively. Additionally, minimum overlaps were consistently above the length for individual piRNAs (~24–31 bp); minimum overlaps were 88 bp, 167 bp, and 29 bp for *A. pulchra*, *P. tuahiniensis*, and *P. evermanni*, respectively, indicating that these associations represented substantial genomic overlaps rather than permissive boundary contacts. In all species, genes that overlapped with piRNA clusters were enriched for GO biological processes related to DNA functions. However, no specific GO terms were enriched across all species. For genes closest to piRNA clusters, *A. pulchra* and *P. tuahiniensis* shared enriched GO biological processes, including regulation of mitochondrial and DNA-templated transcription, DNA

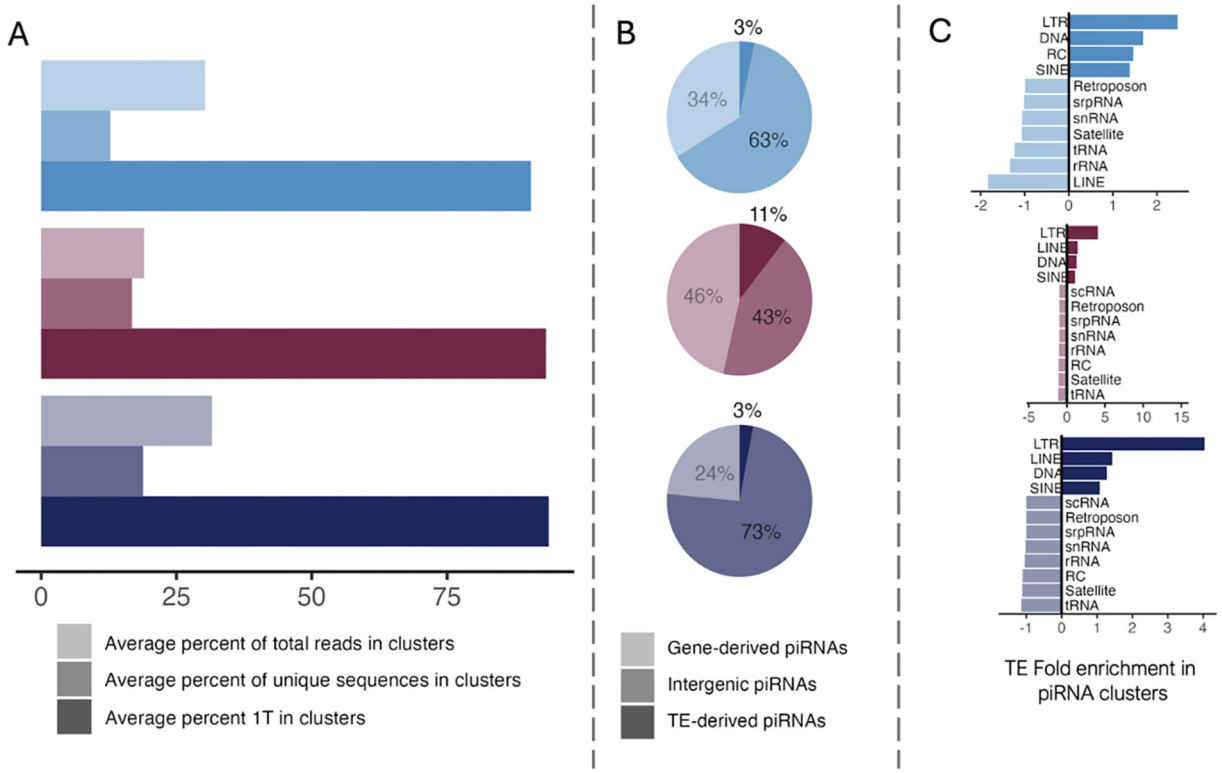

Figure 2 piRNA clusters characteristics in three coral species. (A) Proportion of reads and unique sequences of putative-piRNA included in clusters, and proportion of reads in clusters with thymine (uracil in RNA sequence) in position 1 (5′-end). (B) Proportions of gene-derived, repetitive element-derived, and intergenic putative piRNA sequences in clusters. (C) Transposon enrichment in piRNA clusters respective to their abundance across the genome, indicating which families are more often targeted by piRNA regulation. Alt text: three panels labeled A, B, and C display piRNA cluster statistics, piRNA genomic origins, and piRNA transposable element fold enrichment across three coral species.

recombination, and DNA integration (Fig. 4; Table S8). In contrast, *A. pulchra* and *P. evermanni* shared enriched GO processes related to telomere maintenance, G-quadruplex DNA unwinding, and DNA unwinding during DNA replication (Fig. 4; Table S8). No GO biological processes were shared between *P. evermanni* and *P. tuahiniensis*.

MiRNA target identification

The miRNA regulatory landscape was characterized by high complementarity putative interactions across all three species (Table S9). In *A. pulchra*, 38 miRNAs were predicted to bind 3313 unique targets (mean 105 targets per miRNA), while 37 miRNAs in *P. tuahiniensis* were predicted to bind 3269 unique targets (mean 102). In *P. evermanni*, 45 were predicted to interact with 6027 unique targets (mean 157). While our target prediction utilized a 1000-bp downstream flank in the absence of annotated 3′UTRs, the high degree of sequence complementarity (~75% across all species, spanning covering ~16–18 bp of the miRNA) and high stringency parameters (alignment score ≥ 140; $\Delta G \leq -20$ kcal/mol) suggests that these are robust regulatory candidates (Table S9). This high complementarity aligns with experimental observations in the model cnidarian *Nematostella vectensis*, where miRNAs often mediate target cleavage through extensive base-pairing [53], a mechanism distinct from the seed-based regulation typical of many bilaterians [54].

To identify processes that are potentially regulated by miRNAs in the three species, a GO-enrichment analysis was performed on the predicted gene targets of the miRNAs in each species. *A. pulchra* and *P. tuahiniensis* had the highest number of shared enriched GO biological processes (six terms) including response to bacterium, negative regulation of type I interferon-mediated signaling pathway, motor neuron migration, generation of precursor metabolites and energy, bioluminescence, and axonal fasciculation (Fig. 5; Table S10). *Acropora pulchra* and *P. evermanni* shared two GO biological processes, homophilic cell adhesion via plasma membrane adhesion molecules and defense response to virus (Fig. 5; Table S10). *P. evermanni* and *P. tuahiniensis* shared positive regulation of MAPK cascade and cAMP transport GO terms (Fig. 5; Table S10).

GO enrichment was also evaluated for the miRNAs shared across the three species: mir-100, mir-2023, mir-2025, and mir-2036. Putative targets of mir-100 were enriched for vitamin D receptor signaling pathway, regulation of presynaptic cytosolic calcium ion concentration, and acetyl-CoA biosynthetic processes, in *A. pulchra* and *P. evermanni* (Fig. 6; Table S11). The term sporulation resulting in formation of a cellular spore was enriched in genes targeted by mir-100 in *P. evermanni* and *P. tuahiniensis* (Fig. 6; Table S11). Only *P. evermanni* and *P. tuahiniensis* shared enriched GO terms in the genes targeted by mir-2023; the GO terms negative regulation of T cell mediated immune response to tumor cell, cellular response to forskolin, and cellular response to 2,3,7,8-tetrachlorodibenzodioxine were enriched (Fig. 6; Table S11). For *P. evermanni* and *P. tuahiniensis*, genes targeted by mir-2025 were enriched in processes related to positive regulation of transcription by RNA polymerase II and negative regulation of urine, while mir-2025 targeted genes enriched in coronary artery morphogenesis, and cholesterol catabolic process in *A. pulchra* and *P. tuahiniensis* (Fig. 6; Table S11). Finally, targets of mir-2036 were

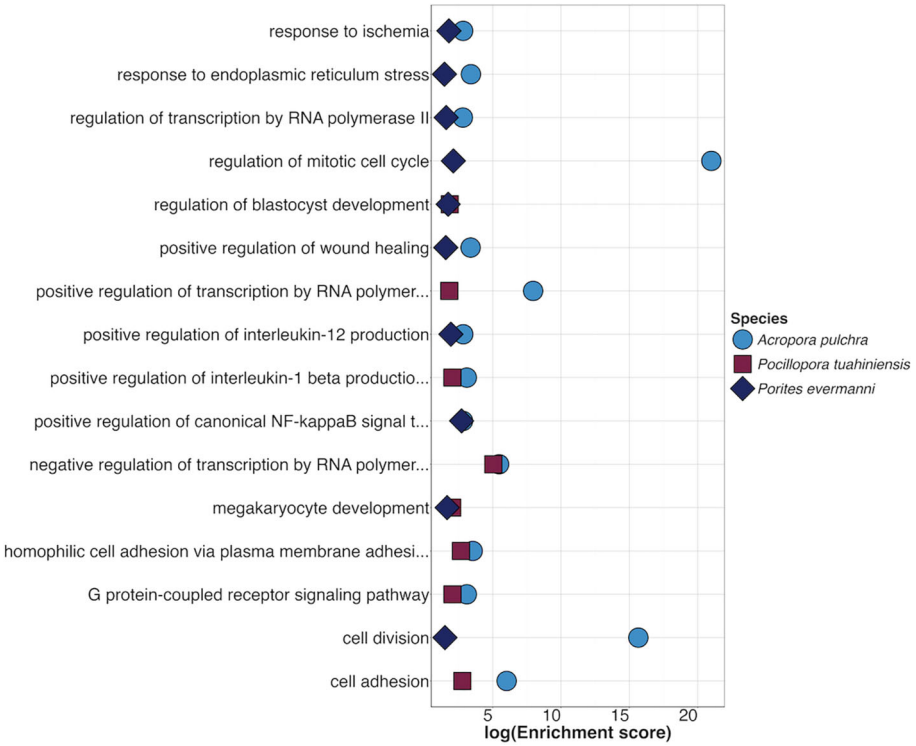

Figure 3 Enriched GO Biological Processes terms in genes closest to lncRNAs in *Acropora pulchra*, *Pocillopora tuahiniensis*, and *Porites evermanni*. The *x*-axis represents the log-transformed enrichment score. Displayed terms include pathways shared between two or more species. Alt text: dot plot displaying log enrichment scores for biological processes in genes closest to lncRNAs across three coral species.

