1 **Genomes of the Caribbean reef-building corals *Colpophyllia natans, Dendrogyra***
2 ***cylindrus,* and *Siderastrea siderea***

3

4 Nicolas S Locatelli[1] and Iliana B Baums[1,2,3,4*]

5 [1]Department of Biology, The Pennsylvania State University, University Park, PA 16802, USA
6 [2]Helmholtz Institute for Functional Marine Biodiversity at the University of Oldenburg (HIFMB),
7 Carl von Ossietzky Universität Oldenburg, Ammerländer, Heerstraße 231, 26129 Oldenburg,
8 Germany
9 [3]Alfred Wegener Institute, Helmholtz-Centre for Polar and Marine Research (AWI), Am
10 Handelshafen Bremerhaven 27570, Germany
11 [4]Institute for Chemistry and Biology of the Marine Environment (ICBM), School of Mathematics
12 and Science, Carl von Ossietzky Universität Oldenburg, Ammerländer Heerstraße 114-118,
13 26129 Oldenburg, Germany
14

15 **Author addresses:**
16 Nicolas Locatelli
17 208 Mueller Laboratory
18 Department of Biology
19 University Park, PA 16802
20 locatelli@psu.edu
21
22 Iliana Baums (**Corresponding author***)
23 Marine Conservation
24 Helmholtz Institute for Functional Marine Biodiversity
25 Ammerländer Heerstraße 231
26 26129 Oldenburg, Germany
27 Iliana.baums@hifmb.de
28

29 **Running head:** Genomes of key Caribbean coral species

30

1

Abstract

Coral populations worldwide are declining rapidly due to elevated ocean temperatures and other human impacts. The Caribbean harbors a high number of threatened, endangered, and critically endangered coral species compared to reefs of the larger Indo-Pacific. The reef corals of the Caribbean are also long diverged from their Pacific counterparts and may have evolved different survival strategies. Most genomic resources have been developed for Pacific coral species which may impede our ability to study the changes in genetic composition of Caribbean reef communities in response to global change. To help fill the gap in genomic resources, we used PacBio HiFi sequencing to generate the first genome assemblies for three Caribbean, reef-building corals, *Colpophyllia natans, Dendrogyra cylindrus*, and *Siderastrea siderea.* We also explore the genomic novelties that shape scleractinian genomes. Notably, we find abundant gene duplications of all classes (e.g., tandem and segmental), especially in *S. siderea.* This species has one of the largest genomes of any scleractinian coral (822Mb) which seems to be driven by repetitive content and gene family expansion and diversification. As the genome size of *S. siderea* was double the size expected of stony corals, we also evaluated the possibility of an ancient whole genome duplication using Ks tests and found no evidence of such an event in the species. By presenting these genome assemblies, we hope to develop a better understanding of coral evolution as a whole and to enable researchers to further investigate the population genetics and diversity of these three species.

1 Introduction

2   Genomic resources are increasingly available for Pacific reef-building corals (e.g. Fuller *et*

3 *al.* 2020; Stephens *et al.* 2022), yet most Caribbean coral species still lack them despite genetic

4 management of populations becoming necessary (Baums et al. 2022). Caribbean reefs represent

5 ecosystems long diverged from Pacific counterparts. During the mid-Miocene, the Mediterranean

6 closed off at both ends and the eastern connection of the Caribbean with the Indo-Pacific basin

7 was severed (Wallace and Rosen 2006). The Isthmus of Panama to the west of the Caribbean

8 remained open until roughly 3 million years ago, after which ocean circulation drastically changed

9 and Caribbean reefs were isolated from Pacific reefs (Burton et al. 1997; O'Dea et al. 2016).

10   Cnidarians diverged early in metazoan evolution roughly 700 Mya (Park et al. 2012) and

11 the three species discussed here represent the two major scleractinian lineages, complex

12 (*Siderastrea siderea)* and robust *(Colpophyllia natans* and *Dendrogyra cylindrus). Dendrogyra*

13 *cylindrus* is a rare Caribbean coral (Hunter and Jones 1996) that has declined sharply in the past

14 two decades due anthropogenic stressors and a highly infectious disease called stony coral tissue

15 loss disease (Brandt *et al.* 2021). *Dendrogyra cylindrus* is extinct in the wild in Florida and is

16 considered critically endangered (Neely et al. 2021; Cavada-Blanco et al. 2022). *Siderastrea*

17 *siderea* and *C. natans* were common reef-building corals that have also experienced significant

18 declines in response to disease and anthropogenic impacts. *Siderastrea siderea* is now listed as

19 critically endangered (Rodriguez-Martinez et al. 2022) and under threat due to acidification,

20 ocean warming (Horvath et al. 2016), and stony coral tissue loss disease (Brandt et al. 2021).

21 *Colpophyllia natans* is also in decline due to stony coral tissue loss disease (Vermeij and Goergen

22 2022; Williamson et al. 2022). Despite their ecological and evolutionary importance, genomic

23 resources are not yet available for these species.

24   Coral genomes are variable in size (e.g., Stephens *et al.* 2022), but have highly conserved

25 gene order (Ying et al. 2018; Locatelli et al. 2024). Anthozoan genomes contain between 13.57%

26 and 52.2% repetitive content (e.g., Shinzato *et al.* 2011; Bongaerts *et al.* 2021) and harbor DNA

27 and retrotransposons that are still active (Chapman et al. 2010; Huang et al. 2012), which can

28 result in gene duplication and movement of genes to disparate regions of the genome.

29 Accumulation of somatic mutations in long-lived coral colonies represents another mechanism

by which coral genomes gain heterozygosity (Devlin-Durante et al. 2016; López and Palumbi 2020) and some of these mutations can be passed on to their sexually produced offspring (Vasquez Kuntz et al. 2022). Development of genomic resources allows for further study of these complex evolutionary mechanisms in metazoans as a whole (Reusch et al. 2021).

To help bridge the gap in genomic resources for Caribbean corals, we present novel PacBio HiFi-derived assemblies for *Colpophyllia natans*, *Dendrogyra cylindrus,* and *Siderastrea siderea.* With these references, we hope to foster an understanding of how corals will respond to environmental change (Bove et al. 2022) and population decline (Cramer et al. 2020), and how the response of Caribbean corals may differ from Indo-Pacific species.

Methods

Tissue sampling

Tissue of *Colpophyllia natans* ([12.1095, -68.95497], database ID 22254) was collected from the Water Factory reef in Curaçao on August 6th, 2022 using a hammer and chisel. *Dendrogyra cylindrus* ([12.0837, -68.89447], database ID 22255) and *Siderastrea siderea* ([12.0839, -68.8944], database ID 22256) were collected from the Sea Aquarium reef in Curaçao on August 12th and 13th, 2022 using hammer and chisel. All collections were made under Curaçao Governmental Permit 2012/48584. All fragments were ca. 12cm^2 in size and were kept alive in coolers filled with seawater during transit prior to being preserved in DNA/RNA Shield (Zymo Research, CA, USA). Samples were stored at -20°C or at -80°C until extraction.

Nucleic acid extraction and sequencing

For all species, DNA was extracted from tissue preserved in DNA/RNA Shield (Zymo Research, CA, USA) using the Qiagen (MD, USA) MagAttract HMW DNA kit, following manufacturer protocols. Following initial extraction, DNA was further purified using a 0.9X AMPure XP (Beckman Coulter, CA, USA) bead cleanup. Purified DNA was then size selected using a Pacific Biosciences (formerly Circulomics) SRE size selection kit. The SRE standard kit selects for DNA predominantly >25kb and a near total depletion of fragments <10kb. Barcoded templates were generated and sequenced by the Huck Institutes of the Life Sciences Genomics Core Facility

1 at Penn State University using a Pacific Biosciences (Menlo Park, CA, USA) Sequel IIe across a total

2 of three SMRTcells (further described below).

3   As RNAseq data was not available for *Dendrogyra cylindrus* or any close relatives for the

4 purposes of gene prediction, RNA was extracted from the same DNA/RNA Shield (Zymo Research,

5 CA, USA) preserved samples as described above using a TriZol and a Qiagen (MD, USA) RNeasy

6 Mini Kit (as in https://openwetware.org/wiki/Haynes:TRIzol_RNeasy). Compared with the RNA

7 sequence data obtained from NCBI SRA for *C. natans* and *S. siderea* (described below in "Gene

8 prediction and functional annotation"), the RNA sample for *D. cylindrus* was of an untreated

9 colony growing in the wild rather than experimental samples exposed to heat and disease-stress.

10 From the extracted total RNA, libraries were prepared and sequenced by the Oklahoma Medical

11 Research Foundation Clinical Genomics Center using the NEBNext® Poly(A) mRNA Magnetic

12 Isolation Module (New England BioLabs Inc., MA, USA), Swift Rapid RNA Library Kit (Swift

13 Biosciences, MI, USA), and 150M read pairs of 2x150bp chemistry on an Illumina (San Diego, CA,

14 USA) NovaSeq 6000 machine.

15 ***Genome assembly***

16   A PacBio library was generated by pooling the barcoded templates for each of the three

17 species in equal proportions and was initially sequenced on two SMRTcells. Prior to genome

18 assembly, k-mer (31-mer) counting was performed on PacBio HiFi data for each species using

19 Jellyfish v2.2.10 (Marçais and Kingsford 2011) for the purpose of haploid genome size estimation.

20 K-mer frequency-based genome-wide heterozygosity and genome size were estimated from 31-

21 mer histograms using GenomeScope2 (Ranallo-Benavidez et al. 2020). With the data from these

22 two initial SMRTcells, a preliminary assembly was performed using hifiasm_meta v0.2 (Feng et al.

23 2022) to assess assembly size and to determine whether the pool balance needed to be adjusted

24 for the third and final SMRTcell run.

25   Because the preliminary assembly and genome size estimate from GenomeScope2 of *S.*

26 *siderea* was larger than the remaining two species, the final SMRTcell was run with a pool balance

27 of 25:25:50 *Colpophyllia*:*Dendrogyra*:*Siderastrea* to provide additional coverage of the larger

28 *Siderastrea* genome. Prior to all stages of data delivery, the sequencing facility used PacBio lima

to demultiplex and remove adapters and unbarcoded sequences. Across all SMRTcells, total sequence yield was 26Gb across 2.8M reads in *Colpophyllia natans*, 25Gb across 2.7M reads in *Dendrogyra cylindrus*, and 32Gb across 3.4M reads in *Siderastrea siderea*. Further breakdown of PacBio yield and read lengths per species per sequencing run can be found in **Table S1**. Utilizing all data, a new set of primary assemblies was generated using hifiasm_meta.

Assembly decontamination, haplotig purging, and repeat annotation

HiFi reads were then mapped to the assembly using minimap2 v2.24 (Li 2018) and BAM files were sorted using samtools v0.1.19 (Danecek et al. 2021). Using blastn v2.14.0 (Camacho et al. 2009), assemblies were searched against a custom database comprised of NCBI's ref_euk_rep_genomes, ref_prok_rep_genomes, ref_viroids_rep_genomes, and ref_viruses_rep_genomes databases combined with dinoflagellate and *Chlorella* genomes (Shoguchi et al. 2013; Hamada et al. 2018; Shoguchi et al. 2018; Beedessee et al. 2020; Shoguchi et al. 2021). All NCBI RefSeq databases were downloaded on March 28th, 2023. Using the mapping and blastn hits files, blobtools v1.1.1 (Laetsch and Blaxter 2017) was used to identify and isolate non-cnidarian contigs. To better identify symbionts within the metagenome assemblies, blastn (Camacho et al. 2009) was used to query putative Symbiodiniaceae contigs against a curated nuclear ribosomal Internal Transcribed Spacer-2 (ITS2) database (Hume et al. 2019). With all non-cnidarian contigs excluded, a repeat database was modeled using RepeatModeler2 v2.0.2a (Flynn et al. 2020). Purge_dups v1.2.6 (Guan et al. 2020) was utilized to identify and remove any remaining putative haplotigs in the respective assemblies. Following haplotig purging, repeats were soft-masked using a filtered repeat library in RepeatMasker4 v4.1.2.p1 (Smit et al.), following recommendations from the Blaxter Lab (https://blaxter-lab-documentation.readthedocs.io/en/latest/filter-repeatmodeler-library.html). Protein references from *Orbicella faveolata* (Prada et al. 2016) and *Fungia sp.* (Ying et al. 2018) were used to filter repeat libraries for the two robust species (*C. natans* and *D. cylindrus*). Protein references from *Acropora millepora* (Fuller et al. 2020)*, Montipora capitata* (Stephens et al. 2022), and *Galaxea fascicularis* (Ying et al. 2018) were used to filter repeat libraries for *S. siderea*.

