

The genome of the cauliflower coral *Pocillopora verrucosa*

Carol Buitrago-López¹, Kiruthiga G. Mariappan¹, Anny Cardenas²,
Hagen M. Gegner^{1,3}, Christian R Voolstra^{1,2,*}

¹ Red Sea Research Center, Division of Biological BESE, King
Abdullah University of Science and Technology (KAUST), 23955
Thuwal, Saudi Arabia

² Department of Biology, University of Konstanz, 78457 Konstanz,
Germany

³ Centre for Organismal Studies (COS), University of Heidelberg,
69120 Heidelberg, Germany

*corresponding author: christian.voolstra@uni-konstanz.de

Abstract

Climate change and ocean warming threaten the persistence of corals worldwide. Genomic resources are critical to study the evolutionary trajectory, adaptive potential, and genetic distinctiveness of coral species. Here we provide a reference genome of the cauliflower coral *Pocillopora verrucosa*, a broadly prevalent reef-building coral with important ecological roles in the maintenance of reefs across the Red Sea, the Indian Ocean, and the Pacific Ocean. The genome has an assembly size of 380,505,698 bp with a scaffold N50 of

333,696 bp and a contig N50 of 75,704 bp. The annotation of the assembled genome returned 27,439 gene models of which 89.88% have evidence of transcription from RNA-Seq data and 97.87% show homology to known genes. A high proportion of the genome (41.22%) is comprised of repetitive elements in comparison to other cnidarian genomes, in particular in relation to the small genome size of *P. verrucosa*.

Keywords: coral reef, genome assembly, reference genome, reference transcriptome, *Pocillopora verrucosa*

Significance Statement

The ongoing destruction of coral reefs worldwide heightens the need to better understand the genomic underpinnings of coral resilience. One critical resource to assist such efforts is the generation and provision of reference genomes to enable population genomics approaches and adaptive evolution studies. Here we generated the genome of the common coral *Pocillopora verrucosa*, a broadly studied and ecologically important species. Our study demonstrates that reference genomes can be obtained utilizing improved assembly algorithms and the availability of highly assembled genomes from close-by species to fill critical gaps for species that lack genomic resources.

Introduction

Coral reefs are among the most biodiverse ecosystems on earth, providing habitat for about a third of all marine species (Plaisance et al. 2011; Spalding & Ravilious 2002). Besides their ecological importance, coral reefs provide numerous goods and ecosystem services to millions of people (Moberg & Folke 1999; Cesar et al. 2003). The well-being of these ecosystems relies on the health of reef-building corals (order Scleractinia), which comprise the foundation species and main ecosystem architects of coral reefs. Despite their importance, corals are severely threatened by local (overfishing, pollution, industrialization) and global (climate change, ocean warming) anthropogenic impacts that continue to decrease coral cover at an alarming rate (Allen et al. 2018). Therefore, better knowledge on coral biology is critical to conceive strategies to mitigate coral reef demise.

The cauliflower coral *Pocillopora verrucosa* (Ellis & Solander, 1786; NCBI Taxonomy ID: 203993; Fig. S1) is one of eight species currently described in the genus *Pocillopora* (Schmidt-Roach et al. 2014) and belongs to the family Pocilloporidae together with the genera *Madracis*, *Seriatopora*, and *Stylophora*. *P. verrucosa* is widely distributed and is prevalent in the Red Sea, across the Indian Ocean and Pacific Ocean, up to the Eastern Tropical Pacific, colonizing an

extensive range of depths that cover shallow to mesophotic settings (Schmidt-Roach et al. 2014; De Palmas et al. 2018; Soto et al. 2018; Ocean Biogeographic Information System 2020). *P. verrucosa* is an important reef framework builder and has an important ecological role as an early colonizer, aiding the recovery of reefs after natural disturbances (Pearson 1981; Bianchi et al. 2006). Despite the cosmopolitan distribution of *P. verrucosa*, which makes it a model for experimental and biological studies covering the fields of histology (Kruger & Schleyer 1998; Hirose et al. 2000), physiology (Ziegler et al. 2014; Sawall et al. 2015; Edmunds & Burgess 2016), phylogenetics (Flot et al. 2008, 2011; Schmidt-Roach et al. 2014; Pinzón et al. 2013), and population genetics (Robitzch et al. 2015; Combosch & Vollmer 2015), a reference genome is not available.