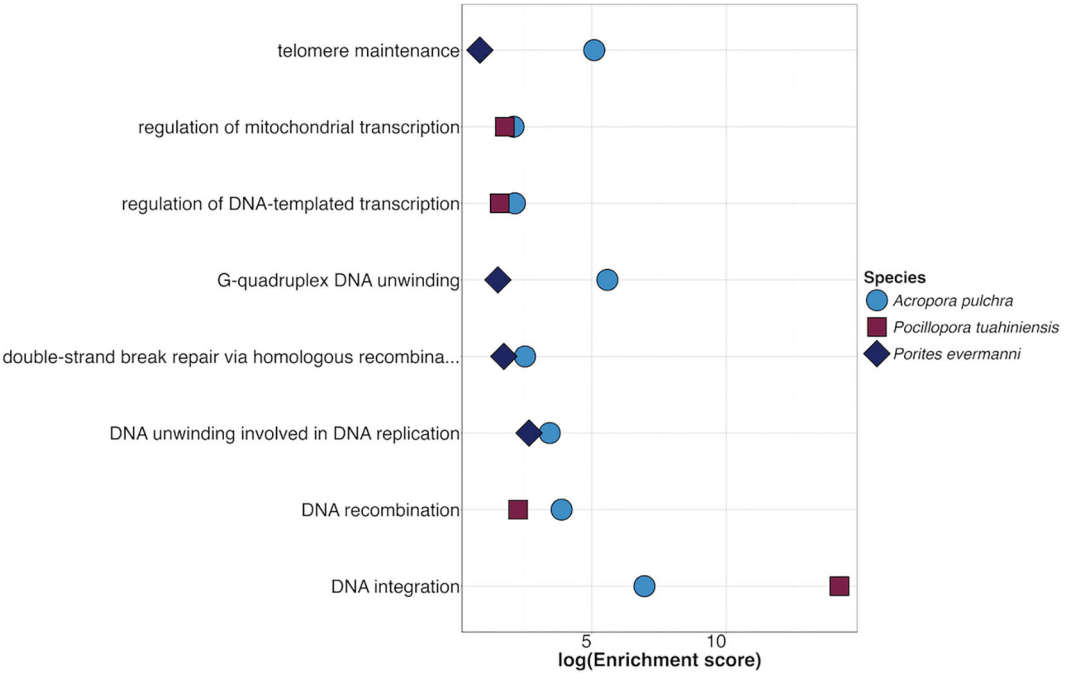

Figure 4 Enriched GO Biological Processes terms in genes overlapping with piRNAs in *Acropora pulchra*, *Pocillopora tuahiniensis*, and *Porites evermanni*. The *x*-axis represents the log-transformed enrichment score. Displayed terms include pathways shared between two or more species. Alt text: dot plot displaying log enrichment scores for biological processes in genes overlapping with piRNAs across three coral species

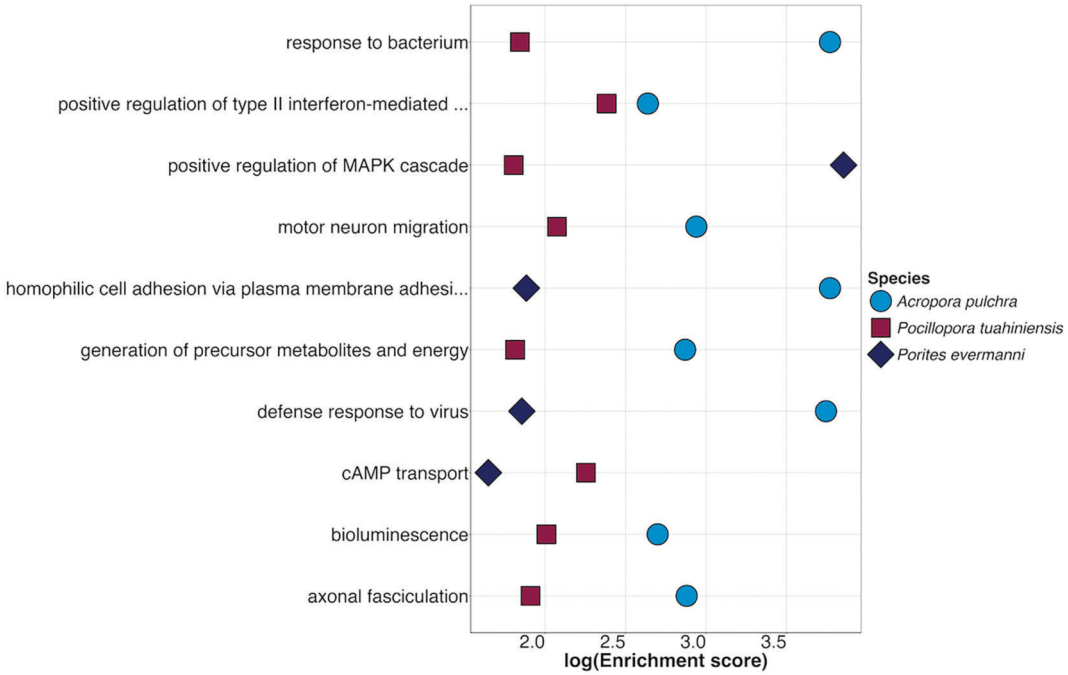

Figure 5 Enriched GO Biological Processes terms in genes putatively targeted by miRNAs in *Acropora pulchra*, *Pocillopora tuahiniensis*, and *Porites evermanni*. The *x*-axis represents the log-transformed enrichment score. Displayed terms include pathways shared between two or more species. Alt text: dot plot displaying log enrichment scores for biological processes for genes putatively targeted by miRNAs across three coral species

enriched in processes including positive regulation of insulin secretion for *A. pulchra* and *P. tuahiniensis*, and homophilic cell adhesion via plasma membrane adhesion molecules for *P. tuahiniensis* and *P. evermanni*.

Discussion

The characterization of diverse ncRNA repertoires in *A. pulchra*, *P. evermanni*, and *P. tuahiniensis* provides new insights into the regulatory RNA landscape of reef-building corals and cnidarians more broadly. The identification of conserved components of ncRNA biogenesis and effector machinery across coral proteomes indicates the molecular framework required for ncRNA-mediated gene regulation is broadly conserved in these taxa. Concurrently, relatively few lncRNA and miRNA sequences were shared across species; while this may be partially impacted by the differing levels of completeness across available reference genomes, it is also indicative of high evolutionary turnover for these ncRNAs, potentially contributing to the unique environmental performance of each coral species. Notably, despite limited sequence conservation, genes associated with ncRNAs were enriched for overlapping functional categories across species, consistent with conservation of regulatory associations rather than conservation of ncRNA sequences themselves.

NcRNA protein machinery present in study species

In all coral species examined in this study, critical miRNA biogenesis proteins, including Drosha, DCGR8, XPO5, and Dicer, were identified (Table 1). The presence of these proteins indicates conservation of canonical miRNA-processing machinery previously reported in corals [34,35] and supports the capacity for production of functional mature miRNA products. AGO2, a key protein in the multiprotein RISC utilized by miRNAs to bind complementary mRNAs for degradation or translation suppression [55], was also present in all coral species. Further, PIP5K1A, a protein that inhibits the transport of pre-miRNAs from the nucleus to the cytoplasm [56], was found, suggesting that multiple levels of post-transcriptional control of miRNAs may be present in corals.

Piwi proteins associated with piRNA pathways were detected in all species examined. In other metazoans, mature piRNAs bind piwi proteins to form piRNA-induced silencing complexes (piRISCs) and silence transposable elements and other targets [7]. While piRNAs were initially described as germline-restricted [52], subsequent studies have found piRNAs and piwi proteins in both somatic and germ cells of the cnidarians *Hydra* and *Nematostella*, indicating putative stem cell properties and/or somatic functionality [57–59]. Because RNA was extracted from whole tissue samples, we were unable to distinguish somatic from germline piRNAs populations in *A. pulchra*, *P. evermanni*, and *P. tuahiniensis*. Given known spawning periods (*A. pulchra* spawns September through November [60,61]; *P. evermanni* spawns December through February [62]; and *P. tuahiniensis* spawns November through December [63]), germ cell abundance was likely low at the time of sampling (March); however, this cannot be conclusively determined. Nevertheless, prevalence of gene-derived piRNAs in our analysis (Fig. 2B) is consistent with the possibility that piRNA pathways in corals may extend beyond canonical germline transpostable element silencing. Further investigation is needed to understand the piRNA-piwi protein functions and dynamics in different coral tissues.

Because lncRNAs arise through diverse biogenesis pathways, identifying proteins specifically associated with lncRNA processing remains challenging [10]. Generally, the biogenesis of lncRNAs is comparable to that of mRNA, in that the lncRNA gets capped, polyadenylated, and spliced utilizing the same proteins as mRNAs [64]. RNase P, while