1 ***Gene prediction and functional annotation***

2   Prior to gene prediction, the hifiasm_meta assemblies were scanned for mitochondrial

3 contamination using MitoFinder v1.4.1 (Allio et al. 2020) and contigs of mitochondrial origin were

4 removed from the assemblies. Nuclear assemblies were annotated using RNAseq data in

5 funannotate v1.8.13 (Palmer and Stajich 2020). *Colpophyllia natans* and *Siderastrea siderea* were

6 annotated using all RNAseq data available on NCBI SRA for the respective species at the time of

7 assembly (see **Table S2**). As no RNAseq data is publicly available for *Dendrogyra cylindrus* or its

8 close relatives, so RNA was extracted as previously described and included within the funannotate

9 annotation process. All RNAseq data was adapter- and quality-trimmed using TrimGalore v0.6.7

10 (Krueger et al. 2021).

11   Briefly, funannotate train was run for all assemblies with a *--max_intronlen* of 100000.

12 Funannotate train is a wrapper that utilizes Trinity (Grabherr et al. 2011) and PASA (Haas et al.

13 2008) for transcript assembly. Upon completion of training, funannotate predict was run to

14 generate initial gene predictions using the arguments *--repeats2evm*, *--organism other*, and *--*

15 *max_intronlen 100000*. Funannotate predict is a wrapper that runs AUGUSTUS (Stanke et al.

16 2006) and GeneMark (Brůna et al. 2020) for gene prediction and EvidenceModeler (Haas et al.

17 2008) to combine gene models. Funannotate update was run to update annotations to be in

18 compliance with NCBI formatting. For problematic gene models, funannotate fix was run to drop

19 problematic IDs from the annotations. Finally, functional annotation was performed using

20 funannotate annotate which annotates proteins using PFAM (Bateman et al. 2004), InterPro

21 (Hunter et al. 2009), EggNog (Huerta-Cepas et al. 2019), UniProtKB (Boutet et al. 2016), MEROPS

22 (Rawlings et al. 2009), CAZyme (Huang et al. 2018), and GO (Harris et al. 2004). For all genes not

23 functionally annotated with gene ontology (GO) terms by funannotate, a single network of

24 ProteInfer (Sanderson et al. 2023) was used to infer functional attributes of genes using pre-

25 trained models.

26   To assess the quality of genome assemblies and annotations, BUSCO v5.8.0 (Manni et al.

27 2021) was run with the metazoa_odb10 lineage dataset. BUSCO was run in genome mode on the

28 full genome assembly, and in protein mode on the predicted proteins dataset output by

29 funannotate.

1 *Mitochondrial genome assembly*

2 To assemble mitochondrial genomes for each samples, MitoHiFi v2.2 (Gabriel et al. 2023)

3 was used on all available HiFi data for each species. For *Siderastrea siderea, Colpophyllia natans,*

4 and *Dendrogyra cylindrus*, accessions NC_008167.1, NC_008162.1, and DQ643832.1 (whole

5 mitogenomes for *Siderastrea radians*, *Colpophyllia natans*, and *Astrangia poculata*), were used

6 as seed sequences for mitochondrial assembly, respectively. For all assemblies, the arguments -a

7 animal and -o 5 were used to indicate that the organism type was an animal and the organism

8 genetic code was invertebrate.

9 *Duplication and orthogroup analysis*

10 To assess the origin of gene duplications, whole genome duplication pipeline and

11 orthogroup analyses were used. The wgd pipeline v1.1 (Zwaenepoel and Van De Peer 2019) was

12 used to investigate duplication and divergence at the whole paranome and anchor-pair levels.

13 The longest, coding CDS transcript of each gene was used as input for wgd. The wgd pipeline acts

14 as a wrapper for a number of programs, and in the case of the analysis here the following

15 programs were run through wgd: blastp (Altschul et al. 1997), MCL (Markov Cluster Process,

16 Hazewinkel and Van Eijck 2000), PAML (Yang 2007), MAFFT (Katoh and Standley 2013), FastTree

17 (Price et al. 2010), and i-ADHoRe 3.0 (Proost et al. 2012). In addition to wgd, OrthoFinder v2.5.4

18 (Emms and Kelly 2019) was run to discover orthologous groups unique to each species and shared

19 between species. For OrthoFinder analyses, the longest peptide isoform for each gene was used

20 as input. A full list of taxa included in OrthoFinder and doubletrouble analyses (described below)

21 can be found in **Table S3.**

22 CAFE5 v5.1.0 (Mendes et al. 2021) was used to discover hierarchical orthogroups from

23 OrthoFinder undergoing phylogenetically significant gene family expansions or contractions. To

24 begin, r8s v1.81 (Sanderson 2003) was used to time-calibrate the phylogeny from OrthoFinder

25 using fossil priors obtained from the PaleoBioDB fossil record (Peters and McClennen 2016). Priors

26 for *Acropora palmata* (5.3Mya, Budd *et al.* 1999), *Porites compressa* (2.588Mya, Faichney *et al.*

27 2011), *Acropora* (59Mya, Vecsei and Moussavian 1997), Faviina (247Mya, Qi 1984), and

28 Scleractinia (268Mya, Gregorio 1930), were used as calibration points. With significantly

29 expanding or contracting hierarchical orthogroups identified by CAFE5, GO terms for genes in

expanding and contracting gene families were extracted and compared to the whole genome background in each species to test for enrichment. Enrichment analyses were performed using GOAtools (Klopfenstein et al. 2018) on genes in expanding gene families not annotated as transposons and transposases. Genes functionally annotated with the transposition GO term (GO:0032196) and its child terms were also excluded. GO term enrichment was also assessed for orthogroups unique to each species (unshared orthogroups). To reduce false discovery, only terms with a Benjamini-Hochberg adjusted p-value < 0.05, depth > 2, and terms present in more than 5 study genes (all genes present in expanding orthogroups) were preserved. As GOAtools propagates child term counts up to parent terms, results can contain high redundancy and semantic similarity. To reduce some of the redundancy in significant GO terms, REVIGO v1.8.1 (Supek et al. 2011) was run using the SimRel semantic similarity measure to simplify enrichment results.

To classify stony coral (Scleractinia) paralogs into duplication types, doubletrouble v1.3.6 (Almeida-Silva and Van de Peer 2025) was run using the longest peptide isoform for each gene and default arguments. Briefly, doubletrouble classifies genes into segmental (SD), tandem (TD), proximal (PD), transposon-derived (TRD), and dispersed duplications (DD) based on collinearity, intron content, and phylogenetic position of paralogs. For instance, duplications are classified as tandem if two paralogs are separated by fewer than ten genes. If the distance between genes is >10, paralogs are classified as proximal duplications. Dispersed duplications (DD) are considered any duplication that is not otherwise classifiable into more specific categories. For all doubletrouble analyses, *Amplexidiscus fenestrafer* (Wang et al. 2017), a member of the naked corals, Corallimorpharia, was used as an outgroup. Not all gene annotations were compatible with the "full" scheme, where transposon-derived duplications are further classified into retrotransposon-derived (rTRD) and DNA transposon-derived (dTRD). As such, the "full" scheme was only utilized for the focal study species here, *Colpophyllia natans, Dendrogyra cylindrus,* and *Siderastrea siderea*. All other species were run using the "extended" scheme.

Results and Discussion

Assembly contiguity, completeness, and heterozygosity

All assemblies exhibit high contiguity (**Table 1**) and are gap-free. The *S. siderea* genome is roughly two times larger than observed in other corals species, with an assembly size of 822M, compared with 526Mb and 399Mb for *D. cylindrus* and *C. natans*, respectively. The assembly size of *S. siderea* is larger than most publicly available coral genome assemblies – only two species have larger assemblies, *Pachyseris speciosa* (Bongaerts et al. 2021) and *Platygyra sinensis* (Pootakham et al. 2021). However, the *Platygyra sinensis* assembly likely contains considerable haplotig duplication, leaving only *Pachyseris speciosa* as a comparable assembly. In addition to being the largest of the three assemblies presented here, the *S. siderea* assembly is the most contiguous assembly (N50=9.1Mb), likely due to the larger read N50 of SMRTcell 3 (see **Table S1**). The genomes of *C. natans* and *D. cylindrus* have N50s of 4.647Mb and 4.902Mb, respectively. Further scaffolding with Hi-C data could help elevate these three references to chromosome-level. Genome-wide GC content is similar across all three species, with 39.81% for *S. siderea*, 38.87% for *C. natans,* and 39.29% for *D. cylindrus*. GC estimates are similar to other published stony coral genomes (e.g. Bongaerts *et al.* 2021).

K-mer duplicity plots from GenomeScope2 (**Fig. 1**) suggest that all species here are diploid in nature, unlike the recent findings in Hawaiian corals (Stephens et al. 2022). All three assemblies exhibited high completeness as determined by BUSCO Metazoa in genome mode (Manni et al. 2021), with *C. natans*, *D. cylindrus*, and *S. siderea* showing 97.2%, 96.5%, and 96.6% completeness, respectively (**Table 1**). In terms of core BUSCO genes, *S. siderea* has the highest number of duplicated genes, with 2.7% of metazoan genes being duplicated. In genome mode, BUSCO showed that all three assemblies had similar percentages of fragmented and missing BUSCO genes. Evaluation of all protein isoforms using BUSCO in protein mode suggested that the complexity of the *S. siderea* genome may have slightly reduced the efficacy of genome annotation compared with the remaining two species, with completeness scores of 95.5%, 94.9%, and 91.8% for *C. natans, D. cylindrus,* and *S. siderea*, respectively. All assemblies are similar to their GenomeScope2 k-mer-based size estimates (**Fig. 1** and **Table 1**). Taken together, these results

1 suggest that the majority of all three genomes were successfully captured and annotated in our

2 assemblies with little remaining haplotig duplication.

3   Genome-wide heterozygosity in corals typically ranges from 1.07% to 1.96% (Shinzato et

4 al. 2021; Stephens et al. 2022; Yu et al. 2022; Young et al. 2024). K-mer frequency-based estimates

5 of genome-wide heterozygosity from GenomeScope2 suggest that *Dendrogyra cylindrus* has the

6 lowest heterozygosity of the three species discussed here (0.799%) and among the lowest in any

7 coral species for which genomic resources are available (Shinzato et al. 2021; Stephens et al. 2022;

8 Yu et al. 2022; Young et al. 2024). The species has been rare throughout history (Hunter and Jones

9 1996; Modys et al. 2023) but with high local abundances in some locations (e.g. St. Thomas in the

10 U.S. Virgin Islands). Recent catastrophic declines due to stony coral tissue loss disease (Neely et

11 al. 2021; Alvarez-Filip et al. 2022) have led to the listing of the species as critically endangered by

12 the International Union for Conservation of Nature (IUCN, Cavada-Blanco *et al.* 2022). *Dendrogyra*

13 *cylindrus* is extinct in the wild in Florida (Neely et al. 2021) and all genets are now in captivity.

14 Captive-based spawning efforts are burgeoning (Craggs et al. 2017; O'Neil et al. 2021) to recover

15 the species. The very low heterozygosity estimate provided here highlights the need for carefully

16 managed breeding (Marhaver et al. 2015) to ensure the persistence of the remaining standing

17 genetic variation and adaptive potential of *D. cylindrus* (Barrett and Schluter 2008; Kardos et al.

18 2021). Of the three species, *Siderastrea siderea* has the highest genome-wide heterozygosity

19 estimate of 1.59% and *C. natans* is intermediate with 0.862%. *Colpophyllia natans* also has low

20 genome-wide heterozygosity compared to other coral species and may require genetic

21 management in the future. However, these genome-wide heterozygosity estimates are generated

22 from singular genets and may not accurately represent the heterozygosity of the wider

23 populations of each species. *Colpophyllia natans* is the only species discussed here that does not

24 have range-wide population genetic information available. As such, further genetic

25 characterization of the species is clearly warranted due to population declines caused by

26 infectious diseases (Alvarez-Filip et al. 2022) and the heterozygosity estimates provided here.

27 ***Repetitive content and transposable elements***

28   The proportion of repeats assigned to each repeat category of *RepeatMasker* was similar

29 across all three species assembled here (**Table 2**). Repetitive content was 47.80%, 40.40%, and

1 23.62% in *S. siderea, D. cylindrus,* and *C. natans*, respectively. The majority of repeats were

2 interspersed, with unclassified repeats being most abundant in all three species (31.91%, 25.57%,

3 and 12.22%). Compared with other cnidarians, these assemblies contain similar levels of

4 repetitive content to jellyfish species such as members of *Clytia, Aurelia,* and *Chrysaora*

5 containing 39-49.5% (Gold et al. 2019; Leclère et al. 2019; Xia et al. 2020) and to other

6 scleractinian corals containing 13.6-58.1% (e.g., Shinzato et al. 2011; Cooke et al. 2020; Locatelli

7 et al. 2024, Bongaerts *et al.* 2021; Stephens *et al.* 2022; Kim *et al.* 2022; Young *et al.* 2024). In our

8 set of three species, we observe a general relationship of increasing repetitive content with

9 increasing with genome size, corroborating that repeat expansion may be important in driving

10 genome size disparities across evolutionary time in stony corals, as similarly observed in

11 Zoantharian and *Hydra* genomes (Wong et al. 2019; Fourreau et al. 2023).

12 In *C. natans*, the DNA transposon class of repeats was reduced by approximately 50%

13 when compared with *S. siderea* and *D. cylindrus* (**Table 2**). Within the DNA transposons, the

14 Maverick subclass represented the most prominent deficits in *C. natans*, representing only 1.41%

15 of the genome compared with 5.20% and 6.74% in *S. siderea* and *D. cylindrus,* respectively. It is

16 unclear whether DNA transposons have contracted in *C. natans* or expanded in *S. siderea* and *D.*

17 *cylindrus*, although both purging and expansions of repetitive content have been implicated in

18 genome size evolution in cnidarians and other organisms (Hawkins et al. 2006; Michael 2014;

19 Roessler et al. 2019; Kon et al. 2024). Further work is required to understand the genomic

20 processes by which repetitive DNA expands and contracts in cnidarian genomes, as well as the

21 overall importance of repetitive content in speciation processes and establishing new lineages.

