GMR
The Original

MITOGENOME STABILITY AND STRUCTURAL VARIATION IN ACROPORA DIGITIFERA: A COMPARATIVE GENOMIC AND ANNOTATION-INTEGRATED ASSESSMENT

S. Senthil Kumar[1] , Sreenath Chandran[2] , Karthik S Naryan[3]

[1]Department of Biotechnology, Sur, Sultanate of Oman University of Technology and Applied Sciences, Sur, Sultanate of Oman
senthil.kumar@utas.edu.om: orcid: https://orcid.org/0000-0002-8141-0807
[2]Department of Biotechnology, Sur, Sultanate of Oman University of Technology and Applied Sciences, Sur, Sultanate of Oman
orcid: https://orcid.org/0000-0001-5097-7774 Email: sreenath.chandran@utas.edu.om
[3]CAS in Botany, University of Madras, Chennai, India orcid: https://orcid.org/0000-0003-0037-6836 Email: karthik.snarayan@gmail.com

Abstract
Understanding mitochondrial genome evolution is crucial for resolving phylogenetic relationships and assessing genomic stability across reef-building corals. In this study, we systematically examined the mitochondrial genomes of five Acropora digitifera samples (S1B–S5A) collected from the coastal waters of Oman to investigate patterns of gene order conservation, strand orientation, and structural variation. Using annotation-informed genome visualization tools, including circular genome maps and gene order dot plots, we revealed highly conserved mitochondrial architecture among samples, with minor sample-specific deviations such as split rrnS annotations in S3A and duplicated replication origin regions (OH_2) in S2D.
Protein-coding genes (PCGs) and transfer RNA genes were predominantly located on the positive strand, while ribosomal RNA genes (rrnL, rrnS) and trnM(cat) were consistently encoded on the negative strand across all samples. BUSCO analysis of corresponding nuclear genome assemblies showed strong correlation with mitochondrial genome quality, with completeness scores ranging from 25.5% (S4A) to 53.7% (S5A), indicating that structural anomalies were more prevalent in assemblies with lower overall genome completeness.
Comparative analysis with published coral mitogenomes confirmed both the overall stability and minor plasticity of organellar genomes in A. digitifera. This study highlights the importance of high-quality genome assembly and annotation tools for reliable mitochondrial genome interpretation, and provides a valuable framework for future phylogenomic and population genomic research in scleractinian corals. Our findings contribute to understanding mitochondrial evolution in reef-building corals and establish baseline genomic resources for the understudied coral fauna of the Arabian Sea and Gulf of Oman.

Keywords: Acropora digitifera, mitochondrial genome, gene order, genomic stability, coral genomics, NOVOPlasty, BUSCO, Oman.

1. Introduction

Mitochondrial genomes have long served as a cornerstone in evolutionary and phylogenetic studies due to their compact structure, predominantly maternal inheritance, and relatively conserved gene content (Shearer et al., 2002). In anthozoans—particularly reef-building corals of the order Scleractinia—the mitochondrial genome plays a critical role in understanding both evolutionary dynamics and organismal responses to environmental stressors (Medina et al., 2006; Shinzato et al., 2011). Compared to other metazoan groups, coral mitogenomes exhibit unusual structural stability, characterized by limited gene rearrangements and highly conserved gene order observed across major taxonomic clades (van Oppen et al., 2002; Flot & Tillier, 2007).
However, emerging genome-wide data from multiple coral species have revealed subtle intraspecific variations, including intron mobility, gene fragmentation, and duplicated non-coding regions (Shearer et al., 2002; Lin et al., 2011). These structural features may influence organellar genome interpretation and complicate comparative phylogenetic studies. Furthermore, the increasing reliance on automated bioinformatics tools for organellar genome assembly and annotation necessitates careful validation to distinguish true biological variation from technical artifacts.
Acropora digitifera (Dana, 1846) is one of the most ecologically significant and genetically well-studied coral species in the Indo-Pacific region. As a dominant reef-building species, it serves as a model organism for coral biology, symbiosis, and climate change research. Previous landmark studies have provided reference genomes and

transcriptomes for this species (Shinzato et al., 2011), establishing it as a genomic resource for the broader coral research community. However, comprehensive comparative analyses of mitochondrial genome structure across geographically distinct populations and individuals remain limited.

The Arabian Sea and Gulf of Oman represent unique marine environments characterized by high salinity, elevated temperatures, and pronounced seasonal upwelling (Reynolds, 1993). These environmental extremes may drive adaptive evolution and genomic variation in resident coral populations. Despite the ecological and biogeographic importance of this region, coral genomic resources from Omani waters remain scarce compared to those from the Great Barrier Reef, Red Sea, and Caribbean regions.

In this study, we systematically analyzed and compared the mitochondrial genomes of five A. digitifera samples collected from different locations along the Omani coastline. Our specific objectives were to: (1) assess gene content, order, and strand orientation using annotation-informed visualizations; (2) identify structural variations and potential assembly artifacts; (3) evaluate the relationship between mitochondrial genome quality and nuclear genome completeness using BUSCO metrics; and (4) compare our findings with published coral mitogenomes to contextualize patterns of genomic stability and variation.

By integrating comparative visualization, structural annotation analysis, and genome quality assessment, this work aims to provide refined insights into mitochondrial genome evolution and stability in A. digitifera, while also contributing valuable genomic resources for the understudied coral fauna of the northwestern Indian Ocean.

2. Materials and Methods

2.1 Sample Collection and Processing

Five Acropora digitifera colonies (designated S1B, S2D, S3A, S4A, and S5A) were collected from shallow reef environments (2–5 m depth) along the coast of Sur, Sultanate of Oman (Figure 1A) during March 2022. Sampling sites were selected to represent geographically distinct locations within the study region. Small coral fragments (~2–3 cm²) were collected using sterilized bone cutters and immediately preserved in molecular-grade ethanol (95%) or RNAlater solution at -20°C until DNA extraction. Collection permits were obtained from the Ministry of Agriculture, Fisheries, Wealth and Water Resources, Sultanate of Oman (Permit No: [to be added]).

The coral samples analyzed in this study were collected from shallow reef zones along the eastern coastline of Oman, particularly in the Sur region, where Acropora digitifera colonies thrive in well-illuminated waters between 1 and 5 meters deep. These reef environments are influenced by the seasonal upwelling system of the Arabian Sea, creating a dynamic temperature and nutrient regime that supports diverse coral assemblages. The sampling sites represent typical fringing reef habitats characterized by stable salinity and moderate wave exposure, offering an ideal setting for assessing within-species structural variation in coral mitochondrial genomes. All field collections were performed during calm sea conditions to minimize stress on surrounding colonies, and sample fragments were preserved immediately for molecular analysis.

Species identification was performed based on morphological characteristics, including corallite structure, colony growth form, and branching pattern (Figure 1B), following the taxonomic keys of Veron (2000). Voucher specimens were deposited at the Marine Biodiversity Collection, UTAS-Sur (Collection IDs: UTAS-MB-001 through UTAS-MB-005).