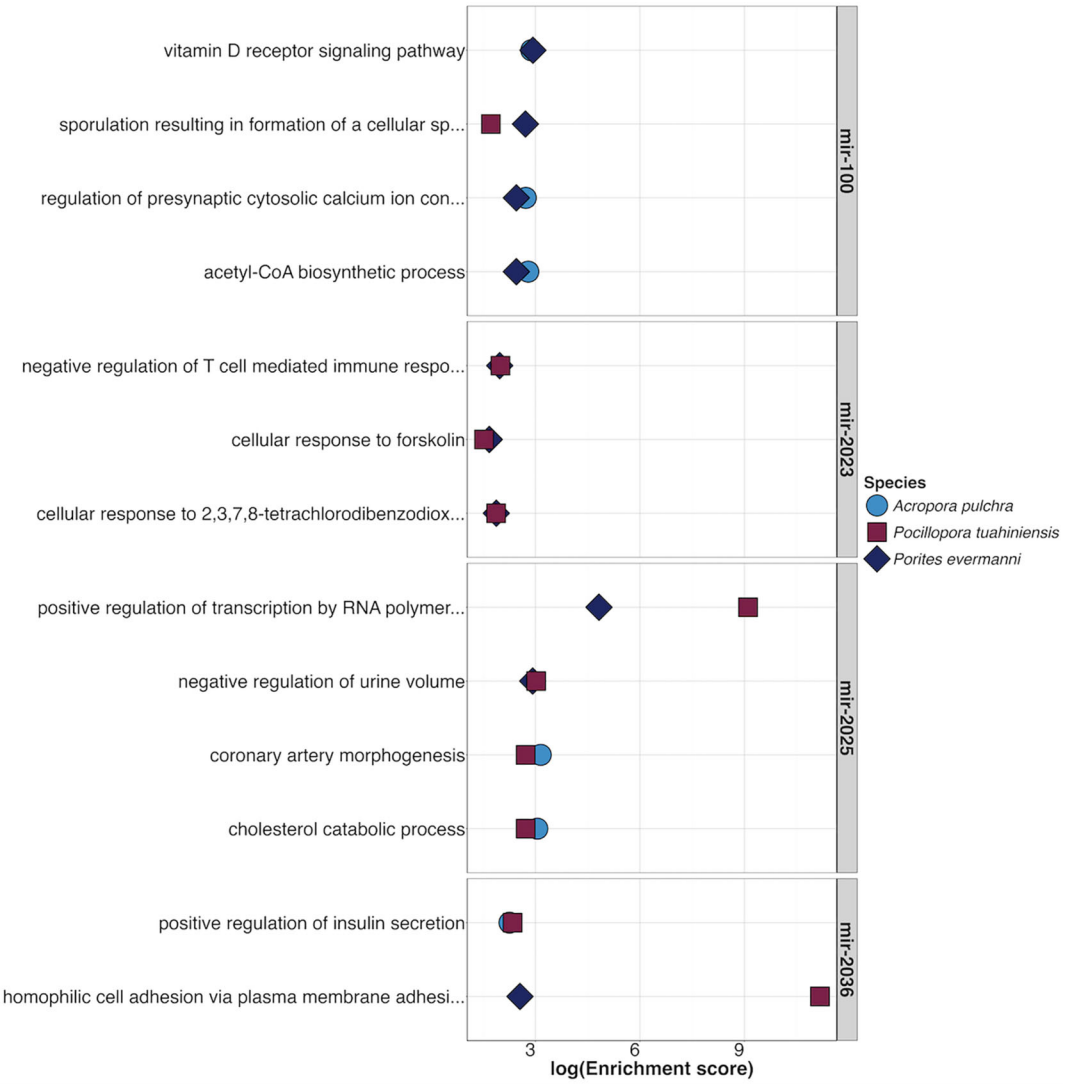

Figure 6 Enriched GO Biological Processes terms in genes putatively targeted by conserved miRNAs (mir-100, mir-2023, mir-2025, and mir-2036) in *Acropora pulchra*, *Pocillopora tuahiniensis*, and *Porites evermanni*. The *x*-axis represents the log-transformed enrichment score. Displayed terms include pathways shared between two or more species. Alt text: dot plot displaying log enrichment scores for biological processes in genes putatively targeted by conserved miRNAs across three coral species

primarily associated with tRNA processing, can also form the mature 3′ end of lncRNAs by cleavage [65]. The identification of RNase P in all three species in this study suggests that noncanonical lncRNA biogenesis pathways may be present in corals.

Evolutionary insights and functional implications of coral lncRNAs

It should be noted that our lncRNA identification workflow was based on the use of polyA selected libraries. While this captures a high-confidence set of polyadenylated lncRNAs, future studies should employ rRNA depletion methods to explain this repertoire to include nonpolyadenylated lncRNAs. Nonetheless, our findings provide a robust baseline for polyadenylated lncRNAs, consistent with existing coral lncRNA datasets [36,37 ,64 ,65]. Our study identified comparable numbers of lncRNAs to other studies in cnidarians: 8117 in *Acropora digitifera* [37], 11 206 in *Protolpalythoa variabilis* [36], 13 240 in *Palythoa caribaeorum* [36], 6328 in *Montipora foliosa* [65], 904 in *Pocillopora damicornis* [38], 7059 in *Pocillopora damicornis* [66], 6798 in *Acropora pulchra* (this study), 2812 in *Porites evermanni* (this study), and 4924 in *Pocillopora tuahiniensis* (this study). Fewer than 1% of lncRNA sequences were shared between *A. pulchra*, *P. evermanni*, and *P. tuahiniensis*, suggesting limited sequence conservation. This pattern contrasts with previous work in the cnidarian order Zoantharia, in which over a third of lncRNAs found in *Palythoa caribaeorum* and *Protolpalythoa variabilis* were shared [36]. The discrepancy between our results and previous findings may stem from the greater evolutionary distance between *A. pulchra*, *P. evermanni*, and *P. tuahiniensis*, compared to the relationship between *P. caribaeorum* and *P. variabilis* [67]. Furthermore, methodological differences likely contributed to this variation. Our comparative BLAST utilized stringent alignment parameters, coupled with expression and coding potential based filtering, which may have led to a conservative estimate of shared lncRNAs across these divergent coral genera. In contrast, the analysis in [66] did not include expression-based filtering, potentially retaining a broader set of low abundance or stochastic transcripts. These findings

reveal significant evolutionary divergence in lncRNAs across diverse coral species and emphasize the importance of expanding comparative research on gene regulation mechanisms within cnidarians.

We identified 4924 lncRNA transcripts in *P. tuahiniensis* using Illumina sequencing technology, a number notably different than the lncRNAs reported for its congener *Pocillopora damicornis*: 904 using Nanopore sequencing [38] and 7059 using PacBio sequencing [66]. These differences may be a result of methodological factors, including extraction protocols, sequencing technologies, lncRNA identification workflows, and stringency of analyses [8]. However, the disparities may also reflect inherent biological differences between *P. tuahiniensis* and *P. damicornis*. Although they are members of the same genus, *P. tuahiniensis* and *P. damicornis* differ in morphology, physiology, and reproductive strategy [47,68], traits that could influence their lncRNA repertoires and their roles in gene expression regulation. Comparative characterization across taxa therefore provides an essential baseline for disentangling methodological variation from true biological diversity.

lncRNAs in *A. pulchra*, *P. evermanni*, and *P. tuahiniensis* were frequently in proximity to genes annotated with immune-related functions. This proximity suggests that lncRNAs may be associated with immune processes in the three species, as proximity to genes can often suggest functional relationships for lncRNA [69,70]. Previous studies have reported associations between lncRNAs and processes, such as thermal resilience [38], bleaching susceptibility [36], and symbiosis [37] in cnidarians. Our findings extend this work by demonstrating that immune-related genes are proximal to lncRNAs across coral species, suggesting that immune-related regulation may be a shared feature of coral lncRNA genomic organization. Experimental validation is required to determine whether these associations reflect direct regulatory interactions, with possible implications for symbiosis and bleaching.

Few miRNAs are conserved across species, but share enriched functions of targets

Four miRNAs (mir-100, mir-2023, mir-2025, and mir-2036) were conserved across all three species of interest and corresponded to miRNAs that have been identified in other cnidarians [54]. mir-100 has been identified in many taxa, including bilaterians and cnidarians [54,71]. In our analyses, genes targeted by mir-100 were enriched in vitamin D receptor signaling pathway, regulation of presynaptic cytosolic calcium ion concentration, and acetyl-CoA biosynthetic process in *A. pulchra* and *P. evermanni*, as well as sporulation resulting in formation of a cellular spore in *P. evermanni* and *P. tuahiniensis*. These enriched processes suggest that while mir-100 is evolutionarily conserved, predicted targets may differ among coral species, highlighting the diverse regulatory roles that mir-100 likely plays in corals. Previous studies have similarly reported divergent targets for miR-100 in cnidarians across multiple life stages [34,35 ,72], underscoring the probable flexibility of miR-100 regulation across species.

Acropora pulchra and *P. tuahiniensis* demonstrated the greatest overlap in the biological processes associated with their predicted target genes. In both species, miRNAs targeted genes with enriched GO terms such as response to bacterium and negative regulation of type I interferon-mediated signaling pathway, suggesting that miRNAs play a role in putative immune response regulation and signaling [39]. In model organisms, miRNAs are increasingly viewed as critical regulators of innate and adaptive immune responses, and their abnormal expression has been linked to human diseases and cancers [73]. In *P. evermanni* and *P. tuahiniensis*, gene targets of miRNAs shared enriched GO terms such as regulation of MAPK cascade and cAMP transport, indicating potential miRNA roles in signal transduction pathways and cellular communication. The identification of diverse yet functionally overlapping GO terms across species indicates that miRNAs may be finely tuned to species-specific needs while still preserving essential functions. This functional diversity could reflect the adaptive role of miRNAs in modulating responses to species-specific environmental challenges, which might include differences in habitat, symbiotic relationships [74,75], or life-history strategies.

Characteristic piRNAs biogenesis and possible regulation of genome maintenance and stability

The putative piRNA clusters identified in *A. pulchra*, *P. tuahiniensis*, and *P. evermanni* had characteristics indicative of piRNA ping-pong biogenesis, including an adenine in the 10th position in the piRNA transcripts and the presence of a uracil on the 5′ end of piRNA transcripts [6]. The piRNA ping-pong mechanism involves an amplification loop in which primary piRNAs, bound to piwi proteins, recognize and cleave target transposon RNAs, generating secondary piRNAs with complementary sequences [7]. The cleavage typically occurs 10 bp downstream of the 5′ end of the primary piRNA, resulting in an adenine in the 10th position [7]. The presence of a 5′ uracil is linked to the binding specificity of the piwi proteins, which preferentially bind to piRNAs with a 5′ uracil, leading to easier loading and stabilization of piRNAs in the piRISC structure [76]. This cycle not only amplifies piRNAs but also provides a robust defense against transposable elements and maintains genomic stability [59]. The conservation of these signatures across coral species suggests a fundamental role for piRNAs in genome defense.

While these biogenesis signatures strongly validate these loci as putative piRNA clusters, their genomic composition showed lower-than-expected TE density. In canonical models, piRNAs are typically enriched for TE sequences to facilitate transposon silencing [7]. The low TE density in our identified clusters likely reflects the inherent challenges of annotating repetitive elements in nonmodel systems, as repetitive TEs are often collapsed or masked during genome assembly, particularly in assemblies with low contiguity [77,78]. Consequently, our results may represent an underestimation of the TE–piRNA relationship in corals. Alternatively, this may suggest that cnidarian piRNAs are utilized for gene regulation rather than exclusive TE defense, a phenomenon increasingly documented in early-diverging metazoans [57,58 ,79 ,80].

