Genomic comparison of the temperate coral *Astrangia poculata* with tropical corals yields insights into winter quiescence, innate immunity, and sexual reproduction

Kathryn H. Stankiewicz[1,2], Nadège Guiglielmoni[3], Sheila A. Kitchen[1,5], Jean-François Flot[3,4], Katie L. Barott[6], Sarah W. Davies[7], John R. Finnerty[7], Sean P. Grace[8], Leslie S. Kaufman[7], Hollie M. Putnam[9], Randi D. Rotjan[7], Koty H. Sharp[10], Esther C. Peters[11], Iliana B. Baums*[1,12,13,14]

***Corresponding author:** iliana.baums@hifmb.de

[1]Department of Biology, The Pennsylvania State University, University Park, PA 16802, USA

[2]Institute for Systems Biology, Seattle, WA 98109, USA

[3]Université libre de Bruxelles, Brussels 1050, Belgium

[4]Interuniversity Institute of Bioinformatics in Brussels, Brussels 1050, Belgium

[5]Texas A&M University at Galveston, Galveston, TX 77554, USA

[6]University of Pennsylvania, Philadelphia, PA 19104, USA

[7]Boston University, Boston, MA 02215, USA

[8]Southern Connecticut State University, Werth Center for Coastal and Marine Studies, New Haven, CT 06515, USA

[9]University of Rhode Island, Kingston, RI 02881, USA

[10]Roger Williams University, Bristol, RI 02809, USA

[11]Department of Environmental Science and Policy, George Mason University, Fairfax, VA 22030, USA

[12]Helmholtz Institute for Functional Marine Biodiversity at the University of Oldenburg (HIFMB), Carl von Ossietzky Universität Oldenburg, Im Technologiepark 5, 26129 Oldenburg, Germany

[13]Alfred Wegener Institute, Helmholtz-Centre for Polar and Marine Research, Am Handelshafen 12, Bremerhaven 27570, Germany

[14]Institute for Chemistry and Biology of the Marine Environment (ICBM), School of Mathematics and Science, Carl von Ossietzky Universität Oldenburg, Ammerländer Heerstraße 114-118, 26129 Oldenburg, Germany

Keywords: Chromosome-level genome assembly, corals, scleractinian, facultative symbiosis, evolution, gene family expansion, whole genome duplication, dormancy, mass spawning

Running Title: Genomics of the temperate coral *Astrangia poculata*

Abstract

Facultatively symbiotic corals provide important experimental models to explore the establishment, maintenance, and breakdown of the mutualism between corals and members of the algal family Symbiodiniaceae. Here, we report the *de novo* chromosome-scale genome assembly and annotation of the facultatively symbiotic, temperate coral *Astrangia poculata*. Though widespread segmental/tandem duplications of genomic regions were detected, we did not find strong evidence of a whole genome duplication (WGD) event. Comparison of the gene arrangement between *A. poculata* and the tropical coral *Acropora millepora* revealed considerable conserved colinearity despite ~415 million years of divergence. Gene families

related to sperm hyperactivation and innate immunity, including lectins, were found to contain more genes in *A. millepora* relative to *A. poculata*. Sperm hyperactivation in *A. millepora* is expected given the extreme requirements of gamete competition during mass spawning events in tropical corals, while lectins are important in the establishment of coral-algal symbiosis. By contrast, gene families involved in sleep promotion, feeding suppression, and circadian sleep/wake cycle processes were expanded in *A. poculata*. These expanded gene families may play a role in *A. poculata*'s ability to enter a dormancy-like state ("winter quiescence") to survive freezing temperatures at the northern edges of the species' range.

Introduction

Anthozoa, the largest class within the phylum Cnidaria, includes some of the most ecologically important and oldest clades of marine metazoans, estimated to have evolved as early as 771 million years ago (Mya) (McFadden et al., 2021). Among these are corals, a diverse group of solitary and colonial organisms that can form a symbiotic association with algae of the family Symbiodiniaceae in the shallow water of the tropic and temperate zones (LaJeunesse et al., 2018). Stony corals of the order Scleractinia contain the engineers of reef ecosystems and are generally divided into two major clades, Complexa and Robusta, which diverged approximately 415 Mya (Kitahara, Cairns, Stolarski, Blair, & Miller, 2010; Romano & Palumbi, 1996; Stolarski et al., 2011). Over the past several decades, tropical coral species have undergone mass mortality due to their sensitivity to bleaching in the face of anthropogenic climate change (Bellwood, Hughes, Folke, & Nyström, 2004; DeCarlo et al., 2017; Hughes et al., 2017). Understanding differences between species of variable temperature tolerance and adaptive strategies has become a focus for conservation efforts. In the "robust" clade, the temperate coral *Astrangia poculata*

(the Northern Star Coral) has recently emerged as a model system for these comparisons due to its facultative symbiosis and temperature tolerance ranging from near freezing to 24°C in Narragansett Bay, the northern part of its distribution (Jacques, Marshall, & Pilson, 1983).

Astrangia poculata, like many other shallow water corals, hosts a photosynthesizing algal symbiont, *Breviolum psygmophilum* (Lajeunesse, Parkinson, & Reimer, 2012). The symbiosis is facultative, with a gradient of symbiont density existing among individual polyps within a single colony of *A. poculata* and between sympatric colonies (Dimond & Carrington, 2007) as revealed by an aposymbiotic 'white' appearance, nearly or entirely devoid of symbionts, or a symbiotic 'brown' appearance (Figure 1a). Unlike many other shallow water coral species, *A. poculata* occurs across a broad temperature and latitudinal gradient from southern Massachusetts to the Gulf of Mexico (Cummings, 1983; Dimond & Carrington, 2007; Peters et al., 1988; Dimond et al 2013). At the northernmost parts of its range, during winter months high in the intertidal zone, *A. poculata* experiences a quiescence characterized by a dormancy state of reduced feeding and growth (Supplementary Figure S1) (Grace, 2017; Jacques et al., 1983). This appears to be an adaptation of *A. poculata* to the intertidal zone, where it is vulnerable to desiccation, predation, and extreme shifts in salinity and temperature. *A. poculata's* hardiness combined with the ability of researchers to experimentally isolate the contributions of the host and the symbiont in aposymbiotic and symbiotic colonies has made this species particularly interesting for the study of coral symbiosis in the face of climate change.

Previous work found signatures of adaptation in thermal tolerance of *A. poculata* across the species' range, with a cold-adapted population exhibiting higher metabolic rates than warm-adapted populations across several temperature treatments (Aichelman, Zimmerman, & Barshis, 2019). Symbiotic and aposymbiotic *A. poculata* have been shown to respond differently to warm

and cold temperatures at both the organismal and transcriptomic levels. Symbiotic colonies significantly outperformed aposymbiotic colonies at 9, 18, and 24°C with respect to wound healing (Burmester et al., 2017), while aposymbiotic colonies responded more strongly transcriptionally to a cold exposure than to a heat exposure treatment, particularly through upregulation of genes involved in the myosin complex, proteasome core, translation regulator activity, nucleic acid binding, extracellular matrix structural constituent, muscle system process, and proteolysis (Wuitchik et al., 2021). In a separate thermal stress experiment, the algal endosymbiont *Breviolum psygmophilum* responded more strongly than the *A. poculata* host to chronic heat stress (Chan et al., 2021). Further, season, rather than symbiont state, was shown to drive the structure of the microbiome in *A. poculata* (Sharp, Pratte, Kerwin, Rotjan, & Stewart, 2017). Thus, across multiple scales of measurement, thermal variation, particularly exposure to cold temperatures, appears to play an important role in the biology and ecology of this coral. However, the evolutionary roots and the genomic mechanisms driving the response to environmental change remain unclear for *A. poculata,* and for many other corals.

The use of comparative genomics has shed light on the evolution of basal metazoans. For example, studies have suggested the potential roles of whole-genome duplication (Mao & Satoh, 2019), horizontal gene transfer (Bhattacharya et al., 2016), and *de novo* biosynthesis pathways (Ying et al., 2018) in coral evolutionary trajectories. Many of these comparisons have focused on complex versus robust lineages, for which there exists a deep evolutionary split based on molecular and phylogenetic evidence (Figure 1b) (Kitahara et al., 2010; Romano & Palumbi, 1996; Stolarski et al., 2011). However, it is not well understood how phylogenetically widespread these genomic traits may be, and the inclusion of other corals representing a wider

array of ecological niches and adaptive strategies is warranted, which requires additional genomic resources.

While over the past several years there has been an increase in the availability of genomic resources for cnidarians, currently few of them are at chromosome-scale (however, see Fuller et al., 2020; Hu et al., 2020; Stephens et al., 2022) and none represent a facultatively symbiotic, temperate coral. To fill this gap, we present here the chromosome-scale assembly of a male *A. poculata* colony. The assembly is among the most contiguous and complete of available coral genome assemblies available to date. We characterized the structure and content of the *A. poculata* genome to investigate potential genomic drivers underlying this species' unique temperature tolerance and flexible symbiotic state. To determine the sex of the sequenced colony, we examined histological sections from the colony. Our aims were to 1) produce a high-quality reference genome for *A. poculata*, 2) use this genome assembly to explore several potential genomic mechanisms that may contribute to the unique plasticity of the species, and 3) characterize the degree of similarity of the gene repertoire and genome organization among *A. poculata* and other corals.

Methods

Genome sequencing

An aposymbiotic colony was collected from Fort Wetherill State Park in Jamestown, Rhode Island, USA in October of 2017. To avoid sequencing the symbiont, a colony with a minimal density of *B. psygmophilum* in its tissue was selected based on the colony's white appearance. A subsample of this colony was collected, frozen, and sent for DNA extraction, library preparation,

and sequencing at Dovetail Genomics. Two subsamples from this colony were maintained in aquaria at Pennsylvania State University, two at Boston University, and two at the University of Rhode Island for future use.

Sequencing involved a multi-strategy approach and included two Illumina datasets of paired-end 150 base-pair (bp) reads: one with 414 million reads and an estimated average insert size of 395 bp, and the second with 235 million reads and an estimated average insert size of 484 bp. Three Hi-C libraries were produced with 198 million, 266 million, and 257 million of 150 bp paired-end reads each. DNA extraction, library preparation, and sequencing of the Illumina and Hi-C data were carried out by Dovetail Genomics (https://dovetailgenomics.com/). All Illumina and Hi-C reads were trimmed using Cutadapt v2.9 with default settings (Martin, 2011).

In preparation for Oxford Nanopore sequencing, extracted genomic DNA was purified with AMPure XP Beads to improve the purity. Short fragments were discarded using Circulomics Short Reads Eliminator XS (Circulomics, Pacific Biosciences). The library was prepared using the Nanopore Ligation Sequencing Kit LSK109, starting with 2.1 µg of DNA, and yielded 1.4 µg of library. The final library was sequenced with a MinION on a R9.4 flow cell with fast base calling (Jain, Olsen, Paten, & Akeson, 2016). The flow cell was washed and reloaded three times (281 ng of DNA for the first load, 187 ng for subsequent loads). A total output of 6.79 Gigabases (Gb) was obtained with an N_{50} of 18 kilobases (kb) and an N_{90} of 5 kb. The reads were trimmed of the adaptors with Porechop (available at https://github.com/rrwick/Porechop), using default parameters. After trimming, the dataset reached 6.77 Gb.

De novo genome assembly

Several assembly strategies were compared to select the most robust approach (Supplementary Table S1). Assemblers that were investigated include Raven (Vaser & Šikić

2019), Flye (Kolmogorov, Yuan, Lin, & Pevzner, 2019), Canu (Koren et al., 2017), and wtdbg2 (Ruan & Li, 2020). Ultimately, the genome was assembled using wtdbg2 v2.5 under default settings. Haplotigs were purged using purge_haplotigs v1.1.1 (Roach, Schmidt, & Borneman, 2018) with default settings (following the recommendations in Guiglielmoni et al. (2021)) with both Illumina datasets mapped using bowtie2 v2.3.5.1 (Langmead & Salzberg, 2012). Polishing was done using HyPo v1.0.3 (Kundu, Casey, & Sung, 2019). Hi-C reads were mapped to the assembly and processed using bowtie2 v2.3.5.1 and hicstuff v2.3.0 (Matthey-Doret et al., 2020) with the parameters --enzyme DpnII --iterative --aligner bowtie2. The draft assembly was then scaffolded using instaGRAAL v0.1.6 no-opengl branch (Baudry et al., 2020), with default parameters (--levels 4, --cycles 100, --coverage-std 1, --neighborhood 5). The output was then refined using the module instaGRAAL-polish. BUSCO v5.4.7 was used to assess the completeness of the assembly (Simão, Waterhouse, Ioannidis, Kriventseva, & Zdobnov, 2015) against the Metazoa odb10 lineage. The genome size was estimated with the Illumina dataset of 235 million reads and the module kmercount.sh from BBmap v38.79 (Bushnell, 2014). The circularized mitochondrial genome was assembled from the Illumina reads using NOVOPlasty v 2.7.2 (Dierckxsens, Mardulyn, & Smits, 2016) with the publicly available *Acropora digitifera* mitochondrial genome as a reference (GenBank: KF448535.1). Our 14 chromosome-scale scaffolds were named Ap1, Ap2... Ap14 on the basis of their homologies with *A. millepora* chromosome-scale scaffolds bearing the same number (in the absence of a published karyotype for *A. poculata*).

