Whole genome assembly and annotation of the endangered Caribbean coral *Acropora cervicornis*

Jason D. Selwyn (iD),* Steven V. Vollmer (iD)

Department of Marine and Environmental Sciences, Northeastern University, Nahant, MA 01908, USA

*Corresponding author: Department of Marine and Environmental Sciences, Northeastern University, 430 Nahant Road, Nahant, MA 01908, USA.
Email: j.selwyn@northeastern.edu

Coral species in the genus *Acropora* are key ecological components of coral reefs worldwide and represent the most diverse genus of scleractinian corals. While key species of Indo-Pacific *Acropora* have annotated genomes, no annotated genome has been published for either of the two species of Caribbean *Acropora*. Here we present the first fully annotated genome of the endangered Caribbean staghorn coral, *Acropora cervicornis*. We assembled and annotated this genome using high-fidelity nanopore long-read sequencing with gene annotations validated with mRNA sequencing. The assembled genome size is 318 Mb, with 28,059 validated genes. Comparative genomic analyses with other *Acropora* revealed unique features in *A. cervicornis*, including contractions in immune pathways and expansions in signaling pathways. Phylogenetic analysis confirms previous findings showing that *A. cervicornis* diverged from Indo-Pacific relatives around 41 million years ago, with the closure of the western Tethys Sea, prior to the primary radiation of Indo-Pacific *Acropora*. This new *A. cervicornis* genome enriches our understanding of the speciose *Acropora* and addresses evolutionary inquiries concerning speciation and hybridization in this diverse clade.

Keywords: *Acropora cervicornis*; acroporidae; scleractinia; de novo assembly; comparative genomics; time-calibrated phylogeny

Introduction

The *Acropora* are one of the most speciose and important genera of reef-building scleractinian corals globally (Wallace 1999). The genus *Acropora* are divided into multiple speciose Indo-Pacific clades and a single depauperate Caribbean clade (Wallace 1999). The two sister species of Caribbean *Acropora*—the Staghorn coral *A. cervicornis* and Elkhorn coral *A. palmata*—and their hybrid—called *A. prolifera* (Vollmer and Palumbi 2002) are thought to have diverged from the Indo-Pacific *Acropora* during the late Eocene after the closure of the western Tethys Sea prior to the rapid diversification in the Indo-Pacific *Acropora* (Wallace 1999; Wallace and Portell 2022). To date, all 16 published de novo assembled and annotated *Acropora* genomes are of Indo-Pacific species (Shinzato et al. 2011, Ying et al. 2019; Fuller et al. 2020; Shinzato et al. 2021; López-Nandam et al. 2023).

Acropora, like all corals, are severely threatened by anthropogenic climate change leading to elevated water temperatures that can cause acute bleaching and subsequently death (Hughes et al. 2018). The Caribbean *Acropora* are also experiencing a secondary range-wide pressure in the form of White Band Disease, which has resulted in ~95% population losses of both species Caribbean wide and is the direct cause of their listing on the Endangered Species List (Aronson and Precht 2001; National Marine Fisheries Service 2006). Because these two species are such important foundational species in the Caribbean reef ecosystem, these losses have likely had tremendous unknown effects on higher order taxa which depend on *Acropora* dominant reefs for survival.

Here we present the first fully annotated genome for the endangered Caribbean staghorn coral *Acropora cervicornis*, importantly representing the first Caribbean species of this diverse clade. This genome was assembled using a combination of long-read nanopore and short-read shotgun sequences and annotated and validated using mRNA sequencing. This reference genome will accelerate genomic research on this endangered coral and address fundamental evolutionary questions about speciation and hybridization in the speciose Acroporids.

Materials and methods
Sample collection and sequencing

High molecular weight genomic DNA was extracted in June 2021 from adult tissue of the K2 genotype maintained in the Coral Restoration Foundation (CRF) Key Largo, Florida nursery. Three libraries were prepared using Oxford Nanopore Technologies (ONT) kit SQK-LSK112. Two libraries were not size selected while the third included 20+kb PippenPrep size-selection. All ONT prepared libraries were sequenced separately on three Minion flow cells (FLO-MIN112). High-quality base-calling was performed using Guppy v6.1.7 (ONT).

Four additional Illumina PCR-free shotgun libraries were constructed using the Discovar protocol to produce libraries with fragments between 400 and 600 bp (Love et al. 2016). KAPA PCR-free library kits were leveraged with the addition of a second round of 0.7× Agencourt AmPure XP SPRI bead cleanup post-adapter ligation. Libraries were multiplexed and sequenced on a single rapid-run HiSeq 2500 flowcell with 250 bp paired-end sequencing. Additionally, a library of paired 150 bp reads was prepared using the Illumina DNA Prep kit and sequenced as part of a NovaSeq S4 run.

To acquire transcriptome data, mRNA sequence data was obtained for 48 individuals (including the K2 genotype) using NEBs unidirectional mRNA library preparations sequenced on an Illumina NEXTSEQ 550 platform and combined with previously published RNA sequencing data from 38 additional *A. cervicornis* (PRJNA222758: Libro and Vollmer 2016; PRJNA423227: Parkinson et al. 2018).

Sequence quality control

Nanopore long-reads (DNA) were quality controlled using PORECHOP (v 0.2.3_seqan2.1.1, https://github.com/rrwick/Porechop) to remove adapter sequences and then quality trimmed into longer assembly reads (minimum average quality 3, minimum length 1,000 bp) and shorter polishing reads (minimum average quality 5, minimum length 500 bp) using NANOFILT (De Coster et al. 2018). Illumina sequenced short-reads (DNA and RNA) were quality controlled initially using FASTP (Chen et al. 2018) to remove adapters and barcodes, filter low quality sequences (PHRED < 30), trim sequences shorter than 140 bp, and PCR artifacts. Contaminants were removed with FASTQ_SCREEN (Wingett and Andrews 2018) by mapping reads against a suite of potential contaminant genomes (e.g. human, viral, and bacterial) as well the 13 available genomes of *Symbiodiniaceae* (Supplementary Table 1) and removing reads which had hits to any potential contaminant genome.

Genome assembly

We estimated the genome size using a k-mer counting approach implemented in JELLYFISH based on the quality-controlled Illumina short reads (Marçais and Kingsford 2011). An initial genome assembly of the ONT sequenced reads was built using FLYE (Kolmogorov et al. 2019) with the nano-hq parameter setting to fit with the chemistry and base-calling method of the sequencing. After initial assembly of the raw nanopore reads, duplicated sequences were removed using purge_dups (Guan et al. 2020; Guiglielmoni et al. 2021). We polished the genome using two rounds of long-read polishing with RACON (Vaser et al. 2017) followed by a round of polishing with MEDAKA v1.7.2 (ONT) and then two rounds of polishing with paired-end Illumina sequences using PILON (Walker et al. 2014). To ensure a final genome assembly of the highest quality and contiguity, we corrected misassembly errors with short reads using MEC (Wu et al. 2020) and remove the mitochondrial genome and other potential genomic contaminants using BLOBTOOLS (Laetsch and Blaxter 2017). This was followed by misassembly correction using the long-read data (Coombe et al. 2021), scaffolding (Coombe et al. 2021), gap closing (Xu et al. 2020), and a final round of long-read polishing with RACON and short-read polishing with PILON (Walker et al. 2014; Vaser et al. 2017). Finally, all contigs shorter than 1,000 bp and/or with no gene annotations (see below) were removed as they did not contain any BUSCOs.