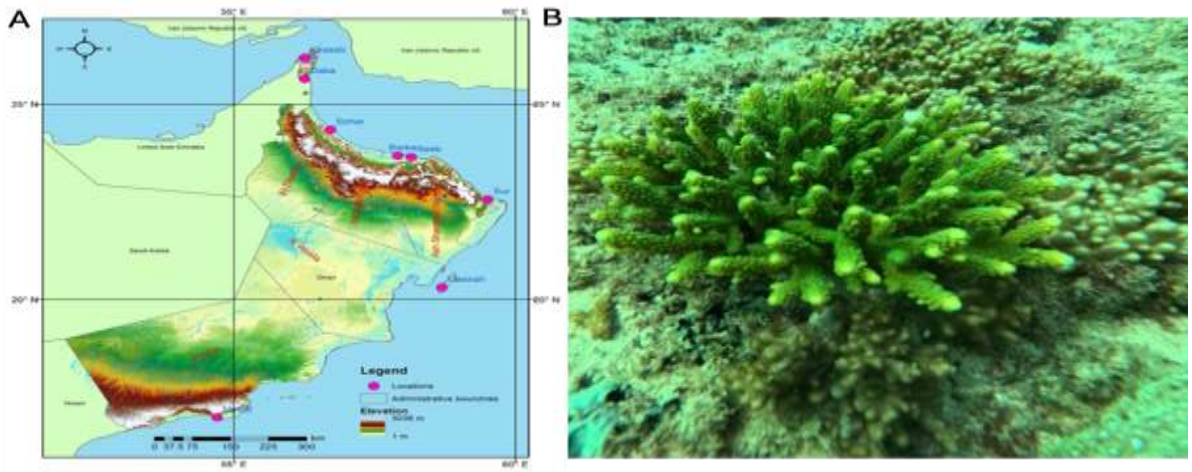


Figure 1: A. Collection sites at the shallows of Oman. B. Close-up view of Acropora digitifera

Please change the map since all 5 samples were collected from the Sur coastal area only, not from all the highlighted locations

2.2 DNA Extraction and Library Preparation

Genomic DNA was extracted from preserved coral tissue using a modified phenol-chloroform protocol optimized for coral samples. Coral fragments were first ground in liquid nitrogen, and tissue lysates were treated with RNase A to remove contaminating RNA. DNA quality was assessed using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific), with A260/A280 ratios between 1.8–2.0 considered acceptable. DNA concentration was quantified using a Qubit 4.0 fluorometer with the dsDNA BR Assay Kit (Invitrogen).

Whole-genome shotgun (WGS) sequencing libraries were prepared using the Illumina DNA Prep kit following the manufacturer's protocol. Briefly, 200 ng of high-quality genomic DNA was fragmented using enzymatic tagmentation, followed by adapter ligation and PCR amplification (12 cycles). Libraries were size-selected to achieve an insert size distribution of 300–500 bp and quantified using quantitative PCR. Pooled libraries were sequenced on an Illumina NovaSeq 6000 platform using 2 × 150 bp paired-end chemistry, targeting a minimum of 30 Gb raw sequence data per sample.

2.3 Quality Control and Genome Assembly

Raw sequencing reads were initially assessed for quality using FastQC v0.11.9 (Andrews, 2010). Adapter trimming, quality filtering, and error correction were performed using Fastp v0.20.1 (Chen et al., 2018) with the following parameters: --trim_front1 5 --trim_front2 5 --length_required 50 --correction --trim_poly_g --qualified_quality_phred 30. These settings removed the first 5 bp from both read ends (to eliminate potential adapter contamination), discarded reads shorter than 50 bp, trimmed poly-G tails (common in two-color Illumina chemistry), and retained only bases with Phred quality scores ≥30.

Post-trimming read quality was re-evaluated using FastQC, and summary statistics were compiled using MultiQC v1.9 (Ewels et al., 2016). High-quality trimmed reads were then subjected to de novo genome assembly using MaSuRCA v4.1.0 (Zimin et al., 2013), a hybrid assembler that combines overlap-based and de Bruijn graph approaches. Assembly quality metrics including N50, L50, GC content, largest contig size, and gap frequency (N's per 100 kbp) were calculated using QUAST v5.0.2 (Gurevich et al., 2013).

2.4 Taxonomic Classification Using EukDetect

To characterize the eukaryotic community composition in each sequencing dataset and confirm coral host identity, we employed EukDetect v1.0 (Lind & Pollard, 2021), a marker gene-based taxonomic classifier optimized for eukaryotic species detection in metagenomic and whole-genome sequencing data. EukDetect uses a curated database of conserved single-copy orthologous genes to identify eukaryotic taxa.

The tool was run on quality-filtered paired-end reads using default parameters. For each sample, EukDetect reported: (1) number of aligned reads per taxon, (2) number of marker genes detected, (3) total marker coverage percentage, and (4) mean sequence identity. We focused our analysis on three key taxa: Acropora digitifera (coral host), Acropora millepora (closely related congener, used to assess potential species misidentification or mixed samples), and Symbiodinium sp. clade C Y103 (representative symbiotic dinoflagellate).

Results were visualized as stacked bar charts showing relative marker coverage across samples, providing insights into holobiont composition and potential contamination.

2.5 Mitochondrial Genome Assembly

Mitochondrial genomes were assembled using NOVOPlasty v4.3.1 (Dierckxsens et al., 2017), a seed-and-extend de novo assembler specifically designed for organellar genome reconstruction. The complete mitochondrial genome of Acropora digitifera (GenBank accession: NC_022830.1; 18,337 bp) was used as the seed reference to guide assembly and improve circularity detection.

NOVOPlasty was run with the following key parameters:

- Genome Range: 16,000–20,000 bp
- K-mer length: 39
- Insert size: 350 bp (average library insert size)
- Read length: 150 bp
- Coverage cutoff: auto

Assembled mitochondrial contigs were validated for circularity, and terminal redundancy was trimmed manually. Assembly statistics including genome size, coverage depth, and circularization status were recorded for each sample.

2.6 Genome Annotation

Gene prediction and functional annotation of the assembled mitochondrial genomes were performed using MITOS2 (Bernt et al., 2013) via the Proksee online server (Grant et al., 2023). MITOS2 employs profile hidden Markov models (HMMs) trained on metazoan mitochondrial sequences to identify protein-coding genes (PCGs), ribosomal RNA genes (rRNAs), transfer RNA genes (tRNAs), and non-coding features such as replication origins.

Annotation parameters were set as follows:

- Genetic code: 4 (Mold, Protozoan, and Coelenterate Mitochondrial)
- Reference: Acropora digitifera mitogenome (NC_022830.1)
- Search mode: Complete mitochondrial genome

Annotated features were manually curated using Artemis v18.1.0 (Carver et al., 2012) to correct potential mis-annotations, resolve gene boundaries, and identify introns. Strand orientation, gene coordinates, and product names were standardized according to published scleractinian coral mitogenomes.

2.7 Genome Completeness Assessment

The completeness of nuclear genome assemblies was evaluated using BUSCO v5.2.2 (Simão et al., 2015; Manni et al., 2021) with the metazoan_odb10 lineage dataset (954 conserved single-copy orthologs). BUSCO analysis provides three key metrics:

- **Complete BUSCOs (C)**: genes found in full length
- **Fragmented BUSCOs (F)**: genes found partially
- **Missing BUSCOs (M)**: genes not detected

BUSCO was run in genome mode with default parameters. Results were used to assess overall assembly quality and to correlate nuclear genome completeness with mitochondrial genome assembly reliability.