Transcriptome assembly

RNA sequencing data from Chan et al. (2021) were used to construct a *de novo* transcriptome. First, to limit the potential for symbiont contamination, only reads for aposymbiotic individuals were used for the host transcriptome assembly. Reads were trimmed using Cutadapt v 3.4 (Martin, 2011) with a quality cut-off of 15 and a minimum read length of 50 nucleotides. The trimmed reads were then assembled into transcripts using rnaSPades v 3.12.0 (Bushmanova, Antipov, Lapidus, & Prjibelski, 2019). To further reduce the possibility of contamination from the algal endosymbiont, a custom database of Symbiodiniaceae protein sequences was assembled that included the following species: *Cladocopium goreaui* (H. Liu et al., 2018), *Breviolum psygmophilum* (Parkinson et al., 2016), *Breviolum minutum* (Shoguchi et al., 2013), *Fugacium kawaguti* (H. Liu et al., 2018), *Symbiodinium fitti* (Reich et al., 2021), *Symbiodinium microadriaticum* (Aranda et al., 2016), and *Symbiodinium tridacnidorum* (González-Pech et al., 2021). A BLAST nucleotide-to-protein alignment (blastx) of the assembled *A. poculata* transcriptome was conducted against this database using BLAST v 2.6.0 (Altschul, Gish, Miller, Myers, & Lipman, 1990; Camacho et al., 2009). Transcripts with at least 80% identity and with at least 100bp mapped length were filtered from the transcriptome.

Genome Annotation

To annotate repetitive content, *de novo* transposable element family identification and modeling was conducted using RepeatModeler v 1.0.1 (Smit & Hubley, 2008). RepeatMasker v 4.0.7 (A. Smit, Hubley, R & Green, P., 2015) was then used to soft-mask repetitive regions prior to gene modeling. Subsequent gene prediction included a multi-tool approach. First, *ab initio*

gene prediction was done using GeneMark-ES v 3.51 (Lukashin & Borodovsky, 1998). Gene prediction using protein-based evidence was conducted with Exonerate v 2.4.0 (Slater & Birney, 2005) using the UniProt eukaryote database (downloaded December 28, 2017). RNA-seq reads from Chan et al. (2021) were aligned to the genome using STAR v 020201 (Dobin et al., 2013) and incorporated into the automated training of gene prediction using Braker2 v 2.1.2 (Brůna, Hoff, Lomsadze, Stanke, & Borodovsky, 2020). From the resulting predictions, high quality gene models (HiQ), here defined as those having >=90% RNAseq coverage support, were extracted. Gene prediction informed by transcriptomic evidence was carried out using PASA v 2.3.3 (Haas et al., 2003) with the flag --TRANSDECODER to keep only the longest open reading frame per group of transcript isoforms (see section 'Transcriptome assembly' above for transcriptome assembly details). Consensus gene predictions were acquired using EvidenceModeler v 1.1.1 (Haas et al., 2008) weighting each line of evidence as follows: *ab initio* predictions: 1; RNA-seq based evidence: 1; protein-based evidence: 1; HiQ: 5; and transcriptomic-based evidence: 10. PASA v 2.3.3 was run a second time to update gene models and add annotations of untranslated regions (UTRs) using the consensus gene prediction from EVidenceModeler. Finally, tRNAscan-Se v 1.3.1 (Chan & Lowe, 2019) was used to annotate transfer RNAs. Genome Annotation Generator (GAG) v 2.0.1 (Geib et al., 2018) was used to extract CDS, protein, and mRNA sequences.

Functional annotation of predicted genes was conducted based on sequence similarity using BLAST v 2.6.0 (Altschul et al., 1990; Camacho et al., 2009) searches via blastp with the settings of eval= 1e-5, max target seqs = 5, and max hsp = 1. Blast queries were conducted against three databases: NCBI nr, UniProt Swiss-Prot, and TrEMBL. Hits to each database were combined and annotated with Gene Ontology (GO) terms using the UniProt-GOA mapping.

Investigating the possibility of whole-genome duplication

Given recent suggestion of a possible whole-genome duplication event in the genus *Acropora* (Mao & Satoh, 2019), we set out to determine whether a similar event may have occurred in *A. poculata*. Detection and classification of duplication in the genome was carried out in several ways. First, estimates were conducted using the tool MCScanX under default settings (Wang et al., 2012). Results were then compared to a run under more relaxed settings (max gap size increased to 50). In both cases, duplication origins were classified using the duplicate_gene_classifier module.

A second method of whole-genome duplication (WGD) detection employed was the tool wgd v1.1.2 (Zwaenepoel & Van de Peer, 2018), which relies on the distributions of synonymous substitutions per synonymous site (*Ks*). To conduct the analysis, genome-wide coding sequences (CDS) were filtered for the longest translatable isoform of each gene. wgd v1.1.2 was then run using the diamond aligner to compute the whole-paranome (the collection of all duplicate genes in a genome). A *Ks* distribution was constructed in pairwise mode and kernel density estimates were subsequently fit to the distribution and visualized. A *Ks* distribution for anchor pairs, defined as paralogs located on colinear duplicated segments, were similarly constructed and visualized. The shape of the *Ks* distribution was inspected for detection of ancient WGD with the expectation of an exponential decay shape in the absence of a WGD event (Zwaenepoel & Van de Peer, 2018).

Lastly, the whole-genome synteny aligner Satsuma2 (available at https://github.com/bioinfologics/satsuma2) was used to detect microhomologous regions between *A. poculata* chromosomes. Syntenic blocks of homologous sequences arranged in a colinear fashion between chromosomes were then then plotted using the tool Orthodotter

(https://github.com/institut-de-genomique/orthodotter) to produce an Oxford grid (Edwards, 1991), an approach used previously to detect whole-genome duplication in arthropods (Schwager et al., 2017).

Comparative genomics: Gene family and conserved synteny analysis

To characterize the genome organization and content of *A. poculata* relative to other cnidarians, we completed several comparative genomic analyses. Phylogenetic orthology inferences were carried out using OrthoFinder2 v 2.4.0 (Emms & Kelly, 2018) on *A. poculata* and 23 other available cnidarian proteomes (Figure 1b; Supplementary Table S2) using default parameters with the species tree inferred using the STAG algorithm (Emms & Kelly, 2018) and rooted using the STRIDE algorithm (Emms & Kelly, 2017). The phylogenetic species tree was visualized using the R package 'ggtree' version 3.6.2 (Xue et al., 2022). Gene Ontology (GO) enrichment of gene families unique to *A. poculata* was conducted using the clusterProfiler package (Yu, Wang, Han, & He, 2012) implemented in R v 4.0.5 with a p-value cutoff of 0.05, a multiple testing correction method of Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995), and a q-value cutoff of 0.2.

Acropora millepora, an obligately symbiotic coral of the complex clade, was selected for more detailed comparison with *A. poculata* as its assembly was similarly complete. The sizes of gene families common to the two species were compared following the methods of González-Pech et al. (2021). Using the orthogroups previously identified by Orthofinder2, size differences of gene families shared between *A. poculata* and *A. millepora* were evaluated using Fisher's exact test with the multiple testing correction method of Benjamini-Hochberg (Benjamini & Hochberg, 1995) and a significance threshold of adjusted $p \leq 0.05$. Gene Ontology (GO)

enrichment of gene families significantly larger in *A. poculata* and in *A. millepora* were conducted separately using the same methods as described above.

Conserved synteny between *A. poculata* and *A. millepora* was assessed with MCScanX_h (Wang et al., 2012) using the homologous genes between the species identified by OrthoFinder2. Default parameters were used to identify colinear blocks (gap size of 25 genes allowable, minimum of 5 genes per colinear block).

Histological examination

Three small fragments from one of the two still-living fragments of the sequenced *A. poculata* colony that were being maintained at Boston University were placed in Z-Fix concentrate:seawater (1:4) fixative in November 2017 and processed in the Histology Laboratory at George Mason University (GMU). Samples were decalcified in 10% ethylenediaminetetraacetic acid (EDTA) pH 7, trimmed into 5-mm strips and placed in labeled cassettes, dehydrated in a series of ethanol solutions (70%–100%), cleared using Clearify (StatLab), and embedded in Paraplast Regular (Leica Microsystems Surgipath). Tissue sections, 4-µm thick, were mounted on clean glass microscope slides and stained with Harris's hematoxylin (Statlab) and eosin-phloxine and Giemsa (StatLab)procedures (Price & Peters, 2018), then covered with Permount (Fisher Scientific) and a glass cover slip and examined using light microscopy.

Results

Genome statistics and assembly quality

Of the four assemblers tried, wtdbg2 (Ruan & Li, 2020) produced the highest quality assembly (Figure 2; Supplementary Table S1). After purging, polishing, and scaffolding, the assembly using wtdbg2 had an N_{50} of 31 Mb and a BUSCO score of 95.5% (93.8% single copies, 1.7% duplicate copies, 2.1% fragmented, and 2.4% missing; Table 1). The assembly size of 458 Mb is in line with BBMap predicted haploid size of 453 Mb (estimated ploidy of 2 and 40.95% repetitive). The 14 chromosome-level scaffolds match the known $2n=28$ formula typical of most scleractinian corals (Figure 2a; Supplementary Table S3) (Flot et al., 2006). Further, the *k*-mer completeness was 52.24%, close to the expected 50% for a haploid assembly (Figure 2b). Gene prediction of the *A. poculata* genome assembly yielded 44,839 gene models, of which 1,613 had functional annotations to transposons, leaving 38,998 protein coding genes excluding repetitive elements (Table 1). RepeatMasker predicted 39.29% of the genomic bases as repetitive elements, with an estimated GC content of 38.50%. Additionally, tRNAscan-Se identified 6,951 transfer RNAs. The assembled transcriptome (used to assist in gene modeling), included 516,681 assembled transcripts and a BUSCO score of 90.8% (Single:32.3%, Duplicated:58.5%, Fragmented:6.4%, Missing:2.8%). The high level of duplicates in the transcriptome result from retaining all isoforms for every gene.

Whole-genome duplication

Duplication analysis via MCScanX (Wang et al., 2012) revealed 88 syntenic blocks with 853 duplicated genes classified as putatively originating from whole genome (WGD) or segmental

duplication (SD; Supplementary Table S4). When relaxing MCScanX parameters by allowing a max gap size of 50 genes within colinear blocks (default gap size is 25), the number of duplications classified as whole genome or segmental increased to 1,005. Detected colinear blocks often involved more than two colinear regions, with some involving three or even four colinear regions. However, a *Ks* based approach using the tool wgd (Zwaenepoel & Van de Peer, 2018) resulted in an exponential decay shape of the distribution of synonymous substitutions per synonymous site, suggesting no signature of a WGD event in *A. poculata* (Figure 3). While we did detect many anchor pairs (colinear paralogs), the anchor *Ks* distribution also declined exponentially. This was consistent with the failure of Orthodotter to detect large colinear regions in the genome of *A. poculata* (Figure 4), suggesting that WGD, if it occurred, was too ancient and the genome had subsequently undergone too much rearrangement to leave an obvious colinearity signature. Rather, these results suggest widespread duplications (tandem, proximal and dispersed) within the *A. poculata* genome, which may have resulted in novel functional gene copies through the process of neofunctionalization (Teshima & Innan, 2008).

Comparative genomics: Gene family size and conserved synteny

Using a gene family analysis involving a total of 24 species, we defined a "core" cnidarian genome that consisted of 2,584 gene families shared among all cnidarians present in the analysis (Figure 5). We found 508 gene families unique to *A. poculata.* Interestingly, only 218 orthogroups were present in all anthozoans included in the analysis. GO enrichment in clusterProfiler (Yu et al., 2012) of the gene families unique to *A. poculata* resulted in 143 enriched GO terms (significance threshold of $p<0.05$ after adjusting for multiple testing; Supplementary Table S5).

Gene families that were common to *A. poculata* and *A. millepora* were assessed for differences in gene numbers using Fisher's Exact test following the methods of González-Pech et al. (2021). In total, 170 gene families were identified as significantly different in gene numbers between the two species (Figure 6a; Supplementary Table S6; adjusted $p \leq 0.05$). This included 73 gene families that were significantly larger in *A. millepora* and 97 gene families that were significantly larger in *A. poculata.* Gene families larger in *A. millepora* that were most different in size compared to *A. poculata* (according to $\log_2$ fold change) included gene families that putatively encoded for zinc finger CCHC domain-containing proteins, cation channel sperm-associated proteins, RING-box proteins, lectins, and serine/threonine-protein kinases. In contrast, gene families larger in *A. poculata* that were most different in size compared to *A. millepora* by $\log_2$ fold change included orthogroups that putatively encoded for transposable elements, G protein-coupled receptors, Zinc finger MYM-type proteins, RNA-directed DNA polymerases, orexin receptors, ATP-dependent DNA helicases, E3 ubiquitin-protein ligases, and histone H3 (Figure 6b; Supplementary Table S6). GO enrichment analysis resulted in 723 and 635 enriched terms for expanded families in *A. poculata* and *A. millepora*, respectively. Enriched terms in *A. millepora* expanded gene families included CatSper complex, sperm principal piece, zinc ion binding, and viral latency (Supplementary Table S7), while enriched terms in expanded *A. poculata* families included transposition, G protein-coupled receptor activity, and peptide receptor activity (Supplementary Table S8).

In addition to gene family size analysis, orthologues common to *A. poculata* and *A. millepora* were evaluated for conserved synteny. MCScanX analysis of colinearity revealed a high level of conserved gene synteny between *A. poculata* and *A. millepora* with 3,719 syntenic blocks

identified of at least 5 colinear genes. In total, 56.63% of orthologous genes were present in the colinear blocks between the two divergent coral species (Figure 6c).