22 However, the genome assemblies presented here echo the standing literature and suggest that

23 losses or gains of certain repetitive classes exist across the diversity of extant stony corals.

24 ### *Gene prediction and unique orthogroups*

25 *S. siderea* is unique amongst the assembled genomes not just for its size and contiguity,

26 but also its gene content. Gene prediction in funannotate identified 61,712 gene models, roughly

27 double the number of genes discovered for *D. cylindrus* and *C. natans* (39,739 and 34,139,

28 respectively; **Table 1**), and compared to other publicly available coral genome assemblies (e.g.,

29 Prada *et al.* 2016; Fuller *et al.* 2020). Of these gene models, 52,473, 34,738, and 29,090 were

predicted to be protein-coding for *S. siderea, D. cylindrus,* and *C. natans*, respectively. *Dendrogyra cylindrus* and *C. natans* fall within the expectations for stony corals in terms of protein-coding gene content. The gene content of *S. siderea* is higher than expected, only comparable to *Montipora capitata* amongst published genomes (Stephens et al. 2022). Of the protein-coding gene models, 1,515, 297, and 287 models in *S. siderea, D. cylindrus,* and *C. natans* contained >=90% repeat-masked bases, suggesting that these models may be derived from repetitive DNA and transposition-related events. A further 79, 47, and 31 gene models in *S. siderea, D. cylindrus,* and *C. natans* were either directly annotated as transposons or transposases or were associated with transposition (GO:0032196) or transposase activity (GO:0004803) related GO terms.

Because of the doubling in overall size and gene content present in the *S. siderea*, Ks tests were performed to test for an ancient whole genome duplication in the evolution of the species. Ks distributions in species having experienced whole genome duplication events exhibit characteristic distributions with a hump (as in Zwaenepoel and Van De Peer 2019), where many gene pairs are derived from a simultaneous duplication event and have all experienced a similar number of synonymous substitutions per synonymous site. Whole genome duplication analyses in wgd did not find Ks ratios indicative of ancient whole genome duplication in any of the species assembled here (**Fig. S1**), suggesting that other processes may be responsible for gain in genome size. BUSCO completeness, as described above, also suggested that duplication of metazoan single copy genes in the *S. siderea* genome is minimal, further reducing support for a whole genome duplication event.

Orthofinder analyses placed 47,786 protein-coding genes into 21,970 orthogroups in *S. siderea*. Of these, 1,004 orthogroups containing 3,849 protein-coding genes were exclusively found in *S. siderea* (**Fig. 2**). An additional 4,687 protein-coding genes could not be binned into orthogroups by Orthofinder. As Orthofinder utilizes DIAMOND (Buchfink et al. 2015) with the --more-sensitive alignment option, orthogroups are only formed if inter- and intraspecies alignments are >=40% in identity. The presence of thousands of unbinned genes and orthogroups unique to *S. siderea* suggests that gene duplication and subsequent diversification is prominent in the lineage. Amongst multi-copy gene families unique to *S. siderea* (**Fig. S2**), the most enriched GO term compared to the genomic background was "bioluminescence" (GO:0008218).

Fluorescent pigment proteins have been shown to undergo rapid evolution and strong selection in corals (Voolstra et al. 2011). These proteins are photoprotective for the coral holobiont (Salih et al. 2000) and can serve to optimize the light environment of symbiotic Symbiodiniaceae (Bollati et al. 2022). Indeed, presence of pink fluorescent pigment in congener *S. stellata* is associated with higher temperatures (Tunala et al. 2023). *Siderastrea siderea* harbors three distinct genetic lineages (Aichelman et al. 2025) of which only one was sequenced here. Additional genome assemblies of the other two lineages may shed light on the taxonomic status of these lineages and what role gene duplication and diversification may have played in their evolution.

Dendrogyra cylindrus had comparatively fewer protein-coding genes placed into a similar number of orthogroups – 32,896 genes in 19,251 orthogroups. Of these, 233 orthogroups were unique to the species, containing a total of 1,061 genes. 1,842 genes remained unbinned. Of the orthogroups unique to *D. cylindrus*, the most enriched GO term was "response to defense of other organism" (GO:0098542, **Fig. S2**). The lack of specificity of the term makes it unclear whether this refers to external organisms (e.g. damage from predation or disease) or internal organisms (e.g., intracellular Symbiodiniaceae symbionts). However, a child term of GO:0098542, "defense response to bacterium" (GO:0042742), is found amongst the genes in gene families unique to *D. cylindrus*, suggesting that immune response to infection is particularly important to the species. Other enriched GO terms such as "apoptotic process" and "regulation of response to external stimulus" further support that response to bacterial infection may be particularly important to the species. Given its lineage age, *Dendrogyra cylindrus* has been suggested to be intrinsically better at fighting infections compared with younger lineages (Pinzón et al. 2014). Recent catastrophic losses of the species due to stony coral tissue loss disease may have broken this long-standing advantage (Alvarez-Filip et al. 2022), although some evidence suggests that the disease may be the result of an infection of the symbiont that cascades to affect the host (Klein et al. 2024), rather than directly infecting the host.

Colpophyllia natans had the fewest orthogroups, with 27,892 protein-coding genes placed into 18,430 orthogroups. 545 genes were placed into 164 orthogroups that were unique to the species and 1,198 genes remained unbinned. There was an enrichment for terms relating to tRNA modification (GO:0002949 and GO:0070525) in the gene families only found in *C. natans* (**Fig. S2**).

1   "Bioluminescence" also appears amongst the most enriched GO terms in orthogroups unique to

2   the species, similar to *S. siderea*. Additionally, several terms relating to growth and development

3   ("blastocyst growth" and "anatomical structure maturation") were found to be enriched amongst

4   orthogroups unique to *C. natans*. *C. natans* is amongst the most quickly developing broadcast

5   spawners in the Caribbean, with settlement and the onset of zooplanktivory occurring in as little

6   as 3-4 days (Geertsma et al. 2022; Yus et al. 2024), possibly due to the enrichment of growth

7   related terms observed here. The "regulation of pH" is also enriched, perhaps allowing the

8   species to survive in environments less conducive to survival in other species. For instance, *C.*

9   *natans* is one of the few coral species able to thrive in unusual habitats such as mangrove canopy

10   environments with comparatively low pH, as well as reef flats and reef slope environments more

11   typically associated with Caribbean reef communities (Stewart et al. 2022).

12

13   ***Mitochondrial genomes***

14   Mitochondrial genomes were successfully assembled for all three species discussed here

15   using MitoHiFi (Gabriel et al. 2023). Both *D. cylindrus* and *C. natans* were of similar size with

16   lengths of 17,299bp and 17,104bp, respectively. *S. siderea* is considerably larger, with a total

17   length of 19,387bp (**Fig. 1**). The *S. siderea* mitogenome is among the largest of all stony coral

18   (Scleractinia). Of all sequenced scleractinians, the mitogenome of *S. siderea* is exceeded in length

19   only by the solitary coral species *Polymyces wellsi* (Flabellidae, NC_082103.1, 19,924bp),

20   *Deltocyathus magnificus* (Deltocyathidae, OR625187.1, 19,736bp), and *Rhombopsammia*

21   *niphada* (Micrabaciidae, MT706034.1, 19,654bp), and colony-forming species *Pseudosiderastrea*

22   *formosa* and *P. tayami* (Siderastreidae, NC_026530.1 and NC_026531.1, 19,475bp). In terms of

23   gene structure, all three mitochondrial genome assemblies consist of thirteen protein-coding

24   genes and two ribosomal RNA (rRNA, rrnL and rrnS) genes with highly conserved gene order

25   (ND5, ATP8, COX1, rrnL, ND1, CYTB, ND2, ND6, ATP6, ND4, rrnS, COX3, COX2, ND4L, and ND3).

26   Both *D. cylindrus* and *C. natans* contain twelve transfer RNA (tRNA) genes while *S. siderea* contains

27   eleven.

Expansion of shared gene families and modes of duplication

Gene ontology (GO) enrichment analyses of shared gene families undergoing phylogenetically significant expansion (as identified by OrthoFinder and CAFE5) may point to the importance of specific functional attributes in the evolution of each of the taxa assembled here (**Fig. 3**). In all three species, there was an enrichment amongst significantly expanding gene families for GO terms relating to nucleosome assembly, chromosome condensation, and chromatin/heterochromatin organization when compared to the genomic background of each species. This suggests that these functional attributes were disproportionately important in the evolution of stony corals. Chromatin accessibility is important for fine-tuning transcriptional response (GO:0006351 and GO:0006366, enriched in *C. natans*), as well as DNA repair (GO:0097510, enriched in *S. siderea*) and recombination (GO:0045910 and GO:0015074, enriched in *D. cylindrus* and *C. natans*, respectively) (Tsompana and Buck 2014). Experiments in the model sea anemones, *Nematostella* and *Aiptasia* further corroborate this hypothesis, demonstrating that chromatin accessibility is dynamic over the course of stressful events such as heat exposure and resulted in expressional changes in pathways related to immune response, oxidative stress, metabolism, and DNA repair (Weizman and Levy 2019; Weizman et al. 2021).

In *Siderastrea siderea,* "phosphatidylserine exposure on apoptotic cell surface" (GO:0070782) is the most enriched GO term in gene families that are significantly expanding compared to the genomic background of the species (**Fig. 3**). Additionally, the terms "engulfment of apoptotic cell" (GO:0043652) and "apoptotic process involved in development" (GO:1902742) were also enriched in *S. siderea*. Stress response in corals involves the activation of apoptotic pathways, particularly in the case of heat stress response (Kvitt et al. 2011; Tchernov et al. 2011; Helgoe et al. 2024) and *S. siderea* is amongst the most heat tolerant corals inhabiting Caribbean reefs (Palacio-Castro et al. 2021). The expansion of apoptosis-regulating gene families may be, in part, responsible for the overall resilience of *S. siderea* to adverse environmental conditions. In addition to apoptotic pathways and chromatin structure, there was also an enrichment of multiple myosin- and muscle-related ontology terms (GO:0031035, GO:0031033, GO:0051146). Myosin proteins have been identified as differentially expressed or differentially concentrated across inshore-offshore gradients, in heat stress conditions, and in diseased tissue across

divergent coral taxa and may also play a role in *S. siderea*'s resilience (DeSalvo et al. 2010; Ricaurte et al. 2016; Wong et al. 2021; Mayfield 2022).

The two GO terms with the highest fold enrichment in *Dendrogyra cylindrus* both involve pathways of the p53 class mediator (GO:1901796 and GO:0043516), one of which involves the regulation of DNA damage response. Sessile, shallow-living marine organisms are exposed to high levels of UV radiation and corals have fast and effective DNA repair mechanisms at all life stages (Reef et al. 2009; Svanfeldt et al. 2014). Melanosome organization (GO:0032438) was also found highly enriched in *D. cylindrus*. Melanin production is important in cnidarian innate immunity (Palmer et al. 2012; Changsut et al. 2022; Van Buren et al. 2024 Jul 18) and may serve to protect shallow living corals from UV exposure (Wall et al. 2018), and also protect their symbionts (Harman et al. 2022). *Dendrogyra cylindrus* is a long-lived species and even colonies in early development with no vertical pillar formation may be older than 30 years (Neely et al. 2021). This longevity may explain the enrichment in processes that reduce UV exposure and repair DNA damage that accumulates during the life of a genet.

Compared with other species in the analysis, gene families most expanded in *Colpophyllia natans* were enriched in functions related to environmental response ("response to osmotic stress", GO:0006970) as well as cell signaling ("Notch signaling pathway", GO:0007219 and "cell-cell adhesion", GO:0098609) and immune response ("defense response", GO:0006952). As described above, *C. natans* is able to persist in habitats such as mangrove stilt roots. These environments are often low in pH and low in salinity, and expansions of gene families relating to osmotic stress response may enable the species to thrive in these challenging habitats. The remainder of the expanded gene families in *C. natans* were involved in chromatin accessibility and nucleosome assembly (**Fig. 3**) as in *S. siderea* and *D. cylindrus*.