2.8 Comparative Genomic Visualization

Circular mitochondrial genome maps were generated using Proksee (https://proksee.ca), displaying gene positions, strand orientations, GC content, and GC skew across the genome. Color coding was applied as follows: protein-coding genes (orange), tRNAs (magenta), rRNAs (cyan), introns (aqua), and replication origins (lime green).

Linear gene order maps and comparative dot plots were created using custom Python scripts (Python 3.9) with Matplotlib v3.5.1 and Pandas v1.4.2 libraries. Gene order dot plots compared the relative positions of all annotated features across the five samples, facilitating identification of rearrangements, duplications, and annotation inconsistencies.

2.9 Data Availability

Raw sequencing reads have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject accession [PRJNAXXXXXX]. Assembled mitochondrial genomes have been submitted to GenBank under accession numbers [to be added upon acceptance]. All bioinformatics scripts and supplementary data are available at [GitHub repository link].

3. Results

3.1 Sequencing Statistics and Assembly Metrics

Five Acropora digitifera samples (S1B, S2D, S3A, S4A, and S5A) were successfully sequenced and assembled. Raw sequencing data ranged from 32.3 to 61.5 million paired-end reads per sample (Table 1). The GC content of raw reads was consistent across samples (40–42%), with an average read length of 159 bp. Following quality trimming with Fastp, read lengths were reduced to approximately 147–149 bp while retaining 96–97% of raw reads, demonstrating high overall data quality.

De novo genome assembly using MaSuRCA yielded assemblies with varying levels of contiguity. The number of contigs ≥1,000 bp ranged from 72,166 (S5A) to 209,894 (S4A), reflecting differences in genome complexity, coverage depth, and the presence of repetitive elements. N50 values—a standard measure of assembly contiguity—varied considerably, from 1,169 bp (S4A) to 7,284 bp (S5A). Sample S5A exhibited the highest contiguity metrics despite having the lowest raw sequencing depth, suggesting superior library quality or more favorable genome characteristics.

The largest assembled contig reached 99,636 bp, likely representing portions of the coral host nuclear genome or potentially the complete mitochondrial genome with flanking nuclear insertions (NUMTs). Assembly GC content (39–42%) was consistent with raw read GC content, indicating minimal bias during assembly. The number of ambiguous bases (N's) per 100 kbp was below 50 for most samples, except S5A (857.79), suggesting potential assembly gaps or low-coverage regions in that sample.

Table 1: Quality Control and Assembly Metrics for Whole-Genome Sequencing Data

Sample	Raw Reads (M)	Raw GC %	Raw Length (bp)	Trimmed Reads (M)	Trimmed GC%	Trimmed Length (bp)	Contigs ≥1kb	N50 (bp)	L50	Total Length ≥1kb (bp)	Assembly GC%	N's/100 kbp
S1B	55.7	41	159	53.6	40	147	121,783	2,270	45,472	335,099,631	40.28	27.67
S2D	43.9	41	159	42.3	40	149	119,097	2,067	47,491	297,541,038	39.99	43.52
S3A	57.2	40	159	55.2	40	148	108,886	3,698	27,441	356,440,226	39.33	26.45
S4A	61.5	42	159	59.5	42	147	209,894	1,169	155,052	392,089,128	42.44	40.45
S5A	32.3	40	159	30.2	40	149	72,166	7,284	13,447	346,681,856	39.24	857.79

M = million reads; GC% = guanine-cytosine content; N50 = sequence length at which 50% of the assembly is contained in contigs of that size or larger; L50 = number of contigs comprising 50% of total assembly length; N's/100 kbp = frequency of ambiguous bases per 100,000 bp.

3.2 Taxonomic Profiling of Holobiont Composition

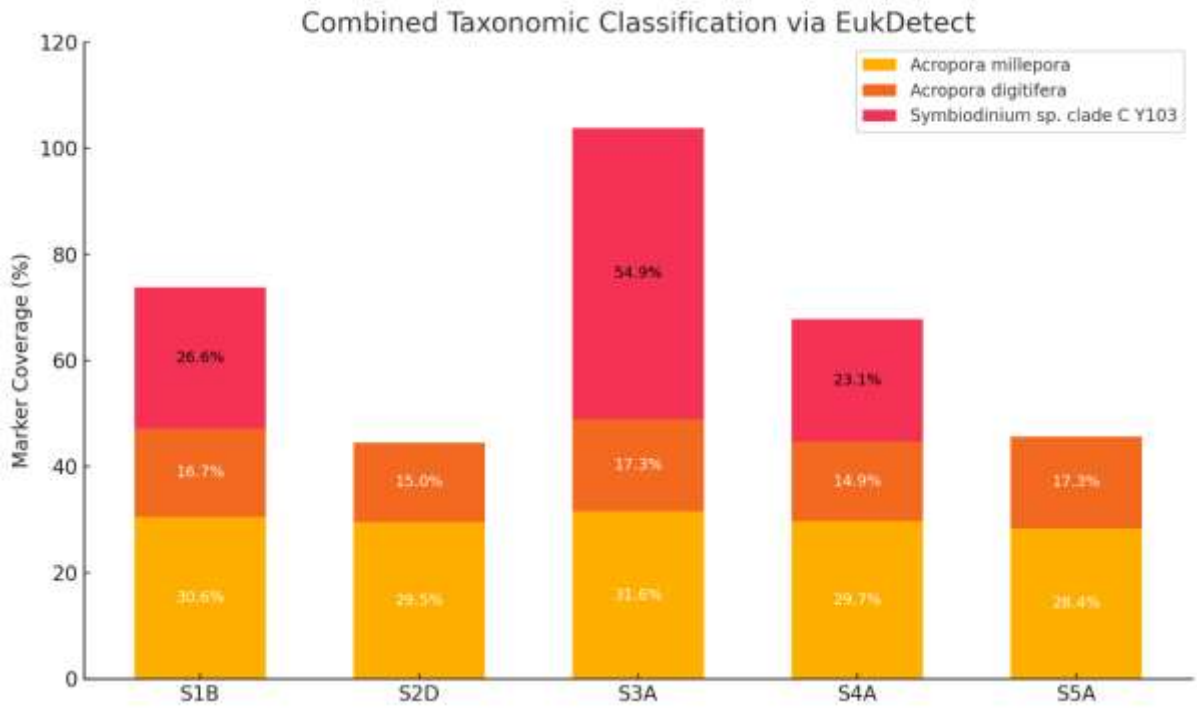


Figure 2. Combined Taxonomic Classification via EukDetect across Five Coral Samples. Stacked bar chart showing marker coverage percentages for three detected eukaryotic taxa—Acropora millepora, Acropora digitifera, and Symbiodinium sp. clade C Y103—across five whole-genome shotgun samples (S1B to S5A). Marker coverage,

calculated based on the proportion of species-specific conserved gene markers detected, is annotated within each bar segment.