Histology

Examination of the histology slides revealed that the sequenced *A. poculata* colony was male, with developing spermaries in Stages II–IV (Szmant, Yevich, & Pilson, 1980) in two of the tissue sections (Figure 7).

Discussion

Astrangia poculata has increasingly been used as a model coral system due to its temperature tolerance and flexibility in symbiont state (Aichelman et al., 2019; Burmester et al 2017; Burmester et al. 2018; Chan et al., 2021; Dimond & Carrington, 2007; Jacques et al., 1983; Peters et al., 1988; Sharp et al., 2017; DiRoberts et al. 2021; Wuitchik et al., 2021). Because *A. poculata* is facultatively symbiotic, the host and the algal symbiont response to manipulation can be distinguished—a study design that is often impossible in adult tropical corals, many of whom do not occur naturally in an aposymbiotic state. In addition to associating with an algal endosymbiont of the family Symbiodiniaceae, *A. poculata* creates a calcium carbonate skeleton similar to reef-building corals (Hayes & Goreau, 1977; Peters et al., 1988). However, corals are a diverse group of organisms with respect to their biology and ecology (e.g., habitat, morphology, and response to environmental change) (Chappell, 1980; Fabricius et al., 2011; Kusumoto et al., 2020; Muir, Wallace, Pichon, & Bongaerts, 2018). Thus, it is important to recognize the differences, as well as the similarities, between *A. poculata* and other corals. Here, we have developed a chromosome-scale genome assembly for *A. poculata*, which has allowed us to

characterize some of these differences and similarities between *A. poculata* and other cnidarians. These results have revealed insight into potential genomic drivers of cnidarian biology, such as *A. poculata's* astounding ability to enter a dormancy state when exposed to extreme cold temperatures, and elucidated the demands placed on sperm during mass spawning events of tropical corals.

No evidence of ancient whole genome duplication, but recent duplications are abundant

Whole-genome duplication (WGD) can provide new genetic material upon which selection or genetic drift may act (Holland & Ocampo Daza, 2018; Ohno, 1970). A possible WGD event in the most recent common ancestor of the genus *Acropora* was suggested using phylogenomic and comparative genomic techniques (Mao & Satoh, 2019). However, it is unknown if such an event may have occurred at other points in the scleractinian lineage where polyploidism is common. Our conservative results from MCScanX indicated that 853 duplicated genes (1.8%) were classified as possibly originating from large-scale duplication events, such as segmental or WGD (Supplementary Table S4). To further identify whether this may indeed be indicative of an ancient WGD event, we examined the distribution of synonymous substitutions per synonymous site (*Ks*). Under a model of a constant rate of duplication and loss, there should be an exponential decay shape to a *Ks* distribution, which is the shape we find in the distribution for *A. poculata* (Figure 3). In contrast, when a WGD event has occurred, it leaves a signature peak in the *Ks* distribution (Zwaenepoel & Van de Peer, 2018). Additionally, we examined the *Ks* distribution of the anchor pairs (paralogs located on colinear duplicated segments). However, many anchors detected represented small *Ks* values (0-0.1) and also followed an exponentially decaying shape. Similarly, a search for microsynteny using SatsumaSynteny followed with colinearity detection

using Orthodotter did not reveal abundant pairs of colinear regions (Figure 4). These results suggest that while large-scale segmental duplications may be present in the *A. poculata* genome, we are unable to detect a strong signature of an ancient WGD event. Because these approaches may not be able to detect very ancient duplications, a phylogenomic approach (Zwaenepoel & Van de Peer, 2018) will be required to further test for an ancient whole-genome duplication event in Scleractinia as additional high-quality genome assemblies representative of each taxonomic group across the cnidarian phylogeny become available.

While there was no evidence of an ancient WGD event in *A. poculata*, we found a high level of large-scale segmental duplications, tandem duplications, and putatively transposon-derived gene models (Table 1). This may explain the higher gene density of *A. poculata* (97.9 genes/Mb; Table 1) relative to the complex coral, *A. millepora* (63.4 genes/Mb). The gene density in *A. poculata* is in-line with those of other Robusta corals (*Pocillipora cf effusa* and *Pocillopora damicornis*) where pervasive tandem duplications were also detected (Noel et al., 2023).

Genome content of temperate *Astrangia poculata* versus tropical *Acropora millepora*

Orthologue analysis resulted in 508 orthogroups unique to *A. poculata* (Figure 5). These gene families were found to be enriched in 143 GO terms, including terms involving ubiquitin E3 ligases, regulation of transcription and gene expression, G protein-coupled receptor activity, and transposition (Supplementary Table S5). To further investigate the unique genomic content of *A. poculata*, we selected one tropical coral to include in a more detailed comparison. The *A. poculata* genome assembly is among the most complete and contiguous coral genomes to date. Many of the other currently available cnidarian genome assemblies remain considerably more fragmented. For this reason, we limited our subsequent genomic analyses to comparisons

between *A. poculata* and *A. millepora*. *Acropora millepora* has a chromosome-scale genome assembly (Fuller et al., 2020), and contrasts with *A. poculata* in its ecology and evolutionary history. While *A. poculata* is a facultatively symbiotic temperate coral of the robust clade, *A. millepora* is an obligately symbiotic tropical coral of the complex clade.

Conserved micro- and macro-synteny between complex and robust clades

Synteny analysis between *A. poculata* and *A. millepora* revealed considerable conserved colinearity (56.63%; Figure 6c) despite approximately 415 Mya of divergence of the two clades, Robusta (*A. poculata*) and Complexa (*A. millepora*) (Stolarski et al., 2011). This surprisingly high level of colinearity is in line with previous work comparing other complex and robust coral species. Ying et al. (2018) found that the extent of conserved gene order within Scleractinia, regardless of clade, was relatively high compared to the level of conserved synteny between sea anemones *Exaiptasia* and *Nematostella* in the order Actinaria. Our results further lend support to this conclusion of consistently high conserved gene order across scleractinians.

Differential gene family expansions related to innate immunity and symbiosis

While gene order analysis highlighted similarities between *A. poculata* and *A. millepora*, gene family size comparisons revealed differences (Figure 6a; Figure 6b; Supplementary Table S6). Gene family expansions are often observed during adaptation in corals (van Oppen & Medina, 2020). Notably, of the gene families expanded in *A. millepora* relative to *A. poculata*, the gene family with the largest $\log_2$ fold change contained several copies of Zinc finger CCHC domain-containing protein 3 (orthogroup OG0000470; Figure 6a; Figure 6b; Supplementary Table S6),

which plays a role in innate immune response to viruses (Lian et al., 2018). Defense against viruses is likely important in *A. millepora*, as previous work has identified massive viral outbreaks in this species (Correa et al., 2016). Further, obligately symbiotic cnidarians have a more advanced innate immunity repertoire relative to non-symbiotic relatives, possibly driven by the dynamic and constant interaction with the algal endosymbiont (Cunning et al., 2018; Shinzato et al., 2011; Shumaker et al., 2019; van Oppen & Medina, 2020; Voolstra et al., 2017). These previous findings relied on comparisons to more distantly related, non-symbiotic anthozoans. However, here we determine that this holds with a comparison to a more closely related facultatively symbiotic coral, as opposed to a non-coral anthozoan relative.

The molecular elements governing the uptake, maintenance, and breakdown of symbiosis in corals still remain largely unclear. Previous work does indicate that features of the innate immune system of the host, notably lectins, play an important role (Fransolet, Roberty, & Plumier, 2012; Hu, Zheng, Fan, & Zheng, 2020; Kvennefors, Leggat, Hoegh-Guldberg, Degnan, & Barnes, 2008; Takeuchi et al., 2021; Zhou et al., 2018). Lectins are pattern recognition proteins that bind to carbohydrates (Goldstein, Hughes, Monsigny, Osawa, & Sharon, 1980). Within the top 15 orthogroups significantly larger in *A. millepora* relative to *A. poculata* were gene families related to innate immunity that have previously been implicated in establishing symbiotic associations in corals, including C-type lectins and macrophage mannose receptors that mediate endocytosis of glycoproteins (orthogroups OG0000360, OG0000958, and OG0001199; Figure 6a; Figure 6b; Supplementary Table S6). Kvennefors et al. (2008) isolated a mannose-binding lectin in *A. millepora* and demonstrated its affinity to binding to both pathogens and algal dinoflagellates of the family Symbiodiniaceae. Since then, comparative genomics and single-cell RNA sequencing have further emphasized the role of lectins in the

cnidarian-algal symbiosis in additional species (Cunning et al., 2018; Hu et al., 2020). Here, we were able to compare two scleractinian corals, one facultatively symbiotic and the other obligately symbiotic. We found that pattern recognition proteins (e.g., lectins) are expanded in *A. millepora* relative to *A. poculata*. Future study characterizing the potential differences in the establishment and maintenance of symbiosis between obligately vs facultatively symbiotic corals is warranted, and the results here highlight lectin-related gene families as excellent targets.

Sexual reproduction in mass spawning tropical corals

In addition to innate immunity, top gene families larger in *A. millepora* compared to *A. poculata* included those related to sexual reproduction (Figure 6b; Figure 6a; Supplementary Table S6), with functions involved in sperm cell hyperactivation (cation channel sperm-associated protein subunit epsilon; orthogroup OG0000900) and meiosis (RING-box protein 1; orthogroup OG0000130). Further, significantly enriched GO terms in the *A. millepora* expanded gene families included CatSper complex (GO:0036128), sperm principal piece (GO:0097228), and sperm flagellum (GO:0036126; Supplementary Table S7). These findings are interesting because of the difference in reproductive modes between *A. millepora* and *A. poculata*. *Astrangia poculata* is a gonochoric species that reproduces via broadcast spawning wherein gametes are released into the water column prior to fertilization (Szmant, Yevich, & Pilson, 1980). Spawning occurs annually from August to September based on the seasonal maximum temperature, with a second cycle sometimes observed in October or November (Peters et al., 1988). *A. poculata* sperm are unlikely to have to compete with many other corals for fertilization of eggs because there are only a few other temperate coral species. We were able to confirm that the sequenced *A. poculata* colony was male as indicated by the presence of spermaries (Figure

7). In contrast, *A. millepora* is a hermaphroditic broadcast spawning species that reproduces during 'mass spawning' events synchronized with the occurrence of the full moon (Kaniewska, Alon, Karako-Lampert, Hoegh-Guldberg, & Levy, 2015). During these annual spawning events, *A. millepora* releases gametes simultaneously with over 100 other coral species, as well as hundreds of other invertebrates over the course of only a few nights (Babcock et al., 1986; Harrison, 2011). This creates competition between gametes, as well as the opportunity for interspecific hybridization. Expansion of these reproductive-related gene families have the potential to soften or maintain species boundaries and would be great targets for future studies examining adaptation to mass spawning in corals.

Genome plasticity in *A. poculata*: transposition and epigenetics

Gene families that were larger in *A. poculata* relative to *A. millepora* included families of transposable elements (Figure 6b; Figure 6a; Supplementary Table S6; orthogroups OG0000015, OG0000418). Further, significantly enriched GO terms in the *A. poculata* expanded gene families included transposition (GO:0032196) and DNA-mediated transposition (GO:0006313; Supplementary Table S8). Retrotransposition has been found to contribute to gene family expansion in Symbiodiniaceae (González-Pech et al., 2021; Lin et al., 2015) and, based on our results, may also play a role in the host *A. poculata*. Further, transposable elements promote adaptation and drive genome plasticity in many species, including bacteria, fungi, plants, and animals (Bennett, 2004; Feschotte & Pritham, 2007; Leitch & Leitch, 2008; Mat Razali, Cheah, & Nadarajah, 2019; Yuan et al., 2021). Epigenetic modification is involved in host genome regulation of transposable elements (Matzke, Mette, & Matzke, 2000). Interestingly, we also found that a family of histone H3 proteins was expanded in *A. poculata* relative to *A. millepora*

(orthogroup OG0000255; Figure 6a, Figure 6b; Supplementary Table S6). In eukaryotes, histone H3 is one of the core histone proteins involved in structuring chromatin (Kornberg, 1977; McGhee & Felsenfeld, 1980). The sequence variants, as well as different modification states of histone H3 are thought to influence gene regulation (Hake et al., 2006; Jiang & Berger, 2017; Klemm, Shipony, & Greenleaf, 2019; Kouzarides, 2007; Loyola & Almouzni, 2007; Maehara et al., 2015; Sarma & Reinberg, 2005). In plants, histone H3 plays a role in development and abiotic stress (Otero, Desvoyes, & Gutierrez, 2014; Yuan, Liu, Luo, Yang, & Wu, 2013). These results suggest that transposable elements and epigenetic modification may play an important role in the plasticity of *A. poculata*.