Subsequent analysis of paralogs using doubletrouble found that proximal duplications (locally duplicated with paralogs separated by ten or more genes) were the most prominent form of classifiable gene duplications in *Siderastrea siderea* (**Fig. S3** and **Table S4**). Previous studies have suggested that tandem duplications drive Scleractinian (stony coral) evolution (Noel et al. 2023). Indeed, tandem duplications appeared to be more abundant in *S. siderea* in comparison

1   with many of the evaluated taxa (**Fig. S3** and **Table S4**). However, duplicate classification is

2   inherently challenging as the order of genes can be the result of many different potential

3   processes. For instance, tandem duplications can be broken apart by dispersed duplications being

4   copied between tandem paralogs. These would resemble proximal duplications according to

5   doubletrouble's classification schema, despite being the result of two separate duplication

6   processes. Additionally, analyses comparing species are somewhat reliant on similarly high-

7   quality annotation and assembly across analyzed taxa. Several of the assemblies evaluated in our

8   duplication analyses are of low contiguity and filled with short-read derived gaps, which could

9   reduce the ability to detect certain forms of duplication. For example, *Orbicella faveolata* (Prada

10   *et al.* 2016) contains no segmental duplications (**Table S4**), potentially because the detection of

11   collinear, duplicated blocks of genes is less likely when the genome is highly fragmented. Further,

12   it may not be possible to assign duplicates as transposon-derived (TRD) with assemblies derived

13   from Nanopore or PacBio CLR data (e.g., *Acropora cervicornis,* Locatelli et al. 2024). Even polished

14   long read assemblies may contain enough error in repetitive proteins such that a single copy of

15   the gene cannot be assigned as ancestral – a requirement for paralogs to be classified as TRDs.

16   Despite the expansion of duplicated genes in Scleractinian species with larger genome

17   sizes (e.g., *Siderastrea siderea* and *Montipora capitata*, **Fig. S3**), tandemly duplicated genes do

18   not appear to have a disproportionate impact on genome size or gene content as suggested

19   previously (Noel et al. 2023). When all duplicates are scaled to a value of 1 (**Fig. S4**), no singular

20   duplication category appears to be most important in governing coral genome size. Instead, the

21   proportion of paralogs assigned to each duplication type is similar across all species (an average

22   of 22.0% tandem, 14.5% proximal, 2.3% segmental, 19.0% transposon-related, and 42.2%

23   dispersed, **Table S4**). This suggests that all duplication types are expanding in synchrony to result

24   in the genome size disparities we see across the phylogeny of Scleractinia. Further expansion of

25   duplication analyses to include assemblies from upcoming efforts of large database projects (e.g.,

26   Reef Genomics, Liew *et al.* 2016; Aquatic Symbiosis Genomics Project, McKenna *et al.* 2021) could

27   help elucidate more fine-scale, lineage-specific duplication processes that we have been unable

28   to capture here.

Symbiont contigs

As metagenome assemblers were utilized in the assembly of the host species, symbiont data was also co-assembled and was of sufficient coverage to identify the prominent symbiont present to at least the genus-level. Both *C. natans* and *D. cylindrus* contained *Breviolum*, with *D. cylindrus* most likely containing *B. dendrogyrum*, as described in (Lewis et al. 2019). However, the top ITS2 hits (determined by e-value, followed by percent identity) for both species do not closely match formally named strains/species in the curated ITS2 database (*C. natans* top symbiont hit B4, 89.89%, e-value 3.33e-24; *D. cylindrus* top hit B1, 97.98%, e-value 2.21e-42). It is possible that the symbionts contained in the genome assembly samples of *C. natans* and *D. cylindrus* are not yet represented in this database.

In the initial separation of host and symbiont contigs using BlobTools, the *S. siderea* genet assembled here was found to be associated with *Cladocopium*, but comparison of contigs with the ITS2 database did not reveal any more specific hits. The psbA region is a more reliable marker for symbiont strain identification than ITS2 (LaJeunesse and Thornhill 2011). However, symbiont reference sequences for psbA are not currently as extensive as ITS2 in strain coverage. As the ITS2 and psbA databases continue to grow, symbiont contigs assembled here could be identified with greater taxonomic resolution.

In addition to eukaryotic algal symbionts, one notable prokaryotic symbiont was recovered. Within the assembly for *C. natans*, a 2.13Mb contig was identified as most closely related to *Prosthecochloris aestuarii*. This bacterium has been proposed as a putatively symbiotic microbe living within coral skeletons (Cai et al. 2017; Chen et al. 2021). Coral metagenomes contain a wealth of symbionts with important functions for the holobiont (Bourne et al. 2009; Thompson et al. 2015; Boilard et al. 2020; Garrido et al. 2021). Further exploration of coral associated microbial communities may identify novel associations that are critical for the survival of the coral host.

Conclusion

Here, we generated novel genome assemblies for key Caribbean reef-building corals, all of which are listed as vulnerable or critically endangered by the IUCN. All genome assemblies are

1 highly complete (>95% BUSCO Metazoa) and contiguous (N50 > 4.6Mb). The genomes of

2 *Dendrogyra cylindrus* and *Colpophyllia natans* fall within nominal expectations of size and gene

3 content based on other published coral genomes. *Siderastrea siderea* is roughly two times larger

4 than expected with twice the number of predicted gene models, despite no evidence for a whole

5 genome duplication event. Repeat and gene family expansions seem to be drivers of the larger *S.*

6 *siderea* genome size. These results align with and expand upon previously published literature

7 which implicated gene duplications as a driving factor of stony coral evolution (Noel et al. 2023).

8 Given the importance of duplications in speciation across corals, further work should explore

9 intraspecific structural polymorphisms (such as copy number variants, CNVs) to understand how

10 structural variation plays a role in structure and adaptation at the population level.

11 These assemblies will help aid the broader research community by enabling high

12 resolution genomic analyses that explore trait variation within species and potentially provide

13 restoration practitioners with useful information to implement in restoration initiatives. As coral

14 populations continue their decline, it is crucial that we develop a thorough understanding of the

15 genomic processes that have driven coral evolution and have allowed them to overcome past

16 extinction events and global stressors. These reference assemblies provide a key stepping stone

17 towards this goal.

Data Availability Statement

Acknowledgments

1 Institutes' Genomics Core Facility (RRID:SCR_023645) for use of the PacBio Sequel IIe sequencing

2 platform.

Conflict of Interest

4 The authors declare no conflict of interest.

Funding

References

Aichelman HE, Benson BE, Gomez-Campo K, Martinez-Rugerio MI, Fifer JE, Tsang L, Hughes AM, Bove CB, Nieves OC, Pereslete AM, et al. 2025. Cryptic coral diversity is associated with symbioses, physiology, and response to thermal challenge. Sci Adv. 11(3):eadr5237. doi:10.1126/sciadv.adr5237.

Allio R, Schomaker-Bastos A, Romiguier J, Prosdocimi F, Nabholz B, Delsuc F. 2020. MitoFinder: Efficient automated large-scale extraction of mitogenomic data in target enrichment phylogenomics. Molecular Ecology Resources. 20(4):892–905. doi:10.1111/1755-0998.13160.

Almeida-Silva F, Van de Peer Y. 2025. doubletrouble: an R/Bioconductor package for the identification, classification, and analysis of gene and genome duplications. Bioinformatics.:btaf043. doi:10.1093/bioinformatics/btaf043.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research. 25(17):3389–3402. doi:10.1086/284325.

Alvarez-Filip L, González-Barrios FJ, Pérez-Cervantes E, Molina-Hernández A, Estrada-Saldívar N. 2022. Stony coral tissue loss disease decimated Caribbean coral populations and reshaped reef functionality. Communications Biology 2022 5:1. 5(1):1–10. doi:10.1038/s42003-022-03398-6.

Barrett RDH, Schluter D. 2008. Adaptation from standing genetic variation. Trends in Ecology & Evolution. 23(1):38–44. doi:10.1016/j.tree.2007.09.008.

Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, Khanna A, Marshall M, Moxon S, Sonnhammer ELL, et al. 2004. The Pfam protein families database. Nucleic Acids Research. 32:D138–D141. doi:10.1093/nar/gkh121.

1   Baums IB, Chamberland VF, Locatelli NS, Conn T. 2022. Maximizing Genetic Diversity in Coral

2   Restoration Projects. In: Van Oppen MJH, Aranda Lastra M, editors. Coral Reef Conservation

3   and Restoration in the Omics Age. Vol. 15. Cham: Springer International Publishing. (Coral

4   Reefs of the World). p. 35–53. [accessed 2024 Apr 22].

5   https://link.springer.com/10.1007/978-3-031-07055-6_3.

6   Beedessee G, Kubota T, Arimoto A, Nishitsuji K, Waller RF, Hisata K, Yamasaki S, Satoh N,

7   Kobayashi J, Shoguchi E. 2020. Integrated omics unveil the secondary metabolic landscape

8   of a basal dinoflagellate. BMC Biology. 18(1):1–16. doi:10.1186/s12915-020-00873-6.

9   Boilard A, Dubé CE, Gruet C, Mercière A, Hernandez-Agreda A, Derome N. 2020. Defining coral

10   bleaching as a microbial dysbiosis within the coral holobiont. Microorganisms. 8(11):1682.

11 Bollati E, Lyndby NH, D'Angelo C, Kühl M, Wiedenmann J, Wangpraseurt D. 2022. Green fluorescent

12   protein-like pigments optimise the internal light environment in symbiotic reef-building

13   corals. Tessmar-Raible K, Schuman MC, Bourgeois D, editors. eLife. 11:e73521.

14   doi:10.7554/eLife.73521.

15 Bongaerts P, Cooke IR, Ying H, Wels D, den Haan S, Hernandez-Agreda A, Brunner CA, Dove S,

16   Englebert N, Eyal G, et al. 2021. Morphological stasis masks ecologically divergent coral

17   species on tropical reefs. Current Biology. 31(11):2286–2298.

18   doi:10.1016/j.cub.2021.03.028.

19 Bourne DG, Garren M, Work TM, Rosenberg E, Smith GW, Harvell CD. 2009. Microbial disease and

20   the coral holobiont. Trends in Microbiology. 17(12):554–562. doi:10.1016/j.tim.2009.09.004.

21 Boutet E, Lieberherr D, Tognolli M, Schneider M, Bansal P, Bridge AJ, Poux S, Bougueleret L, Xenarios

22   I. 2016. Uniprotkb/swiss-prot, the manually annotated section of the uniprot

23   knowledgebase. In: Methods in Molecular Biology. Vol. 1374. p. 23–54. [accessed 2023 Mar

24   29]. http://www.expasy.org/sprot/ppap/.

1   Bove CB, Mudge L, Bruno JF. 2022. A century of warming on Caribbean reefs. PLOS Climate.

2   1(3):e0000002. doi:10.1371/journal.pclm.0000002.

3   Brandt ME, Ennis RS, Meiling SS, Townsend J, Cobleigh K, Glahn A, Quetel J, Brandtneris V,

4   Henderson LM, Smith TB. 2021. The Emergence and Initial Impact of Stony Coral Tissue

5   Loss Disease (SCTLD) in the United States Virgin Islands. Frontiers in Marine Science. 8.

6   [accessed 2023 Jun 4]. https://www.frontiersin.org/articles/10.3389/fmars.2021.715329.

7   Brůna T, Lomsadze A, Borodovsky M. 2020. GeneMark-EP+: Eukaryotic gene prediction with self-

8   training in the space of genes and proteins. NAR Genomics and Bioinformatics. 2(2).

9   doi:10.1093/nargab/lqaa026. [accessed 2023 Mar 29].

10   https://academic.oup.com/nargab/article/2/2/lqaa026/5836691.

11   Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. Nat

12   Methods. 12(1):59–60. doi:10.1038/nmeth.3176.

13   Budd AF, Johnson KG, Stemann TA, Tompkins B. 1999. Pliocene to Pleistocene reef coral

14   assemblages in the Limon Group of Costa Rica. Bulletins of American Paleontology.

15   113(357):119–158.

16   Burton KW, Ling H-F, O'Nions RK. 1997. Closure of the Central American Isthmus and its effect on

17   deep-water formation in the North Atlantic. Nature. 386(6623):382–385.

18   doi:10.1038/386382a0.

19   Cai L, Zhou G, Tian R-M, Tong H, Zhang W, Sun J, Ding W, Wong YH, Xie JY, Qiu J-W, et al. 2017.

20   Metagenomic analysis reveals a green sulfur bacterium as a potential coral symbiont. Sci

21   Rep. 7(1):9320. doi:10.1038/s41598-017-09032-4.

22   Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+:

23   Architecture and applications. BMC Bioinformatics. 10(1):1–9. doi:10.1186/1471-2105-10-

24   421.

1 Cavada-Blanco F, Croquer A, Vermeij M, Goergen L, Rodriguez-Martinez R. 2022. Dendrogyra

2   cylindrus. IUCN Red List of Threatened Species. [accessed 2023 Jun 4].

3   e.T133124A129721366.

4 Changsut I, Womack HR, Shickle A, Sharp KH, Fuess LE. 2022. Variation in symbiont density is

5   linked to changes in constitutive immunity in the facultatively symbiotic coral, Astrangia

6   poculata. Biology Letters. 18(11):20220273. doi:10.1098/rsbl.2022.0273.