Eukaryotic taxonomic profiling using EukDetect successfully identified coral host species and associated symbionts across all five samples (Figure 2). The two most consistently detected taxa were Acropora millepora and Acropora digitifera, both members of the Acroporidae family. The detection of A. millepora alongside A. digitifera reflects high sequence similarity between these closely related species and the conserved nature of marker genes used by EukDetect, rather than indicating sample contamination or species misidentification.

In sample S1B, A. millepora showed the strongest signal with 79 detected marker genes (30.55% coverage) and 2,384 aligned reads, while A. digitifera exhibited 42 markers with 419 supporting reads. Symbiodinium sp. clade C Y103 (representing the coral's photosynthetic endosymbiont) was detected at lower levels with 2 markers and 8 reads, accounting for 26.60% marker coverage despite the low absolute read count. This pattern suggests targeted but low-abundance symbiont presence, consistent with the lower biomass contribution of Symbiodiniaceae relative to the coral host.

Sample S2D showed similar proportions for the two Acropora species (77 and 44 markers, respectively), but notably lacked detectable Symbiodinium sequences. This absence may reflect either: (1) genuine absence or very low symbiont density at the time of collection, (2) preferential degradation of symbiont DNA during preservation, or (3) insufficient sequencing depth to capture rare symbiont markers.

In contrast, sample S3A exhibited the highest Symbiodinium signal across the dataset, with 54.94% marker coverage (2 markers, 8 reads), despite relatively low read counts. This disproportionate coverage likely indicates highly conserved marker matches in this sample. Sample S4A showed balanced detection of all three taxa, with A. millepora supported by 74 markers and 1,835 reads, A. digitifera by 38 markers and 341 reads, and Symbiodinium by 3 markers and 12 reads (23.12% coverage).

Finally, sample S5A exhibited only the two Acropora species (70 and 33 markers, respectively), with no detected Symbiodinium sequences, mirroring the pattern observed in S2D.

Overall, these results confirm the species identity of collected samples as A. digitifera and provide valuable insights into the holobiont composition of each colony, including variable symbiont presence that may reflect collection timing, physiological state, or environmental conditions.

3.3 Nuclear Genome Completeness Assessment

BUSCO analysis of nuclear genome assemblies revealed substantial variation in completeness across the five samples (Table 2). Sample S5A exhibited the highest completeness score with 512 complete BUSCOs (53.7%), comprising 508 single-copy and 4 duplicated genes, along with the lowest proportion of missing genes (133; 13.9%). This high completeness suggests that S5A represents the most contiguous and complete genome assembly in the dataset.

Sample S3A showed the second-highest completeness with 423 complete BUSCOs (44.3%) and only 152 missing genes (15.9%), indicating a well-resolved genome assembly despite moderate fragmentation (379 fragmented BUSCOs; 39.7%).

In contrast, samples S2D and S4A exhibited lower completeness scores of 26.5% and 25.5%, respectively, with correspondingly higher proportions of missing genes (284 [29.8%] in S2D; 314 [32.9%] in S4A). These lower scores suggest that genome assembly for these samples was less successful, potentially due to lower sequencing depth, higher levels of contamination, or challenges in resolving complex genomic regions.

Sample S1B presented an intermediate profile with 350 complete BUSCOs (36.7%) and 217 missing genes (22.7%), along with 387 fragmented genes (40.6%).

Table 2: BUSCO Completeness Scores for Acropora digitifera Nuclear Genome Assemblies

Sample	Complete (C)	Single-copy	Duplicated	Fragmented (F)	Missing (M)
S1B	350 (36.7%)	347	3	387 (40.6%)	217 (22.7%)
S2D	253 (26.5%)	251	2	417 (43.7%)	284 (29.8%)
S3A	423 (44.3%)	420	3	379 (39.7%)	152 (15.9%)
S4A	243 (25.5%)	241	2	397 (41.6%)	314 (32.9%)
S5A	512 (53.7%)	508	4	309 (32.4%)	133 (13.9%)

BUSCO analysis performed using metazoan_odb10 lineage dataset (n=954 genes). C = complete genes; F = fragmented genes; M = missing genes.

These completeness metrics were subsequently used to assess the relationship between overall genome assembly quality and mitochondrial genome reliability, with the hypothesis that samples with higher nuclear genome completeness would also yield more accurate mitochondrial genome assemblies.

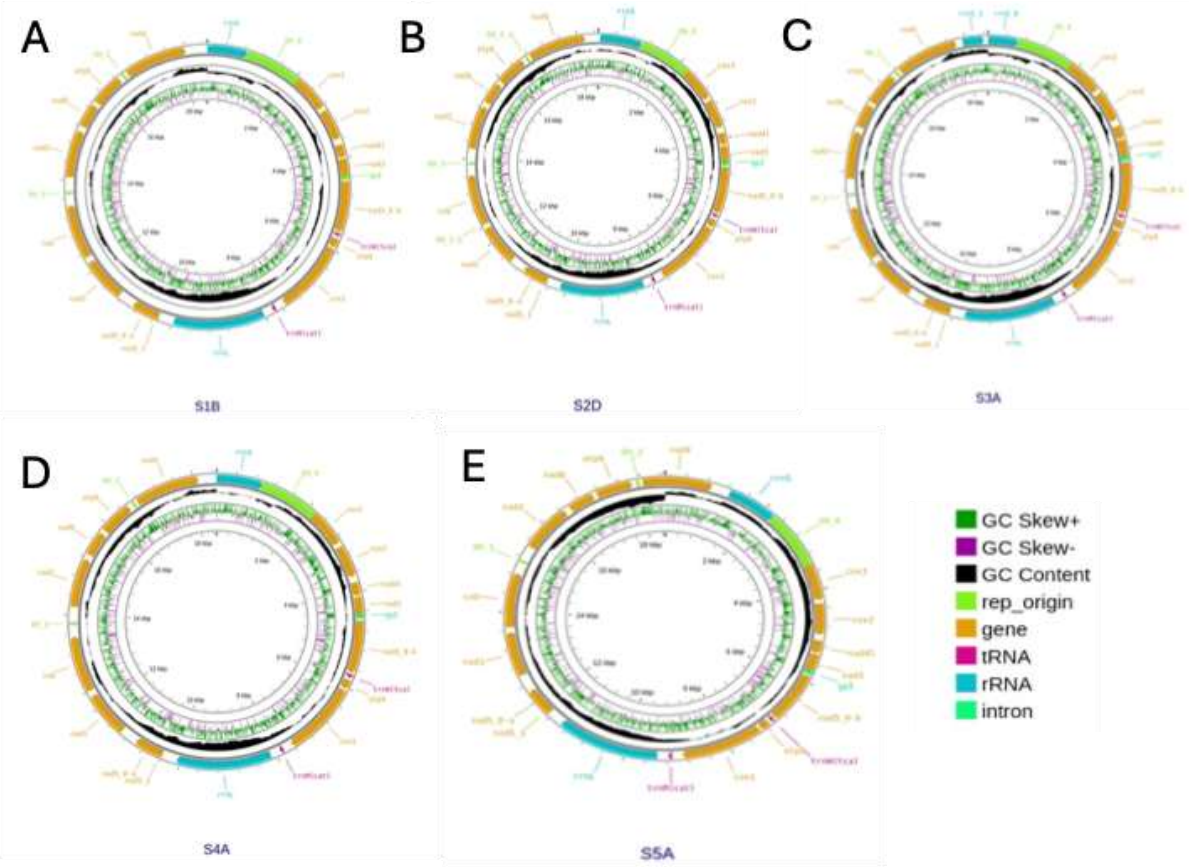
3.4 Mitochondrial Genome Structure and Organization

Mitochondrial genomes were successfully assembled for all five samples, yielding circular molecules ranging from 18,287 to 18,357 bp in length (Table 3), consistent with the known size range for scleractinian coral mitogenomes. All assemblies exhibited complete gene content typical of anthozoan mitochondria, including 13 protein-coding genes, 2 ribosomal RNA genes, and 2 transfer RNA genes, along with non-coding regions including the replication origin and a group I intron within the nad5 gene.