Transcriptome assembly

A preliminary transcriptome was assembled using a genome-guided assembly in TRINITY (Grabherr et al. 2011). mRNA sequencing reads were splice-aware mapped to the genome using GSNAP (Wu et al. 2016). TRINITY was used to perform a genome guided assembly of the transcriptome with an assumed max intron length of 100,000 bp.

Genome annotation

Genome annotation was performed using MAKER (Cantarel et al. 2008; Holt and Yandell 2011). Repetitive elements were identified and masked using REPEATMODELER (Flynn et al. 2020) and REPEATMASKER (Smit et al. 2015). Evidence-based gene annotation was performed using the assembled transcriptome and proteins identified in either all Acroporids in the UniProt database (The UniProt Consortium 2023) or the closest reference proteomes in UniRef from *Stylophora pistillata* (Voolstra et al. 2017), *Pocillopora damicornis* (Cunning et al. 2018), *Actinia tenebrosa* (Surm et al. 2019), and *Nematostella vectensis* (Putnam et al. 2007). This initial round of annotation was used to train the ab initio gene identification models of Augustus (Stanke et al. 2006), Snap (Korf 2004), and Genemark-ES (Lomsadze et al. 2005). After the initial round of annotation based solely on protein and RNA evidence, we performed four subsequent rounds of annotation with the results of the previous round being used to train the ab initio gene predictors run in the subsequent round of annotation. Genes were functionally annotated using EnTAP (Hart et al. 2020) and Interproscan (Zdobnov and Apweiler 2001), and formatted the annotations for NCBI using GAG and ANNIE (Tate et al. 2014; Geib et al. 2018).

Mitochondrial assembly

The mitochondrial genome was assembled using the quality-controlled Illumina short-reads using MITOZ (Meng et al. 2019). Briefly, this was done by first assembling a subset of reads into initial contigs which are then identified as mitogenome sequences using a profile Hidden Markov Model (Wheeler and Eddy 2013; Xie et al. 2014; Nurk et al. 2017). Contigs were then annotated to find the 13 protein-coding mitochondrial genes along with tRNAs and rRNAs with any contigs not containing any annotations removed (Birney et al. 2004; Gertz et al. 2006; Li and Durbin 2009; Jühling et al. 2012; Nawrocki and Eddy 2013). Finally, retained contigs were assembled and circularized into the complete mitochondrial genome and visualized (Gertz et al. 2006; Krzywinski et al. 2009; Meng et al. 2019).

Phylogenetic analysis

Sixteen published *Acropora* genomes with structural genome annotations along with two *Montipora* species, *Montipora capitata* and *Montipora efflorescens*; two Pocilloporids, *P. damicornis* and *S. pistillata*; and three anemones, *N. vectensis*, *A. tenebrosa*, and *Exaiptasia diaphana* were downloaded from NCBI (Table 1). Protein sequences for all annotated genes were extracted and clustered into orthogroups derived from a single gene in the last common ancestor of the group using ORTHOFINDER (Emms and Kelly 2019). Orthogroups were used to infer rooted gene and species trees to develop a phylogenetic hypothesis for the group (Emms and Kelly 2017, 2018). Specifically, we used the STAG (Emms and Kelly 2018) algorithm to infer the species tree from 7,110 multicopy gene trees which had all species present, each created using DendroBLAST (Kelly and Maini 2013), this species tree was then rooted using the STRIDE algorithm (Emms and Kelly 2017). To infer divergence times, we time-calibrated the species tree using least-squares dating and 1,000 bootstraps to estimate divergence time confidence intervals (To et al. 2016). Ancestral dates were gathered from the Fossilworks database (Supplementary Table 2; Behrensmeyer and Turner 2013).

Comparative genomics

Structural gene annotations for all species included in the phylogenetic hypothesis were functionally annotated using BLAST (Altschul et al. 1990; Camacho et al. 2009) against the Swiss-Prot curated portion of the UniProt database (The UniProt Consortium 2023). To identify KEGG orthologs for all orthogroups, we matched KEGG gene annotations to orthogroups across

Table 1. National Center for Biotechnology Information accession numbers and citations for coral genomes with structural gene annotations used to build phylogeny and for comparative genomic analysis.

Species	Genome size (Mb)	Number of genes	Accession number	Citation
Acropora acuminata	394.7	26,151	GCA_014633975	Shinzato et al. (2021)
Acropora awi	428.8	26,801	GCA_014634005	Shinzato et al. (2021)
Acropora cervicornis	308	28,059	GCA_032359415	(This Study)
Acropora cytherea	426.3	27,327	GCA_014634045	Shinzato et al. (2021)
Acropora digitifera	415.8	25,278	GCA_014634065	Shinzato et al. (2021)
Acropora echinata	401.5	26,170	GCA_014634105	Shinzato et al. (2021)
Acropora florida	442.8	27,573	GCA_014634605	Shinzato et al. (2021)
Acropora gemmifera	401	26,269	GCA_014634125	Shinzato et al. (2021)
Acropora hyacinthus	447.2	27,215	GCA_014634145	Shinzato et al. (2021)
Acropora intermedia	416.9	26,982	GCA_014634585	Shinzato et al. (2021)
Acropora microphthalma	383.9	26,384	GCA_014634165	Shinzato et al. (2021)
Acropora millepora	475.4	41,860	GCA_013753865	Fuller et al. (2020)
Acropora muricata	420.7	27,409	GCA_014634545	Shinzato et al. (2021)
Acropora nasuta	416.4	27,379	GCA_014634205	Shinzato et al. (2021)
Acropora selago	392.9	27,036	GCA_014634525	Shinzato et al. (2021)
Acropora tenuis	403.1	27,236	GCA_014633955	Shinzato et al. (2021)
Acropora yongei	438	27,452	GCA_014634225	Shinzato et al. (2021)
Montipora cactus	652.7	29,158	GCA_014634245	Shinzato et al. (2021)
Montipora efflorescens	643.3	29,424	GCA_014634505	Shinzato et al. (2021)
Nematostella vectensis	356.6	34,311	GCF_000209225	Putnam et al. (2007)
Pocillopora damicornis	234.3	25,183	GCF_003704095	Cunning et al. (2018)
Stylophora pistillata	397.6	33,252	GCF_002571385	Voolstra et al. (2017)
Actinia tenebrosa	238.2	27,037	GCF_009602425	Surm et al. (2019)
Exaiptasia diaphana	256.1	27,753	GCF_001417965	Baumgarten et al. (2015)

species and found the consensus KEGG ortholog across all species in which the orthogroup was found. KEGG ortholog membership within KEGG pathways was identified using KEGGREST (Tenenbaum and Volkening 2023).

We used a logistic regression model to test if there are differences in the percentage of orthogroups with KEGG annotations across taxa. Post-hoc tests were used to compare the rate of annotation in *A. cervicornis* to other *Acropora* sp. and make pairwise comparisons between genera. To identify systematic differences in the distribution of genes in pathways within the *Acropora*, we used χ^2 tests followed by a post-hoc analysis with FDR correction to identify the species and KEGG pathways with significantly more or less gene copies (i.e. orthogroups) than would be expected compared to the other *Acropora* (Beasley and Schumacker 1995; Benjamini and Hochberg 1995).