No enriched GO terms were shared across all species in genes closest to piRNA clusters. However, genes closest to piRNA clusters in *A. pulchra* and *P. tuahiniensis* shared enriched GO terms, including regulation of mitochondrial and DNA-templated transcription, DNA recombination, and DNA integration. Additionally, enriched functions in genes closest to piRNA clusters in *A. pulchra* and *P. evermanni* included telomere maintenance, G-quadruplex DNA unwinding, and DNA unwinding involved in DNA replication. piRNAs bind to piwi proteins to either silence transposons or genes, indicating potential piRNA-dependent gene regulations [81]. Thus, piRNAs in *A. pulchra*, *P. tuahiniensis*, and *P. evermanni* may be important in regulating genes related to genomic stability, as seen in other taxa [82,83]. The enrichment of shared terms such as DNA recombination and integration in *A. pulchra* and *P. tuahiniensis* and telomere maintenance and

G-quadruplex DNA unwinding in *A. pulchra* and *P. evermanni* suggests that there is a conserved role for piRNAs in maintaining genomic stability and integrity across species [7]. The overlap of piRNAs with genes associated with transcription, DNA maintenance, and genomic stability highlights the potential for piRNA-mediated regulation of these essential processes, potentially safeguarding the genomes of diverse coral species.

Study limitations

While we provide the most comprehensive overview of the ncRNAs in multiple species of coral to date, we acknowledge the limitations of the analyses presented. Our study species presented varying degrees of genomic quality and completeness, which may have contributed to differences observed among species. For *A. pulchra* and *P. tuahiniensis*, we utilized high-quality congeneric genomes (*A. millepora* and *P. meandrina*, respectively). While these species are closely related, the use of proxy genomes may mask species-specific ncRNA variations. Furthermore, the reference genomes varied significantly in assembly contiguity and annotation depth. These discrepancies, inherent to the current state of coral genomics, likely influenced our discovery and functional pipelines.

There was a large difference in the amount of known miRNAs among our taxa, with the highest reported in *A. pulchra*. Using our curated cnidarian database and miRBase, *A. pulchra* had more known miRNAs (24 known miRNAs; 68% of total miRNAs identified) than *P. evermanni* (nine known miRNAs; 19% of total miRNAs identified) or *P. tuahiniensis* (nine known miRNAs; 24% of total miRNAs identified). This higher rate of database matching in *A. pulchra* is potentially related to the composition of our curated reference database. Our curated reference of known cnidarian miRNAs (see Table S2) contained 90 *Acropora* spp. miRNAs, but zero known miRNAs identified in *Porites* or *Pocillopora*. Further, miRBase does not contain known sequences from any of these three study genera.

We acknowledge that several enriched GO terms (e.g. "motor neuron migration," "type I interferon-mediated signaling," and "axonal fasciculation") appear vertebrate-centric and physiologically incongruent with coral biology. These terms result from the transfer of vertebrate-centric databases (SwissProt) to coral homologs during functional annotation. These results are not evidence of specific vertebrate processes, but instead as indicators of conserved molecular elements, such as axonal development and signaling being co-opted for general cell migration or interferon terms representing ancestral innate immune pathways. We frame these functional assignments as homology-based hypotheses and note that the manifestation of these pathways in corals remains to be experimentally validated.

Methods to characterize gene expression in corals are well established, but standardized methodologies and tools to describe gene expression regulation in corals are still in their early stages. Most tools for ncRNA identification or analysis are designed for model systems, which may rely on assumptions that may not apply to ncRNA dynamics in corals. Additionally, the choice of identification method or tool can affect the number of valid ncRNAs identified, as different methodologies often apply varying filters and classification thresholds. Here, we applied stringent filters across all tools to ensure greater confidence in the ncRNA candidates identified and to provide a robust framework for further analysis in coral systems.

Conclusion

Our comprehensive characterization of ncRNA repertoires in three diverse coral species expands the current understanding of regulatory RNA landscapes in corals. The conservation of ncRNA protein machinery and shared ncRNA sequences across species highlight a common molecular framework underlying ncRNA-mediated gene expression regulation, while pronounced sequence divergence accentuates lineage-specific regulatory architectures. Enrichment of genes associated with ncRNAs across functional categories such as immunity, metabolism, and genome maintenance suggests persistent regulatory contexts that may be conserved across coral taxa. Together, these findings greatly expand our current understanding of coral molecular biology and provide essential baseline information regarding coral ncRNAs.

ncRNAs are integral components of epigenetic regulatory systems and may interact with other mechanisms, including DNA methylation and chromatin modification, to influence gene expression [84,85]. Although such interactions are unexplored in corals, the comparative resource presented here provides a foundational framework for future studies aimed at linking ncRNAs to developmental, physiological, and environmental contexts. While our study does not include comparative RNA-seq contrasts across environmental or developmental contexts, it provides essential empirical validation of predicted ncRNAs, which would otherwise rely on computational models. By establishing baseline ncRNA repertoires across coral taxa, this work supports future experimental and comparative efforts to elucidate the roles of regulatory ncRNAs in coral biology.

Materials and methods

Study site and collection

Sample fragments of three coral species (*A. pulchra*, *P. tuahiniensis*, and *P. evermanni*) were collected from the lagoon backreef in Mo'orea, French Polynesia on 5 March 2020 on the north shore (−17.476872, −149.80594). These genera are ecologically common on French Polynesian reefs [86,87] with great importance as habitat builders [88]. During sampling, $n = 1$ fragment (1 cm) was collected from five colonies of each species using bone clippers, snap frozen in liquid nitrogen and then put into 1.5 ml tubes with 600 μl of Zymo DNA/RNA Shield (Zymo CAT R1100-250) and stored at −40°C at the University of California Berkeley Richard B. Gump South Pacific Research Station until transportation to the University of Rhode Island for further processing. Samples were exported under CITES FR2198700194-E on 6 April 2022.

Extractions and sequencing

DNA and RNA were extracted using the Zymo Quick-DNA/RNA Miniprep Plus Kit (Zymo CAT R1057) according to the manufacturer's instructions. DNA and RNA integrity and quantity were assessed using a 1.5% agarose gel (nondenaturing) and a Qubit fluorometer (CAT Q33216, Thermo Fisher/Invitrogen), respectively. Library preparation and sequencing was performed by Azenta Life Sciences. For RNA sequencing (which identifies lncRNAs), strand-specific RNA sequencing libraries were prepared using NEBNext Ultra II Directional RNA Library Prep Kit (NEBCAT:7490) with the polyA mRNA workflow for Illumina following the manufacturer's instructions. Libraries were sequenced using a 2 × 150 bp Paired End (PE) configuration on Illumina

HiSeq 4000. For short RNA sequencing, sequencing libraries were prepared using NEB Small RNA Library Prep Kit (NEBCAT: E7560S) following the manufacturer's instructions. Libraries were sequenced using a 2 × 150 bp Paired End (PE) configuration on Illumina HiSeq 4000.

Species identification

Cryptic species are common in reef-building corals [89–91]. *Acropora pulchra* has historically reliably been identified by morphology and was transplanted to the reef site from a coral nursery, so a single sample was sequenced for validation. One *A. pulchra* colony sample was sequenced using Sanger sequencing by amplifying the *Pax*-C 46/47 intron in the nuclear genome region as described by [49,50] as well as the mitochondrial putative control region (933+ bp) plus 83 bp of cytochrome oxidase III as described by [51]. *Pocillopora* spp. and *Porites* spp. are known to include cryptic species, requiring genetic identification [89–91]. *Pocillopora* species were identified by amplifying the mitochondrial open reading frame (mtORF) region as described by [90,92]. *Porites* species were identified using the coral nuclear histone region spanning H2A–H4 (i.e. H2) [48]. Species identification methods can be found in the Supplemental Methods . Sanger sequences have been deposited at https://osf.io/aw53f/.

Genomic resources

Genomes and protein and gene sequences were downloaded from the following sources described below. Annotations were generated by querying protein (v2.11; blastp) and RNA (v2.11; blastx) fasta files against the SwissProt database (accessed in August 2024; e-value $1E^{-05}$) [93]. At the time of analysis, *Acropora pulchra* did not have a published genome, so genomic information from *Acropora millepora* was used for analysis. *Acropora millepora* genomic information was downloaded from NCBI (PRJNA633778) and has an estimated genome size of 475 Mbp, N50 of 19.8 Mbp, BUSCO score of 96% completeness, and 30 136 protein coding genes, 29 893 of which were annotated (99% of genes annotated) [94]. While *A. millepora* and *A. pulchra* are morphologically distinct, their relatively recent divergence (˜3.5–4 million years ago; [95]) and ability to interbreed [96,97] support the use of the *A. millepora* genome as a reference for *A. pulchra* analyses. Currently to our knowledge, *P. tuahiniensis* does not have a published genome, so genomic information from *P. meandrina*, a closely related species, was used for analysis. *Pocillopora meandrina* genomic information was downloaded from http://cyanophora.rutgers.edu/Pocillopora_meandrina/ with an estimated genome size of 377 Mbp, N50 of 10 Mbp, BUSCO score of 94.9% completeness, and 25 642 protein coding genes, 18 565 of which were annotated (72% of genes annotated) [98]. *Pocillopora meandrina* and *P. tuahiniensis* are morphologically similar, and despite being in different clades, mitochondrial and nuclear genomic data demonstrate that they are congeners; this evolutionary relationship supports the use of the *P. meandrina* genome as a reference for *P. tuahiniensis* [47]. While *P. verrucosa* is more closely related to *P. tuahiniensis* [47] and has a published genome [99], our analysis chose to utilize the *P. meandrina* genome due to its higher assembly contiguity (*P. meandrina* N50 = 10 Mbp, *P. verrucosa* N50 = 0.334 Mbp). *Porites evermanni* genomic information (generated from a sample from Mo'orea, [100]) was downloaded from https://www.genoscope.cns.fr/corals/genomes.html with an estimated genome size of 497 Mbp, N50 of 0.17 Mbp, BUSCO score of 94.5% completeness, and 40 389 protein coding genes, 24 303 of which were annotated (60% of genes annotated) [100]. We note that the reference genomes used in this study vary in assembly contiguity and annotation completeness, differences that are inherent to the current state of coral genomic resources [101]. To mitigate possible bias associated with unequal annotation, comparative analyses utilized SwissProt-validated GO annotations generated in this study to ensure functional comparisons were based on a consistent annotation database rather than genus-specific predicted models. Versions of the genomic resources used in this study are archived at https://osf.io/aw53f/.