7 Chapman JA, Kirkness EF, Simakov O, Hampson SE, Mitros T, Weinmaier T, Rattei T,

8   Balasubramanian PG, Borman J, Busam D, et al. 2010. The dynamic genome of Hydra.

9   Nature. 464(7288):592–596. doi:10.1038/nature08830.

10 Chen Y-H, Yang S-H, Tandon K, Lu C-Y, Chen H-J, Shih C-J, Tang S-L. 2021. Potential syntrophic

11   relationship between coral-associated Prosthecochloris and its companion sulfate-

12   reducing bacterium unveiled by genomic analysis. Microb Genom. 7(5):000574.

13   doi:10.1099/mgen.0.000574.

14 Cooke I, Ying H, Forêt S, Bongaerts P, Strugnell JM, Simakov O, Zhang J, Field MA, Rodriguez-Lanetty

15   M, Bell SC, et al. 2020. Genomic signatures in the coral holobiont reveal host adaptations

16   driven by Holocene climate change and reef specific symbionts. Sci Adv. 6(48):eabc6318.

17   doi:10.1126/sciadv.abc6318.

18 Craggs J, Guest JR, Davis M, Simmons J, Dashti E, Sweet M. 2017. Inducing broadcast coral

19   spawning ex situ: Closed system mesocosm design and husbandry protocol. Ecology and

20   Evolution. 7(24):11066–11078. doi:10.1002/ece3.3538.

21 Cramer KL, Jackson JBC, Donovan MK, Greenstein BJ, Korpanty CA, Cook GM, Pandolfi JM. 2020.

22   Widespread loss of Caribbean acroporid corals was underway before coral bleaching and

23   disease outbreaks. Science Advances. 6(17). doi:10.1126/SCIADV.AAX9395. [accessed

24   2022 Jun 30]. https://www.science.org/doi/10.1126/sciadv.aax9395.

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, et al. 2021. Twelve years of SAMtools and BCFtools. GigaScience. 10(2):1–4. doi:10.1093/gigascience/giab008.

DeSalvo MK, Sunagawa S, Voolstra CR, Medina M. 2010. Transcriptomic responses to heat stress and bleaching in the elkhorn coral Acropora palmata. Marine Ecology Progress Series. 402:97–113. doi:10.3354/meps08372.

Devlin-Durante MK, Miller MW, Group CAR, Precht WF, Baums IB. 2016. How old are you? Genet age estimates in a clonal animal. Molecular Ecology. 25(22):5628–5646. doi:10.1111/mec.13865.

Emms DM, Kelly S. 2019. OrthoFinder: Phylogenetic orthology inference for comparative genomics. Genome Biology. 20(1):1–14. doi:10.1186/s13059-019-1832-y.

Faichney IDE, Webster JM, Clague DA, Braga JC, Renema W, Potts DC. 2011. The impact of the Mid-Pleistocene Transition on the composition of submerged reefs of the Maui Nui Complex, Hawaii. Palaeogeography, Palaeoclimatology, Palaeoecology. 299(3):493–506. doi:10.1016/j.palaeo.2010.11.027.

Feng X, Cheng H, Portik D, Li H. 2022. Metagenome assembly of high-fidelity long reads with hifiasm-meta. Nature Methods. 19(6):671–674. doi:10.1038/s41592-022-01478-3.

Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, Smit AF. 2020. RepeatModeler2 for automated genomic discovery of transposable element families. Proceedings of the National Academy of Sciences of the United States of America. 117(17):9451–9457. doi:10.1073/PNAS.1921046117.

Fourreau CJL, Kise H, Santander MD, Pirro S, Maronna MM, Poliseno A, Santos MEA, Reimer JD. 2023. Genome sizes and repeatome evolution in zoantharians (Cnidaria: Hexacorallia: Zoantharia). PeerJ. 11:e16188. doi:10.7717/peerj.16188.

1 Fuller ZL, Mocellin VJL, Morris LA, Cantin N, Shepherd J, Sarre L, Peng J, Liao Y, Pickrell J, Andolfatto
2   P, et al. 2020. Population genetics of the coral Acropora millepora: Toward genomic
3   prediction of bleaching. Science. 369(6501). doi:10.1126/SCIENCE.ABA4674. [accessed
4   2021 Dec 7]. https://www.science.org/doi/abs/10.1126/science.aba4674.
5 Gabriel J, Ferreira RN, Krasheninnikova K, Uliano-Silva DM, Formenti G, Abueg L, Torrance J, Myers
6   EW, Durbin R, Mccarthy SA, et al. 2023 Jan 10. MitoHiFi: a python pipeline for mitochondrial
7   genome assembly from PacBio High Fidelity reads. bioRxiv.:2022.12.23.521667.
8   doi:10.1101/2022.12.23.521667.
9 Garrido AG, Machado LF, Zilberberg C, Leite DC de A. 2021. Insights into 'Symbiodiniaceae
10   phycosphere' in a coral holobiont. Symbiosis. 83(1):25–39. doi:10.1007/s13199-020-00735-
11   3.
12 Geertsma RC, Wijgerde T, Latijnhouwers KRW, Chamberland VF. 2022. Onset of zooplanktivory and
13   optimal water flow rates for prey capture in newly settled polyps of ten Caribbean coral
14   species. Coral Reefs. 41(6):1651–1664. doi:10.1007/s00338-022-02310-2.
15 Gold DA, Katsuki T, Li Y, Yan X, Regulski M, Ibberson D, Holstein T, Steele RE, Jacobs DK, Greenspan
16   RJ. 2019. The genome of the jellyfish Aurelia and the evolution of animal complexity. Nat
17   Ecol Evol. 3(1):96–104. doi:10.1038/s41559-018-0719-8.
18 Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury
19   R, Zeng Q, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a
20   reference genome. Nature Biotechnology. 29(7):644–652. doi:10.1038/nbt.1883.
21 Gregorio A de. 1930. Sul Permiano di Sicilia (Fossili del calcare con Fusulina di palazzo adriano non
22   descritti del Prof. G. Gemmellaro conservati nel mio private Gabinetto). Annales Géol
23   Paléontol. 52:18–32.

Guan D, Guan D, McCarthy SA, Wood J, Howe K, Wang Y, Durbin R, Durbin R. 2020. Identifying and removing haplotypic duplication in primary genome assemblies. Bioinformatics. 36(9):2896–2898. doi:10.1093/BIOINFORMATICS/BTAA025.

Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, White O, Robin CR, Wortman JR. 2008. Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. Genome Biology. 9(1):1–22. doi:10.1186/gb-2008-9-1-r7.

Hamada M, Schröder K, Bathia J, Kürn U, Fraune S, Khalturina M, Khalturin K, Shinzato C, Satoh N, Bosch TCG. 2018. Metabolic co-dependence drives the evolutionarily ancient Hydra–Chlorella symbiosis. eLife. 7. doi:10.7554/eLife.35122. [accessed 2022 Dec 19]. https://pubmed.ncbi.nlm.nih.gov/29848439/.

Harman TE, Barshis DJ, Hauff Salas B, Hamsher SE, Strychar KB. 2022. Indications of symbiotic state influencing melanin-synthesis immune response in the facultative coral Astrangia poculata. Dis Aquat Organ. 151:63–74. doi:10.3354/dao03695.

Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, Eilbeck K, Lewis S, Marshall B, Mungall C, et al. 2004. The Gene Ontology (GO) database and informatics resource. Nucleic Acids Research. 32:D258–D261. doi:10.1093/nar/gkh036.

Hawkins JS, Kim H, Nason JD, Wing RA, Wendel JF. 2006. Differential lineage-specific amplification of transposable elements is responsible for genome size variation in Gossypium. Genome Res. 16(10):1252–1261. doi:10.1101/gr.5282906.

Hazewinkel M, Van Eijck J. 2000. Graph clustering by flow simulation. University of Utrecht Dissertation. [accessed 2023 Apr 17]. https://dspace.library.uu.nl/handle/1874/848.

Helgoe J, Davy SK, Weis VM, Rodriguez-Lanetty M. 2024. Triggers, cascades, and endpoints: connecting the dots of coral bleaching mechanisms. Biological Reviews. 99(3):715–752. doi:10.1111/brv.13042.

1   Horvath KM, Castillo KD, Armstrong P, Westfield IT, Courtney T, Ries JB. 2016. Next-century ocean

2   acidification and warming both reduce calcification rate, but only acidification alters

3   skeletal morphology of reef-building coral Siderastrea siderea. Sci Rep. 6(1):29613.

4   doi:10.1038/srep29613.

5   Huang CRL, Burns KH, Boeke JD. 2012. Active Transposition in Genomes. Annu Rev Genet. 46:651–

6   675. doi:10.1146/annurev-genet-110711-155616.

7   Huang L, Zhang H, Wu P, Entwistle S, Li X, Yohe T, Yi H, Yang Z, Yin Y. 2018. DbCAN-seq: A database

8   of carbohydrate-active enzyme (CAZyme) sequence and annotation. Nucleic Acids

9   Research. 46(D1):D516–D521. doi:10.1093/nar/gkx894.

10   Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR,

11   Letunic I, Rattei T, Jensen LJ, et al. 2019. EggNOG 5.0: A hierarchical, functionally and

12   phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses.

13   Nucleic Acids Research. 47(D1):D309–D314. doi:10.1093/nar/gky1085.

14   Hume BCC, Smith EG, Ziegler M, Warrington HJM, Burt JA, LaJeunesse TC, Wiedenmann J, Voolstra

15   CR. 2019. SymPortal: A novel analytical framework and platform for coral algal symbiont

16   next-generation sequencing ITS2 profiling. Molecular Ecology Resources. 19(4):1063–1080.

17   doi:10.1111/1755-0998.13004.

18   Hunter IG, Jones B. 1996. Coral associations of the Pleistocene Ironshore Formation, Grand

19   Cayman. Coral Reefs. 15:249–267.

20   Hunter S, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Bork P, Das U, Daugherty L,

21   Duquenne L, et al. 2009. InterPro: The integrative protein signature database. Nucleic Acids

22   Research. 37:D211–D215. doi:10.1093/nar/gkn785.

Kardos M, Armstrong EE, Fitzpatrick SW, Hauser S, Hedrick PW, Miller JM, Tallmon DA, Funk WC. 2021. The crucial role of genome-wide genetic variation in conservation. Proceedings of the National Academy of Sciences. 118(48):e2104642118. doi:10.1073/pnas.2104642118.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Molecular Biology and Evolution. 30(4):772–780. doi:10.1093/molbev/mst010.

Kim J, Choi JP, Kim MS, Jo Y, Min WG, Woo S, Yum S, Bhak J. 2022. Comparative Genome and Evolution Analyses of an Endangered Stony Coral Species Dendrophyllia cribrosa Near Dokdo Islands in the East Sea. Genome Biology and Evolution. 14(9):evac132. doi:10.1093/gbe/evac132.

Klein AM, Sturm AB, Eckert RJ, Walker BK, Neely KL, Voss JD. 2024. Algal symbiont genera but not coral host genotypes correlate to stony coral tissue loss disease susceptibility among Orbicella faveolata colonies in South Florida. Front Mar Sci. 11. doi:10.3389/fmars.2024.1287457. [accessed 2024 Nov 20]. https://www.frontiersin.org/journals/marine-science/articles/10.3389/fmars.2024.1287457/full.

Klopfenstein DV, Zhang L, Pedersen BS, Ramírez F, Warwick Vesztrocy A, Naldi A, Mungall CJ, Yunes JM, Botvinnik O, Weigel M, et al. 2018. GOATOOLS: A Python library for Gene Ontology analyses. Sci Rep. 8(1):10872. doi:10.1038/s41598-018-28948-z.

Kon T, Kon-Nanjo K, Simakov O. 2024. Expansion of a single Helitron subfamily in Hydractinia symbiolongicarpus suggests a shared mechanism of cnidarian chromosomal extension. :2024.11.21.624632. doi:10.1101/2024.11.21.624632. [accessed 2024 Dec 19]. https://www.biorxiv.org/content/10.1101/2024.11.21.624632v1.

1   Krueger F, James F, Ewels P, Afyounian E, Schuster-Boeckler B. 2021. FelixKrueger/TrimGalore:

2   v0.6.7. doi:10.5281/zenodo.5127899.

3   Kvitt H, Rosenfeld H, Zandbank K, Tchernov D. 2011. Regulation of Apoptotic Pathways by

4   Stylophora pistillata (Anthozoa, Pocilloporidae) to Survive Thermal Stress and Bleaching.

5   PLOS ONE. 6(12):e28665. doi:10.1371/journal.pone.0028665.

6   Laetsch DR, Blaxter ML. 2017. BlobTools: Interrogation of genome assemblies. F1000Research

7   2017 6:1287. 6:1287. doi:10.12688/f1000research.12232.1.