Winter quiescence in *A. poculata*

Among the top gene families expanded in *A. poculata* was a family of G-coupled protein receptors, including RYamide receptors, orexin receptors, and neuropeptide SIFamide receptors (orthogroup OG0000211; Figure 6a; Figure 6b; Supplementary Table S6). Expanded *A. poculata* gene families were also enriched for G protein-coupled receptor activity (GO:0004930; Supplementary Table S8). In *Drosophila*, RYamide receptors are possibly associated with feeding suppression (Ida et al., 2011), while neuropeptide SIFamide receptors have been associated with promotion of sleep (Park, Sonn, Oh, Lim, & Choe, 2014). Similarly, orexin receptors are known to regulate circadian sleep/wake cycles in mammals (Chemelli et al., 1999). The expansion of this gene family in *A. poculata* relative to *A. millepora* may explain *A. poculata*'s ability to enter a dormant state when exposed to near freezing temperatures. During winter months in the intertidal and subtidal regions at the northernmost edges of the species' range, *A. poculata* enters this dormancy, referred to as "winter quiescence", when water

temperatures plummet to below 6°C (Grace, 2017). During this state, polyps are retracted and oral discs (Peters, 2016) are puffed out, while feeding is reduced or ceased entirely (Grace, 2017; Jacques et al., 1983) (Supplementary Figure S1). This finding has relevance to tropical systems, in which some Mediterranean anthozoans enter a similar "summer dormancy" state, including corals (Caroselli, Falini, Goffredo, Dubinsky, & Levy, 2015; Coma, Ribes, Gili, & Zabala, 2000). Overall, these results indicate that gene family expansions may have contributed to adaption of *A. poculata* to the high variance in environmental conditions that this species experiences temporally and spatially across its range.

Conclusion

In this study, we present the first chromosome-scale assembly of the facultatively symbiotic, temperate coral *A. poculata*. Our contribution of a high-quality genome resource for *A. poculata* advances the use of this species as an experimental model and lays the groundwork for numerous future studies, as *A. poculata* is an important emerging model for coral health (Neff, 2020). Further, comparison of the *A. poculata* genome to the tropical, obligately symbiotic coral *A. millepora* uncovered potential genomic drivers of unique features of not only *A. poculata*, but of *A. millepora*, as well. Taken together, these results have generated genomic targets for future study of adaptation in these species and emphasize the power of comparative genomics to reveal novel insights into the biology of corals.

Data Availability Statement

All raw sequence data and the genome assembly have been submitted to NCBI under BioProject PRJNA1123198. The genome assembly and annotation files are also available at https://zenodo.org/records/14110456. Supplemental Files are available at https://zenodo.org/records/14226509. The scripts associated data analysis presented here are available at https://github.com/kate-stankiewicz/apoculata_genome_assembly.

Acknowledgements

We thank all authors who made their coral genome assemblies available for the comparative genomic analyses in this work. Special thanks to Sam Piorkowski and Claire Klippel for their computational and aquaria support; and to Amanda Drill for histology assistance.

Funding

This work was made possible by NSF grant OCE-1537959 to IBB, NIH grant T32: Computation, Bioinformatics, and Statistics (CBIOS) Training Program to KHS, and the Pennsylvania State University Biology Department. IBB acknowledges support by the Open Access publication fund of Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung.

Competing Interests

The authors declare no conflict of interest.

References

Aichelman, H. E., Zimmerman, R. C., & Barshis, D. J. (2019). Adaptive signatures in thermal performance of the temperate coral *Astrangia poculata*. *Journal of Experimental Biology, 222*(5). doi:10.1242/jeb.189225

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology, 215*(3), 403-410. doi:https://doi.org/10.1016/S0022-2836(05)80360-2

Aranda, M., Li, Y., Liew, Y. J., Baumgarten, S., Simakov, O., Wilson, M. C., . . . Voolstra, C. R. (2016). Genomes of coral dinoflagellate symbionts highlight evolutionary adaptations conducive to a symbiotic lifestyle. *Scientific Reports, 6*(1), 39734. doi:10.1038/srep39734

Babcock, R. C., Bull, G. D., Harrison, P. L., Heyward, A. J., Oliver, J. K., Wallace, C. C., & Willis, B. L. (1986). Synchronous spawnings of 105 scleractinian coral species on the Great Barrier Reef. *Marine Biology, 90*(3), 379-394. doi:10.1007/BF00428562

Baudry, L., Guiglielmoni, N., Marie-Nelly, H., Cormier, A., Marbouty, M., Avia, K., . . . Koszul, R. (2020). instaGRAAL: chromosome-level quality scaffolding of genomes using a proximity ligation-based scaffolder. *Genome Biology, 21*(1), 148. doi:10.1186/s13059-020-02041-z

Baumgarten, S., Simakov, O., Esherick, L. Y., Liew, Y. J., Lehnert, E. M., Michell, C. T., ... & Voolstra, C. R. (2015). The genome of *Aiptasia*, a sea anemone model for coral symbiosis. *Proceedings of the National Academy of Sciences*, *112*(38), 11893-11898.

Bellwood, D. R., Hughes, T. P., Folke, C., & Nyström, M. (2004). Confronting the coral reef crisis. *Nature*, *429*(6994), 827-833. doi:10.1038/nature02691

Benjamini, Y. and Hochberg, Y. (1995), Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57: 289-300. doi: 10.1111/j.2517-6161.1995.tb02031.x

Bennett, P. M. (2004). Genome Plasticity. In N. Woodford & A. P. Johnson (Eds.), *Genomics, Proteomics, and Clinical Bacteriology: Methods and Reviews* (pp. 71-113). Totowa, NJ: Humana Press.

Bhattacharya, D., Agrawal, S., Aranda, M., Baumgarten, S., Belcaid, M., Drake, J. L., . . . Falkowski, P. G. (2016). Comparative genomics explains the evolutionary success of reef-forming corals. *eLife, 5*, e13288. doi:10.7554/eLife.13288

Brůna, T., Hoff, K. J., Lomsadze, A., Stanke, M., & Borodovsky, M. (2020). BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. *bioRxiv*, 2020.2008.2010.245134. doi:10.1101/2020.08.10.245134

Buitrago-López, C., Mariappan, K. G., Cárdenas, A., Gegner, H. M., & Voolstra, C. R. (2020). The genome of the cauliflower coral *Pocillopora verrucosa*. *Genome biology and evolution, 12*(10), 1911-1917.

Burmester, E. M., Breef-Pilz, A., Lawrence, N. F., Kaufman, L., Finnerty, J. R., & Rotjan, R. D. (2018). The impact of autotrophic versus heterotrophic nutritional pathways on colony health and wound recovery in corals. *Ecology and evolution, 8*(22), 10805–10816. https://doi.org/10.1002/ece3.4531

Burmester E. M., Finnerty J. R., Kaufman L., Rotjan R. D. (2017) Temperature and symbiotic state impact healing in experimentally wounded corals. *Marine Ecology Progress Series* 570, 87-99. doi:10.3354/meps12114

Bushmanova, E., Antipov, D., Lapidus, A., & Prjibelski, A. D. (2019). rnaSPAdes: a de novo transcriptome assembler and its application to RNA-Seq data. *GigaScience, 8*(9). doi:10.1093/gigascience/giz100

Bushnell, B. (2014). BBMap: a fast, accurate, splice-aware aligner, 9th Annual Genomics of Energy & Environment Meeting, Walnut Creek, CA, March 17-20, 2014

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

Caroselli, E., Falini, G., Goffredo, S., Dubinsky, Z., & Levy, O. (2015). Negative response of photosynthesis to natural and projected high seawater temperatures estimated by pulse amplitude modulation fluorometry in a temperate coral. *Frontiers in Physiology, 6*. doi:10.3389/fphys.2015.00317

Chan, A. N., González-Guerrero, L. A., Iglesias-Prieto, R., Burmester, E. M., Rotjan, R. D., Finnerty, J. R., & Baums, I. B. (2021). An algal symbiont *Breviolum psygmophilum* responds more strongly to chronic high temperatures than its facultatively symbiotic coral host *Astrangia poculata*. *bioRxiv*, 2021.2002.2008.430325. doi:10.1101/2021.02.08.430325

Chan, P. P., & Lowe, T. M. (2019). tRNAscan-SE: Searching for tRNA genes in genomic sequences. *Methods in Molecular Biology , 1962*, 1-14. doi:10.1007/978-1-4939-9173-0_1

Chapman, J. A., Kirkness, E. F., Simakov, O., Hampson, S. E., Mitros, T., Weinmaier, T., ... & Steele, R. E. (2010). The dynamic genome of *Hydra*. *Nature*, *464*(7288), 592-596.

Chappell, J. (1980). Coral morphology, diversity and reef growth. *Nature, 286*(5770), 249-252. doi:10.1038/286249a0

Chemelli, R. M., Willie, J. T., Sinton, C. M., Elmquist, J. K., Scammell, T., Lee, C., . . . Yanagisawa, M. (1999). Narcolepsy in orexin knockout mice: molecular genetics of sleep regulation. *Cell, 98*(4), 437-451. doi:10.1016/S0092-8674(00)81973-X

Coma, R., Ribes, M., Gili, J.-M., & Zabala, M. (2000). Seasonality in coastal benthic ecosystems. *Trends in Ecology & Evolution, 15*(11), 448-453. doi: 10.1016/S0169-5347(00)01970-4

Conlan, J. A., Bay, L. K., Jones, A., Thompson, A., & Francis, D. S. (2020). Seasonal variation in the lipid profile of *Acropora millepora* at Halfway Island, Great Barrier Reef. Coral Reefs 39:1753–1765. doi: 10.1007/s00338-020-02001-w

Conti-Jerpe, I. E., Thompson, P. D., Wong, C. W. M., Oliveira, N. L., Duprey, N. N, Moynihan, M. A., & Baker, D. M. (2020). Trophic strategy and bleaching resistance in reef-building corals. *Science Advances 6*, eaaz5443.

Cooke, I., Ying, H., Forêt, S., Bongaerts, P., Strugnell, J. M., Simakov, O., Zhang, J., Field, M. A., Rodriguez-Lanetty, M., Bell, S. C., Bourne, D. G., van Oppen, M. J. H., Ragan, M. A., & Miller, D. J. (2020). Genomic signatures in the coral holobiont reveal host adaptations driven by Holocene climate change and reef specific symbionts. *Science Advances*. https://doi.org/10.1126/sciadv.abc6318

Correa, A. M. S., Ainsworth, T. D., Rosales, S. M., Thurber, A. R., Butler, C. R., & Vega Thurber, R. L. (2016). Viral outbreak in corals associated with an *in situ* bleaching event: atypical herpes-like viruses and a new megavirus infecting *Symbiodinium*. *Frontiers in Microbiology, 7*. doi:10.3389/fmicb.2016.00127

Cummings, C. E. (1983). *The biology of Astrangia danae.* (Ph.D. Dissertation). University of Rhode Island, Kingston, RI.

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

DeCarlo, T. M., Cohen, A. L., Wong, G. T. F., Davis, K. A., Lohmann, P., & Soong, K. (2017). Mass coral mortality under local amplification of 2 °C ocean warming. *Scientific Reports, 7*(1), 44586. doi:10.1038/srep44586

Dierckxsens, N., Mardulyn, P., & Smits, G. (2016). NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res, 45*(4), e18-e18. doi:10.1093/nar/gkw955

Dimond, J., & Carrington, E. (2007). Temporal variation in the symbiosis and growth of the temperate scleractinian coral *Astrangia poculata*. *Marine Ecology Progress Series, 348*, 161-172. doi: 10.3354/meps07050.

Dimond, J.L., Kerwin, A.H., Rotjan, R. *et al.* (2013). A simple temperature-based model predicts the upper latitudinal limit of the temperate coral *Astrangia poculata*. *Coral Reefs* **32**, 401–409. doi: 10.1007/s00338-012-0983-z

DiRoberts, L.E., Dudek, A., Ray, N.E., Fulweiler, R.W. and Rotjan, R.D. (2021). Testing assumptions of nitrogen cycling between a temperate, model coral host and its facultative symbiont: symbiotic contributions to dissolved inorganic nitrogen assimilation. *Marine Ecology Progress Series, 670*, 61-74. doi: 10.3354/meps13731

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., . . . Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England), 29*(1), 15-21. doi:10.1093/bioinformatics/bts635

Duman, J. G., & Newton, S. S. (2020). Antifreeze proteins in other species. In: Ramløv, H., Friis, D. (eds) *Antifreeze Proteins* Volume 1. Springer, Cham. https://doi.org/10.1007/978-3-030-41929-5_8

Edwards J. H. (1991). The Oxford Grid. *Annals of human genetics*, *55*(1), 17–31. doi:1111/j.1469-1809.1991.tb00394.x

Emms, D. M., & Kelly, S. (2017). STRIDE: Species tree root inference from gene duplication events. Molecular Biology and Evolution, 34(12), 3267–3278.