NcRNA protein machinery identification

To validate the presence of specific ncRNAs in these coral species, homologs of ncRNA-related proteins of interest were identified in the target coral species genomes, including: ribonuclease III (Drosha), microprocessor complex subunit (DCGR8), exportin-5 (XPO5), endoribonuclease (Dicer), argonaute-2 (AGO2), piwi, phosphatidylinositol-4-phosphate 5-kinase type 1 alpha (PIP5K1A), and ribonuclease P (RNase P). The reference sequences of proteins of interest were obtained from six species (*Pocillopora verrucosa*, *Orbicella faveolata*, *Acropora millepora*, *Stylophora pistillata*, *Nematostella vectensis*, and *Homo sapiens)* available on NCBI (Table S1), and were compared against the protein fasta for each of the three target species genomes above using blastp (v2.11; e-value $1E^{-40}$) [93].

Long ncRNA characterization

RNA sequencing read quality was initially assessed using FastQC (v0.12.1; [102]) and MultiQC (v1.14; [119]). RNA sequencing reads underwent quality trimming using Fastp (v0.23.1; [110]), which involved the identification of read pairs, automatic detection and removal of adapter sequences (—detect_adapter_for_pe), trimming of poly-G tails (—trim_poly_g), and removal of the first 20 bases from the 5′ end to eliminate low-quality bases (—trim_front1). Trimmed reads were aligned to respective genomes with HISAT2 [103] and assembled with Stringtie (v2.2.1; [120]) to obtain read counts and merge sample GTF files. Sample GTF files were compared to genome annotations using gffcompare (v0.12.6; 71). Transcripts were retained if they were >199 bp and had one of the following gffcompare class codes: "i" (intronic), "u" (intergenic), "x" (antisense), or "y" (sense-overlapping; [104]). Initial coding potential was assessed using CPC2 (v1.0.1), which applies a species-neutral support vector machine model trained on human coding and non-coding transcripts [105]. Transcripts classified as "coding" were removed from analysis. To reduce the inclusion of spurious transcripts, individual lncRNAs were only retained if they possessed above five raw read counts, as generated using featureCounts (v2.0.2; [106]), in at least three biological replicates per species. To further ensure the removal of transcripts with coding potential, coding probability was assessed using PLEK (v1.2; [107]), which uses a k-mer based support vector machine to distinguish between noncoding and coding sequences. Finally, transcripts were searched against the Pfam-A database (v35.0; [108]) using HMMscan (HMMER v3.3.2; [109]), and any sequences containing known protein domains were excluded from the final dataset. For all software, default parameters were used. This robust filtering pipeline ensures that only lncRNAs that pass a stringent threshold are retained for analysis and labeled as putative lncRNAs. In order to evaluate sequence similarity across species, lncRNAs from each species were compared to the other species using blastn (v2.11; e-value $1E^{-40}$) [93].

Short ncRNA: miRNA characterization

To characterize miRNAs, the quality of short RNA sequence reads was assessed using FastQC (v0.12.1; [102]) and MultiQC (v1.14; [119]). Fastp (v0.23.1) default parameters were used for read quality trimming [110], which involved the removal of adapter sequences (—adapter_fasta, sequences provided from the NEB Small RNA Library Prep Kit) and polyG sequences (—trim_poly_g). To specifically target miRNAs and piRNAs while minimizing noise from larger RNA products, reads with a length greater than 31 bp were discarded (—length_limit 31). Read 1 and Read 2 were then merged, with 17 bp as the minimum overlap (—length_limit 31—merge).

To identify previously described miRNAs, a database containing mature miRNA sequences from miRBase (v22; https://www.mirbase.org/; [111]) and cnidarian miRNA sequences manually collected from published articles (Table S2) was used as a reference. Previously described and candidate novel miRNAs were identified from alignments of trimmed and filtered reads to respective reference genomes using ShortStack (v4.0.2) with the parameter –dn_mirna, which identifies putative novel miRNAs by searching read-genome alignments for "islands" of significant coverage and evaluating neighboring regions based on a stringent set of miRNA precursor gene characteristics, including hairpin structure [112,113]. To identify miRNAs conserved within all or a subset of the three study species, blastn was run with the –task blastn-short option (v2.11; e-value $1E^{-40}$) [93].

Short ncRNA: Piwi RNA characterization

For piRNA characterization, quality filtered and trimmed short reads from the "*Short ncRNA: miRNA characterization*" section were filtered to keep reads between 24 and 31 bp. Low complexity sequences were removed using NGS toolbox scripts [114]. To remove non-piRNA, short reads were mapped to a track of other noncoding RNAs with SortMeRNA (v4.3.6; [121]). Reference tRNA, rRNA, miRNA and siRNA sequences were acquired or generated for each species. For *A. pulchra*, tRNA and rRNA sequences were obtained from annotation files from the *A. millepora* genomic resources; for *P. tuahiniensis* and *P. evermanni*, tRNA and rRNA sequences were generated *de novo* using tRNAscan-SE (v.2.0.9; [117]) and Barrnap (v.0.9; https://github.com/tseemann/barrnap) with default parameters for eukaryotes. Hairpin, star, and mature miRNAs, and siRNAs annotations were obtained from ShortStack output as described above. Specifically, siRNAs were defined as 20–24 bp small RNAs that exhibited Dicer-processing length but lacked the structural requirements for miRNA annotation. Sequences that did not align to these "known" ncRNAs were mapped to the respective genome using sRNAmappper.pl from the NGSToolbox using the option -alignments best, and origin assignment for reads with multiple mappings was determined using reallocate.pl [114,115]. The resulting sequences were considered putative piRNAs and used for downstream analyses.

The following analysis was used to identify putative piRNA clusters in *A. pulchra*, *P. tuahiniensis*, and *P. evermanni*. ProTrac (v2.4.4) was used to predict piRNA clusters for each sample [115]. Resulting piRNA clusters were merged by species if they were present in at least two replicates, with a minimum overlap of 50% of the cluster length. PPmeter (v0.4; [122]) was used with default parameters to quantify ping-pong amplification signatures, a hallmark of secondary piRNA biogenesis [6]. In the ping-pong cycle, piRNAs are generated through reciprocal cleavage events, which creates the characteristic 10 base sequence overlap between piRNAs [6].

In eukaryotes, piRNAs typically are located in transposable element (TE) regions. To investigate if piRNA clusters were located in TE in corals, TE annotation was performed in the reference genome of each species using RepeatMasker (v4.1.0, http://repeatmasker.org) with default parameters. TE enrichment in piRNA clusters was calculated respectively to each genome by quantifying annotated TE overlap with piRNA clusters and contrasting piRNA cluster-TE overlaps with respect to TE coverage across each genome using bedtools (v2.30.0; [116]).

LncRNA and piRNA genomic proximity to genes

To investigate the proximity of ncRNAs (lncRNAs and piRNA clusters) to genes, bedtools closest (v2.30.0) was performed using the feature files (GFFs) downloaded for each species and the ncRNA feature files (GFFs or BEDs) generated from our analyses [116]. The percentage of ncRNAs that did and did not overlap with a gene were calculated. If an ncRNA did overlap with a gene, average overlap in base pairs (bp) was calculated. If an ncRNA did not overlap with a gene, average distance between the two features in bp was calculated.

Functional enrichment was performed using gene ontology (GO) enrichment analyses with the R package topGO (v.2.60.1; [118]) on the genes that were closest to or overlapping ncRNAs. The use of "closest" will herein refer to either the closest gene or the overlapping gene, depending on where the ncRNA is in the genome relative to the gene. GO enrichment was run using Fisher's exact test (weight01 algorithm, threshold of $P < .05$) to assess GO enrichment in Biological Processes for each species. Enriched GO terms were compared across species to investigate shared enriched gene functions related to Biological Processes.

MiRNA target identification

In cnidarians, miRNAs typically require binding to the 3′ untranslated region (3′UTR) of mRNAs with almost perfect complementarity to repress or downgrade translation [72]. Therefore, target prediction is less prone to false positives in this group. Given the lack of 3′UTR annotation on these genomes, 1000 bp on the right flank of mRNA tracks was classified as the 3′UTR. If a 3′UTR was overlapping a contiguous gene, the overlapping portions of those 3′UTRs were subtracted using bedtools (v2.30.0; [116]). 3′UTR sequences were annotated from the respective genomes using bedtools getfasta (v2.30.0; [116]). MiRanda (v3.3) was then used to predict miRNA gene targets using the constructed 3′UTR sequences and mature miRNA sequences as input with strict seed binding invoked [117]. To minimize false positives associated with unannotated 3′UTRs, an energy threshold of $\Delta G_{duplex} \leq -20$ kcal/mol and a minimum alignment score of 140 was required for all putative interactions. Similar to the previous GO enrichment analysis, topGO with Fisher's exact test (weight01 algorithm, threshold of $P < .05$) was used to assess GO enrichment in Biological Processes on the predicted miRNA-targeted genes [118]. GO enrichment analyses were also conducted on the genes targeted by the miRNAs shared across species (mir-100, mir-2023, mir-2025, and mir-2036). Enriched GO terms were compared across species to investigate shared enriched gene functions related to Biological Processes for the miRNA targets.

Acknowledgments

We thank Kristina Terpis for laboratory assistance. The authors acknowledge use of the resources of the URI Center for Computational Re-

search and the FIU Instructional and Research Computing Center for this work. As guests, we honor and acknowledge the Mā'ohi community, we give thanks for the land and water resources of Polynesia and the traditional custodians of the land on which this experimental work was conducted on the island of Mo'orea. Māuruuru roa. With respect to the spelling of Tahitian words, we endeavored to follow the Te Fare Vāna'a, transcription system that is adhered to by a large segment of the Tahitian community, but also recognize other community members follow the Raapoto transcription system where the island name of Moorea is, for example, spelled without the 'eta (i.e. Moorea).