Here we report on the sequencing, assembly, and annotation of the genome of the cauliflower coral *Pocillopora verrucosa* from the Red Sea (Fig. S1), which we obtained through initial assembly of a high throughput sequenced single paired-end short-read library and subsequently reference scaffolding to the available genome of *Pocillopora damicornis* (Cunning et al. 2018). The generated draft genome is about 380 Mb in size (scaffold N50 of 333 kb, contig N50 of 75 kb) and features a total of 27,529 predicted protein-coding genes. The genome of

P. verrucosa will be of value for comparative studies (Bhattacharya et al. 2016; Woolstra et al. 2017; Cunning et al. 2018; Woolstra et al. 2015), as a genomic reference to enable population genomic approaches (e.g. RAD-Seq), and as a foundation for molecular biology studies (e.g., RNA-Seq).

Materials & Methods

Coral sampling and isolation of genomic DNA

On April 26th, 2018, a medium-size fragment (10 cm length) of a coral colony of *Pocillopora verrucosa* was collected at 6 m depth in the Red Sea Al Fahal forereef (22°15.100N, 38°57.386E) and transferred to the Coastal and Marine Resources Core Laboratory (CMOR) aquaria facilities in KAUST, Saudi Arabia (Extended Materials & Methods). The Saudi Coastguard Authority issued sailing permits to the site that included coral collection, following the Nagoya protocol. Coral genomic DNA (gDNA) was extracted using the Blood & Tissue kit (Qiagen, Hilden, Germany).

Pocillopora verrucosa lineage determination

Previous phylogenetic analyses have identified two different genetic lineages (type 3 and type 7) within *P. verrucosa* in the Red Sea (Pinzón et al. 2013; Schmidt-Roach et al. 2014) with type 3 more commonly found than type 7. To determine the

genetic lineage of our *P. verrucosa* specimen, we used the primers FATP6.1 (5'-TTTGGGSATTCGTTTAGCAG-3') and RORF (5'-SCCAATATGTTAAACASCATGTCA-3') (Flot et al. 2008) to amplify the mitochondrial open reading frame (mtORF) region. Phylogenetic analysis showed that our *P. verrucosa* specimen clustered within the mtORF lineage type 3 sequences (Pinzón, Sampayo, Cox, Chauka, Chen, Voolstra & LaJeunesse 2013), Clade 2 (Schmidt-Roach et al. 2014). Consequently, the genome represents the most common Red Sea *P. verrucosa* lineage (Fig. S2, File S1, Extended Materials & Methods).

Physical genome size estimation

We determined the physical genome size by measurement of the *P. verrucosa* nuclei size using chicken erythrocyte nuclei (CEN) as a reference (DNA QC Particles kit, BD Biosciences, San Jose, CA, USA; Extended Materials & Methods; Fig. S3). The coral genome size estimation was based on the diploid DNA content of chicken erythrocytes of 2.5 pg +/- 0.04 per cell and was calculated as follows: sample genome size [pg] = 1.25 x/y (x: fluorescence intensity of unknown sample; y: fluorescence intensity of CEN). After calculating the mean DNA content per copy of genetic information (1 C), the genome size was determined by considering that 1 pg DNA equals 978 Mb (Doležel et al. 2003).

Sequencing library, read filtering, genome assembly, and scaffolding

A single genomic DNA library was constructed using the NEBNext Ultra II DNA Library Prep Kit for Illumina according to the manufacturer's instructions (New England Biolabs, Ipswich, Massachusetts, USA). The paired-end library was sequenced on the Illumina HiSeq2500 platform (rapid run - 500cycles) at the KAUST Bioscience Core Lab (BCL). The sequencing yielded 141,993,203 read pairs (>174x coverage assuming a genome size of ~400 Mb), which were trimmed, quality assessed, and filtered to remove unwanted sequences according to the pipeline detailed in the Supplement.