The time-calibrated phylogeny along with the number of genes found within orthogroups for each species was used to estimate the rate of gene family evolution across the phylogeny using CAFE5 (Hahn et al. 2005; Mendes et al. 2020) and identify KEGG orthologs which exhibit significant expansions/contractions at each node of the phylogeny. In the CAFE5 analysis we assumed a single rate of evolution across all gene families using the error rate estimating model and Poisson prior distribution (De Bie et al. 2006; Han et al. 2013). We then performed an overrepresentation analysis using Fisher's exact test to determine if any pathways in *A. cervicornis* showed significant expansions/contractions. All statistical analyses were performed using R v4.2.1 (R Core Team 2022).

Results and discussion
### Genome/transcriptome assembly and annotation
Nanopore sequencing resulted in ~6.6 million reads containing 15.5 Gb with an N50 of 5,072 bp with 3.3 million high-quality reads after filtering used in the initial long-read assembly containing 13.3 Gb of DNA with an N50 of 6,078 bp and a polishing set of 4.5 million reads containing 13.3 Gb of DNA with an N50 of

Table 2. *A. cervicornis* assembly statistics.

Cumulative scaffold length (bp)	307.4
Number of scaffolds	398
Number of contigs	415
G + C content (%)	38.95
Number of Ns/100 kbp	37.23
Largest scaffold (Mb)	8.337
Scaffold N50 (Mb)	2.8
Scaffold L50	35
Largest contig (Mb)	8.337
Contig N50 (Mb)	2.7
Contig L50 (Mb)	36
BUSCO complete single-copy	878
BUSCO complete multicopy	4
BUSCO fragmented	34
BUSCO missing	38

5,269 bp. 92 million paired-end Illumina reads (totally 40 Gb) were retained after filtering and decontamination for x polishing. mRNA sequencing from 86 corals totaled 1.4 billion single-end reads and 174 Gb of mRNA sequencing data for gene annotation.

Using the filtered paired-end short-read sequences, we estimated the genome size of *A. cervicornis* to be ~318 Mb, somewhat smaller than the Pacific Acroporids (384–475 Mb; Shinzato et al. 2011, 2021; Ying et al. 2019; Fuller et al. 2020), and obtained ~42× genomic coverage of the long-read nanopore sequences for the genomic assembly. The initial genome assembly measured 328 Mb with 3,014 contigs (N50 = 0.94 Mb; L50 = 99) with the longest contig being 4.9 Mb and a BUSCO completeness of 92.8% (1.2% duplicated, 3.7% fragmented). After genomic post-processing, polishing, and scaffolding, the final genome assembly measured 307 Mb (96.5% the estimated genome size, Table 2) with 398 scaffolded contigs (N50 = 2.8 Mb; L50 = 35) with the longest scaffold measuring 8.3 Mb and a BUSCO completeness of 92.4% (0.4% duplicated, 3.6% fragmented, Fig. 1). The pooled transcriptome contained 374,749 transcripts with a total N50 of 3,659 and longest isoform N50 of 1,483 with a BUSCO completeness of 93.4% (78% duplicated, 3.5% fragmented). The high degree of transcriptome
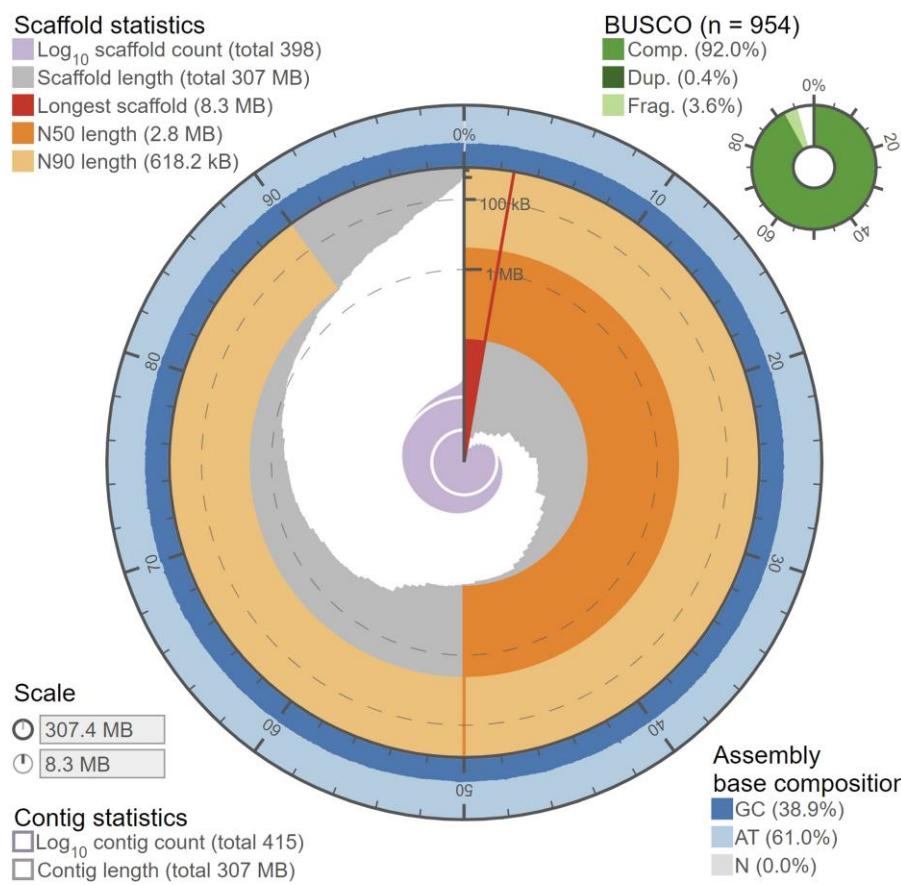
Fig. 1. Genome snail plot showing genome contiguity and completeness statistics. For an interactive version of the figure see: https://jdselwyn.github.io/assembly-stats/. Created using Challis (2017).

duplication in the BUSCO value is likely due to the pooling of individuals to create the transcriptome. Similar to other *Acropora* genomes (Shinzato et al. 2021), *A. cervicornis* contained 39% interspersed repeats (Supplementary Table 3) with the most common identifiable class of repeat being short interspersed nuclear elements, though the plurality (17%) of repeats in the genome were unable to be classified and may be taxon specific (Shinzato et al. 2021). *A. cervicornis* had 28,059 validated genes (Table 3), intermediate among other Acroporids (25,278–41,860; Shinzato et al. 2011; Ying et al. 2019; Fuller et al. 2020; Shinzato et al. 2021), with 53% having a Swiss-Prot annotation at an e-value $< 10^{-6}$ and a BUSCO completeness of 81.6% (2.8% duplicated, 10% fragmented).

Mitochondrial genome

The *A. cervicornis* mitochondrial genome was assembled into a single circular structure containing 18,259 bp, 13 protein-coding genes, two tRNA genes (tRNA-Met and tRNA-Trp), and two rRNA genes (16 and 12 s; Fig. 2, GenBank accession number: OQ772303). Like all *Acropora*, the *A. cervicornis* mitochondrial genome has a particular dearth of tRNA coding sequences when compared to other metazoans, likely due to this diversification occurring after the split between cnidarians and metazoans (van Oppen et al. 1999; van Oppen, Catmull, *et al.* 2002; Liu et al. 2015; Tian and Niu 2017; Colin et al. 2021). The gene order found in *A. cervicornis* is the same as that found in other Acroporids, including ND5 being split into two portions with a large intron containing all genes except the two tRNA coding genes, the large ribosomal

Table 3. *A. cervicornis* annotation statistics.

|	
Number of genes	28,059
Cumulative length of CDSs (bp)	36,143,406
Median gene length (bp)	4,072
Median CDS length (bp)	119
Median exon length (bp)	124
Median intron length (bp)	569
Number of intronless genes	4,591
Median number of exons per gene	4
Median number of exons per multiexon gene	4
BUSCO complete single-copy	752
BUSCO complete multicopy	27
BUSCO fragmented	95
BUSCO missing	80

subunit rRNA, ATP8, and COI (Figure Y; van Oppen, Catmull, *et al.* 2002; Liu et al. 2015; Zhang et al. 2016; Colin et al. 2021).