Author contributions

Jill Ashey (Formal Analysis [lead], Investigation [equal], Methodology [equal], Resources [equal], Software [equal], Validation [equal], Visualization [equal], Writing – original draft [lead], Writing – review & editing [lead]), Javier A Rodríguez-Casariego (Data curation [equal], Formal Analysis [equal], Investigation [equal], Resources [equal], Software [equal], Validation [equal], Visualization [equal], Writing – review & editing [equal]), Kathleen M. Durkin (Data curation [equal], Formal Analysis [equal], Investigation [equal], Resources [equal], Software [equal], Validation [equal], Visualization [equal], Writing – review & editing [equal]), Zachary Bengtsson (Formal Analysis [equal], Investigation [equal], Resources [equal], Software [equal], Validation [equal], Visualization [equal], Writing – review & editing [equal]), Samuel J. White (Data curation [equal], Methodology [equal], Resources [equal], Software [equal]), Ariana S Huffmyer (Investigation [equal], Resources [equal], Software [equal], Writing – review & editing [equal]), Danielle M. Becker (Investigation [equal], Writing – review & editing [equal]), Jose M. Eirin Lopez (Conceptualization [equal], Funding acquisition [equal], Investigation [equal], Methodology [equal], Project administration [equal], Resources [equal], Supervision [equal], Writing – review & editing [equal]), Hollie M Putnam (Conceptualization [equal], Funding acquisition [equal], Investigation [equal], Methodology [equal], Project administration [equal], Resources [equal], Supervision [equal], Writing – review & editing [equal]), and Steven B Roberts (Conceptualization [equal], Funding acquisition [equal], Investigation [equal], Methodology [equal], Project administration [equal], Resources [equal], Supervision [equal], Validation [equal], Writing – review & editing [equal])

Supplementary data

Supplementary data is available at *EnvEpig* online.

Conflict of interest

None declared.

Funding

This work was supported by the National Science Foundation (NSF) Rules of Life-Epigenetics Awards to H.M.P. (grant number 1921465), J.E.L. (grant number 1921402), and S.B.R. (grant number 1921149), as well as NSF Graduate Research Fellowships to J.A. and D.M.B., and supported by resources from NSF-OCE 2224354 to the Mo'orea Coral Reef LTER, as well as a generous gift from the Gordon and Betty Moore Foundation.

Data availability

Data and scripts are available on GitHub at https://github.com/urol-e5/deep-dive. ncRNA feature and genomic visualization for the three species is available on our genome browser platform at https://urol-e5.github.io/deep-dive-genome-browser/. Sanger sequences and archived genomic references have been deposited at https://osf.io/aw53f/. RNA sequences and small RNA sequences are stored on NCBI under BioProjects PRJNA1201098 and PRJNA1236666, respectively.

References

1. Eddy SR. Computational genomics of noncoding RNA genes. *Cell* 2002;**109**:137–40. https://doi.org/10.1016/S0092-8674(02)00727-4
2. Hüttenhofer A, Schattner P, Polacek N. Non-coding RNAs: hope or hype? *Trends Genet* 2005;**21**:289–97. https://doi.org/10.1016/j.tig.2005.03.007
3. Gomes A, Nolasco S, Soares H. Non-coding RNAs: multi-tasking molecules in the cell. *Int J Mol Sci* 2013;**14**:16010–39. https://doi.org/10.3390/ijms140816010
4. Bohnsack MT, Sloan KE. Modifications in small nuclear RNAs and their roles in spliceosome assembly and function. *Biol Chem* 2018;**399**:1265–76. https://doi.org/10.1515/hsz-2018-0205
5. Kaikkonen MU, Lam MTY, Glass CK. Non-coding RNAs as regulators of gene expression and epigenetics. *Cardiovasc Res* 2011;**90**:430–40. https://doi.org/10.1093/cvr/cvr097
6. Czech B, Hannon GJ. One loop to rule them all: the ping-pong cycle and piRNA-guided silencing. *Trends Biochem Sci* 2016;**41**:324–37. https://doi.org/10.1016/j.tibs.2015.12.008
7. Czech B, Munafò M, Ciabrelli F *et al.* piRNA-guided genome defense: from biogenesis to silencing. *Annu Rev Genet* 2018;**52**:131–57. https://doi.org/10.1146/annurev-genet-120417-031441
8. Mattick JS, Amaral PP, Carninci P *et al.* Long non-coding RNAs: definitions, functions, challenges and recommendations. *Nat Rev Mol Cell Biol* 2023;**24**:430–47. https://doi.org/10.1038/s41580-022-00566-8
9. Fu XD. Non-coding RNA: a new frontier in regulatory biology. *Natl Sci Rev* 2014;**1**:190–204. https://doi.org/10.1093/nsr/nwu008
10. Chao H, Hu Y, Zhao L *et al.* Biogenesis, functions, interactions, and resources of non-coding RNAs in plants. *Int J Mol Sci* 2022;**23**:3695. https://doi.org/10.3390/ijms23073695
11. Pagliuso DC, Bodas DM, Pasquinelli AE. Recovery from heat shock requires the microRNA pathway in *Caenorhabditis elegans*. *PLos Genet* 2021;**17**:e1009734. https://doi.org/10.1371/journal.pgen.1009734
12. Liu G, Liu F, Wang Y *et al.* A novel long noncoding RNA CIL1 enhances cold stress tolerance in Arabidopsis. *Plant Sci* 2022;**323**:111370. https://doi.org/10.1016/j.plantsci.2022.111370
13. Belicard T, Jareosettasin P, Sarkies P. The piRNA pathway responds to environmental signals to establish intergenerational adaptation to stress. *BMC Biol* 2018;**16**:103. https://doi.org/10.1186/s12915-018-0571-y
14. Eirin-Lopez JM, Putnam HM. Marine environmental epigenetics. *Annu Rev Mar Sci* 2018;**11**:335–68. doi:https://doi.org/10.1146/annurev- marine-010318095114
15. Li Y, Hui JHL. Small RNAs in Cnidaria: a review. *Evol Appl* 2023;**16**:354–64. https://doi.org/10.1111/eva.13445
16. Miller D, Ball E, Technau U. Cnidarians and ancestral genetic complexity in the animal kingdom. *Trends Genet* 2005;**21**:536–39. https://doi.org/10.1016/j.tig.2005.08.002
17. Baird AH, Guest JR, Willis BL. Systematic and biogeographical patterns in the reproductive biology of scleractinian corals. *Annu Rev*

Ecol Evol Syst 2009;**40**:551–71. https://doi.org/10.1146/annurev.ecolsys.110308.120220
18. Marshall DJ, Morgan SG. Ecological and evolutionary consequences of linked life-history stages in the sea. *Curr Biol* 2011;**21**:R718–25. https://doi.org/10.1016/j.cub.2011.08.022
19. Costanza R, De Groot R, Sutton P *et al.* Changes in the global value of ecosystem services. *Glob Environ Chang* 2014;**26**:152–58. https://doi.org/10.1016/j.gloenvcha.2014.04.002
20. Fitt WK, McFarland FK, Warner ME *et al.* Seasonal patterns of tissue biomass and densities of symbiotic dinoflagellates in reef corals and relation to coral bleaching. *Limnol Oceanogr* 2000;**45**:677–85. https://doi.org/10.4319/lo.2000.45.3.0677
21. Thornhill DJ, Rotjan RD, Todd BD *et al.* A connection between colony biomass and death in caribbean reef-building corals. *PLoS One* 2011;**6**:e29535. https://doi.org/10.1371/journal.pone.0029535
22. Scheufen T, Krämer WE, Iglesias-Prieto R *et al.* Seasonal variation modulates coral sensibility to heat-stress and explains annual changes in coral productivity. *Sci Rep* 2017;**7**:4937. https://doi.org/10.1038/s41598-017-04927-8
23. Barshis DJ, Ladner JT, Oliver TA *et al.* Genomic basis for coral resilience to climate change. *Proc Natl Acad Sci* 2013;**110**:1387–92. https://doi.org/10.1073/pnas.1210224110
24. Safaie A, Silbiger NJ, McClanahan TR *et al.* High frequency temperature variability reduces the risk of coral bleaching. *Nat Commun* 2018;**9**:1671. https://doi.org/10.1038/s41467-018-04074-2
25. Brown KT, Mello-Athayde MA, Sampayo EM *et al.* Environmental memory gained from exposure to extreme pCO2 variability promotes coral cellular acid–base homeostasis. *Proc R Soc B Biol Sci* 2022;**289**:20220941. https://doi.org/10.1098/rspb.2022.0941
26. Ashey J, Scucchia F, Huffmyer AS *et al.* Thermal and acidification gradients reveal tolerance thresholds in Pocillopora acuta recruits. *Mol Ecol* 2025; **34**. https://doi.org/10.1111/mec.70116
27. Grigg RW, Dollar SJ. Natural and anthropogenic disturbance on coral reefs. In: *A Natural History of the Hawaiian Islands: Selected Readings II*. Honolulu: University of Hawaii Press, 1994, 490–503. https://doi.org/10.1515/9780824844264-042
28. Sweet MJ, Brown BE. Coral responses to anthropogenic stress in the twenty- first century: an ecophysiological perspective. In: *Oceanography and Marine Biology*. 1st edn. Boca Raton: CRC Press, 2016. https://doi.org/10.1201/9781315368597
29. Richmond RH, Tisthammer KH, Spies NP. The effects of anthropogenic stressors on reproduction and recruitment of corals and reef organisms. *Front Mar Sci* 2018;**5**:226. https://doi.org/10.3389/fmars.2018.00226
30. Ellis JI, Jamil T, Anlauf H *et al.* Multiple stressor effects on coral reef ecosystems. *Global Change Biol* 2019;**25**:4131–46. https://doi.org/10.1111/gcb.14819
31. Gutierrez L, Polidoro B, Obura D *et al.* Half of Atlantic reef-building corals at elevated risk of extinction due to climate change and other threats. *PLoS One* 2024;**19**:e0309354. https://doi.org/10.1371/journal.pone.0309354
32. Mellin C, Brown S, Cantin N *et al.* Cumulative risk of future bleaching for the world's coral reefs. *Sci Adv* 2024;**10**:eadn9660. https://doi.org/10.1126/sciadv.adn9660
33. Gates RD, Baghdasarian G, Muscatine L. Temperature stress causes host cell detachment in symbiotic cnidarians: implications for coral bleaching. *Biol Bull* 1992;**182**:324–32. https://doi.org/10.2307/1542252
34. Liew YJ, Aranda M, Carr A *et al.* Identification of MicroRNAs in the coral *Stylophora pistillata*. *PLoS One* 2014;**9**:e91101. https://doi.org/10.1371/journal.pone.0091101
35. Gajigan AP, Conaco C. A microRNA regulates the response of corals to thermal stress. *Mol Ecol* 2017;**26**:3472–83. https://doi.org/10.1111/mec.14130
36. Huang C, JÉRL M, Cai J *et al.* Identification of long non-coding RNAs in two anthozoan species and their possible implications for coral bleaching. *Sci Rep* 2017;**7**:5333. https://doi.org/10.1038/s41598-017-02561-y
37. Huang C, Leng D, Sun S *et al.* Re-analysis of the coral *Acropora digitifera* transcriptome reveals a complex lncRNAs-mRNAs interaction network implicated in Symbiodinium infection. *BMC Genomics* 2019;**20**:48. https://doi.org/10.1186/s12864-019-5429-3
38. Yu X, Yu K, Chen B *et al.* Nanopore long-read RNAseq reveals regulatory mechanisms of thermally variable reef environments promoting heat tolerance of scleractinian coral *Pocillopora damicornis*. *Environ Res* 2021;**195**:110782. https://doi.org/10.1016/j.envres.2021.110782
39. Despard BA, Selwyn JD, Shupp AN *et al.* A network approach to white band disease challenged staghorn coral *Acropora cervicornis* microRNAs and their targets. *Ecol Evol* 2025;**15**:e71351. https://doi.org/10.1002/ece3.71351
40. Kitahara MV, Cairns SD, Stolarski J *et al.* A comprehensive phylogenetic analysis of the Scleractinia (Cnidaria, Anthozoa) based on mitochondrial CO1 sequence data. *PLoS One* 2010;**5**:e11490. https://doi.org/10.1371/journal.pone.0011490
41. Stolarski J, Kitahara MV, Miller DJ *et al.* The ancient evolutionary origins of Scleractinia revealed by azooxanthellate corals. *BMC Evol Biol* 2011;**11**:316. https://doi.org/10.1186/1471-2148-11-316
42. Huffmyer AS, Strand EL, Hackerott S *et al.* Seasonal plasticity of symbiotic strategies clarifies coral holobiont resistance and resilience. bioRxiv 2025. https://doi.org/10.1101/2025.11.20.689604
43. Loya Y, Sakai K, Yamazato K *et al.* Coral bleaching: the winners and the losers. *Ecol Lett* 2001;**4**:122–31. https://doi.org/10.1046/j.1461-0248.2001.00203.x
44. Putnam HM, Stat M, Pochon X *et al.* Endosymbiotic flexibility associates with environmental sensitivity in scleractinian corals. *Proc R Soc B Biol Sci* 2012;**279**:4352–61. https://doi.org/10.1098/rspb.2012.1454
45. Starko S, Fifer JE, Claar DC *et al.* Marine heatwaves threaten cryptic coral diversity and erode associations among coevolving partners. *Sci Adv* 2023;**9**:eadf0954. doi:https://doi.org/10.1126/sciadv.adf0954
46. Grupstra CGB, Meyer-Kaiser KS, Bennett MJ *et al.* Holobiont traits shape climate change responses in cryptic coral lineages. *Global Change Biol* 2024;**30**:e17578. doi:https://doi.org/10.1111/gcb.17578
47. Johnston EC, Burgess SC. *Pocillopora tuahiniensis*: a new species of scleractinian coral (Scleractinia, Pocilloporidae) from French Polynesia. *Zootaxa* 2023;**5369**:117–24. https://doi.org/10.11646/zootaxa.5369.1.5
48. Tisthammer KH, Forsman ZH, Toonen RJ *et al.* Genetic structure is stronger across human-impacted habitats than among islands in the coral *Porites lobata*. *PeerJ* 2020 Feb 18;**8**:e8550. https://doi.org/10.7717/peerj.8550
49. Van Oppen MJH, McDonald BJ, Willis B *et al.* The evolutionary history of the coral genus *Acropora* (Scleractinia, Cnidaria) based on a mitochondrial and a nuclear marker: reticulation, incomplete lineage sorting, or morphological convergence? *Mol Biol*