Emms, D. M., & Kelly, S. (2018). OrthoFinder2: fast and accurate phylogenomic orthology analysis from gene sequences. bioRxiv, 466201. doi:10.1101/466201

Emms, D. M., & Kelly, S. (2018). STAG: Species tree inference from all genes. In Evolutionary Biology (No. biorxiv;267914v1). bioRxiv. https://www.biorxiv.org/content/10.1101/267914v1

Fabricius, K. E., Langdon, C., Uthicke, S., Humphrey, C., Noonan, S., De'ath, G., . . . Lough, J. M. (2011). Losers and winners in coral reefs acclimatized to elevated carbon dioxide concentrations. *Nature Climate Change, 1*(3), 165-169. doi:10.1038/nclimate1122

Feschotte, C., & Pritham, E. J. (2007). DNA transposons and the evolution of eukaryotic genomes. *Annual review of genetics, 41*, 331-368. doi:10.1146/annurev.genet.40.110405.090448

Fletcher C, Pereira da Conceicoa L, Natural History Museum Genome Acquisition Lab et al. (2023). The genome sequence of the starlet sea anemone, *Nematostella vectensis* (Stephenson, 1935) [version 1; peer review: 1 approved, 1 approved with reservations, 1

not approved]. *Wellcome Open Res 2023, 8:79* https://doi.org/10.12688/wellcomeopenres.18991.1

Flot, J.-F., Ozouf-Costaz, C., Tsuchiya, M., & Van Woesik, R. (2006). *Comparative coral cytogenetics. Proceedings of the 10th International Coral Reef Symposium*, 1, 4-8

Fransolet, D., Roberty, S., & Plumier, J.-C. (2012). Establishment of endosymbiosis: The case of cnidarians and *Symbiodinium*. *Journal of Experimental Marine Biology and Ecology, 420-421*, 1-7. doi:10.1016/j.jembe.2012.03.015

Fuller, Z. L., Mocellin, V. J. L., Morris, L. A., Cantin, N., Shepherd, J., Sarre, L., . . . Przeworski, M. (2020). Population genetics of the coral *Acropora millepora*: Toward genomic prediction of bleaching. *Science, 369*(6501), eaba4674. doi:10.1126/science.aba4674

Geib, S. M., Hall, B., Derego, T., Bremer, F. T., Cannoles, K., & Sim, S. B. (2018). Genome Annotation Generator: a simple tool for generating and correcting WGS annotation tables for NCBI submission. *GigaScience, 7*(4). doi:10.1093/gigascience/giy018

Goldstein, I. J., Hughes, R. C., Monsigny, M., Osawa, T., & Sharon, N. (1980). What should be called a lectin? *Nature, 285*(5760), 66-66. doi:10.1038/285066b0

González-Pech, R. A., Stephens, T. G., Chen, Y., Mohamed, A. R., Cheng, Y., Shah, S., . . . Chan, C. X. (2021). Comparison of 15 dinoflagellate genomes reveals extensive sequence and structural divergence in family Symbiodiniaceae and genus *Symbiodinium*. *BMC Biology, 19*(1), 73. doi:10.1186/s12915-021-00994-6

Grace, S. (2017). Winter quiescence, growth rate, and the release from competition in the temperate scleractinian coral *Astrangia poculata* (Ellis & Solander 1786). *Northeastern Naturalist* 24:B119–B134. doi:10.1656/045.024.s715

Guiglielmoni, N., Houtain, A., Derzelle, A. *et al.* Overcoming uncollapsed haplotypes in long-read assemblies of non-model organisms. *BMC Bioinformatics* **22**, 303 (2021). doi: 10.1186/s12859-021-04118-3

Haas, B. J., Delcher, A. L., Mount, S. M., Wortman, J. R., Smith, R. K., Jr., Hannick, L. I., . . . White, O. (2003). Improving the *Arabidopsis* genome annotation using maximal transcript alignment assemblies. *Nucleic Acids Res, 31*(19), 5654-5666. doi:10.1093/nar/gkg770

Haas, B. J., Salzberg, S. L., Zhu, W., Pertea, M., Allen, J. E., Orvis, J., . . . Wortman, J. R. (2008). Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biology, 9*(1), R7. doi:10.1186/gb-2008-9-1-r7

Hake, S. B., Garcia, B. A., Duncan, E. M., Kauer, M., Dellaire, G., Shabanowitz, J., . . . Hunt, D. F. (2006). Expression patterns and post-translational modifications associated with mammalian histone H3 variants. *Journal of Biological Chemistry, 281*(1), 559-568.

Harrison, P. L. (2011). Sexual reproduction of scleractinian corals. In Z. Dubinsky & N. Stambler (Eds.), *Coral Reefs: An Ecosystem in Transition* (pp. 59-85). Dordrecht: Springer Netherlands.

Hayes, R. L., & Goreau, N. I. (1977). Intracellular crystal-bearing vesicles in the epidermis of scleractinian corals, *Astrangia danae* (Agassiz) and *Porites porites* (Pallas). *Biol Bull, 152*(1), 26-40. doi:10.2307/1540724

Holland, L. Z., & Ocampo Daza, D. (2018). A new look at an old question: when did the second whole genome duplication occur in vertebrate evolution? *Genome Biology, 19*(1), 209. doi:10.1186/s13059-018-1592-0

Hu, M., Zheng, X., Fan, C.-M., & Zheng, Y. (2020). Lineage dynamics of the endosymbiotic cell type in the soft coral *Xenia*. *Nature, 582*(7813), 534-538. doi:10.1038/s41586-020-2385-7

Hughes, T. P., Kerry, J. T., Álvarez-Noriega, M., Álvarez-Romero, J. G., Anderson, K. D., Baird, A. H., . . . Wilson, S. K. (2017). Global warming and recurrent mass bleaching of corals. *Nature, 543*, 373. doi:10.1038/nature21707

Ida, T., Takahashi, T., Tominaga, H., Sato, T., Kume, K., Ozaki, M., . . . Kojima, M. (2011). Identification of the novel bioactive peptides dRYamide-1 and dRYamide-2, ligands for a neuropeptide Y-like receptor in Drosophila. *Biochem Biophys Res Commun, 410*(4), 872-877. doi:10.1016/j.bbrc.2011.06.081

Imbs, A. B., Latyshev, N. A., Dautova, T. N., & Latypov, Y. Y. (2010). Distribution of lipids and fatty acids in corals by their taxonomic position and presence of zooxanthellae. *Mar. Ecol. Prog. Ser. 409*, 65-75. doi: 10.3354/meps08622

Jacques, T. G., Marshall, N., & Pilson, M. E. Q. (1983). Experimental ecology of the temperate scleractinian coral *Astrangia danae*. *Marine Biology, 76*(2), 135-148. doi:10.1007/BF00392730

Jain, M., Olsen, H. E., Paten, B., & Akeson, M. (2016). The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome Biology, 17*(1), 239. doi:10.1186/s13059-016-1103-0

Jiang, D., & Berger, F. (2017). Histone variants in plant transcriptional regulation. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms, 1860*(1), 123-130. doi: 10.1016/j.bbagrm.2016.07.002

Jiang, J. B., Quattrini, A. M., Francis, W. R., Ryan, J. F., Rodríguez, E., & McFadden, C. S. (2019). A hybrid de novo assembly of the sea pansy (*Renilla muelleri*) genome. *GigaScience, 8*(4), giz026.

Kaniewska, P., Alon, S., Karako-Lampert, S., Hoegh-Guldberg, O., & Levy, O. (2015). Signaling cascades and the importance of moonlight in coral broadcast mass spawning. *eLife, 4*, e09991. doi:10.7554/eLife.09991

Kitahara, M. V., Cairns, S. D., Stolarski, J., Blair, D., & Miller, D. J. (2010). A comprehensive phylogenetic analysis of the Scleractinia (Cnidaria, Anthozoa) based on mitochondrial CO1 sequence data. *PLoS One, 5*(7), e11490. doi:10.1371/journal.pone.0011490

Klemm, S. L., Shipony, Z., & Greenleaf, W. J. (2019). Chromatin accessibility and the regulatory epigenome. *Nature Reviews Genetics, 20*(4), 207-220. doi:10.1038/s41576-018-0089-8

Kolmogorov, M., Yuan, J., Lin, Y., & Pevzner, P. A. (2019). Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnology, 37*(5), 540-546. doi:10.1038/s41587-019-0072-8

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., & Phillippy, A. M. (2017). Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Research, 27*(5), 722-736. doi:10.1101/gr.215087.116

Kornberg, R. D. (1977). Structure of chromatin. *Annual review of biochemistry, 46*(1), 931-954.

Kouzarides, T. (2007). Chromatin modifications and their function. *Cell, 128*(4), 693-705.

Kundu, R., Casey, J., & Sung, W.-K. (2019). HyPo: super fast & accurate polisher for long read genome assemblies. *bioRxiv*, 2019.2012.2019.882506. doi:10.1101/2019.12.19.882506

Kusumoto, B., Costello, M. J., Kubota, Y., Shiono, T., Wei, C.-L., Yasuhara, M., & Chao, A. (2020). Global distribution of coral diversity: Biodiversity knowledge gradients related to spatial resolution. *Ecological Research, 35*(2), 315-326. doi:10.1111/1440-1703.12096

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

LaJeunesse, T. C., Parkinson, J. E., Gabrielson, P. W., Jeong, H. J., Reimer, J. D., Voolstra, C. R., & Santos, S. R. (2018). Systematic revision of Symbiodiniaceae highlights the antiquity and diversity of coral endosymbionts. *Current Biology, 28*(16), 2570-2580.e2576. doi:10.1016/j.cub.2018.07.008

Lajeunesse, T. C., Parkinson, J. E., & Reimer, J. D. (2012). A genetics-based description of *Symbiodinium minutum* sp. nov. and *S. psygmophilum* sp. nov. (Dinophyceae), two dinoflagellates symbiotic with Cnidaria. *Journal of Phycology, 48*(6), 1380-1391. doi:10.1111/j.1529-8817.2012.01217.x

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods, 9*(4), 357-359. doi:10.1038/nmeth.1923

Leitch, A. R., & Leitch, I. J. (2008). Genomic plasticity and the diversity of polyploid plants. *Science, 320*(5875), 481-483. doi:10.1126/science.1153585

Li, Y., Liao, X., Wang, X., Li, Y., Zhao, H., Zhao, Y., Chen, J., He, C., & Lu, Z. (2023). Polyp-Canal reconstruction reveals evolution toward complexity in corals. *Research, 6*, 0166, doi: 10.34133/research.0166.

Lian, H., Zang, R., Wei, J., Ye, W., Hu, M. M., Chen, Y. D., . . . Shu, H. B. (2018). The zinc-finger protein ZCCHC3 binds RNA and facilitates viral RNA sensing and activation of the RIG-I-like receptors. *Immunity, 49*(3), 438-448.e435. doi:10.1016/j.immuni.2018.08.014

Lin, S., Cheng, S., Song, B., Zhong, X., Lin, X., Li, W., . . . Morse, D. (2015). The *Symbiodinium kawagutii* genome illuminates dinoflagellate gene expression and coral symbiosis. *Science, 350*(6261), 691-694. doi:10.1126/science.aad0408

Liu, D., Hunt, M., & Tsai, I. J. (2018). Inferring synteny between genome assemblies: a systematic evaluation. *BMC bioinformatics, 19*(1), 26. doi:10.1186/s12859-018-2026-4

Liu, H., Stephens, T. G., González-Pech, R. A., Beltran, V. H., Lapeyre, B., Bongaerts, P., . . . Chan, C. X. (2018). *Symbiodinium* genomes reveal adaptive evolution of functions related to coral-dinoflagellate symbiosis. *Communications Biology, 1*(1), 95. doi:10.1038/s42003-018-0098-3

Locatelli, N. S., Kitchen, S. A., Stankiewicz, K. H., Osborne, C. C., Dellaert, Z., Elder, H., Kamel, B., Koch, H. R., Fogarty, N. D., & Baums, I. B. (2024). Chromosome-level genome assemblies and genetic maps reveal heterochiasmy and macrosynteny in endangered Atlantic *Acropora*. BMC Genomics, 25(1), 1119.

López-Nandam EH, Albright R, Hanson EA, Sheets EA, Palumbi SR. Mutations in coral soma and sperm imply lifelong stem cell renewal and cell lineage selection. Proc Biol Sci. 2023 Jan 25;290(1991):20221766. doi: 10.1098/rspb.2022.1766. Epub 2023 Jan 18. PMID: 36651044; PMCID: PMC9846893.

Loyola, A., & Almouzni, G. (2007). Marking histone H3 variants: How, when and why? *Trends in Biochemical Sciences, 32*(9), 425-433. doi:10.1016/j.tibs.2007.08.004

Lukashin, A. V., & Borodovsky, M. (1998). GeneMark.hmm: New solutions for gene finding. *Nucleic Acids Res, 26*(4), 1107-1115. doi:10.1093/nar/26.4.1107

Maehara, K., Harada, A., Sato, Y., Matsumoto, M., Nakayama, K. I., Kimura, H., & Ohkawa, Y. (2015). Tissue-specific expression of histone H3 variants diversified after species separation. *Epigenetics & Chromatin, 8*(1), 35. doi:10.1186/s13072-015-0027-3

Mao, Y., & Satoh, N. (2019). A likely ancient genome duplication in the speciose reef-building coral genus, *Acropora*. *iScience, 13*, 20-32. doi:10.1016/j.isci.2019.02.001

Mapleson, D., Garcia Accinelli, G., Kettleborough, G., Wright, J., & Clavijo, B. J. (2017). KAT: a K-mer analysis toolkit to quality control NGS datasets and genome assemblies. Bioinformatics (Oxford, England), 33(4), 574–576.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011, EMBnet.journal, 17*(1), 3. doi:10.14806/ej.17.1.200

Mat Razali, N., Cheah, B. H., & Nadarajah, K. (2019). Transposable elements adaptive role in genome plasticity, pathogenicity and evolution in fungal phytopathogens. *International Journal of Molecular Sciences, 20*(14), 3597. doi:10.3390/ijms20143597

Matthey-Doret, C., Baudry, L., Bignaud, A., Cournac, A., Remi, M., Guiglielmoni, N., . . . Scolari, V. F. (2020). hicstuff: Simple library/pipeline to generate and handle Hi-C data (Version v2.3.1): Zenodo. Retrieved from http://doi.org/10.5281/zenodo.4066363

Matzke, M. A., Mette, M. F., & Matzke, A. J. M. (2000). Transgene silencing by the host genome defense: implications for the evolution of epigenetic control mechanisms in plants and vertebrates. *Plant Molecular Biology, 43*(2), 401-415. doi:10.1023/A:1006484806925

McFadden, C. S., Quattrini, A. M., Brugler, M. R., Cowman, P. F., Dueñas, L. F., Kitahara, M. V., . . . Rodríguez, E. (2021). Phylogenomics, origin, and diversification of anthozoans (Phylum Cnidaria). *Systematic Biology, 70*(4), 635-647. doi:10.1093/sysbio/syaa103

McGhee, J. D., & Felsenfeld, G. (1980). Nucleosome structure. *Annual Review of Biochemistry, 49*(1), 1115-1156.