8   LaJeunesse TC, Thornhill DJ. 2011. Improved Resolution of Reef-Coral Endosymbiont

9   (Symbiodinium) Species Diversity, Ecology, and Evolution through psbA Non-Coding Region

10   Genotyping. PLOS ONE. 6(12):e29013. doi:10.1371/journal.pone.0029013.

11   Leclère L, Horin C, Chevalier S, Lapébie P, Dru P, Peron S, Jager M, Condamine T, Pottin K, Romano

12   S, et al. 2019. The genome of the jellyfish Clytia hemisphaerica and the evolution of the

13   cnidarian life-cycle. Nat Ecol Evol. 3(5):801–810. doi:10.1038/s41559-019-0833-2.

14   Lewis AM, Chan AN, LaJeunesse TC. 2019. New Species of Closely Related Endosymbiotic

15   Dinoflagellates in the Greater Caribbean have Niches Corresponding to Host Coral

16   Phylogeny. Journal of Eukaryotic Microbiology. 66(3):469–482. doi:10.1111/jeu.12692.

17   Li H. 2018. Minimap2: Pairwise alignment for nucleotide sequences. Bioinformatics. 34(18):3094–

18   3100. doi:10.1093/bioinformatics/bty191.

19   Liew YJ, Aranda M, Voolstra CR. 2016. Reefgenomics.Org - a repository for marine genomics data.

20   Database.

21   Locatelli NS, Kitchen SA, Stankiewicz KH, Osborne CC, Dellaert Z, Elder H, Kamel B, Koch HR,

22   Fogarty ND, Baums IB. 2024. Chromosome-level genome assemblies and genetic maps

23   reveal heterochiasmy and macrosynteny in endangered Atlantic Acropora. BMC Genomics.

24   25(1):1119. doi:10.1186/s12864-024-11025-3.

López EH, Palumbi SR. 2020. Somatic Mutations and Genome Stability Maintenance in Clonal Coral Colonies. Molecular Biology and Evolution. 37(3):828–838. doi:10.1093/molbev/msz270.

Manni M, Berkeley MR, Seppey M, Sim~ Ao FA, Zdobnov EM. 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. doi:10.1093/molbev/msab199.

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.

Marhaver KL, Vermeij MJA, Medina MM. 2015. Reproductive natural history and successful juvenile propagation of the threatened Caribbean Pillar CoralDendrogyra cylindrus. BMC Ecology. 15(1). doi:10.1186/s12898-015-0039-7.

Mayfield AB. 2022. Machine-Learning-Based Proteomic Predictive Modeling with Thermally-Challenged Caribbean Reef Corals. Diversity. 14(1):33. doi:10.3390/d14010033.

McKenna V, Archibald JM, Beinart R, Dawson MN, Hentschel U, Keeling PJ, Lopez JV, Martín-Durán JM, Petersen JM, Sigwart JD, et al. 2021. The Aquatic Symbiosis Genomics Project: probing the evolution of symbiosis across the tree of life. Wellcome Open Research. 6(254):254.

Mendes FK, Vanderpool D, Fulton B, Hahn MW. 2021. CAFE 5 models variation in evolutionary rates among gene families. Bioinformatics. 36(22–23):5516–5518. doi:10.1093/bioinformatics/btaa1022.

Michael TP. 2014. Plant genome size variation: bloating and purging DNA. Briefings in Functional Genomics. 13(4):308–317. doi:10.1093/bfgp/elu005.

Modys AB, Toth LT, Mortlock RA, Oleinik AE, Precht WF. 2023. Discovery of a rare pillar coral (Dendrogyra cylindrus) death assemblage off southeast Florida reveals multi-century

persistence during the late Holocene. Coral Reefs. 42(4):801–807. doi:10.1007/s00338-023-02387-3.

Neely KL, Lewis CL, Lunz KS, Kabay L. 2021. Rapid Population Decline of the Pillar Coral Dendrogyra cylindrus Along the Florida Reef Tract. Frontiers in Marine Science. 8:434. doi:10.3389/fmars.2021.656515.

Noel B, Denoeud F, Rouan A, Buitrago-López C, Capasso L, Poulain J, Boissin E, Pousse M, Da Silva C, Couloux A, et al. 2023. Pervasive tandem duplications and convergent evolution shape coral genomes. Genome Biology. 24(1):123. doi:10.1186/s13059-023-02960-7.

O'Dea A, Lessios HA, Coates AG, Eytan RI, Restrepo-Moreno SA, Cione AL, Collins LS, de Queiroz A, Farris DW, Norris RD, et al. 2016. Formation of the Isthmus of Panama. Science Advances. 2(8). doi:10.1126/sciadv.1600883. [accessed 2023 Aug 11]. https://www.science.org/doi/10.1126/sciadv.1600883.

O'Neil KL, Serafin RM, Patterson JT, Craggs JRK. 2021. Repeated ex situ Spawning in Two Highly Disease Susceptible Corals in the Family Meandrinidae. Front Mar Sci. 8. doi:10.3389/fmars.2021.669976. [accessed 2024 May 10]. https://www.frontiersin.org/articles/10.3389/fmars.2021.669976.

Palacio-Castro AM, Dennison CE, Rosales SM, Baker AC. 2021. Variation in susceptibility among three Caribbean coral species and their algal symbionts indicates the threatened staghorn coral, Acropora cervicornis, is particularly susceptible to elevated nutrients and heat stress. Coral Reefs. 40(5):1601–1613. doi:10.1007/s00338-021-02159-x.

Palmer CV, Bythell JC, Willis BL. 2012. Enzyme activity demonstrates multiple pathways of innate immunity in Indo-Pacific anthozoans. Proceedings of the Royal Society B: Biological Sciences. 279(1743):3879–3887. doi:10.1098/rspb.2011.2487.

Palmer JM, Stajich J. 2020. nextgenusfs/funannotate: funannotate v1.8.13 (Version 1.8.13). Zenodo. doi:10.5281/zenodo.1134477.

Park E, Hwang D-S, Lee J-S, Song J-I, Seo T-K, Won Y-J. 2012. Estimation of divergence times in cnidarian evolution based on mitochondrial protein-coding genes and the fossil record. Mol Phylogenet Evol. 62(1):329–345. doi:10.1016/j.ympev.2011.10.008.

Peters SE, McClennen M. 2016. The Paleobiology Database application programming interface. Paleobiology. 42(1):1–7. doi:10.1017/pab.2015.39.

Pinzón JHC, Beach-Letendre J, Weil E, Mydlarz LD. 2014. Relationship between Phylogeny and Immunity Suggests Older Caribbean Coral Lineages Are More Resistant to Disease. PLoS ONE. 9(8):e104787. doi:10.1371/journal.pone.0104787.

Pootakham W, Sonthirod C, Naktang C, Kongjandtre N, Putchim L, Sangsrakru D, Yoocha T, Tangphatsornruang S. 2021. De novo Assembly of the Brain Coral Platygyra sinensis Genome. Frontiers in Marine Science. 8. [accessed 2023 May 2]. https://www.frontiersin.org/articles/10.3389/fmars.2021.732650.

Prada C, Hanna B, Budd AF, Woodley CM, Schmutz J, Grimwood J, Iglesias-Prieto R, Pandolfi JM, Levitan D, Johnson KG, et al. 2016. Empty Niches after Extinctions Increase Population Sizes of Modern Corals. Current Biology. 26(23):3190–3194. doi:10.1016/j.cub.2016.09.039.

Price MN, Dehal PS, Arkin AP. 2010. FastTree 2 - Approximately maximum-likelihood trees for large alignments. PLoS ONE. 5(3):e9490. doi:10.1371/journal.pone.0009490.

Proost S, Fostier J, De Witte D, Dhoedt B, Demeester P, Van De Peer Y, Vandepoele K. 2012. i-ADHoRe 3.0-fast and sensitive detection of genomic homology in extremely large data sets. Nucleic Acids Research. 40(2). doi:10.1093/nar/gkr955. [accessed 2023 Apr 17]. https://pubmed.ncbi.nlm.nih.gov/22102584/.

Qi W. 1984. An Anisian coral fauna in Guizhou, South China. Palaeontographica Americana. 54:187–190.

Ranallo-Benavidez TR, Jaron KS, Schatz MC. 2020. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. Nat Commun. 11(1). doi:10.1038/s41467-020-14998-3. [accessed 2023 Sep 21]. https://www.nature.com/articles/s41467-020-14998-3.

Rawlings ND, Barrett AJ, Bateman A. 2009. MEROPS: The peptidase database. Nucleic Acids Research. 38:D227–D233. doi:10.1093/nar/gkp971.

Reef R, Dunn S, Levy O, Dove S, Shemesh E, Brickner I, Leggat W, Hoegh-Guldberg O. 2009. Photoreactivation is the main repair pathway for UV-induced DNA damage in coral planulae. J Exp Biol. 212(17):2760–2766. doi:10.1242/jeb.031286.

Reusch TBH, Baums IB, Werner B. 2021. Evolution via somatic genetic variation in modular species. Trends in Ecology & Evolution. 36(12):1083–1092. doi:10.1016/j.tree.2021.08.011.

Ricaurte M, Schizas NV, Ciborowski P, Boukli NM. 2016. Proteomic analysis of bleached and unbleached *Acropora palmata*, a threatened coral species of the Caribbean. Marine Pollution Bulletin. 107(1):224–232. doi:10.1016/j.marpolbul.2016.03.068.

Rodriguez-Martinez R, Vermeij M, Kitahara MV, Alvarez-Filip L. 2022. Siderastrea siderea. IUCN Red List of Threatened Species. [accessed 2023 Jun 4]. e.T133389A165863977.

Roessler K, Muyle A, Diez CM, Gaut GRJ, Bousios A, Stitzer MC, Seymour DK, Doebley JF, Liu Q, Gaut BS. 2019. The genome-wide dynamics of purging during selfing in maize. Nat Plants. 5(9):980–990. doi:10.1038/s41477-019-0508-7.

Salih A, Larkum A, Cox G, Kühl M, Hoegh-Guldberg O. 2000. Fluorescent pigments in corals are photoprotective. Nature. 408(6814):850–853. doi:10.1038/35048564.

1   Sanderson MJ. 2003. r8s: inferring absolute rates of molecular evolution and divergence times in

2   the absence of a molecular clock. Bioinformatics. 19(2):301–302.

3   doi:10.1093/bioinformatics/19.2.301.

4   Sanderson T, Bileschi ML, Belanger D, Colwell LJ. 2023. ProteInfer, deep neural networks for protein

5   functional inference. Dötsch V, Staller MV, editors. eLife. 12:e80942.

6   doi:10.7554/eLife.80942.

7   Shinzato C, Shoguchi E, Kawashima T, Hamada M, Hisata K, Tanaka M, Fujie M, Fujiwara M,

8   Koyanagi R, Ikuta T, et al. 2011. Using the Acropora digitifera genome to understand coral

9   responses to environmental change. Nature. 476(7360):320–323. doi:10.1038/nature10249.

10 Shinzato C, Takeuchi T, Yoshioka Y, Tada I, Kanda M, Broussard C, Iguchi A, Kusakabe M, Marin F,

11   Satoh N, et al. 2021. Whole-Genome Sequencing Highlights Conservative Genomic

12   Strategies of a Stress-Tolerant, Long-Lived Scleractinian Coral, Porites australiensis

13   Vaughan, 1918. Genome Biology and Evolution. 13(12):evab270. doi:10.1093/gbe/evab270.

14 Shoguchi E, Beedessee G, Hisata K, Tada I, Narisoko H, Satoh N, Kawachi M, Shinzato C. 2021. A

15   New Dinoflagellate Genome Illuminates a Conserved Gene Cluster Involved in Sunscreen

16   Biosynthesis. Genome biology and evolution. 13(2). doi:10.1093/gbe/evaa235. [accessed

17   2021 Oct 28]. https://pubmed.ncbi.

18 Shoguchi E, Beedessee G, Tada I, Hisata K, Kawashima T, Takeuchi T, Arakaki N, Fujie M, Koyanagi R,

19   Roy MC, et al. 2018. Two divergent Symbiodinium genomes reveal conservation of a gene

20   cluster for sunscreen biosynthesis and recently lost genes. BMC Genomics. 19(1):1–11.

21   doi:10.1186/s12864-018-4857-9.

22 Shoguchi E, Shinzato C, Kawashima T, Gyoja F, Mungpakdee S, Koyanagi R, Takeuchi T, Hisata K,

23   Tanaka M, Fujiwara M, et al. 2013. Draft assembly of the symbiodinium minutum nuclear

genome reveals dinoflagellate gene structure. Current Biology. 23(15):1399–1408.

doi:10.1016/j.cub.2013.05.062.

Smit A, Hubley R, Green P. RepeatMasker Open-4.0. http://www.repeatmasker.org.

Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B. 2006. AUGUSTUS: ab initio

prediction of alternative transcripts. Nucleic Acids Research. 34. doi:10.1093/NAR/GKL200.

[accessed 2022 May 1]. https://academic.oup.com/nar/article/34/suppl_2/W435/2505582.

Stephens TG, Lee J, Jeong Y, Yoon HS, Putnam HM, Majerová E, Bhattacharya D. 2022. High-quality

genome assemblies from key Hawaiian coral species. GigaScience. 11.

doi:10.1093/gigascience/giac098. [accessed 2023 Jun 4].

https://doi.org/10.1093/gigascience/giac098.