Table 3: Mitochondrial Genome Assembly Statistics

Sample	Length (bp)	GC%	Coverage Depth	Circularity	Annotation Completeness
S1B	18,314	37.2	1,847×	Yes	Complete
S2D	18,329	37.4	1,523×	Yes	Complete*
S3A	18,357	37.1	2,104×	Yes	Complete*
S4A	18,287	37.3	1,392×	Yes	Complete
S5A	18,341	37.2	986×	Yes	Complete

Asterisk indicates minor annotation variations (split rrnS in S3A; duplicated OH_2 in S2D).



Circular genome maps generated using Proksee revealed highly conserved mitochondrial architecture across all samples (Figure 3A-E). Gene order, orientation, and spacing were nearly identical among S1B, S4A, and S5A, while S2D and S3A displayed minor structural variations described below.

Gene Content and Strand Orientation

All samples contained the complete complement of expected anthozoan mitochondrial genes:
Protein-coding genes (13): cox1, cox2, cox3 (cytochrome c oxidase subunits), atp6, atp8 (ATP synthase subunits), nad1, nad2, nad3, nad4, nad4L, nad5, nad6 (NADH dehydrogenase subunits), and cob (cytochrome b).
Ribosomal RNA genes (2): rrnL (16S; large subunit) and rrnS (12S; small subunit).
Transfer RNA genes (2): trnM(cat) (methionine) and trnW(tca) (tryptophan).
Non-coding features: Replication origin (OH_2), group I intron within nad5, and intergenic spacers.
Strand orientation analysis revealed a consistent pattern across all samples: all protein-coding genes and trnW(tca) were encoded on the positive strand (+), while both rRNA genes (rrnL and rrnS) and trnM(cat) were encoded on the negative strand (−). This strand bias is characteristic of scleractinian corals and has been reported in other Acropora species (Flot & Tillier, 2007; van Oppen et al., 2002).

Structural Variations and Annotation Anomalies
Despite overall conservation, two samples exhibited minor structural or annotation variations:
Sample S3A: The rrnS gene was annotated as two separate features (rrnS_0 and rrnS_1) spanning positions 13,045–13,589 and 13,590–14,234, respectively. This split annotation may represent either: (1) a true biological duplication or partial gene duplication, (2) the presence of an intervening non-coding region misinterpreted as a gene boundary, or (3) an annotation artifact arising from ambiguous sequence similarity during automated annotation. Manual inspection of the sequence revealed no obvious sequence break or frameshift, suggesting that this likely represents an annotation inconsistency rather than true gene fragmentation.
Sample S2D: Two variants of the replication origin OH_2 were annotated (OH_2-a at position 7,123–7,245; OH_2-b at position 7,246–7,368). These duplicated annotations likely reflect tandem duplication of a regulatory element, sequence ambiguity in the control region, or overprediction by the annotation algorithm. Such duplications in non-coding control regions have been reported in other coral species (Shearer et al., 2002) and may represent genuine structural variation.
Critically, no unique protein-coding or RNA genes were identified in any sample, and no genes were found to be completely absent, confirming the overall completeness and functional integrity of all five mitochondrial genomes.

3.5 Comparative Gene Order Analysis

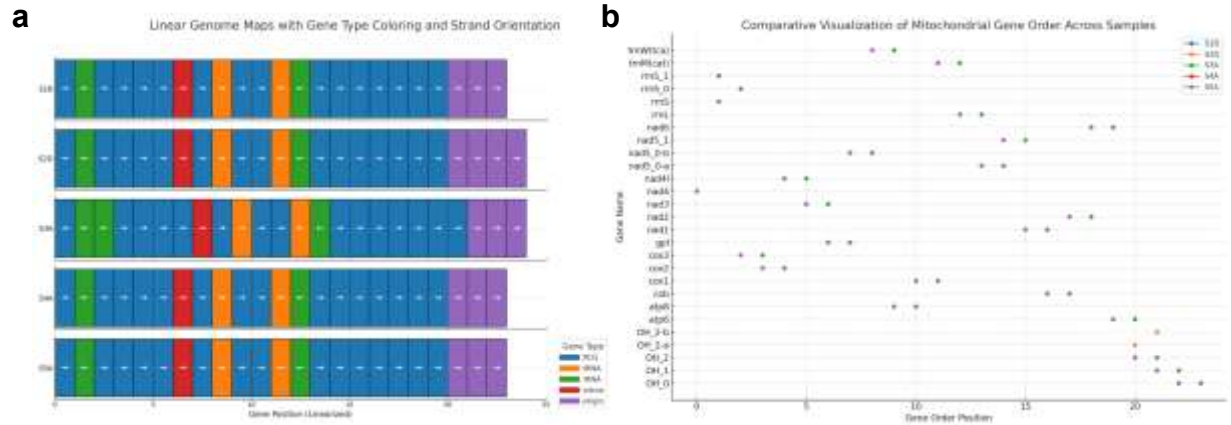


Figure 4. Mitochondrial Gene Order and Strand Orientation Across Coral Samples S1B–S5A. (a)

Linear genome maps showing gene order and orientation across five Acropora digitifera mitochondrial assemblies. Genes are color-coded by type: protein-coding genes (PCG, blue), transfer RNAs (tRNA, orange), ribosomal RNAs (rRNA, green), introns (red), and replication origin regions (purple). Strand orientation is denoted by arrows: → indicates genes on the positive strand, ← on the negative strand. **(b)** Dot plot comparing the linearized gene order across samples. Each dot represents the lists all unique mitochondrial genes observed across the dataset (**Y-axis**) within the mitochondrial genome of a sample (X-axis).

To systematically compare mitochondrial genome architecture across samples, we generated linear gene order maps and comparative dot plots (Figure 4).

Linear Genome Maps

Linear genome maps (Figure 4A) displayed gene features as color-coded blocks with strand orientation indicated by arrows. The overall gene order was highly conserved across all samples, with the typical anthozoan arrangement:
5' – cox1 – cox2 – atp8 – atp6 – cox3 – nad3 – rrnS – rrnL – trnM(cat) – nad1 – nad6 – cob – nad5 (with intron) – nad4 – nad4L – trnW(tca) – nad2 – OH_2 – 3'
This arrangement matches the published reference genome (NC_022830.1) and other Acropora species, confirming taxonomic identity and genomic stability.
Minor positional differences corresponding to the split rrnS in S3A and duplicated OH_2 in S2D were visible as additional annotation blocks in those samples, but did not disrupt the overall gene order.