Comparative genomics

The 28,059 genes found in *A. cervicornis* belong to 15,191 distinct orthogroups of which 54.4% had a KEGG annotation. The percent of orthogroups with KEGG annotations varied significantly by species ($\chi^2_{(23)} = 595.8$, $P < 0.0001$) due primarily to differences between genera (*Acropora* vs *Montipora*, $Z = 12.9$, $P < 0.0001$) and higher-order taxonomic comparisons outside of the *Acropora* (*Acropora* vs outgroups, $Z = -17.6$, $P < 0.0001$) rather than within *Acropora* which all show a similar degree of annotation (*A. cervicornis* vs other *Acropora*, $Z = -0.87$, $P = 0.38$). The number of orthogroups per KEGG pathway were similar for all *Acropora* sp. ($\chi^2_{(4,976)} = 478.1$, $P = 1$)
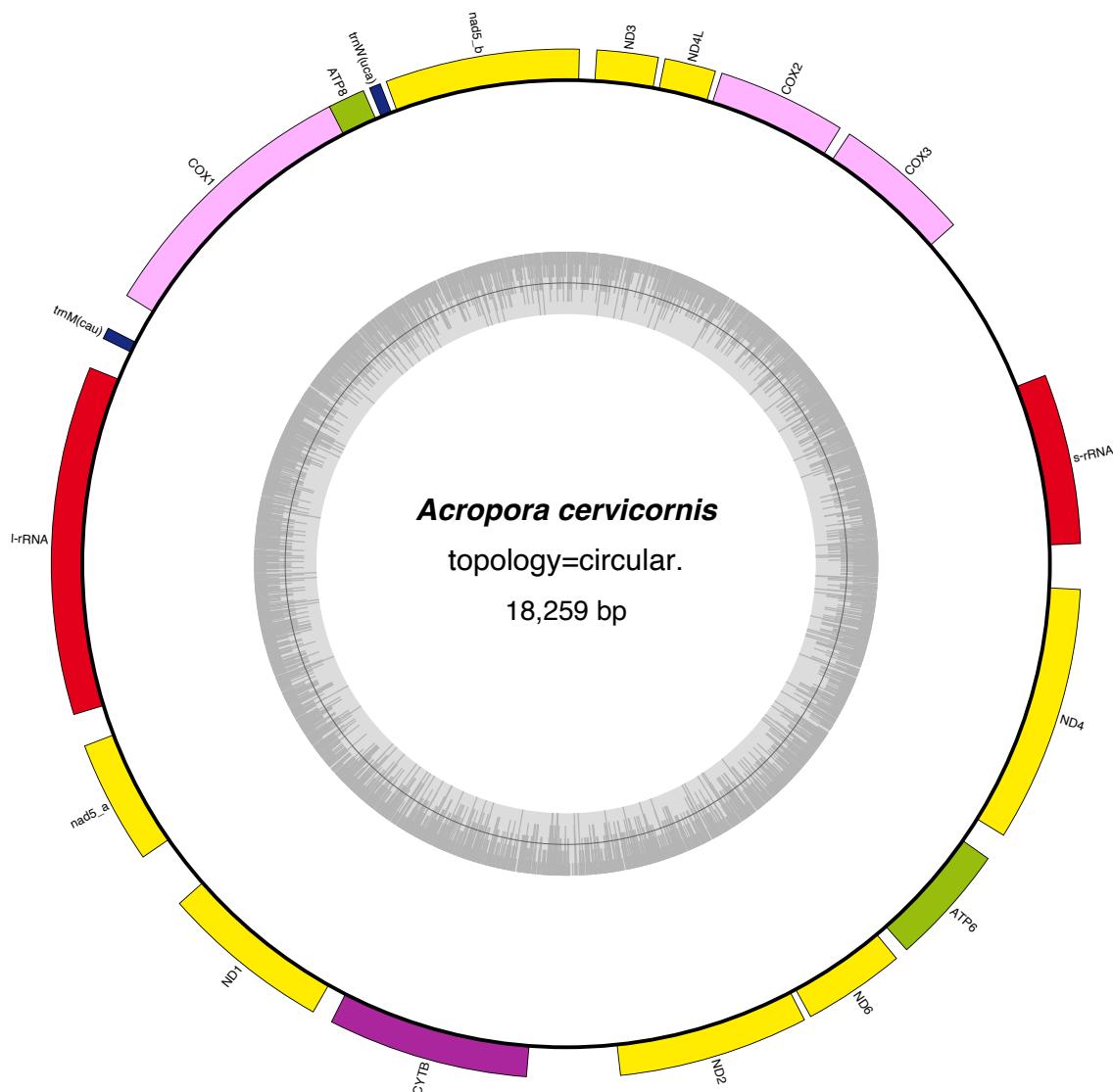
Fig. 2. Mitochondrial genome plot showing the relative locations and orientations of genes on the mitochondria with gene groups color coded by type. Plot made using Lohse et al. 2013).

indicating that no Acropora genomes had broad gains or losses of genes. Amongst five major KEGG categories, 25.7% of the orthogroups were found in pathways involved in organismal systems and 24.8% were involved in environmental information processing. The remaining orthogroups were split between cellular processes (18.8%), metabolism-related processes (16.7%), and genetic information processing (14.1%). *A. cervicornis* possessed 135 unique orthogroups, more than the other *Acropora* species (17–70), possibly as a result of the separate evolutionary history in the Caribbean. An alternative hypothesis for this discrepancy could be that the ab initio gene models for most other *Acropora* were trained on *A. digitifera* rather than being trained for each species independently (Shinzato et al. 2021).