Stewart HA, Wright JL, Carrigan M, Altieri AH, Kline DI, Araújo RJ. 2022. Novel coexisting mangrove-

coral habitats: Extensive coral communities located deep within mangrove canopies of

Panama, a global classification system and predicted distributions. PLOS ONE.

17(6):e0269181. doi:10.1371/journal.pone.0269181.

Supek F, Bošnjak M, Škunca N, Šmuc T. 2011. REVIGO Summarizes and Visualizes Long Lists of

Gene Ontology Terms. PLOS ONE. 6(7):e21800. doi:10.1371/journal.pone.0021800.

Svanfeldt K, Lundqvist L, Rabinowitz C, Sköld HN, Rinkevich B. 2014. Repair of UV-induced DNA

damage in shallow water colonial marine species. Journal of Experimental Marine Biology

and Ecology. 452:40–46. doi:10.1016/j.jembe.2013.12.003.

Tchernov D, Kvitt H, Haramaty L, Bibby TS, Gorbunov MY, Rosenfeld H, Falkowski PG. 2011.

Apoptosis and the selective survival of host animals following thermal bleaching in

zooxanthellate corals. Proceedings of the National Academy of Sciences. 108(24):9905–

9909. doi:10.1073/pnas.1106924108.

1 Thompson JR, Rivera HE, Closek CJ, Medina M. 2015. Microbes in the coral holobiont: partners

2 through evolution, development, and ecological interactions. Frontiers in Cellular and

3 Infection Microbiology. 4(January):1–20. doi:10.3389/fcimb.2014.00176.

4 Tsompana M, Buck MJ. 2014. Chromatin accessibility: a window into the genome. Epigenetics &

5 Chromatin. 7(1):33. doi:10.1186/1756-8935-7-33.

6 Tunala LP, Guerra CR, de Menezes RG, Diogo CP, de Almeida TB, Duarte HM, de Souza Tâmega FT,

7 Coutinho R. 2023. Photobiology of Symbiodiniaceae hosted on *Siderastrea stellata* in the

8 southwestern Atlantic. Regional Studies in Marine Science. 63:103027.

9 doi:10.1016/j.rsma.2023.103027.

10 Van Buren EW, Ponce IE, Beavers KM, Stokes A, Cornelio MN, Emery M, Mydlarz LD. 2024 Jul 18.

11 Structural and Evolutionary Relationships of Melanin Cascade Proteins in Cnidarian Innate

12 Immunity. Integrative and Comparative Biology.:icae115. doi:10.1093/icb/icae115.

13 Vasquez Kuntz KL, Kitchen SA, Conn TL, Vohsen SA, Chan AN, Vermeij MJA, Page C, Marhaver KL,

14 Baums IB. 2022. Inheritance of somatic mutations by animal offspring. Sci Adv. 8(35).

15 doi:10.1126/sciadv.abn0707.

16 Vecsei A, Moussavian E. 1997. Paleocene reefs on the Maiella Platform margin, Italy: An example of

17 the effects of the Cretaceous/Tertiary boundary events on reefs and carbonate platforms.

18 Facies.(36):123–140.

19 Vermeij M, Goergen L. 2022. Colpophyllia natans. IUCN Red List of Threatened Species. [accessed

20 2023 Jun 4]. e.T132884A165613226.

21 Voolstra CR, Sunagawa S, Matz MV, Bayer T, Aranda M, Buschiazzo E, DeSalvo MK, Lindquist E,

22 Szmant AM, Coffroth MA, et al. 2011. Rapid Evolution of Coral Proteins Responsible for

23 Interaction with the Environment. PLOS ONE. 6(5):e20392.

24 doi:10.1371/journal.pone.0020392.

Wall CB, Ricci CA, Foulds GE, Mydlarz LD, Gates RD, Putnam HM. 2018. The effects of environmental history and thermal stress on coral physiology and immunity. Mar Biol. 165(3):56. doi:10.1007/s00227-018-3317-z.

Wallace CC, Rosen BR. 2006. Diverse staghorn corals (Acropora) in high-latitude Eocene assemblages: implications for the evolution of modern diversity patterns of reef corals. Proc Biol Sci. 273(1589):975–982. doi:10.1098/rspb.2005.3307.

Wang X, Liew YJ, Li Y, Zoccola D, Tambutte S, Aranda M. 2017. Draft genomes of the corallimorpharians Amplexidiscus fenestrafer and Discosoma sp. Molecular Ecology Resources. 17(6):e187–e195. doi:10.1111/1755-0998.12680.

Weizman E, Levy O. 2019. The role of chromatin dynamics under global warming response in the symbiotic coral model Aiptasia. Commun Biol. 2(1):1–12. doi:10.1038/s42003-019-0543-y.

Weizman E, Rinsky M, Simon-Blecher N, Lampert-Karako S, Yaron O, Tarrant AM, Levy O. 2021. Chromatin Dynamics and Gene Expression Response to Heat Exposure in Field-Conditioned versus Laboratory-Cultured Nematostella vectensis. International Journal of Molecular Sciences. 22(14):7454. doi:10.3390/ijms22147454.

Williamson OM, Dennison CE, O'Neil KL, Baker AC. 2022. Susceptibility of Caribbean Brain Coral Recruits to Stony Coral Tissue Loss Disease (SCTLD). Frontiers in Marine Science. 9. [accessed 2023 Jun 4]. https://www.frontiersin.org/articles/10.3389/fmars.2022.821165.

Wong WY, Simakov O, Bridge DM, Cartwright P, Bellantuono AJ, Kuhn A, Holstein TW, David CN, Steele RE, Martínez DE. 2019. Expansion of a single transposable element family is associated with genome-size increase and radiation in the genus Hydra. Proceedings of the National Academy of Sciences. 116(46):22915–22917. doi:10.1073/pnas.1910106116.

Wong Y-H, Zhang Y, Lun JCY, Qiu J-W. 2021. A proteomic analysis of skeletal tissue anomaly in the brain coral *Platygyra carnosa*. Marine Pollution Bulletin. 164:111982. doi:10.1016/j.marpolbul.2021.111982.

Xia W, Li Haorong, Cheng W, Li Honghui, Mi Y, Gou X, Liu Y. 2020. High-Quality Genome Assembly of Chrysaora quinquecirrha Provides Insights Into the Adaptive Evolution of Jellyfish. Front Genet. 11. doi:10.3389/fgene.2020.00535. [accessed 2024 Jul 7]. https://www.frontiersin.org/journals/genetics/articles/10.3389/fgene.2020.00535/full.

Yang Z. 2007. PAML 4: Phylogenetic analysis by maximum likelihood. Molecular Biology and Evolution. 24(8):1586–1591. doi:10.1093/molbev/msm088.

Ying H, Cooke I, Sprungala S, Wang W, Hayward DC, Tang Y, Huttley G, Ball EE, Forêt S, Miller DJ. 2018. Comparative genomics reveals the distinct evolutionary trajectories of the robust and complex coral lineages. Genome Biology. 19(1):175. doi:10.1186/s13059-018-1552-8.

Young BD, Williamson OM, Kron NS, Andrade Rodriguez N, Isma LM, MacKnight NJ, Muller EM, Rosales SM, Sirotzke SM, Traylor-Knowles N, et al. 2024. Annotated genome and transcriptome of the endangered Caribbean mountainous star coral (Orbicella faveolata) using PacBio long-read sequencing. BMC Genomics. 25(1):226. doi:10.1186/s12864-024-10092-w.

Yu Y, Nong W, So WL, Xie Y, Yip HY, Haimovitz J, Swale T, Baker DM, Bendena WG, Chan TF, et al. 2022. Genome of elegance coral Catalaphyllia jardinei (Euphylliidae). Frontiers in Marine Science. 9. [accessed 2023 Nov 30]. https://www.frontiersin.org/articles/10.3389/fmars.2022.991391.

Yus J, Nixon EN, Li J, Noriega Gimenez J, Bennett M-J, Flores D, Marhaver KL, Wegley Kelly L, Espinosa-Marzal RM, Wagoner Johnson AJ. 2024. Composite substrates for coral larval settlement and reef restoration based on natural hydraulic lime and inorganic strontium and

1   magnesium compounds. Ecological Engineering. 202:107236.

2   doi:10.1016/j.ecoleng.2024.107236.

3   Zwaenepoel A, Van De Peer Y. 2019. Wgd-simple command line tools for the analysis of ancient

4   whole-genome duplications. Bioinformatics. 35(12):2153–2155.

5   doi:10.1093/bioinformatics/bty915.

6

1   **Tables**

Table 1. Assembly summary statistics for *Colpophyllia natans, Dendrogyra cylindrus,* and *Siderastrea siderea.*

	Siderastrea siderea	*Dendrogyra cylindrus*	*Colpophyllia natans*
Contig Total (Mb)	822.514	526.444	398.943
Gap Percentage	0%	0%	0%
Number of contigs	265	301	174
Contig N50	9.1Mb	4.647Mb	4.902Mb
Largest contig	25.215Mb	21.044Mb	14.745Mb
GC Content (%)	39.81	39.29	38.87
% of k-mer estimate recovered	105.61	105.99	103.47
BUSCO Metazoa, complete (%) – Genome Mode	96.6	96.5	97.2
Single copy (Genome)	93.9	95.3	96.1
Duplicated (Genome)	2.7	1.3	1.0
Fragmented (%) (Genome)	0.9	0.8	1.0
Missing (%) (Genome)	2.4	2.6	1.8
BUSCO Metazoa, complete (%) – Protein Mode	91.8	94.9	95.5
Single copy (Protein)	87.1	88.2	88.8
Duplicated (Protein)	4.7	6.7	6.7
Fragmented (%) (Protein)	3.2	1.8	2.0
Missing (%) (Protein)	4.9	3.4	2.5
Gene models	61,712	39,739	34,139
Protein-coding gene models	52,473	34,738	29,090

2

Table 2. Repetitive content and transposable elements identified by RepeatMasker (Smit et al.)
across *Siderastrea siderea, Dendrogyra cylindrus,* and *Colpophyllia natans*. The top three repeat
families (e.g. Maverick) within each major repeat class (DNA, LINE, LTR, SINE, and RNA repeats)
are presented in this table.

		Siderastrea siderea			*Dendrogyra cylindrus*			*Colpophyllia natans*		
DNA	Total	138,779	68,819,835	8.39%	82,872	47,945,358	9.09%	68,722	17,971,690	4.52%
	Maverick	22,353	42,786,635	5.20%	13,811	35,496,289	6.74%	5,091	5,626,678	1.41%
	Sola-3	17,315	6,171,939	0.75%	4,051	1,802,393	0.34%	3,249	1,019,211	0.26%
	PIF-Harbinger	9,247	1,246,387	0.15%	9,092	1,764,101	0.34%	5,361	821,524	0.21%
	Academ-1	4,662	1,873,178	0.23%	2,578	872,091	0.17%	1,886	777,963	0.20%
LINE	Total	104,185	32,929,736	4.02%	53,305	17,696,634	3.36%	46,166	15,626,622	3.92%
	Penelope	38,138	11,157,756	1.36%	16,584	5,217,438	0.99%	21,897	5,971,804	1.50%
	L1-Tx1	13,255	8,443,143	1.03%	9,745	5,273,753	1.00%	6,122	3,626,933	0.91%
	L2	29,218	6,994,614	0.85%	16,810	3,452,973	0.66%	11,440	3,468,236	0.87%
	RTE-BovB	4,796	946,003	0.12%	2,582	1,414,380	0.27%	2,717	1,144,012	0.29%
LTR	Total	36,888	17,315,718	2.09%	11,593	7,626,581	1.44%	10,698	7,645,820	1.92%
	Gypsy	14,784	5,919,375	0.72%	4,918	2,999,690	0.57%	5,407	4,098,188	1.03%
	Pao	4,227	4,683,913	0.57%	3,053	2,945,104	0.56%	1,809	1,501,851	0.38%
	DIRS	3,419	2,371,017	0.29%	1,323	882,015	0.17%	2,095	1,383,508	0.35%
	Ngaro	7,195	3,053,083	0.37%	869	469,818	0.09%	1,057	496,392	0.12%
SINE	Total	16,023	2,146,747	0.26%	4,485	541,212	0.10%	5,616	674,504	0.17%
	tRNA-V	3,912	567,988	0.07%	2,623	350,202	0.07%	3,166	512,682	0.13%
	MIR	8,147	1,173,564	0.14%	0	0	0.00%	0	0	0.00%
	tRNA-RTE	1,612	150,650	0.02%	1,010	100,831	0.02%	0	0	0.00%
	Alu	1,485	131,235	0.02%	0	0	0.00%	0	0	0.00%
Low complexity		455	76,303	0.01%	282	53,301	0.01%	123	26,366	0.01%
Retroposon	L1-dep	175	32,571	0.00%	0	0	0.00%	0	0	0.00%
Rolling circle	Helitron	2,701	1,113,967	0.14%	837	186,063	0.04%	5,529	2,007,025	0.50%
Satellites		476	226,250	0.03%	1,120	113,510	0.02%	1,795	189,861	0.05%
Simple repeats		17,276	2,670,182	0.32%	11,761	1,849,810	0.35%	6,765	1,167,479	0.29%
RNA repeats	Total	24,191	5,349,132	0.65%	21,009	2,054,088	0.39%	727	135,302	0.03%
	tRNA	23,498	5,198,251	0.63%	20,541	1,915,611	0.36%	395	44,379	0.01%
	rRNA	693	150,881	0.02%	468	138,477	0.03%	260	82,379	0.02%
	snRNA	0	0	0.00%	0	0	0.00%	72	8,544	0.00%
Unclassified		1,000,623	262,395,404	31.91%	635,176	134,635,802	25.57%	262,664	48,767,859	12.22%
Total		**1,341,772**	**393,075,845**	**47.80%**	**822,440**	**212,702,359**	**40.40%**	**408,805**	**94,212,528**	**23.62%**

1 **Figures**

Figure 1. K-mer multiplicity plots (left panes) from GenomeScope2 (Ranallo-Benavidez et al. 2020) for a kmer size of 31 for A) *Siderastrea siderea,* B) *Dendrogyra cylindrus,* and C) *Colpophyllia natans*. Mitochondrial genome gene order (right panes) in *Siderastrea siderea, Dendrogyra cylindrus,* and *Colpophyllia natans.* Mitogenomes assembled using MitoHiFi (Gabriel et al. 2023).