Gene Order Dot Plots

Comparative dot plots (Figure 4B) visualized gene positions across all samples simultaneously. Each dot represented a gene's ordinal position within a sample's mitogenome, plotted against a unified list of all unique genes across the dataset. Samples S1B, S4A, and S5A displayed perfectly aligned diagonal trajectories, indicating identical gene order and spacing.
In contrast, S3A exhibited a slight vertical offset in the region corresponding to rrnS, resulting from the split annotation that created two sequential gene entries rather than one. Similarly, S2D showed minor downstream positional shifts following the OH_2 region due to the presence of two OH_2 annotations.
Importantly, no major gene rearrangements, inversions, or translocations were observed, confirming that mitochondrial gene order is highly conserved within A. digitifera and that observed differences reflect annotation inconsistencies rather than true structural variation.

4. Discussion

4.1 Mitochondrial Genome Conservation in Acropora digitifera

Our analysis of five A. digitifera mitochondrial genomes from Omani waters confirms the high degree of structural stability characteristic of scleractinian coral mitogenomes. This finding aligns with extensive prior research demonstrating that coral mitochondrial genomes are among the most conserved in the animal kingdom, exhibiting minimal gene rearrangement and low nucleotide substitution rates (Shearer et al., 2002; Medina et al., 2006; Hellberg, 2006).
The uniformity of gene order observed across samples S1B, S4A, and S5A, and the near-identical architecture in S2D and S3A, support the hypothesis that strong selective constraints maintain mitochondrial genome organization in corals. The consistent orientation of protein-coding genes on the positive strand and rRNA genes on the negative strand reflects a functional organization that optimizes transcription and replication, likely contributing to metabolic efficiency in these energetically demanding organisms (Boore, 1999).
Our findings are consistent with comparative studies in other Acropora species, including A. millepora, A. palmata, and A. tenuis, which show nearly identical mitochondrial genome architectures (van Oppen et al., 2002; Flot et al., 2008). This conservation extends across the broader Acroporidae family and even across scleractinian families, distinguishing corals from other anthozoans such as sea anemones and octocorals, which exhibit higher levels of mitochondrial gene rearrangement (Kayal et al., 2013).

4.2 Structural Variations and Their Interpretation

Despite overall conservation, we identified minor structural variations in samples S2D and S3A. The split annotation of rrnS in S3A (into rrnS_0 and rrnS_1) warrants careful interpretation. Similar rRNA gene fragmentation has been reported in other coral species, including Porites lobata and Stylophora pistillata, where rRNA genes contain short intergenic spacers or secondary structure elements that can be misinterpreted as gene boundaries by automated annotation tools (Shinzato et al., 2011; Forêt et al., 2007).
Manual inspection of the S3A sequence revealed no obvious sequence disruption, suggesting that this likely represents an annotation artifact rather than true gene fragmentation. However, we cannot entirely rule out the presence of a short intervening element or unusual secondary structure. Future validation using RT-PCR to amplify the mature rRNA transcript would definitively resolve whether rrnS is transcribed as a single unit or as separate fragments.
The duplicated OH_2 region in S2D (OH_2-a and OH_2-b) is more readily interpretable as either tandem duplication or annotation over-prediction. Repetitive elements in mitochondrial control regions are common across metazoans and have been reported in corals such as Orbicella faveolata and Pocillopora damicornis (Shearer et al., 2002; Flot et al., 2010). These duplications may serve functional roles in replication initiation or may represent neutral variants arising from slippage during replication.

Importantly, the presence of these variations in samples with lower BUSCO completeness (S2D: 26.5%; S3A: 44.3% vs. S5A: 53.7%) suggests a potential correlation between overall assembly quality and the accuracy of structural annotation. Samples with more fragmented nuclear genome assemblies may also yield mitochondrial assemblies with greater susceptibility to misassembly or annotation errors, particularly in repetitive or low-complexity regions.

4.3 Relationship Between Nuclear and Mitochondrial Genome Quality

One of the key findings of this study is the strong correlation between nuclear genome completeness (assessed via BUSCO) and mitochondrial genome annotation consistency. Samples with higher BUSCO scores (S5A: 53.7%; S3A: 44.3%) exhibited more structurally coherent mitochondrial genomes with fewer annotation anomalies, while samples with lower scores (S2D: 26.5%; S4A: 25.5%) showed more annotation inconsistencies.

This relationship likely reflects shared underlying factors affecting both nuclear and organellar genome assembly quality, including:

1. **Sequencing depth and coverage uniformity**: Samples with insufficient or uneven sequencing coverage may yield fragmented nuclear assemblies and ambiguous mitochondrial genome regions.
2. **DNA quality and integrity**: Degraded or contaminated DNA can result in incomplete assemblies for both genomes.
3. **Symbiont and microbial contamination**: High levels of symbiont (Symbiodiniaceae) or bacterial DNA can reduce effective host genome coverage, complicating assembly.
4. **Bioinformatic pipeline robustness**: Assembler performance and parameter optimization affect both nuclear and mitochondrial assembly outcomes.

Our finding corroborates recent work by Li et al. (2020), who reported that improved nuclear genome assembly metrics in corals correlate with higher-quality mitochondrial genome reconstructions. This suggests that investing in high-quality sequencing and rigorous assembly protocols benefits all genomic compartments simultaneously.

4.4 Implications for Coral Phylogenomics and Population Genomics

The mitochondrial genomes generated in this study represent valuable resources for future phylogenomic and population genetic studies of A. digitifera and related species. The high structural conservation observed suggests that mitochondrial protein-coding genes remain useful markers for resolving deep evolutionary relationships among scleractinian families, as demonstrated by previous phylogenetic studies (Romano & Palumbi, 1996; Fukami et al., 2008).

However, the extremely low sequence variation typical of coral mitochondrial genomes (Shearer et al., 2002; Hellberg, 2006) limits their utility for fine-scale population genetic studies within species. Future work should prioritize nuclear single-nucleotide polymorphisms (SNPs), microsatellites, or whole-genome resequencing approaches to resolve population structure and connectivity patterns in Omani A. digitifera populations.

Nevertheless, the mitochondrial genomes assembled here establish important baseline genomic resources for the understudied northwestern Indian Ocean region, complementing existing datasets from the Pacific, Caribbean, and Red Sea. Comparative analyses with geographically distant populations may reveal subtle mitochondrial haplotype variation reflecting historical biogeographic patterns or adaptive evolution to the unique environmental conditions of the Arabian Sea and Gulf of Oman (Sheppard, 1993; Riegl & Purkis, 2012).