The filtered paired-end reads were assembled with DISCOVAR *de novo* (Weisenfeld et al. 2014). To assess the level of heterozygosity, the filtered paired-end reads were used to obtain k-mer counts of lengths 25 and 31 using the software jellyfish version 2.2.6 (Marçais & Kingsford 2011) (details available in the Supplement). Based on k-mer distributions of lengths 31 and 25, the obtained heterozygosity rate was 1.21% and 1.32% (Table S1), respectively. This is in line with the high heterozygosity rates (1-2%) reported for other coral genomes (Bellis et al. 2016; Helmkampf et al. 2019; Robbins et al. 2019). Consequently, we applied a hierarchical filtering strategy as follows: 1) circular scaffolds, as well as contigs

of mitochondrial origin, were identified and removed, 2) contigs containing potential sequences from dinoflagellate, bacterial, or viral origin were identified with BLASTN (Altschul et al. 1990; Camacho et al. 2009) applying a 90 % identity cutoff over 50 % query length and removed (Voolstra et al. 2017). After that, the assembly was scaffolded using a reference-based approach using CSAR (Chen et al. 2017) and the available genome of the closely related species *Pocillopora damicornis* (Cunning et al. 2018). Gaps were filled with GapFiller version 1.11 (Boetzer & Pirovano 2012) using the filtered reads of the paired-end sequencing library (see above). As a final filtering step 3), the scaffolded assembly was processed with the Haplomerger2 pipeline v20180514 (Huang et al. 2017) to further improve the assembly.

Basic summary statistics of the initial DISCOVAR *de novo* assembly as well as the putative haploid final genome assembly were estimated by QUAST version 5.0.2 (Gurevich et al. 2013) (Table S2). QUAST version 5.0.2 was also used to estimate genome statistics for the genomes of *Pocillopora damicornis* (Cunning et al. 2018) and *Stylophora pistillata* (Voolstra et al. 2017). Both genomes were downloaded from the reefgenomics.org database (Liew et al. 2016) at <http://pdam.reefgenomics.org/> and <http://spis.reefgenomics.org>, respectively.

Genome assembly completeness analysis

Completeness of the *P. verrucosa* genome was assessed by searching for 978 universal metazoan Single-Copy Orthologs (metazoa_odb9) using BUSCO version 3 in mode "genome". For comparative purposes, we performed the same genome completeness assessment for the genomes of *P. damicornis* and *S. pistillata* genomes (Table S3).

Annotation of repetitive elements

Repetitive elements in the genome assembly of *P. verrucosa* were identified *de novo* using RepeatScout version 1.0.5 (Price et al. 2005) with an 1-mer size of 16bp and default settings (Extended Materials & Methods; Table S4).

Reference transcriptome sequencing and assembly

An RNA-Seq library was generated from 500 ng of total RNA using the TruSeq Stranded mRNA Library prep Kit (Illumina, San Diego, California, USA) as per the manufacturer's instructions (further details in the Supplement). The constructed library (median size 340 bp) was indexed and paired-end sequenced on the Illumina HiSeq4000 platform (300 cycles) at the KAUST Bioscience Core Facility (KAUST, Thuwal, KSA). RNA sequencing yielded 121,056,948 read pairs (i.e., 242 mio. paired-end

reads), which were trimmed and further processed as detailed in the Supplement.

Gene model prediction

Coral transcripts from the reference transcriptome (n = 49,384, Extended Materials & Methods) were mapped to the genome assembly and filtered by PASA version 2.3.3 (Haas et al. 2008) to create a training set for AUGUSTUS version 2.5.5 (Stanke & Morgenstern 2005).

Protein set annotation

The final set of predicted proteins was generated by selecting the longest isoform per gene. The protein sequences were annotated following a hierarchical approach using the UniProt (SwissProt and TrEMBL) and the NCBI 'nr' databases *sensu* Baumgarten et al. (2015). This approach led to the annotation of 19,427 in SwissProt, 7,145 proteins in TrEMBL, and 283 proteins using the NCBI 'nr' database; 584 protein had no hits in either of the reference databases (Extended Materials & Methods; File S2).