Evol 2001;**18**:1315–29. https://doi.org/10.1093/oxfordjournals.molbev.a003916
50. Van Oppen MJH, Koolmees EM, Veron JEN. Patterns of evolution in the scleractinian coral genus *Montipora* (Acroporidae). *Mar Biol* 2004 Jan;**144**:9–18. https://doi.org/10.1007/s00227-003-1188-3
51. Vollmer SV, Palumbi SR. Hybridization and the evolution of reef coral diversity. *Science* 2002 Jun 14;**296**:2023–25. https://doi.org/10.1126/science.1069524
52. Siomi MC, Sato K, Pezic D *et al.* PIWI-interacting small RNAs: the vanguard of genome defence. *Nat Rev Mol Cell Biol* 2011;**12**:246–58. https://doi.org/10.1038/nrm3089
53. Admoni Y, Fridrich A, Weavers PK *et al.* miRNA-target complementarity in cnidarians resembles its counterpart in plants. *EMBO Rep* 2025;**26**:836–59. https://doi.org/10.1038/s44319-024-00350-z
54. Praher D, Zimmermann B, Dnyansagar R *et al.* Conservation and turnover of miRNAs and their highly complementary targets in early branching animals. *Proc R Soc B Biol Sci* 2021;**288**:20203169. https://doi.org/10.1098/rspb.2020.3169
55. Höck J, Meister G. The Argonaute protein family. *Genome Biol* 2008;**9**:210. https://doi.org/10.1186/gb-2008-9-2-210
56. Li C, Yoon B, Stefani G *et al.* Lipid kinase PIP5K1A regulates *let-7* microRNA biogenesis through interacting with nuclear export protein XPO5. *Nucleic Acids Res* 2023;**51**:9849–62. https://doi.org/10.1093/nar/gkad709
57. Juliano CE, Reich A, Liu N *et al.* PIWI proteins and PIWI-interacting RNAs function in *Hydra* somatic stem cells. *Proc Natl Acad Sci* 2014;**111**:337–42. https://doi.org/10.1073/pnas.1320965111
58. Praher D, Zimmermann B, Genikhovich G *et al.* Characterization of the piRNA pathway during development of the sea anemone *Nematostella vectensis*. *RNA Biol* 2017;**14**:1727–41. https://doi.org/10.1080/15476286.2017.1349048
59. Teefy BB, Siebert S, Cazet JF *et al. RNA* 2020;**26**:550–63. https://doi.org/10.1261/rna.072835.119
60. Carroll A, Harrison P, Adjeroud M. Sexual reproduction of *Acropora* reef corals at Moorea, French Polynesia. *Coral Reefs* 2006;**25**:93–97. https://doi.org/10.1007/s00338-005-0057-6
61. Becker D. *Learning from the Past: Environmental Impacts on Coral Performance, Reproduction, and Epigenetic Inheritance*. Kingston: University of Rhode Island, 2024. https://digitalcommons.uri.edu/oa_diss/1718 https://10.23860/diss-1718 (11 March 2025, date last accessed).
62. Adjeroud M, Penin L, Carroll A. Spatio-temporal heterogeneity in coral recruitment around Moorea, French Polynesia: implications for population maintenance. *J Exp Mar Biol Ecol* 2007;**341**:204–18. https://doi.org/10.1016/j.jembe.2006.10.048
63. Harnay P, Turner A, Burgess S *et al.* Daytime broadcast spawning in cryptic *Pocillopora* species at Mo'orea, French Polynesia. bioRxiv 2023. http://biorxiv.org/lookup/doi/10.1101/2023.10.05.558016
64. Liu Y, Ding W, Yu W *et al.* Long non-coding RNAs: biogenesis, functions, and clinical significance in gastric cancer. *Mol Ther Oncol* 2021;**23**:458–76. https://doi.org/10.1016/j.omto.2021.11.005
65. Liu Y, Liao X, Han T *et al.* Full-length transcriptome sequencing of the scleractinian coral *Montipora foliosa* reveals the gene expression profile of coral–Zooxanthellae holobiont. *Biology* 2021;**10**:1274. https://doi.org/10.3390/biology10121274
66. Guo Z, Liao X, Han T *et al.* Full-length transcriptomics reveal the gene expression profiles of reef-building coral *Pocillopora damicornis* and symbiont Zooxanthellae. *Diversity* 2021;**13**:543. https://doi.org/10.3390/d13110543
67. Reimer JD, Ono S, Takishita K *et al.* Molecular evidence suggesting species in the zoanthid genera Palythoa and Protopalythoa (Anthozoa: hexacorallia) are congeneric. *Zoolog Sci* 2006;**23**:87–94. https://doi.org/10.2108/zsj.23.87
68. Schmidt-Roach S, Miller KJ, Lundgren P *et al.* With eyes wide open: a revision of species within and closely related to the *Pocillopora damicornis* species complex (Scleractinia; Pocilloporidae) using morphology and genetics: *pocillopora* Species. *Zool J Linnean Soc* 2014;**170**:1–33. https://doi.org/10.1111/zoj.12092
69. Ma L, Bajic VB, Zhang Z. On the classification of long non-coding RNAs. *RNA Biol* 2013;**10**:924–33. https://doi.org/10.4161/rna.24604
70. Marchese FP, Raimondi I, Huarte M. The multidimensional mechanisms of long noncoding RNA function. *Genome Biol* 2017;**18**:206. https://doi.org/10.1186/s13059-017-1348-2
71. Grimson A, Srivastava M, Fahey B *et al.* Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. *Nature* 2008;**455**:1193–97. https://doi.org/10.1038/nature07415
72. Moran Y, Fredman D, Praher D *et al.* Cnidarian microRNAs frequently regulate targets by cleavage. *Genome Res* 2014;**24**:651–63. https://doi.org/10.1101/gr.162503.113
73. Raisch J, Darfeuille-Michaud A, Nguyen HTT. Role of microRNAs in the immune system, inflammation and cancer. *World J Gastroenterol* 2013;**19**:2985. https://doi.org/10.3748/wjg.v19.i20.2985
74. Baumgarten S, Bayer T, Aranda M *et al.* Integrating microRNA and mRNA expression profiling in *Symbiodinium microadriaticum*, a dinoflagellate symbiont of reef-building corals. *BMC Genomics* 2013;**14**:704. https://doi.org/10.1186/1471-2164-14-704
75. Baumgarten S, Cziesielski MJ, Thomas L *et al.* Evidence for miNA-mediated modulation of the host transcriptome in cnidarian–dinoflagellate symbiosis. *Mol Ecol* 2018;**27**:403–18. https://doi.org/10.1111/mec.14452
76. Stein CB, Genzor P, Mitra S *et al.* Decoding the 5′ nucleotide bias of PIWI-interacting RNAs. *Nat Commun* 2019;**10**:828. https://doi.org/10.1038/s41467-019-08803-z
77. Platt RN, Blanco-Berdugo L, DA R. Accurate transposable element annotation is vital when analyzing new genome assemblies. *Genome Biol Evolut* 2016;**8**:403–10. https://doi.org/10.1093/gbe/evw009
78. Lima L, Sinaimeri B, Sacomoto G *et al.* Playing hide and seek with repeats in local and global de novo transcriptome assembly of short RNA-seq reads. *Algorit Mol Biol* 2017;**12**:2. https://doi.org/10.1186/s13015-017-0091-2
79. Krishna S, Nair A, Cheedipudi S *et al.* Deep sequencing reveals unique small RNA repertoire that is regulated during head regeneration in *Hydra magnipapillata*. *Nucleic Acids Res* 2013;**41**:599–616. https://doi.org/10.1093/nar/gks1020
80. Urbarova I, Patel H, Forêt S *et al.* Elucidating the small regulatory RNA repertoire of the sea anemone *Anemonia viridis* based on whole genome and small RNA sequencing. *Genome Biol Evol* 2018;**10**:410–26. https://doi.org/10.1093/gbe/evy003
81. piRNAs SM. master regulators of gene expression. *Cell Res* 2014;**24**:779–80. https://doi.org/10.1038/cr.2014.78
82. Van Wolfswinkel JC, Ketting RF. The role of small non-coding RNAs in genome stability and chromatin organization. *J Cell Sci* 2010;**123**:1825–39. https://doi.org/10.1242/jcs.061713
83. Moyano M, Stefani G. piRNA involvement in genome stability and human cancer. *J Hematol Oncol* 2015;**8**:38. https://doi.org/10.1186/s13045-015-0133-5
84. Costa FF. Non-coding RNAs, epigenetics and complexity. *Gene* 2008;**410**:9–17. https://doi.org/10.1016/j.gene.2007.12.008
85. Gibney ER, Nolan CM. Epigenetics and gene expression. *Heredity* 2010;**105**:4–13. https://doi.org/10.1038/hdy.2010.54