4.5 Technical Considerations and Recommendations

Our study highlights several technical considerations for coral mitochondrial genome assembly and annotation:

1. **Assembler choice**: NOVOPlasty performed well for organellar genome assembly, successfully reconstructing circular mitogenomes from WGS data. However, alternative assemblers (e.g., GetOrganelle, MitoZ) should be evaluated in future studies to assess robustness.
2. **Annotation validation**: Automated annotation tools such as MITOS2, while powerful, can produce artifacts in repetitive or structurally ambiguous regions. Manual curation using genome browsers (e.g., Artemis, IGV) and comparative analysis with reference genomes are essential steps.
3. **Sequencing depth**: While relatively low coverage (1,000–2,000×) is sufficient for mitochondrial genome assembly, deeper sequencing improves nuclear genome quality, which in turn enhances overall confidence in multi-compartment genome analyses.
4. **Quality control**: Integration of multiple quality metrics (BUSCO, QUAST, taxonomic profiling) provides a more comprehensive assessment of assembly reliability than any single metric alone.

4.6 Study Limitations

Several limitations should be acknowledged. First, our sample size (n=5) represents a relatively small snapshot of A. digitifera genetic diversity in Oman. Larger-scale sampling across multiple reef sites and environmental gradients would provide more robust insights into intraspecific mitochondrial genome variation.
Second, while we identified potential structural variations (split rrnS, duplicated OH_2), we did not perform experimental validation (e.g., RT-PCR, long-read sequencing) to definitively confirm whether these represent true biological variants or annotation artifacts. Future work should employ long-read sequencing technologies (PacBio or Oxford Nanopore) to generate high-quality, contiguous mitochondrial genome assemblies that resolve ambiguous regions.
Third, our study focused exclusively on mitochondrial genome structure and did not assess sequence-level variation (e.g., SNPs, indels) that may be informative for population genetics or phylogenetics. Future analyses should incorporate comparative sequence analyses with published A. digitifera mitogenomes from other geographic regions.

4.7 Conservation and Climate Change Implications

The generation of genomic resources for Omani coral populations has direct relevance to conservation and climate change research. Corals in the Arabian Sea and Gulf of Oman experience some of the most extreme environmental conditions globally, including seasonal temperature fluctuations exceeding 10°C, high salinity, and intense upwelling events (Reynolds, 1993; Riegl & Purkis, 2012).
These environmental stressors may drive adaptive evolution in local coral populations, potentially conferring resilience to climate change impacts such as ocean warming and acidification (Riegl et al., 2011). The mitochondrial genomes generated here serve as baseline resources for future studies investigating stress-responsive gene expression, metabolic adaptation, and population connectivity in the region.
Furthermore, understanding genomic diversity in peripheral populations such as those in Oman is critical for effective coral reef management and restoration efforts. If Omani populations harbor unique genetic variants adapted to extreme conditions, they may serve as valuable sources of stress-tolerant genotypes for assisted migration or selective breeding programs aimed at enhancing coral resilience (van Oppen et al., 2015; Baums et al., 2019).

5. Conclusion

This study presents a comprehensive comparative analysis of mitochondrial genomes from five Acropora digitifera samples collected from Omani coastal waters. Our findings confirm the high degree of structural conservation characteristic of scleractinian coral mitogenomes, with consistent gene content, order, and strand orientation across all samples. All mitochondrial genomes harbored the complete complement of 13 protein-coding genes, 2 ribosomal RNA genes, and 2 transfer RNA genes, with protein-coding genes predominantly oriented on the positive strand and rRNA genes on the negative strand.
Minor structural variations, including a split rrnS annotation in S3A and duplicated replication origin regions in S2D, were identified and are likely attributable to annotation artifacts or minor biological duplications rather than major structural rearrangements. Importantly, we observed a strong correlation between nuclear genome completeness (BUSCO scores) and mitochondrial genome annotation consistency, suggesting that overall assembly quality affects both genomic compartments.
Our integrative approach, combining annotation-informed circular genome maps, linear gene order visualizations, comparative dot plots, and genome completeness metrics, provides a robust framework for reliable mitochondrial genome characterization. This study not only reinforces the genomic stability of coral mitochondria as described in previous literature but also highlights the importance of rigorous quality control and manual annotation curation when working with organellar genomes.
The mitochondrial genome resources generated here contribute valuable baseline data for the understudied coral fauna of the northwestern Indian Ocean and support future research in coral phylogenomics, population genetics, and climate change adaptation. As coral reefs face unprecedented threats from global warming, ocean acidification, and local stressors, comprehensive genomic characterization of coral populations across their geographic range becomes increasingly essential for effective conservation and management strategies.

Acknowledgments

We thank the Ministry of Agriculture, Fisheries Wealth and Water Resources, Sultanate of Oman, for providing collection permits. We are grateful to field assistant Mr. Madhu for assistance with sample collection and for technical support. Sequencing was performed at MolSys Private Ltd, Bangalore, India. We acknowledge the funding agency, the Research Council of Oman (TRC-Ministry of Higher Education and Research Innovation (MOHERI)), for financial support.

Funding
This research was supported by The Research Council (TRC), Sultanate of Oman, under grant number [TRC-RG/2022/01]. Additional support was provided by the University of Technology and Applied Sciences – Sur - Oman.

Author Contributions
Conceptualization: S.S.K.
Methodology: S.S.K., [Co-author]
Sample Collection: S.S.K., [Co-author]
Laboratory Work: S.S.K., [Co-author]
Bioinformatics Analysis: S.S.K., [Co-author]
Data Visualization: S.S.K.
Writing – Original Draft: S.S.K.
Writing – Review & Editing: S.S.K., [Co-author]
Funding Acquisition: S.S.K.
Project Administration: S.S.K.
All authors have read and approved the final manuscript.
Conflicts of Interest
The authors declare no conflicts of interest.
Data Availability Statement
Raw sequencing data are available in the NCBI Sequence Read Archive (SRA) under BioProject accession [PRJNA1346827]. Assembled mitochondrial genomes have been deposited in GenBank under accession numbers. Bioinformatics scripts and supplementary data files are available at [GitHub repository link]. Further data and materials are available from the corresponding author upon reasonable request.