Six gene pathways were significantly overrepresented in *A. cervicornis* among the significantly expanded and contracted KEGG orthologs (Supplementary Table 4). Two immune system pathways—NOD-like receptor signaling (map04621, OR = 19.7, $P < 0.001$, $p_{adj} < 0.001$) and Neutrophil extracellular trap formation (map04613, OR = 16.9, $P < 0.001$, $p_{adj} = 0.018$)—and the related cellular growth and death pathway Necroptosis (map04217, OR = 17.6, $P < 0.001$, $p_{adj} = 0.018$) were significantly overrepresented due to one expansion and three contractions in NACHT, LRR, and PYD domain-containing proteins, three contractions in histone proteins (H2A, H3, and H4), and one expansion in a cation channel protein (TRPM7). In corals, NOD-like receptor signaling pathway has been found to be an integral part of the innate immune system
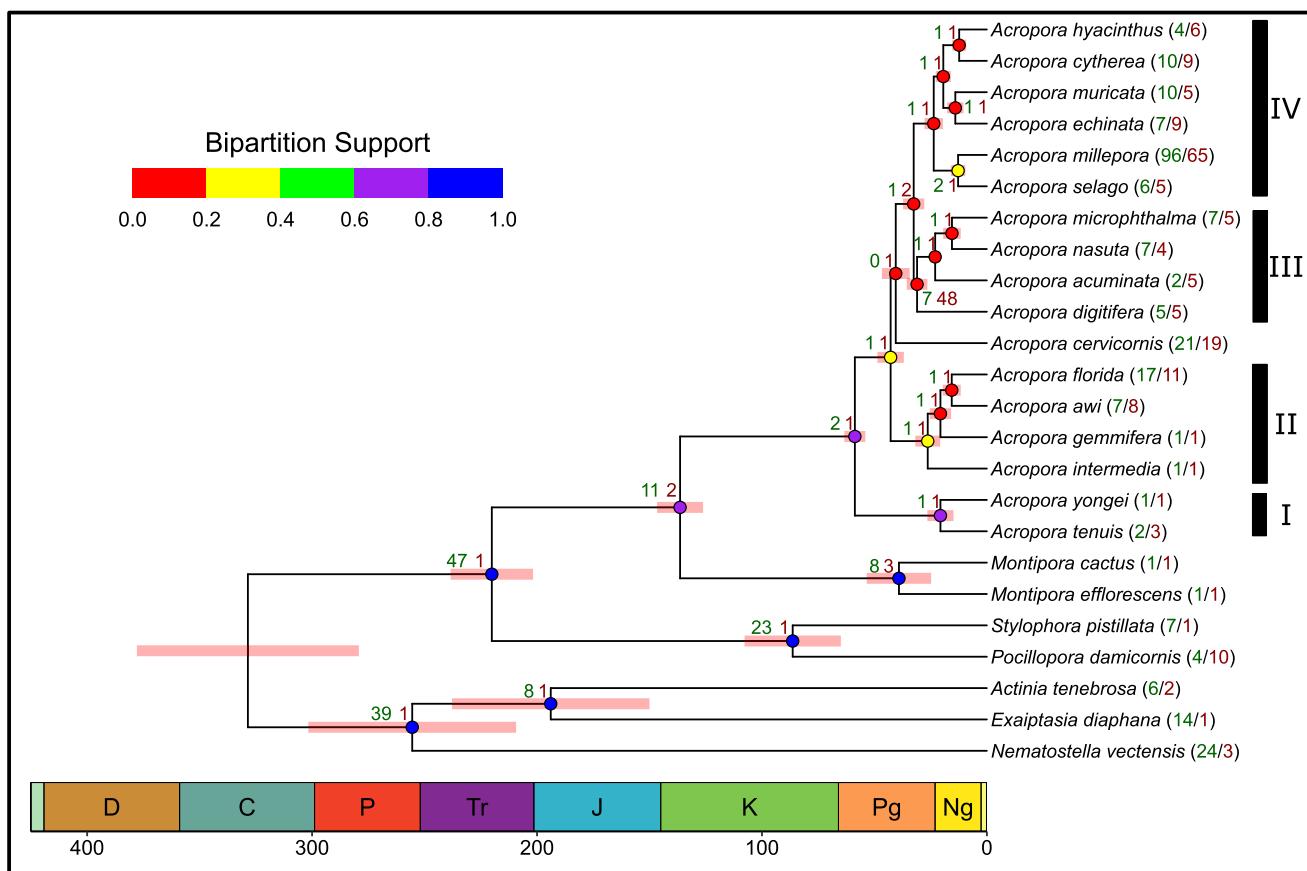
Fig. 3. Species phylogeny based on identified orthogroups. Node color indicates the proportion of single-locus gene trees supporting each bipartition (i.e. gene concordance factor, a more stringent metric of node support than bootstrap support, Emms and Kelly 2018; Minh et al. 2020). Red bars at each node indicate the 95% confidence interval of the estimated time of divergence marked in geologic periods on the x axis. Green/red node numbers indicate the number of KEGG orthologs which significantly expanded/contracted at each divergence point. Clade groupings from Shizato et al. (2021).

(Hamada et al. 2013) and histone phosphorylation has been shown to be a key response to nutrient stress and regulating the dinoflagellate symbionts (Rodriguez-Casariego et al. 2018).

The second major group of overrepresented pathways are signaling pathways, particularly the calcium signaling pathway (map04020, OR = 9.93, $P < 0.001$, $p_{adj} = 0.025$) and the neuroactive ligand–receptor interaction (map04080, OR = 39.8, $P < 0.001$, $p_{adj} < 0.001$). While the immune-related pathways were predominantly gene contractions in these pathways, all KEGG orthologs showed expansions (Supplementary Table 4). The calcium signaling and neuroactive ligand–receptor pathways have been found to be important in regulation reproduction (Hilton et al. 2012; Rosenberg et al. 2017), nematocyst regulation (Russell and Watson 1995), and biomineralization (Reyes-Bermudez et al. 2009). Further, these pathways appear to be differentially expressed between the growth tips and bases in the Caribbean *Acropora* (Hemond et al. 2014).

The final KEGG pathway overrepresented among the rapidly evolving KEGG orthologs is the Taurine and hypotaurine metabolism pathway (map00430, OR = 91.6, $P < 0.001$, $p_{adj} = 0.04$). This pathway was driven by the loss of a single KEGG orthogroup, cysteine dioxygenase (−2, K00456). Unlike other scleractinian corals, Indo-Pacific *Acropora* lack the cystathionine beta-synthase gene involved in cysteine biosynthesis (Shinzato et al. 2011) suggesting that *Acroporid* corals must rely on their algal symbionts to supply this vital amino acid (Shinzato et al. 2014). However, an alternative cysteine biosynthesis pathway has recently been discovered in *Acropora loripes* suggesting *Acropora* corals can natively

synthesize cysteine through this alternate pathway (Salazar et al. 2022). Similar to the other studied *Acropora*, *A. cervicornis* lacks genes coding for cystathionine beta-synthase but does possess the genes required for the alternate cysteine biosynthesis pathway (Salazar et al. 2022). The loss of cysteine dioxygenase in *A. cervicornis* may suggest that the alternative pathway does not fully compensate the biosynthesis of cysteine resulting in a lack of excess cysteine needing to be metabolized.

Phylogeny

A. cervicornis is estimated to have diverged from the other *Acropora* 41 million years ago (mya) (35–47, 95% CI) during the Paleogene, approximately coinciding with the initial closure of the Tethys Sea (van Oppen et al. 2001; Wallace and Portell 2022). The addition of *A. cervicornis* as a Caribbean *Acropora* to the published *Acropora* phylogenomic tree (Shinzato et al. 2021) suggests that the *Acropora* radiation in the Indo-Pacific occurred in at least two stages with an initial split occurring between 58 and 68 mya resulting in clades I and II prior to the divergence of the Caribbean and Indo-Pacific *Acropora*, followed by a second set of radiations of the Indo-Pacific *Acropora* 35–47 mya and resulting in the more speciose clades III and IV (Fig. 3). Low bipartition support values within *Acropora* could result from the relatively rapid radiation within *Acropora*, incomplete lineage sorting and/or introgressive hybridization (Vollmer and Palumbi 2002; van Oppen, Willis, et al. 2002; Minh et al. 2020). The Acroporids (*Montipora* and *Acropora*) likely diverged from the Pocilloporids (*S. pistillata* and *P. damicornis*) during the Triassic period 220 mya (188–225, 95% CI) with *Acropora*

and *Montipora* diverging during the early Cretaceous 136 mya (116–136, 95% CI).