Muir, P. R., Wallace, C. C., Pichon, M., & Bongaerts, P. (2018). High species richness and lineage diversity of reef corals in the mesophotic zone. *Proceedings of the Royal Society B: Biological Sciences, 285*(1893), 20181987. doi:10.1098/rspb.2018.1987

Neff, E.P. (2020). The quest for an animal model of coral health and disease. *Lab Animal, 49*(2), 37-41. doi: 10.1038/s41684-019-0467-7

Nilsen, T. W., & Graveley, B. R. (2010). Expansion of the eukaryotic proteome by alternative splicing. *Nature, 463*(7280), 457-463. doi:10.1038/nature08909

Noel, B., Denoeud, F., Rouan, A., Buitrago-López, C., Capasso, L., Poulain, J., Boissin, E., Pousse, M., Da Silva, C., Couloux, A., Armstrong, E., Carradec, Q., Cruaud, C., Labadie, K., Lê-Hoang, J., Tambutté, S., Barbe, V., Moulin, C., Bourdin, G., … Aury, J.-M. (2023). Pervasive tandem duplications and convergent evolution shape coral genomes. Genome Biology, 24(1), 123.

Ohdera, A., Ames, C. L., Dikow, R. B., Kayal, E., Chiodin, M., Busby, B., ... & Ryan, J. F. (2019). Box, stalked, and upside-down? Draft genomes from diverse jellyfish (Cnidaria, Acraspeda) lineages: *Alatina alata* (Cubozoa), *Calvadosia cruxmelitensis* (Staurozoa), and *Cassiopea xamachana* (Scyphozoa). *GigaScience*, *8*(7), giz069.

Ohno, S. (1970). *Evolution by gene duplication*. New York: Springer-Verlag.

Otero, S., Desvoyes, B., & Gutierrez, C. (2014). Histone H3 dynamics in plant cell cycle and development. *Cytogenetic and Genome Research, 143*(1-3), 114-124.

Park, S., Sonn, J. Y., Oh, Y., Lim, C., & Choe, J. (2014). SIFamide and SIFamide receptor defines a novel neuropeptide signaling to promote sleep in *Drosophila*. *Mol Cells, 37*(4), 295-301. doi:10.14348/molcells.2014.2371

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

Peters, E. C., Cairns, Stephen D., Pilson, M. E. Q., Wells, J. W., Jaap, W. C., Lang, J. C., Vasleski, C. E., and St Pierre Gollahon, L. (1988). Nomenclature and biology of *Astrangia poculata* (=*A. danae*, =*A. astreiformis*) (Cnidaria: Anthozoa). *Proceedings of the Biological Society of Washington, 2*(101), 234–250.

Prada, C., Hanna, B., Budd, A. F., Woodley, C. M., Schmutz, J., Grimwood, J., ... & Medina, M. (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, K. L, Peters, E. C. 2018. *Histological Techniques for Corals*. Glen Muir Technologies and Pathobiology Consulting Services, Annapolis, MD, and Annandale, VA.

Putnam, N. H., Srivastava, M., Hellsten, U., Dirks, B., Chapman, J., Salamov, A., ... & Rokhsar, D. S. (2007). Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science, 317*(5834), 86-94. doi: 10.1126/science.1139158

Reich, H. G., Kitchen, S. A., Stankiewicz, K. H., Devlin-Durante, M., Fogarty, N. D., & Baums, I. B. (2021). Genomic variation of an endosymbiotic dinoflagellate (*Symbiodinium 'fitti'*) among closely related coral hosts. *Molecular Ecology, 30*(14), 3500-3514. doi:10.1111/mec.15952

Roach, M. J., Schmidt, S. A., & Borneman, A. R. (2018). Purge Haplotigs: allelic contig reassignment for third-gen diploid genome assemblies. *BMC bioinformatics, 19*(1), 460. doi:10.1186/s12859-018-2485-7

Robbins, S. J., Singleton, C. M., Chan, C. X., Messer, L. F., Geers, A. U., Ying, H., ... & Bourne, D. G. (2019). A genomic view of the reef-building coral *Porites lutea* and its microbial symbionts. *Nature Microbiology, 4*(12), 2090-2100. doi: 10.1038/s41564-019-0532-4

Romano, S. L., & Palumbi, S. R. (1996). Evolution of scleractinian corals inferred from molecular systematics. *Science, 271*(5249), 640-642. doi:doi:10.1126/science.271.5249.640

Ruan, J., & Li, H. (2020). Fast and accurate long-read assembly with wtdbg2. *Nature Methods, 17*(2), 155-158. doi:10.1038/s41592-019-0669-3

Sarma, K., & Reinberg, D. (2005). Histone variants meet their match. *Nature Reviews Molecular Cell Biology, 6*(2), 139-149. doi:10.1038/nrm1567

Schwager, E.E., Sharma, P.P., Clarke, T. *et al.* The house spider genome reveals an ancient whole-genome duplication during arachnid evolution. *BMC Biol* **15**, 62 (2017). doi: 10.1186/s12915-017-0399-x

Sharp, K. H., Pratte, Z. A., Kerwin, A. H., Rotjan, R. D., & Stewart, F. J. (2017). Season, but not symbiont state, drives microbiome structure in the temperate coral *Astrangia poculata*. *Microbiome, 5*(1), 120. doi:10.1186/s40168-017-0329-8

Shinzato, C., Shoguchi, E., Kawashima, T., Hamada, M., Hisata, K., Tanaka, M., . . . Satoh, N. (2011). Using the *Acropora digitifera* genome to understand coral responses to environmental change. *Nature, 476*, 320. doi:10.1038/nature10249

Shinzato, C., Khalturin, K., Inoue, J., Zayasu, Y., Kanda, M., Kawamitsu, M., ... & Satoh, N. (2021). Eighteen coral genomes reveal the evolutionary origin of *Acropora* strategies to accommodate environmental changes. *Molecular biology and evolution*, *38*(1), 16-30. doi: 10.1093/molbev/msaa216

Shoguchi, E., Shinzato, C., Kawashima, T., Gyoja, F., Mungpakdee, S., Koyanagi, R., . . . Satoh, N. (2013). Draft assembly of the *Symbiodinium minutum* nuclear genome reveals dinoflagellate gene structure. *Curr Biol, 23*(15), 1399-1408. doi:10.1016/j.cub.2013.05.062

Shumaker, A., Putnam, H. M., Qiu, H., Price, D. C., Zelzion, E., Harel, A., . . . Bhattacharya, D. (2019). Genome analysis of the rice coral *Montipora capitata*. *Scientific Reports, 9*(1), 2571. doi:10.1038/s41598-019-39274-3

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. doi:10.1093/bioinformatics/btv351

Simakov, O., Bredeson, J., Berkoff, K., Marletaz, F., Mitros, T., Schultz, D. T., O'Connell, B. L., Dear, P., Martinez, D. E., Steele, R. E., Green, R. E., David, C. N., & Rokhsar, D. S. (2022). Deeply conserved synteny and the evolution of metazoan chromosomes. Science Advances, 8(5), eabi5884.

Slater, G. S. C., & Birney, E. (2005). Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics, 6*, 31. doi:10.1186/1471-2105-6-31.

Smit, A., Hubley, R & Green, P. (2015). RepeatMasker Open-4.0. Retrieved from http://www.repeatmasker.org

Smit, A., Hubley, R. (2008). RepeatModeler Open-1.0. Retrieved from http://www.repeatmasker.org

Stephens, T. G., Lee, J., Jeong, Y., Yoon, H. S., Putnam, H. M., Majerová, E., & Bhattacharya, D. (2022). High-quality genome assembles from key Hawaiian coral species. *GigaScience*, *11*, giac098. https://doi.org/10.1093/gigascience/giac098

Stolarski, J., Kitahara, M. V., Miller, D. J., Cairns, S. D., Mazur, M., & Meibom, A. (2011). The ancient evolutionary origins of Scleractinia revealed by azooxanthellate corals. *BMC Evolutionary Biology*, *11*(1), 316. doi:10.1186/1471-2148-11-316

Szmant, A., Yevich, P., & Pilson, M. (1980). Gametogenesis and early development of the temperate coral *Astrangia danae* (Anthozoa: Scleractinia). *Biological Bulletin, 158*. doi:10.2307/1540935

Takeuchi, R., Jimbo, M., Tanimoto, F., Iijima, M., Yamashita, H., Suzuki, G., . . . Watabe, S. (2021). N-Acetyl-d-glucosamine-binding lectin in *Acropora tenuis* attracts specific Symbiodiniaceae cell culture strains. *Marine Drugs, 19*(3), 146. doi: 10.3390/md19030146

Teshima, K. M., & Innan, H. (2008). Neofunctionalization of duplicated genes under the pressure of gene conversion. *Genetics, 178*(3), 1385–1398. doi: 10.1534/genetics.107.082933

Van de Plas, R., Yang, J., Spraggins, J., & Caprioli, R. M. (2015). Image fusion of mass spectrometry and microscopy: a multimodality paradigm for molecular tissue mapping. *Nature Methods 12*, 366–372. doi: 10.1038/nmeth.3296

van Oppen, M. J. H., & Medina, M. (2020). Coral evolutionary responses to microbial symbioses. *Philosophical Transactions of the Royal Society B: Biological Sciences, 375*(1808), 20190591. doi:10.1098/rstb.2019.0591

Vaser, R., & Šikić , M. (2019). Yet another de novo genome assembler. *2019 11th International Symposium on Image and Signal Processing and Analysis (ISPA)*, 147-151. doi: 10.1109/ISPA.2019.8868909

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

Wang, Y., Tang, H., Debarry, J. D., Tan, X., Li, J., Wang, X., . . . Paterson, A. H. (2012). MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res, 40*(7), e49-e49. doi:10.1093/nar/gkr1293

Wang, X., Liew, Y. J., 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.

Wuitchik, D. M., Almanzar, A., Benson, B. E., Brennan, S., Chavez, J. D., Liesegang, M. B., . . . Davies, S. W. (2021). Characterizing environmental stress responses of aposymbiotic *Astrangia poculata* to divergent thermal challenges. *Molecular Ecology, 30*(20), 5064-5079. doi: 10.1111/mec.16108

Xu, S., Li, L., Luo, X., Chen, M., Tang, W., Zhan, L., Dai, Z., Lam, T. T., Guan, Y., & Yu, G. (2022). Ggtree: A serialized data object for visualization of a phylogenetic tree and annotation data. iMeta, 1(4), e56.

Ying, H., Cooke, I., Sprungala, S., Wang, W., Hayward, D. C., Tang, Y., . . . Miller, D. J. (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

Yu, G., Wang, L.-G., Han, Y., & He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics: a journal of integrative biology, 16*(5), 284-287. doi:10.1089/omi.2011.0118

Yuan, J., Zhang, X., Wang, M., Sun, Y., Liu, C., Li, S., . . . Li, F. (2021). Simple sequence repeats drive genome plasticity and promote adaptive evolution in penaeid shrimp. *Communications Biology, 4*(1), 186. doi:10.1038/s42003-021-01716-y

Yuan, L., Liu, X., Luo, M., Yang, S., & Wu, K. (2013). Involvement of histone modifications in plant abiotic stress responses. *Journal of Integrative Plant Biology, 55*(10), 892-901. doi: 10.1111/jipb.12060

Zhou, Z., Zhao, S., Ni, J., Su, Y., Wang, L., & Xu, Y. (2018). Effects of environmental factors on C-type lectin recognition to zooxanthellae in the stony coral *Pocillopora damicornis*. *Fish Shellfish Immunology, 79*, 228-233. doi:10.1016/j.fsi.2018.05.026

Zwaenepoel, A., & Van de Peer, Y. (2018). wgd—simple command line tools for the analysis of ancient whole-genome duplications. *Bioinformatics, 35*(12), 2153-2155. doi:10.1093/bioinformatics/bty91

Table 1. Assembly summary statistics for the *A. poculata* genome

Metric	Value
Assembly size (Mb)	458
Number of contigs	488
N50 (Mb)	31
Genome BUSCO % (singles; duplicates; missing; fragmented)	95.5 (93.8; 1.7; 2.1; 2.4)
k-mer completeness (%)	52.24
Number of genes	44,839
Gene density (genes/Mb)	97.9
Average gene length (bp)	5,204
Average exon length (bp)	244
Average intron length (bp)	1,071
Average CDS length (bp)	1,268
Gene model BUSCO % (singles; duplicates; missing; fragmented)	92.9 (87.1; 5.8; 2.7; 4.4)

Figures

Figure 1. Photograph of *Astrangia poculata* and phylogenetic tree. **a)** Underwater photograph of *A. poculata* colony with extended tentacles exhibiting aposymbiotic (white appearance, left) and symbiotic (brown appearance, right) states. **b)** Phylogeny of cnidarians included in the comparative genomic analyses of this study. Scleractinian coral clades are highlighted (yellow = Complexa; blue = Robusta). *A. poculata* and *A. millepora* are highlighted in red. The species tree was inferred using the STAG algorithm (Emms & Kelly, 2018) and rooted using the STRIDE algorithm. Branch lengths represent the number of substitutions per site.