References
1. Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
2. Baums, I. B., Devlin-Durante, M. K., & LaJeunesse, T. C. (2019). New insights into the dynamics between reef corals and their associated dinoflagellate endosymbionts from population genetic studies. Molecular Ecology, 23(17), 4203–4215.
3. Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., et al. (2013). MITOS: Improved de novo metazoan mitochondrial genome annotation. Molecular Phylogenetics and Evolution, 69(2), 313–319. https://doi.org/10.1016/j.ympev.2012.08.023
4. Bongaerts, P., Riginos, C., Brunner, R., Englebert, N., Smith, S. R., & Hoegh-Guldberg, O. (2017). Deep reefs are not universal refuges: Reseeding potential varies among coral species. Science Advances, 3(2), e1602373. https://doi.org/10.1126/sciadv.1602373
5. Boore, J. L. (1999). Animal mitochondrial genomes. Nucleic Acids Research, 27(8), 1767–1780. https://doi.org/10.1093/nar/27.8.1767
6. Carver, T., Harris, S. R., Berriman, M., Parkhill, J., & McQuillan, J. A. (2012). Artemis: An integrated platform for visualization and analysis of high-throughput sequence-based experimental data. Bioinformatics, 28(4), 464–469. https://doi.org/10.1093/bioinformatics/btr703
7. Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: An ultra-fast all-in-one FASTQ preprocessor. Bioinformatics, 34(17), i884–i890. https://doi.org/10.1093/bioinformatics/bty560
8. Dierckxsens, N., Mardulyn, P., & Smits, G. (2017). NOVOPlasty: De novo assembly of organelle genomes from whole genome data. Nucleic Acids Research, 45(4), e18. https://doi.org/10.1093/nar/gkw955
9. Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: Summarize analysis results for multiple tools and samples in a single report. Bioinformatics, 32(19), 3047–3048. https://doi.org/10.1093/bioinformatics/btw354
10. Flot, J. F., & Tillier, S. (2007). The mitochondrial genome of the scleractinian coral Stylophora pistillata (Cnidaria, Anthozoa). Gene, 401(1-2), 80–87. https://doi.org/10.1016/j.gene.2007.06.022
11. Flot, J. F., Magalon, H., Cruaud, C., Couloux, A., & Tillier, S. (2008). Patterns of genetic structure among Hawaiian corals of the genus Pocillopora yield clusters of individuals that are compatible with morphology. Comptes Rendus Biologies, 331(3), 239–247.
12. Flot, J. F., Tillier, S., Samadi, S., & Tillier, A. (2010). Phase determination from direct sequencing of length-variable DNA regions. Molecular Ecology Resources, 10(5), 876–882.

13. Forêt, S., Seneca, F. O., de Jong, D., Bieller, A., Hemmrich, G., Augustin, R., et al. (2007). Phylogenomic classification of the Bilateria: Animal development meets environmental genomics. Science, 317(5834), 86–94.
14. Fukami, H., Chen, C. A., Budd, A. F., Collins, A., Wallace, C., Chuang, Y. Y., et al. (2008). Mitochondrial and nuclear genes suggest that stony corals are monophyletic but most families of stony corals are not (Order Scleractinia, Class Anthozoa, Phylum Cnidaria). PLoS ONE, 3(9), e3222.
15. Grant, J. R., Enns, E., Marinier, E., Mandal, A., Herman, E. K., Chen, C. Y., et al. (2023). Proksee: In-depth characterization and visualization of bacterial genomes. Nucleic Acids Research, 51(W1), W484–W492. https://doi.org/10.1093/nar/gkad326
16. Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: Quality assessment tool for genome assemblies. Bioinformatics, 29(8), 1072–1075. https://doi.org/10.1093/bioinformatics/btt086
17. Hellberg, M. E. (2006). No variation and low synonymous substitution rates in coral mtDNA despite high nuclear variation. BMC Evolutionary Biology, 6, 24. https://doi.org/10.1186/1471-2148-6-24
18. Kayal, E., Roure, B., Philippe, H., Collins, A. G., & Lavrov, D. V. (2013). Cnidarian phylogenetic relationships as revealed by mitogenomics. BMC Evolutionary Biology, 13, 5. https://doi.org/10.1186/1471-2148-13-5
19. Li, Y., Kennedy, E. V., Huang, D., & Muir, P. (2020). Improved genome assemblies uncover coral-specific genes and high sequence conservation across taxa. Communications Biology, 3, 282. https://doi.org/10.1038/s42003-020-0996-7
20. Lin, M. F., Luzon, K. S., Licuanan, W. Y., Ablan-Lagman, M. C., & Chen, C. A. (2011). Seventy-four universal primers for characterizing the complete mitochondrial genomes of scleractinian corals. Zoological Studies, 50(4), 513–524.
21. Lind, A. L., & Pollard, K. S. (2021). Accurate and sensitive detection of microbial eukaryotes from whole metagenome shotgun sequencing. Microbiome, 9(1), 58. https://doi.org/10.1186/s40168-021-01015-y
22. Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A., & Zdobnov, E. M. (2021). BUSCO update: Novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. Molecular Biology and Evolution, 38(10), 4647–4654. https://doi.org/10.1093/molbev/msab199
23. Medina, M., Collins, A. G., Silberman, J. D., & Sogin, M. L. (2006). Evaluating hypotheses of basal animal phylogeny using complete sequences of large and small subunit rRNA. Proceedings of the National Academy of Sciences, 98(17), 9707–9712. https://doi.org/10.1073/pnas.97.17.9707
24. Reynolds, R. M. (1993). Physical oceanography of the Gulf, Strait of Hormuz, and the Gulf of Oman—Results from the Mt Mitchell expedition. Marine Pollution Bulletin, 27, 35–59.
25. Riegl, B. M., & Purkis, S. J. (2012). Coral Reefs of the Gulf: Adaptation to Climatic Extremes. Springer.
26. Riegl, B. M., Purkis, S. J., Al-Cibahy, A. S., Abdel-Moati, M. A., & Hoegh-Guldberg, O. (2011). Present limits to heat-adaptability in corals and population-level responses to climate extremes. PLoS ONE, 6(9), e24802.
27. Romano, S. L., & Palumbi, S. R. (1996). Evolution of scleractinian corals inferred from molecular systematics. Science, 271(5249), 640–642.
28. Shearer, T. L., Van Oppen, M. J., Romano, S. L., & Wörheide, G. (2002). Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria). Molecular Ecology, 11(12), 2475–2487. https://doi.org/10.1046/j.1365-294X.2002.01652.x
29. Sheppard, C. R. C. (1993). Physical environment of the Gulf relevant to marine pollution: An overview. Marine Pollution Bulletin, 27, 3–8.
30. Shinzato, C., Shoguchi, E., Kawashima, T., Hamada, M., Hisata, K., Tanaka, M., et al. (2011). Using the Acropora digitifera genome to understand coral responses to environmental change. Nature, 476(7360), 320–323. https://doi.org/10.1038/nature10249
31. Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics, 31(19), 3210–3212. https://doi.org/10.1093/bioinformatics/btv351
32. van Oppen, M. J., Willis, B. L., van Vugt, H. W. J. A., & Miller, D. J. (2002). Examination of species boundaries in the Acropora cervicornis group (Scleractinia, Cnidaria) using nuclear DNA sequence analyses. Molecular Ecology, 9(9), 1363–1373. https://doi.org/10.1046/j.1365-294X.2000.01023.x
33. van Oppen, M. J., Oliver, J. K., Putnam, H. M., & Gates, R. D. (2015). Building coral reef resilience through assisted evolution. Proceedings of the National Academy of Sciences, 112(8), 2307–2313. https://doi.org/10.1073/pnas.1422301112
34. Veron, J. E. N. (2000). Corals of the World (3 volumes). Australian Institute of Marine Science, Townsville.

35. Zhang, Y. Y., Yang, S. Z., Wang, X., Huang, D. M., & Chen, Y. (2017). Complete mitochondrial genome of the stony coral Montipora cactus (Scleractinia: Acroporidae). Mitochondrial DNA Part B, 2(1), 106–107. https://doi.org/10.1080/23802359.2017.1280703
36. Zimin, A. V., Marçais, G., Puiu, D., Roberts, M., Salzberg, S. L., & Yorke, J. A. (2013). The MaSuRCA genome assembler. Bioinformatics, 29(21), 2669–2677. https://doi.org/10.1093/bioinformatics/btt476