Conclusion

In summary, this study presents the first fully annotated genome of the endangered Caribbean staghorn coral, *A. cervicornis*. Comparative genomics highlights distinctive genetic traits, including immune pathway contractions and signaling pathway expansions, suggesting further research into the species response to environmental stressors, particularly White Band Disease. Phylogenetic analysis places *A. cervicornis* within the broader *Acropora* genus, dating its divergence to around 41 mya, aligning previous morphological findings and supporting the hypothesis of divergence coincident with the closure of the western Tethys Sea. This annotated genome serves as a valuable resource for future research, facilitating conservation and restoration efforts for Caribbean coral reefs, and deepening our understanding of speciation and adaptation within *Acropora*.

Data availability

Genome assembly and associated Nanopore and short-read DNA sequencing data can be accessed from NCBI BioProject: PRJNA948411 with the mitochondrial genome assembly accessible at NCBI GenBank: OQ772303. RNA sequencing data used for annotation can be accessed from NCBI BioProject: PRJNA949884. Genome assembly and annotation pipeline code available from GitHub: https://github.com/VollmerLab/Acerv_Genome.
    Supplemental material available at G3 online.

Acknowledgment

Funding

Conflicts of interest statement

The authors declare no conflict of interest.

Literature cited

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol. 215(3):403–410. doi:10.1016/S0022-2836(05)80360-2.

Aronson RB, Precht WF. 2001. White-band disease and the changing face of Caribbean coral reefs. In: Porter JW, editors. The Ecology and Etiology of Newly Emerging Marine Diseases. Developments in Hydrobiology. Dordrecht: Springer. p. 25–38.

Baumgarten S, Simakov O, Esherick LY, Liew YJ, Lehnert EM, Michell CT, Li Y, Hambleton EA, Guse A, Oates ME, *et al.* 2015. The genome of *Aiptasia*, a sea anemone model for coral symbiosis. Proc Natl Acad Sci U S A. 112(38):11893–11898. doi:10.1073/pnas.1513318112.

Beasley TM, Schumacker RE. 1995. Multiple regression approach to analyzing contingency tables: post hoc and planned comparison procedures. J Exp Educ. 64(1):79–93. doi:10.1080/00220973.1995.9943797.

Behrensmeyer AK, Turner A. Taxonomic occurrences of Suidae recorded in the Paleobiology Database. Fossilworks [Internet]. 2013. Available from: http://fossilworks.org.2013.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B: Stat Methodol. 57:289–300. doi:10.1111/j.2517-6161.1995.tb02031.x.

Birney E, Clamp M, Durbin R. 2004. Genewise and genomewise. Genome Res. 14(5):988–995. doi:10.1101/gr.1865504.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. BMC Bioinformatics. 10(1):421. doi:10.1186/1471-2105-10-421.

Cantarel BL, Korf I, Robb SMC, Parra G, Ross E, Moore B, Holt C, Sánchez Alvarado A, Yandell M. 2008. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Res. 18(1):188–196. doi:10.1101/gr.6743907.

Challis R. 2017 rjchallis/assembly-stats 17.02.

Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 34(17):i884–i890. doi:10.1093/bioinformatics/bty560.

Colin L, Yesson C, Head C, Colin L, Yesson C, Head C, Colin L, Yesson C, Head C, Colin L, *et al.* 2021. Complete mitochondrial genomes of three reef forming *Acropora* corals (acroporidae, scleractinia) from chagos archipelago, Indian ocean. Biodivers Data J. 9: e72762. doi:10.3897/BDJ.9.e72762.

Coombe L, Li JX, Lo T, Wong J, Nikolic V, Warren RL, Birol I. 2021. LongStitch: high-quality genome assembly correction and scaffolding using long reads. BMC Bioinformatics. 22(1):534. doi:10.1186/s12859-021-04451-7.

Cunning R, Bay RA, Gillette P, Baker AC, Traylor-Knowles N. 2018. Comparative analysis of the *Pocillopora damicornis* genome highlights role of immune system in coral evolution. Sci Rep. 8(1): 16134. doi:10.1038/s41598-018-34459-8.

De Bie T, Cristianini N, Demuth JP, Hahn MW. 2006. CAFE: a computational tool for the study of gene family evolution. Bioinformatics. 22(10):1269–1271. doi:10.1093/bioinformatics/btl097.

De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. Bioinformatics. 34(15):2666–2669. doi:10.1093/bioinformatics/bty149.

Emms DM, Kelly S. 2017. STRIDE: species tree root inference from gene duplication events. Mol Biol Evol. 34(12):3267–3278. doi:10.1093/molbev/msx259.

Emms DM, Kelly S. 2018. STAG: Species Tree Inference from All Genes. 267914.

Emms DM, Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 20(1):238. doi:10.1186/s13059-019-1832-y.

Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, Smit AF. 2020. RepeatModeler2 for automated genomic discovery of transposable element families. Proc Natl Acad Sci U S A. 117(17):9451–9457. doi:10.1073/pnas.1921046117.

Fuller ZL, Mocellin VJL, Morris LA, Cantin N, Shepherd J, Sarre L, Peng J, Liao Y, Pickrell J, Andolfatto P, *et al.* 2020. Population genetics of the coral *Acropora millepora*: toward genomic prediction of bleaching. Science. 369(6501):eaba4674. doi:10.1126/science.aba4674.

Geib SM, Hall B, Derego T, Bremer FT, Cannoles K, Sim SB. 2018. Genome annotation generator: a simple tool for generating and correcting WGS annotation tables for NCBI submission. GigaScience. 7(4):giy018. doi:10.1093/gigascience/giy018.

Gertz EM, Yu Y-K, Agarwala R, Schäffer AA, Altschul SF. 2006. Composition-based statistics and translated nucleotide searches: improving the TBLASTN module of BLAST. BMC Biol. 4(1):41. doi: 10.1186/1741-7007-4-41.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, *et al.* 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 29(7):644–652. doi:10.1038/nbt.1883.

Guan D, McCarthy SA, Wood J, Howe K, Wang Y, Durbin R. 2020. Identifying and removing haplotypic duplication in primary genome assemblies. Bioinformatics. 36(9):2896–2898. doi:10.1093/bioinformatics/btaa025.

Guiglielmoni N, Houtain A, Derzelle A, Van Doninck K, Flot J-F. 2021. Overcoming uncollapsed haplotypes in long-read assemblies of non-model organisms. BMC Bioinformatics. 22(1):303. doi:10.1186/s12859-021-04118-3.

Hahn MW, De Bie T, Stajich JE, Nguyen C, Cristianini N. 2005. Estimating the tempo and mode of gene family evolution from comparative genomic data. Genome Res. 15(8):1153–1160. doi:10.1101/gr.3567505.

Hamada M, Shoguchi E, Shinzato C, Kawashima T, Miller DJ, Satoh N. 2013. The complex NOD-like receptor repertoire of the coral *Acropora digitifera* includes novel domain combinations. Mol Biol Evol. 30(1):167–176. doi:10.1093/molbev/mss213.

Han MV, Thomas GWC, Lugo-Martinez J, Hahn MW. 2013. Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. Mol Biol Evol. 30(8):1987–1997. doi:10.1093/molbev/mst100.

Hart AJ, Ginzburg S, Xu M, Fisher CR, Rahmatpour N, Mitton JB, Paul R, Wegrzyn JL. 2020. EnTAP: bringing faster and smarter functional annotation to non-model eukaryotic transcriptomes. Mol Ecol Resour. 20(2):591–604. doi:10.1111/1755-0998.13106.