Figure 2. *A.strangia poculata* genome assembly. **a)** Proximity ligation sequencing data (Hi-C) contact map displaying the 14 chromosome-level scaffolds of the *A. poculata* assembly. Interaction points between chromosomes are represented by red dots with binning = 100 kb. Chromosomes are ordered by size from smallest to largest. **b)** KAT plot (Mapleson et al., 2017)

of distinct *k*-mer duplicity and the number of times these *k*-mers are represented in the final *A. poculata* genome assembly, with *k*=27.

Figure 3. Distribution of synonymous substitutions per synonymous site (*Ks*) for all inferred duplications in the *A. poculata* genome with light grey representing all paralogous gene pairs and black representing anchor gene pairs.

Figure 4. Oxford grid representing pairs of homologous regions detected by SatsumaSynteny across the genome of *A. poculata*. On this grid (not drawn to scale), each point represents a pair of identical or nearly identical 4096-bp regions. Ap1, Ap2… Ap14 represent the 14 chromosome-scale scaffolds in the assembly of the genome of *A. poculata*.

Figure 5. An UpSet plot representing the number of orthogroups containing each species included in the analysis. Each dot represents the presence of a given species in orthogroups, with orthogroups unique to *A. poculata* (green), shared amongst cnidarians (orange), and shared among anthozoans (gold).

Figure 6. Comparison of *Astrangia poculata* and *Acropora millepora* genomes. **a)** Volcano plot of gene family size comparison using Fisher's Exact test between *A. poculata* and *A. millepora*. Points are colored according to whether they significantly larger in *A. poculata* (blue), significantly larger in *A. millepora* (gold), or not significantly different (grey) with adjusted $p > 0.05$. The top 20 (note: equal rank (i.e. equal $\log_2$ fold change values) orthogroups are all depicted, leading to 23 gene families for *A. millepora*) significant gene families for each species

are labeled with orthogroup IDs in red. For readability, orthogroups are labeled with leading "OG" and zeros removed from their IDs. **b)** The number of genes for these top gene families that are significantly larger in *A. poculata* (panel 'A.poculata') and *A. millepora* (panel 'A.millepora'). Each bar represents the number of genes in the gene family for *A. poculata* (blue) and *A. millepora* (gold). Gene families with putative functions in reproduction, symbiosis, innate immunity, transposition, and quiescence are highlighted in red. **c)** A bar plot representing the conserved synteny between the genomes of *A. poculata* (AP) and *A. millepora* (AM). Each chromosome of *A. poculata* is painted with the color of the *A. millepora* chromosome with which there is conserved synteny. White spaces indicate regions where colinear blocks were not detected.

Figure 7. Representative section from the sequenced *Astrangia poculata* colony stained with hematoxylin and eosin. The scale bar = 20 µm. Mesentery showing developing spermaries (SP) at black arrows.

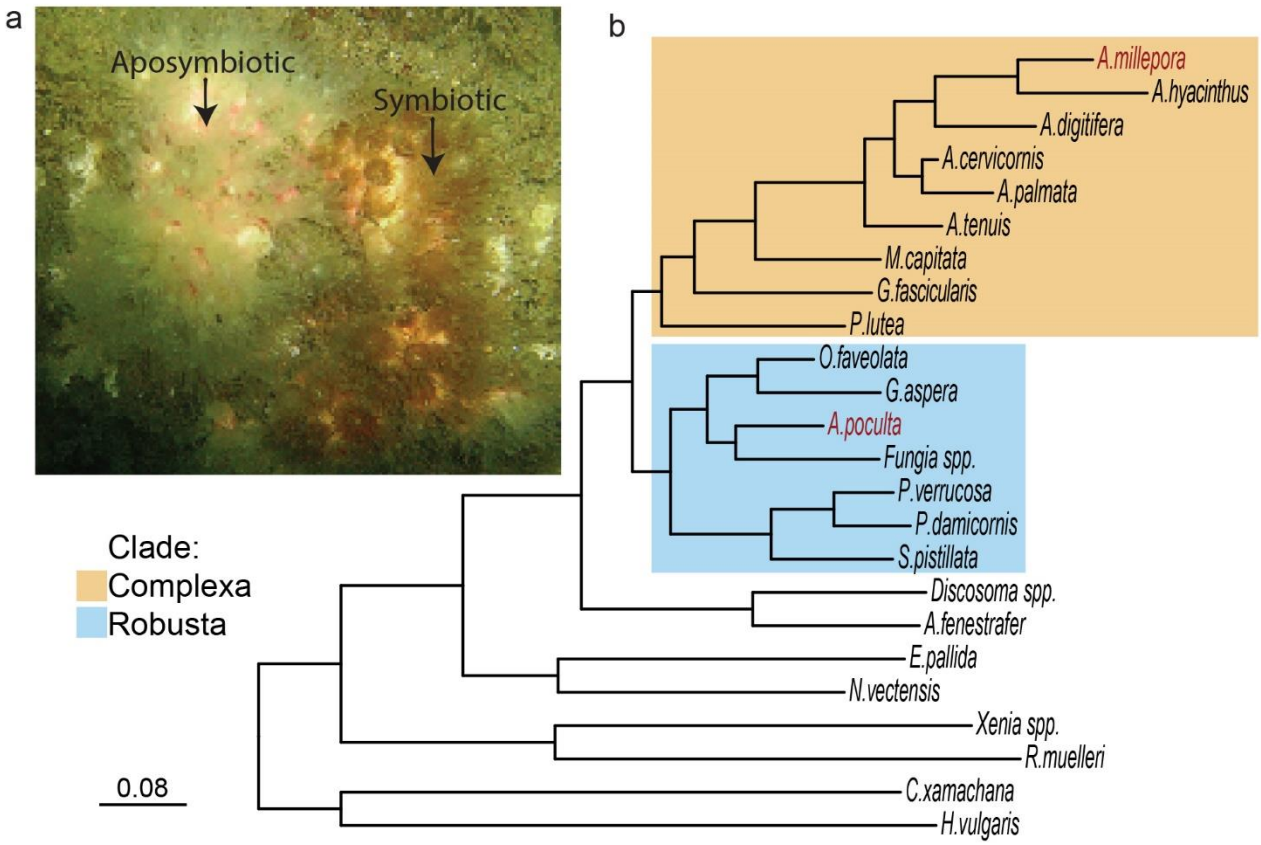

Figure 1

170x114 mm (x DPI)

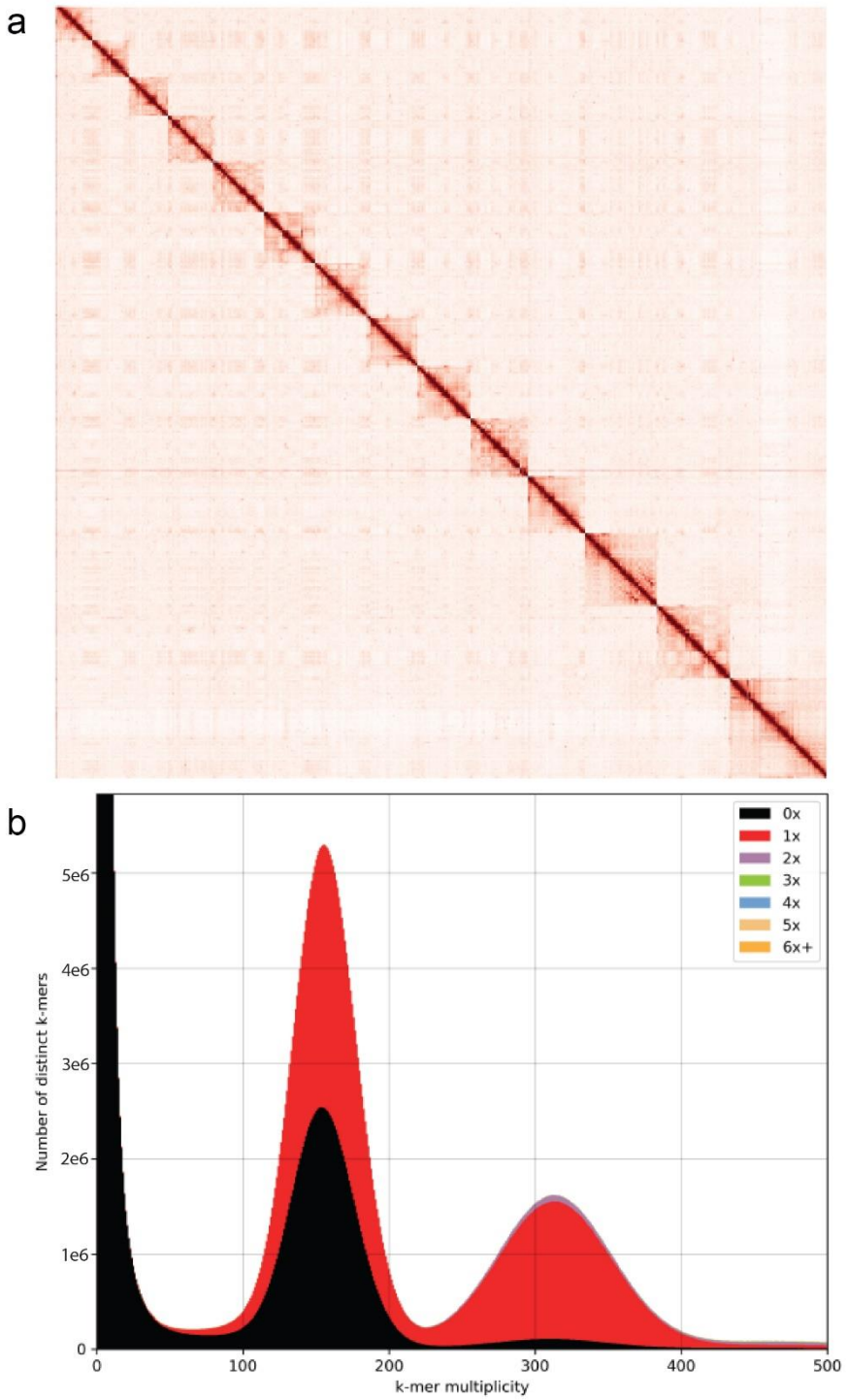

Figure 2

164x245 mm (x DPI)

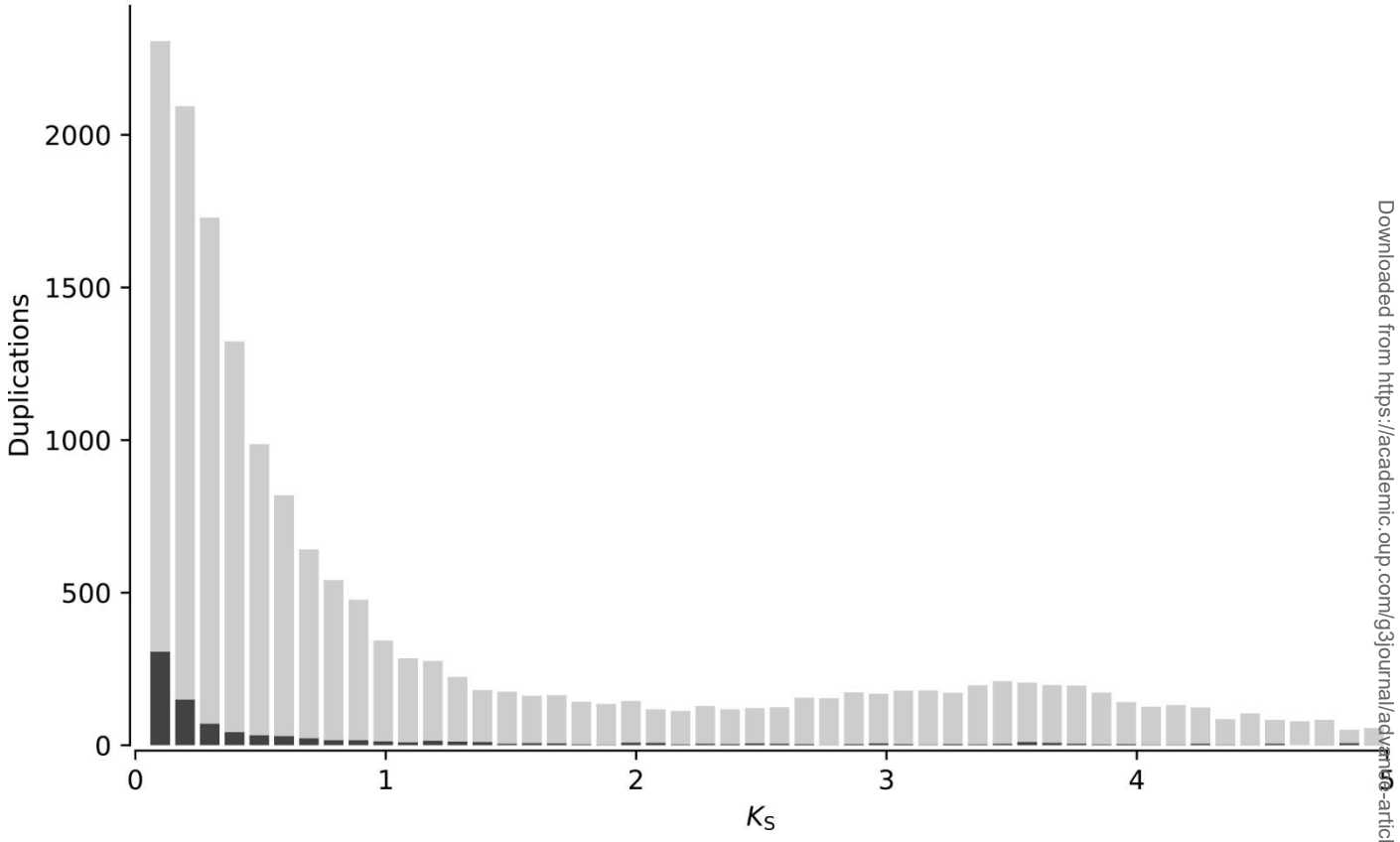

Figure 3

185x111 mm (x DPI)