86. Edmunds PJ, Leichter JJ. Spatial scale-dependent vertical zonation of coral reef community structure in French Polynesia. *Ecosphere* 2016;**7**:e01342. https://doi.org/10.1002/ecs2.1342
87. Moorea Coral Reef LTER, Edmunds PJ. MCR LTER: coral reef: long-term population and community dynamics: corals, ongoing since 2005. Environmental Data Initiative, 2025. https://portal.edirepository.org/nis/mapbrowse?scope=knb-lter-mcr&identifier=4 (20 December 2025, date last accessed).
88. Vercelloni J, Kayal M, Chancerelle Y *et al.* Exposure, vulnerability, and resiliency of French Polynesian coral reefs to environmental disturbances. *Sci Rep* 2019;**9**:1027. https://doi.org/10.1038/s41598-018-38228-5
89. Forsman ZH, Barshis DJ, Hunter CL *et al.* Shape-shifting corals: molecular markers show morphology is evolutionarily plastic in Porites. *BMC Evol Biol* 2009;**9**:45. https://doi.org/10.1186/1471-2148-9-45
90. Burgess SC, Johnston EC, Wyatt ASJ *et al.* Response diversity in corals: hidden differences in bleaching mortality among cryptic *Pocillopora* species. *Ecology* 2021;**102**:e03324. https://doi.org/10.1002/ecy.3324
91. Burgess SC, Turner AM, Johnston EC. Niche breadth and divergence in sympatric cryptic coral species (*Pocillopora* spp.) across habitats within reefs and among algal symbionts. *Evol Appl* 2024;**17**:e13762. https://doi.org/10.1111/eva.13762
92. Johnston EC, Forsman ZH, Toonen RJ. A simple molecular technique for distinguishing species reveals frequent misidentification of Hawaiian corals in the genus *Pocillopora*. *PeerJ* 2018 Feb 8;**6**:e4355. https://doi.org/10.7717/peerj.4355
93. Altschul SF, Gish W, Miller W *et al.* Basic local alignment search tool. *J Mol Biol* 1990;**215**:403–10. doi:https://doi.org/10.1016/S0022-2836(05)80360-2
94. Fuller ZL, Mocellin VJL, Morris LA *et al.* Population genetics of the coral *Acropora millepora* : toward genomic prediction of bleaching. *Science* 2020;**369**:eaba4674. https://doi.org/10.1126/science.aba4674
95. Richards ZT, Miller DJ, Wallace CC. Molecular phylogenetics of geographically restricted Acropora species: implications for threatened species conservation. *Mol Phylogenet Evol* 2013;**69**:837–51. https://doi.org/10.1016/j.ympev.2013.06.020
96. Wallace CC, Willis BL. Systematics of the coral genus *Acropora* : implications of new biological findings for species concepts. *Annu Rev Ecol Syst* 1994;**25**:237–62. https://doi.org/10.1146/annurev.es.25.110194.001321
97. Willis BL, Babcock RC, Harrison PL. *et al.* Experimental hybridization and breeding incompatibilities within the mating systems of mass spawning reef corals. *Coral Reefs* 1997;**16**:S53–65. doi:https://doi.org/10.1007/s003380050242
98. Stephens TG, Lee J, Yoon HS *et al.* High-quality genome assembles from key Hawaiian coral species. *Curr Biol* 2022;**21**:R676–77. https://doi.org/10.1016/j.cub.2011.06.013
99. Buitrago-López C, Mariappan KG, Cárdenas A *et al.* The genome of the cauliflower coral *Pocillopora verrucosa*. *Genome Biol Evolut* 2020;**12**:1911–17. https://doi.org/10.1093/gbe/evaa184
100. Noel B, Denoeud F, Rouan A *et al.* Pervasive tandem duplications and convergent evolution shape coral genomes. *Genome Biol* 2023;**24**:123. https://doi.org/10.1186/s13059-023-02960-7
101. Cowen LJ, Putnam HM. Bioinformatics of corals: investigating heterogeneous omics data from coral holobionts for insight into reef health and resilience. *Ann Rev Biomed Data Sci* 2022;**5**:205–31. https://doi.org/10.1146/annurev-biodatasci-122120-030732
102. Andrews SR. FastQC: a quality control tool for high throughput sequence data. Babraham Bioinformatics, 2010. http://www.bioinformatics.babraham.ac.uk/projects/fastqc (7 March 2025, date last acceessed).
103. Kim D, Paggi JM, Park C *et al.* Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* 2019;**37**:907–15. https://doi.org/10.1038/s41587-019-0201-4
104. Pertea G, Pertea M. GFF utilities: gffRead and GffCompare. *F1000Research* 2020;**9**, 304. https://doi.org/10.12688/f1000research.23297.2
105. Kang YJ, Yang DC, Kong L *et al.* CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. *Nucleic Acids Res* 2017;**45**:W12–16. https://doi.org/10.1093/nar/gkx428
106. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 2014;**30**:923–30. https://doi.org/10.1093/bioinformatics/btt656
107. Li A, Zhang J, Zhou Z. PLEK: a tool for predicting long noncoding RNAs and messenger RNAs based on an improved k-mer scheme. *BMC Bioinf* 2014;**15**:311. https://doi.org/10.1186/1471-2105-15-311
108. Finn RD, Bateman A, Clements J *et al.* Pfam: the protein families database. *Nucleic Acids Res* 2014;**42**:D222–30. https://doi.org/10.1093/nar/gkt1223
109. Finn RD, Clements J, Eddy SR. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res* 2011;**39**:W29–37. https://doi.org/10.1093/nar/gkr367
110. Chen S, Zhou Y, Chen Y *et al.* fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 2018;**34**:i884–90. https://doi.org/10.1093/bioinformatics/bty560
111. Kozomara A, Birgaoanu M, miRBase G-JS.: from microRNA sequences to function. *Nucleic Acids Res* 2019;**47**:D155–62. https://doi.org/10.1093/nar/gky1141
112. ShortStack AMJ.: Comprehensive annotation and quantification of small RNA genes. *RNA* 2013;**19**:740–51. https://doi.org/10.1261/rna.035279.112
113. Shahid S, Axtell MJ. Identification and annotation of small RNA genes using ShortStack. *Methods* 2014;**67**:20–27. https://doi.org/10.1016/j.ymeth.2013.10.004
114. Rosenkranz D, Han CT, Roovers EF *et al.* Piwi proteins and piRNAs in mammalian oocytes and early embryos: from sample to sequence. *Genomics Data* 2015;**5**:309–13. https://doi.org/10.1016/j.gdata.2015.06.026
115. Rosenkranz D, Zischler H. proTRAC—a software for probabilistic piRNA cluster detection, visualization and analysis. *BMC Bioinf* 2012;**13**:5. https://doi.org/10.1186/1471-2105-13-5
116. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010;**26**:841–42. https://doi.org/10.1093/bioinformatics/btq033
117. Chan PP, Lin BY, Mak AJ *et al.* tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. *Nucleic Acids Res* 2021;**49**:9077–96. https://doi.org/10.1093/nar/gkab688
118. Alexa A, topGO RJ Enrichment analysis for gene ontology. Gene Ontology Resource, 2024. https://bioconductor.org/packages/topGO (30 April 2025, date last accessed).

119. Ewels P, Magnusson M, Lundin S *et al.* MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 2016;**32**:3047–48. https://doi.org/10.1093/bioinformatics/btw354
120. Pertea M, Pertea GM, Antonescu CM *et al.* StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol* 2015;**33**:290–95. https://doi.org/10.1038/nbt.3122
121. Kopylova E, Noé L, Touzet H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* 2012;**28**:3211–17. https://doi.org/10.1093/bioinformatics/bts611
122. Jehn J, Gerbert D Jehn J, Gerbert D, Pipilescu F *et al.* PIWI genes and piRNAs are ubiquitously expressed in mollusks and show patterns of lineage-specific adaptation. *Commun Biol* 2018;**1**. https://doi.org/10.1038/s42003-018-0141-4.

Received: 31 December 2025. **Revised:** 7 May 2026. **Accepted:** 1 June 2026