Hemond EM, Kaluziak ST, Vollmer SV. 2014. The genetics of colony form and function in Caribbean *Acropora* corals. BMC Genomics. 15(1):1133. doi:10.1186/1471-2164-15-1133.

Hilton JD, Brady AK, Spaho SA, Vize PD. 2012. Photoreception and signal transduction in corals: proteomic and behavioral evidence for cytoplasmic calcium as a mediator of light responsivity. Biol Bull. 223(3):291–299. doi:10.1086/BBLv223n3p291.

Holt C, Yandell M. 2011. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. BMC Bioinformatics. 12(1):491. doi:10.1186/1471-2105-12-491.

Hughes TP, Anderson KD, Connolly SR, Heron SF, Kerry JT, Lough JM, Baird AH, Baum JK, Berumen ML, Bridge TC, *et al.* 2018. Spatial and temporal patterns of mass bleaching of corals in the anthropocene. Science. 359(6371):80–83. doi:10.1126/science.aan8048.

Jühling F, Pütz J, Bernt M, Donath A, Middendorf M, Florentz C, Stadler PF. 2012. Improved systematic tRNA gene annotation allows new insights into the evolution of mitochondrial tRNA structures and into the mechanisms of mitochondrial genome rearrangements. Nucleic Acids Res. 40(7):2833–2845. doi:10.1093/nar/gkr1131.

Kelly S, Maini PK. 2013. DendroBLAST: approximate phylogenetic trees in the absence of multiple sequence alignments. PLoS One. 8(3):e58537. doi:10.1371/journal.pone.0058537.

Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol. 37(5):540–546. doi:10.1038/s41587-019-0072-8.

Korf I. 2004. Gene finding in novel genomes. BMC Bioinformatics. 5(1):59. doi:10.1186/1471-2105-5-59.

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. 2009. Circos: an information aesthetic for comparative genomics. Genome Res. 19(9):1639–1645. doi:10.1101/gr.092759.109.

Laetsch DR, Blaxter ML. 2017. BlobTools: Interrogation of genome assemblies. F1000Research. 6. doi:10.12688/f1000research.12232.1.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–wheeler transform. Bioinformatics. 25(14):1754–1760. doi:10.1093/bioinformatics/btp324.

Libro S, Vollmer SV. 2016. Genetic signature of resistance to white band disease in the Caribbean staghorn coral *Acropora cervicornis*. PLoS One. 11(1):e0146636. doi:10.1371/journal.pone.0146636.

Liu S-YV, Chan C-LC, Hsieh HJ, Fontana S, Wallace CC. 2015. Massively parallel sequencing (MPS) assays for sequencing mitochondrial genomes: the phylogenomic implications for *Acropora* staghorn corals (Scleractinia; Acroporidae). Mar Biol. 162(6):1383–1392. doi:10.1007/s00227-015-2657-1.

Lohse M, Drechsel O, Kahlau S, Bock R. 2013. OrganellarGenomeDRAW—a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. Nucleic Acids Res. 41(W1):W575–W581. doi:10.1093/nar/gkt289.

Lomsadze A, Ter-Hovhannisyan V, Chernoff YO, Borodovsky M. 2005. Gene identification in novel eukaryotic genomes by self-training algorithm. Nucleic Acids Res. 33(20):6494–6506. doi:10.1093/nar/gki937.

López-Nandam EH, Albright R, Hanson EA, Sheets EA, Palumbi SR. 2023. Mutations in coral soma and sperm imply lifelong stem cell renewal and cell lineage selection. Proc Biol Sci. 290:20221766. doi:10.1098/rspb.2022.1766.

Love RR, Weisenfeld NI, Jaffe DB, Besansky NJ, Neafsey DE. 2016. Evaluation of DISCOVAR de novo using a mosquito sample for cost-effective short-read genome assembly. BMC Genomics. 17(1):187. doi:10.1186/s12864-016-2531-7.

Marçais G, Kingsford C. 2011. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. Bioinformatics. 27(6):764–770. doi:10.1093/bioinformatics/btr011.

Mendes FK, Vanderpool D, Fulton B, Hahn MW. 2020. CAFE 5 models variation in evolutionary rates among gene families. Bioinformatics. 36(22–23):5516–5518. doi:10.1093/bioinformatics/btaa1022.

Meng G, Li Y, Yang C, Liu S. 2019. MitoZ: a toolkit for animal mitochondrial genome assembly, annotation and visualization. Nucleic Acids Res. 47(11):e63. doi:10.1093/nar/gkz173.

Minh BQ, Hahn MW, Lanfear R. 2020. New methods to calculate concordance factors for phylogenomic datasets. Mol Biol Evol. 37(9):2727–2733. doi:10.1093/molbev/msaa106.

National Marine Fisheries Service. 2006. Endangered and Threatened Species: Final Listing Determinations for Elkhorn Coral and Staghorn Coral. Silver Spring, MD: National Marine Fisheries Service (NMFS), 71(FR26852). p. 26852–26872.

Nawrocki EP, Eddy SR. 2013. Infernal 1.1: 100-fold faster RNA homology searches. Bioinformatics. 29(22):2933–2935. doi:10.1093/bioinformatics/btt509.

Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. Genome Res. 27(5):824–834. doi:10.1101/gr.213959.116.

Parkinson JE, Bartels E, Devlin-Durante MK, Lustic C, Nedimyer K, Schopmeyer S, Lirman D, LaJeunesse TC, Baums IB. 2018. Extensive transcriptional variation poses a challenge to thermal stress biomarker development for endangered corals. Mol Ecol. 27(5):1103–1119. doi:10.1111/mec.14517.

Putnam NH, Srivastava M, Hellsten U, Dirks B, Chapman J, Salamov A, Terry A, Shapiro H, Lindquist E, Kapitonov VV, *et al.* 2007. Sea Anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. Science. 317(5834):86–94. doi:10.1126/science.1139158.

R Core Team. 2022. R: A Language and Environment for Statistical Computing.

Reyes-Bermudez A, DeSalvo MK, Voolstra CR, Sunagawa S, Szmant AM, Iglesias-Prieto R, Medina M. 2009. Gene expression microarray analysis encompassing metamorphosis and the onset of calcification in the scleractinian coral *Montastraea faveolata*. Mar Genom. 2(3–4):149–159. doi:10.1016/j.margen.2009.07.002.

Rodriguez-Casariego JA, Ladd MC, Shantz AA, Lopes C, Cheema MS, Kim B, Roberts SB, Fourqurean JW, Ausio J, Burkepile DE, *et al.*

2018. Coral epigenetic responses to nutrient stress: histone H2A.X phosphorylation dynamics and DNA methylation in the staghorn coral *Acropora cervicornis*. Ecol Evol. 8(23):12193–12207. doi:10.1002/ece3.4678.

Rosenberg Y, Doniger T, Harii S, Sinniger F, Levy O. 2017. Canonical and cellular pathways timing gamete release in *Acropora digitifera*, Okinawa, Japan. Mol Ecol. 26(10):2698–2710. doi:10.1111/mec.14062.

Russell TJ, Watson GM. 1995. Evidence for intracellular stores of calcium ions involved in regulating nematocyst discharge. J Exp Zool. 273(3):175–185. doi:10.1002/jez.1402730302.

Salazar OR, Arun PN, Cui G, Bay LK, van Oppen MJH, Webster NS, Aranda M. 2022. The coral *Acropora loripes* genome reveals an alternative pathway for cysteine biosynthesis in animals. Sci Adv. 8(38):eabq0304. doi:10.1126/sciadv.abq0304.