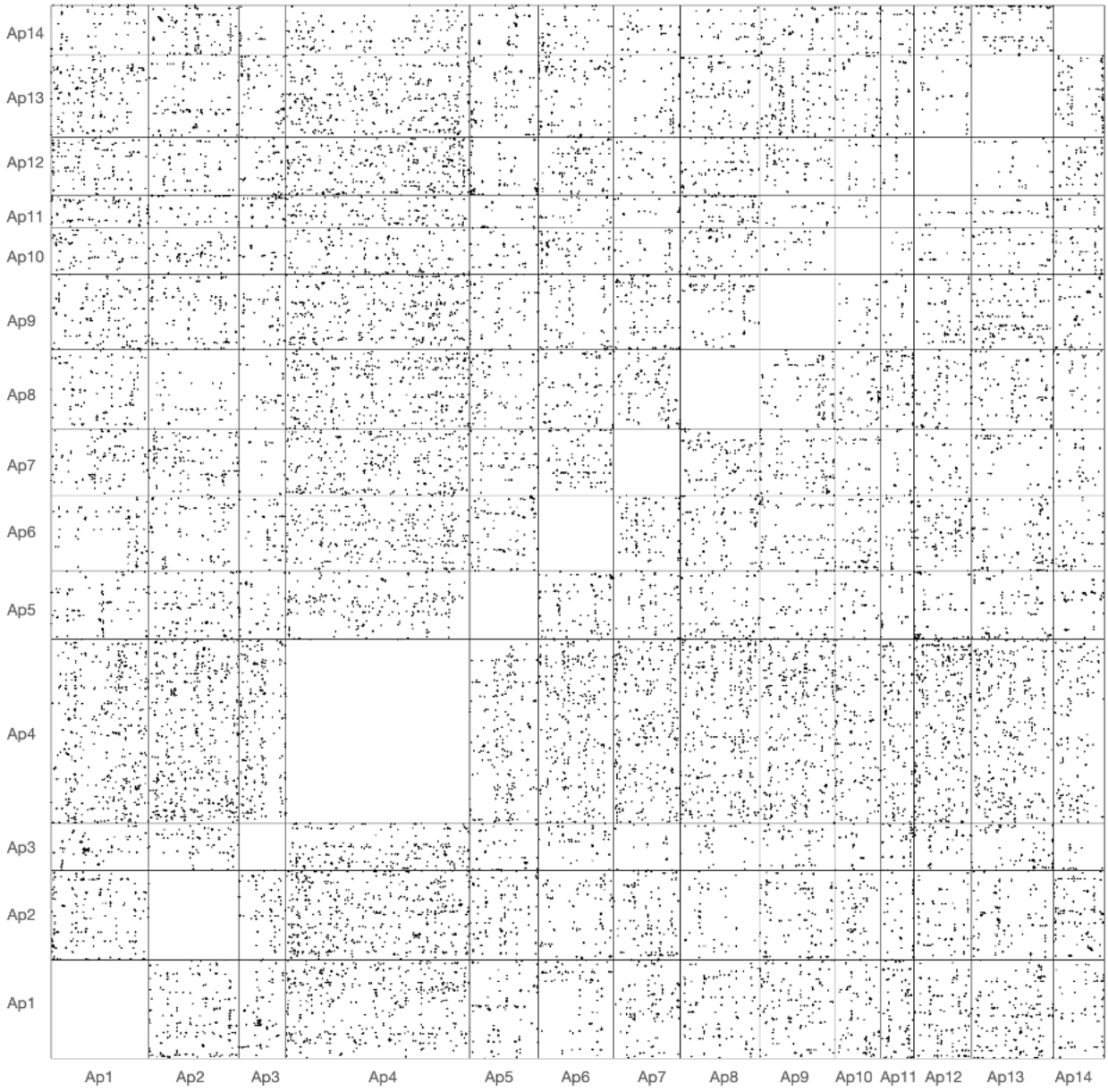

Figure 4

165x163 mm (x DPI)

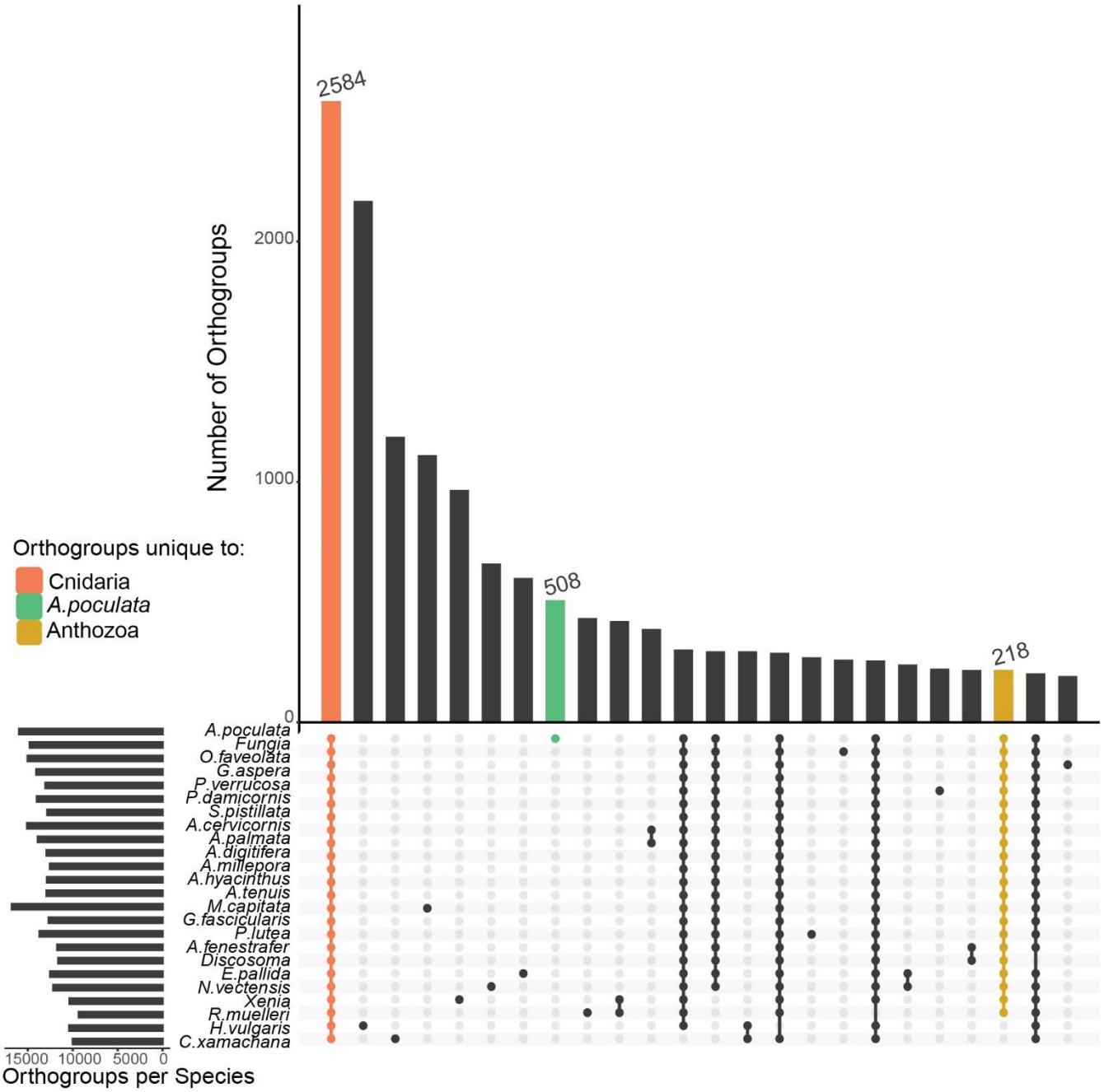

Figure 5

181x176 mm (x DPI)

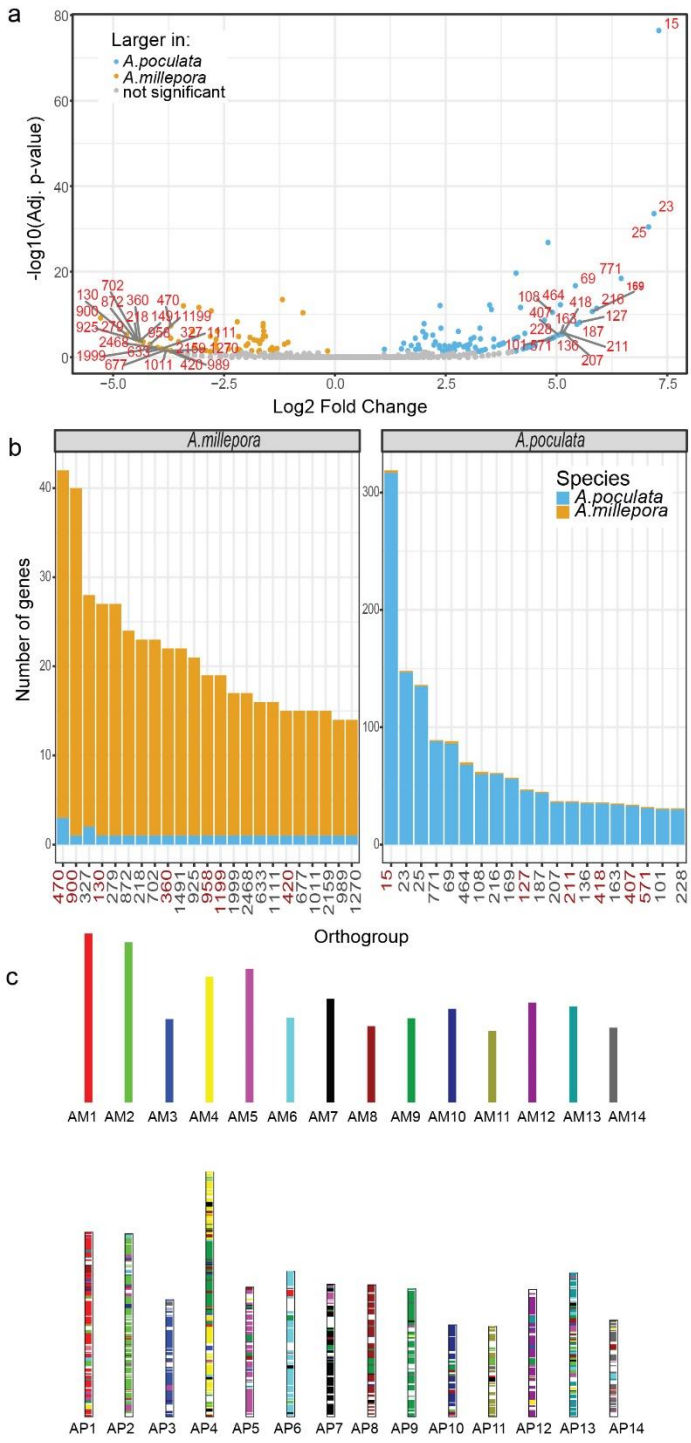

Figure 6

127x247 mm (x DPI)

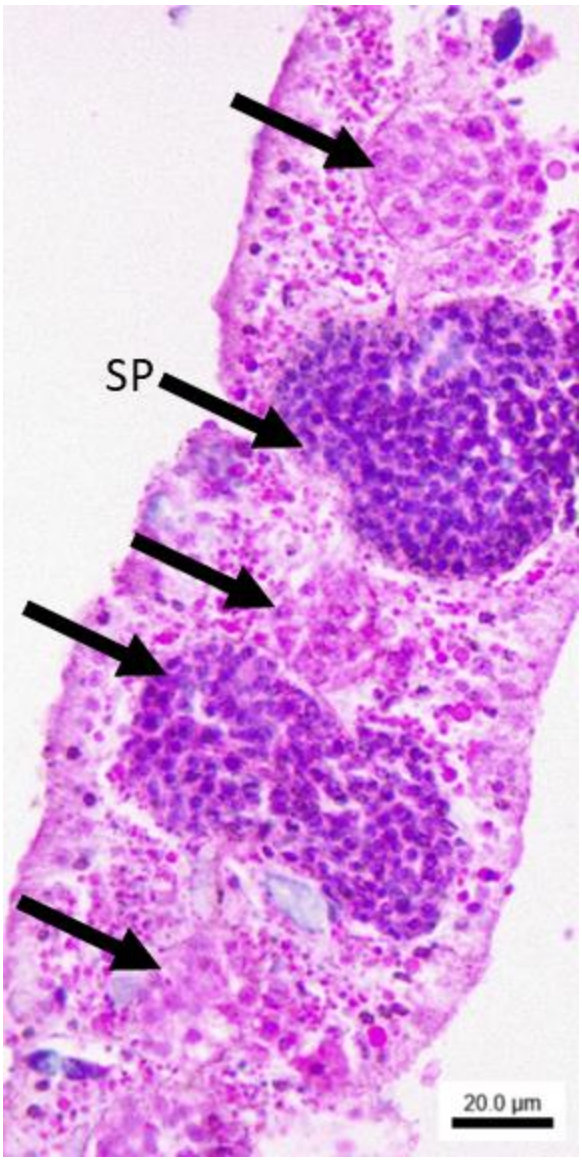

Figure 7
76x152 mm (x DPI)