Figure 2. Upset plot describing unique and shared orthogroups across scleractinian corals and an outgroup, Corallimorpharia. Gene models were assigned to orthogroups using OrthoFinder (Emms and Kelly 2019). All included taxa are listed in **Table S3**. The focal taxa assembled in the present study are indicated by bold font and asterisks (*).

Figure 3. Top 10 gene ontology (GO) terms enriched in orthogroups undergoing phylogenetically significant expansion in *Siderastrea siderea, Dendrogyra cylindrus,* and *Colpophyllia natans*. Orthogroups were assigned using OrthoFinder (Emms and Kelly 2019). Gene families undergoing phylogenetically significant expansion were identified using CAFE5 (Mendes et al. 2021). GO enrichment analyses were performed in GOATools (Klopfenstein et al. 2018).
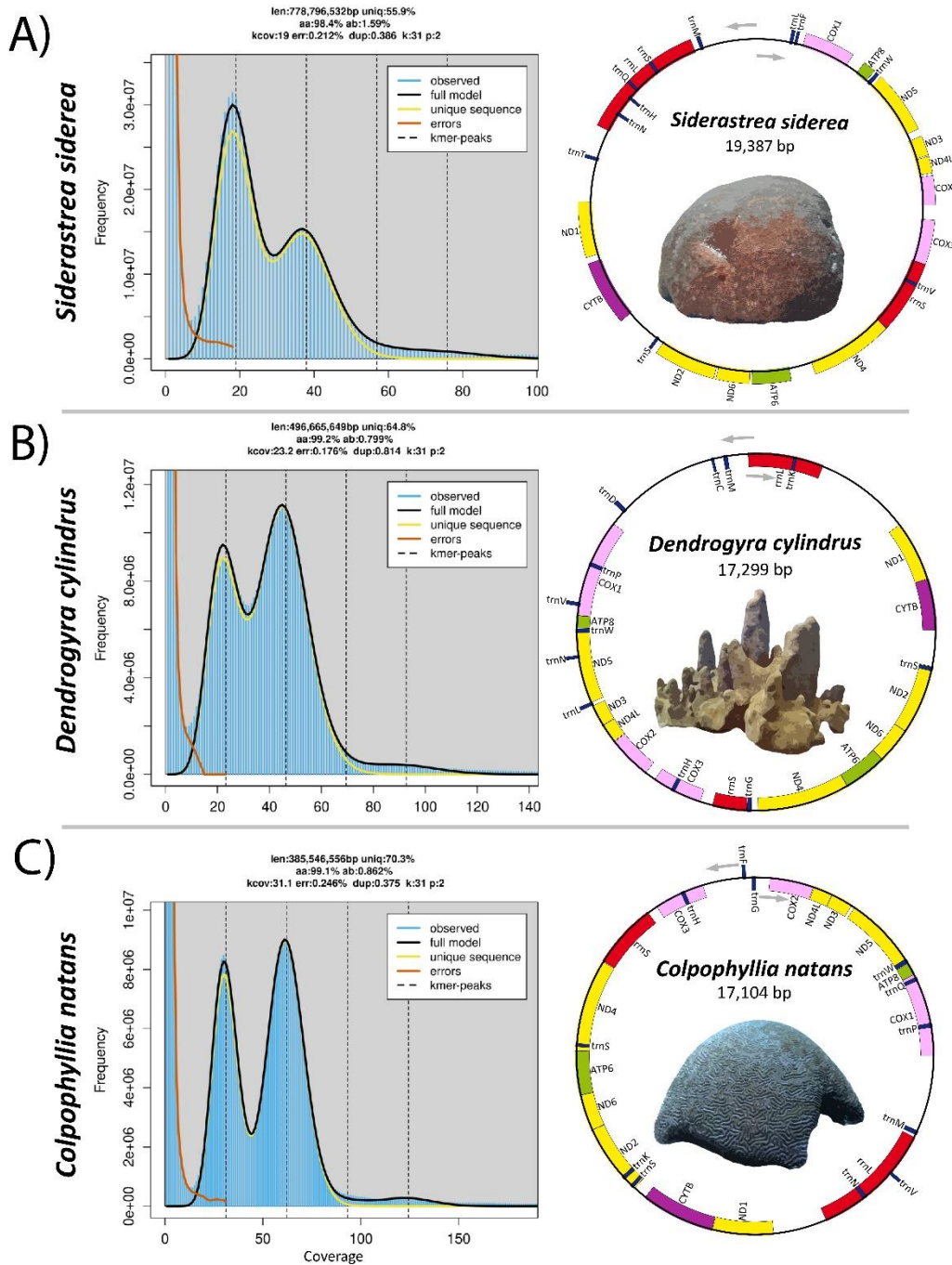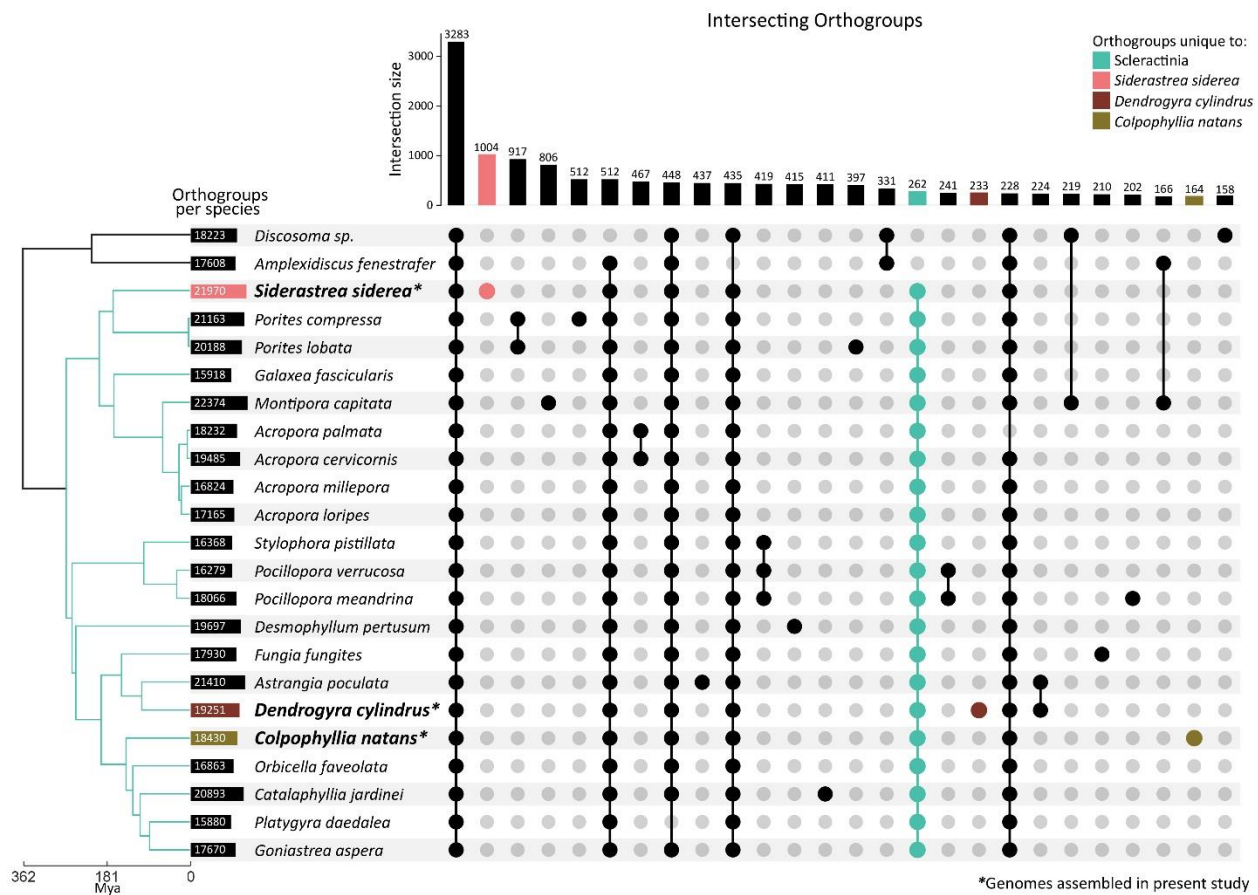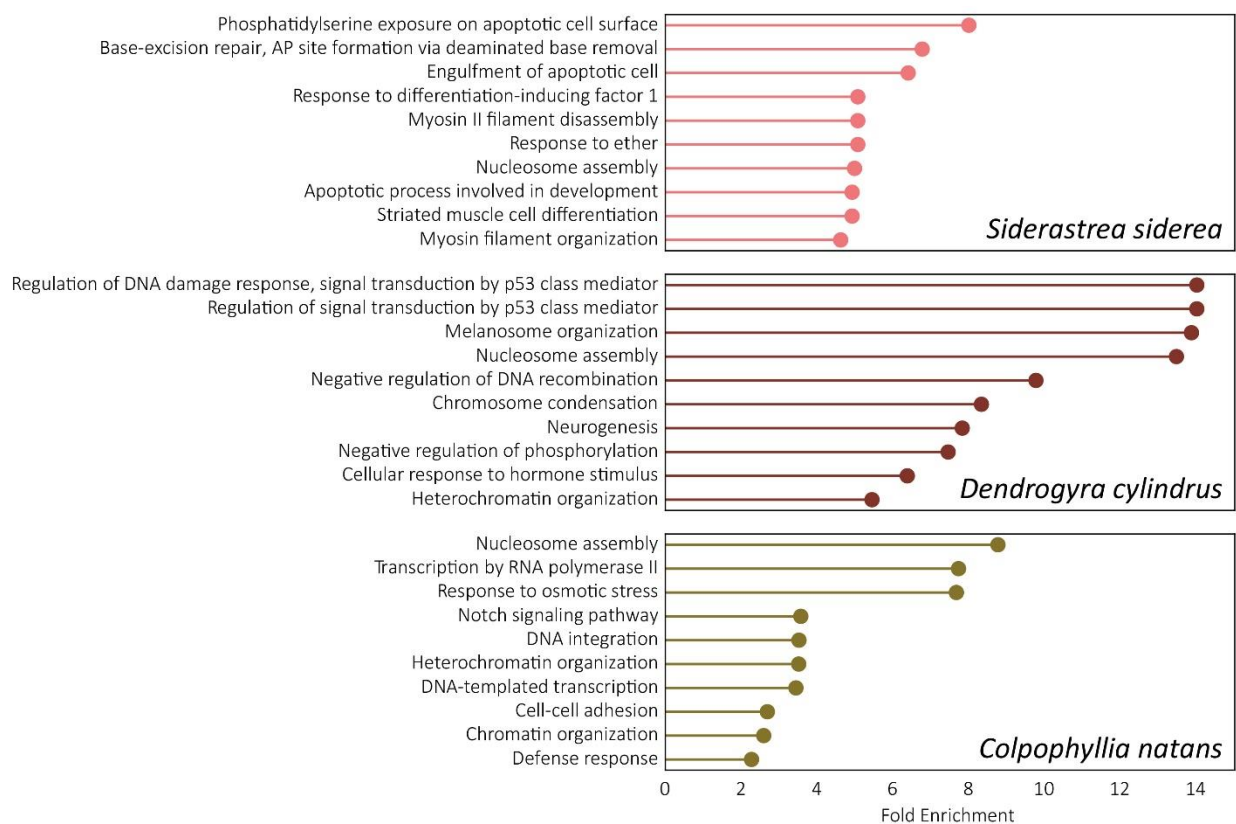
2

Figure 1
428x559 mm (x DPI)

1

2

Figure 2
354x253 mm (x DPI)

1

Figure 3
259x171 mm (x DPI)