Shinzato C, Khalturin K, Inoue J, Zayasu Y, Kanda M, Kawamitsu M, Yoshioka Y, Yamashita H, Suzuki G, Satoh N. 2021. Eighteen coral genomes reveal the evolutionary origin of *Acropora* strategies to accommodate environmental changes. Mol Biol Evol. 38(1):16–30. doi:10.1093/molbev/msaa216.

Shinzato C, Mungpakdee S, Satoh N, Shoguchi E. 2014. A genomic approach to coral-dinoflagellate symbiosis: studies of *Acropora digitifera* and *Symbiodinium minutum*. Front Microbiol. 5:336. doi:10.3389/fmicb.2014.00336.

Shinzato C, Shoguchi E, Kawashima T, Hamada M, Hisata K, Tanaka M, Fujie M, Fujiwara M, Koyanagi R, Ikuta T, *et al.* 2011. Using the *Acropora digitifera* genome to understand coral responses to environmental change. Nature. 476(7360):320–323. doi:10.1038/nature10249.

Smit AFA, Hubley R, Green P. 2015. RepeatMasker Open-4.0. 2013–2015.

Stanke M, Schöffmann O, Morgenstern B, Waack S. 2006. Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. BMC Bioinformatics. 7(1):62. doi:10.1186/1471-2105-7-62.

Surm JM, Stewart ZK, Papanicolaou A, Pavasovic A, Prentis PJ. 2019. The draft genome of *Actinia tenebrosa* reveals insights into toxin evolution. Ecol Evol. 9(19):11314–11328. doi:10.1002/ece3.5633.

Tate R, Hall B, DeRego T, Geib S. 2014 Annie: the ANNotation Information Extractor (Version 1.0). 2014.

Tenenbaum D, Volkening J. 2023. KEGGREST: Client-side REST access to the Kyoto Encyclopedia of Genes and Genomes (KEGG).

The UniProt Consortium. 2023. Uniprot: the universal protein knowledgebase in 2023. Nucleic Acids Res. 51(D1):D523–D531. doi:10.1093/nar/gkac1052.

Tian P, Niu W. 2017. The complete mitochondrial genome of the *Acropora pruinosa*. Mitochondrial DNA B Resour. 2(2):652–653. doi:10.1080/23802359.2017.1375882.

To T-H, Jung M, Lycett S, Gascuel O. 2016. Fast dating using least-squares criteria and algorithms. Syst Biol. 65(1):82–97. doi:10.1093/sysbio/syv068.

van Oppen MJH, Catmull J, Mcdonald BJ, Hislop NR, Hagerman PJ, Miller DJ. 2002. The mitochondrial genome of *Acropora tenuis* (Cnidaria; Scleractinia) contains a large group I intron and a candidate control region. J Mol Evol. 55(1):1–13. doi:10.1007/s00239-001-0075-0.

van Oppen MJ, Hislop NR, Hagerman PJ, Miller DJ. 1999. Gene content and organization in a segment of the mitochondrial genome of the scleractinian coral *Acropora tenuis*: major differences in gene order within the anthozoan subclass zoantharia. Mol Biol Evol. 16(12):1812–1815. doi:10.1093/oxfordjournals.molbev.a026094.

van Oppen MJH, McDonald BJ, Willis B, Miller DJ. 2001. 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. 18(7):1315–1329. doi:10.1093/oxfordjournals.molbev.a003916.

van Oppen MJH, Willis BL, Van Rheede T, Miller DJ. 2002. Spawning times, reproductive compatibilities and genetic structuring in the *Acropora aspera* group: evidence for natural hybridization and semi-permeable species boundaries in corals. Mol Ecol. 11(8):1363–1376. doi:10.1046/j.1365-294X.2002.01527.x.

Vaser R, Sović I, Nagarajan N, Šikić M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res. 27(5):737–746. doi:10.1101/gr.214270.116.

Vollmer SV, Palumbi SR. 2002. Hybridization and the evolution of reef coral diversity. Science. 296(5575):2023–2025. doi:10.1126/science.1069524.

Voolstra CR, Li Y, Liew YJ, Baumgarten S, Zoccola D, Flot J-F, Tambutté S, Allemand D, Aranda M. 2017. Comparative analysis of the genomes of *Stylophora pistillata* and *Acropora digitifera* provides evidence for extensive differences between species of corals. Sci Rep. 7(1):17583. doi:10.1038/s41598-017-17484-x.

Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, *et al.* 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One. 9(11):e112963. doi:10.1371/journal.pone.0112963.

Wallace CC. 1999. Staghorn Corals of the World: A Revision of the Genus Acropora. Collingwood (VIC): Csiro Publishing.

Wallace CC, Portell RW. 2022. Earliest western Atlantic staghorn corals (*Acropora*) from the lower Oligocene Suwannee Limestone of Florida, USA, and their significance for modern coral distribution. J Paleontol. 96(6):1390–1399. doi:10.1017/jpa.2022.47.

Wheeler TJ, Eddy SR. 2013. nhmmer: DNA homology search with profile HMMs. Bioinformatics. 29(19):2487–2489. doi:10.1093/bioinformatics/btt403.

Wingett SW, Andrews S. 2018. FastQ Screen: a tool for multi-genome mapping and quality control. F1000Res. 7:1338. doi:10.12688/f1000research.15931.1.

Wu B, Li M, Liao X, Luo J, Wu F-X, Pan Y, Wang J. 2020. MEC: misassembly error correction in contigs based on distribution of paired-End reads and statistics of GC-contents. IEEE/ACM Transact Comput Biol Bioinform. 17(3):847–857. doi:10.1109/TCBB.2018.2876855.

Wu TD, Reeder J, Lawrence M, Becker G, Brauer MJ. 2016. GMAP And GSNAP for genomic sequence alignment: enhancements to speed, accuracy, and functionality. In: Mathé E, Davis S, editors. Statistical Genomics: Methods and Protocols. Methods in Molecular Biology. New York: Springer. p. 283–334.

Xie Y, Wu G, Tang J, Luo R, Patterson J, Liu S, Huang W, He G, Gu S, Li S, *et al.* 2014. SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads. Bioinformatics. 30(12):1660–1666. doi:10.1093/bioinformatics/btu077.

Xu M, Guo L, Gu S, Wang O, Zhang R, Peters BA, Fan G, Liu X, Xu X, Deng L, *et al.* 2020. TGS-GapCloser: a fast and accurate gap closer for large genomes with low coverage of error-prone long reads. GigaScience. 9(9):giaa094. doi:10.1093/gigascience/giaa094.

Ying H, Hayward DC, Cooke I, Wang W, Moya A, Siemering KR, Sprungala S, Ball EE, Forêt S, Miller DJ. 2019. The whole-genome sequence of the coral *Acropora millepora*. Genome Biol Evol. 11(5):1374–1379. doi:10.1093/gbe/evz077.

Zdobnov EM, Apweiler R. 2001. Interproscan – an integration platform for the signature-recognition methods in InterPro. Bioinformatics. 17(9):847–848. doi:10.1093/bioinformatics/17.9.847.

Zhang Y, Yu X, Zhou Z, Huang B. 2016. The complete mitochondrial genome of *Acropora aculeus* (cnidaria, scleractinia, acroporidae). Mitochondrial DNA A DNA Mapp Seq Anal. 27(6):4276–4277. doi:10.3109/19401736.2015.1082092.

Editor: K. Vogel