Results & Discussion

Genome size estimation

We estimated the genome size of *P. verrucosa* using FACS of propidium-iodide-stained nuclei in addition to a k-mer based approach. The physical estimate of the genome yielded a size of approximately 407 Mb (Fig. S3) showing a good agreement with the k-mer based estimate, which obtained a size of 406 Mb and 418 Mb for k-mers 25 and 31, respectively (Table S1). This genome size is larger than the estimated genome size of *Pocillopora damicornis* (349 Mb assessed using k-mer counts, Cunning et al. 2018) and smaller than that of *Stylophora pistillata* (434 Mb assessed using FACS, Voolstra et al. 2017), which are the two other coral species in the family Pocilloporidae with published genomes.

Genome assembly and statistics

We obtained 141,993,203 paired-end read pairs (2 x250 bp) from a single genomic DNA library. About 88% of read pairs were kept after sequence filtering and used for a *de novo* genome assembly. Read coverage was estimated to be 153X based on the physical genome size (407 Mb), which was reported to be in the optimal range for DISCOVAR *de novo* genome assemblies (Love et al. 2016). After assembly, we applied a filtering strategy to remove non-nuclear and non-coral contigs before scaffolding

using the reference genome of the closely related coral *P. damicornis*. To address the issue of high levels of heterozygosity, we applied the Haplomerger2 pipeline to consolidate incompletely merged loci/alleles and produce a haploid reference genome assembly with a total length of ~380 Mb assembled in 25,605 contigs and 18,268 scaffolds with a corresponding contig N50 of 75,704 bp and scaffold N50 of 333,696 bp, respectively (Table 1). The assembly statistics of the scaffolded haploid genome compare well to the genome assemblies of the available genomes in the family Pocilloporiadae, namely *P. damicornis* (contig N50 = 25,987, scaffold N50 = 326,133) and *S. pistillata* (contig N50 = 20,518, scaffold N50 = 457,453).

Table 1. Summary statistics of the *Pocillopora verrucosa* genome assembly.

| Genome assembly statistics | Length | | Number | |
|----------------------------|-------------|-------------|--------|----------|
| | Contig | Scaffold | Contig | Scaffold |
| Total [bp] | 379,998,154 | 380,505,698 | 25,605 | 18,268 |
| Max [bp] | 571,091 | 2,095,917 | - | - |
| ≥5000 [bp] | 353,178,732 | 358,580,702 | 8,537 | 3,932 |
| N50 [bp] | 75,704 | 333,696 | 1,401 | 326 |
| N75 [bp] | 30,836 | 125,371 | 3,336 | 766 |
| N's per 100kbp | 7.89 | 134.40 | | |

To compare the assembly before and after reference genome scaffolding, we computed basic statistics of the DISCOVAR *de novo* assembly and of the final reference scaffolded and filtered assembly (after DISCOVAR assembly) (Table S2). As

expected, the scaffolded haploid genome was a substantial improvement over the *de novo* assembly, highlighting the utility of having close-by genomes available for the generation of reference genomes from so far unconsidered species.

Genome assembly completeness was assessed by searching for 978 universal metazoan single-copy orthologs using BUSCO. We could identify 902 (92.22%) complete metazoan single-copy orthologs (of which 3.17% were duplicated), 18 (1.84%) were present but fragmented, and 58 (5.93%) were missing. Comparison to *P. damicornis* and *S. pistillata* showed that *P. damicornis* had 862 (88.14%) and *S. pistillata* had 861 (88.03%) complete metazoan single-copy orthologs (Table S3). The number of fragmented benchmarking universal genes is comparable among the three pocilloporid species, while the number of missing single-copy benchmarking universal genes is smaller in *P. verrucosa* (58, 5.93%) than in *P. damicornis* (87, 8.90%) and *S. pistillata* (89, 9.10%).

Genomic repeats

The identification and annotation of repetitive elements showed that 41.22% of the genome comprised repetitive elements. Out of the 15,905 repeat motifs identified, we could annotate 4,915 motifs, which were further classified in 6 groups encompassing

63 superfamilies/clades (Table S4). Transposable elements (TEs), such as DNA transposons, LTR, and non-LTR retrotransposons (among others) comprised the largest portion of the *P. verrucosa* genome (17.22%), followed by unclassified interspersed repeats (13.47%), and simple repeats (10.28%). Given that unclassified interspersed repeats comprise putative species-specific TEs, we considered them TEs for the purpose of comparison between *P. verrucosa* and *S. pistillata*, as previously done for other coral genomes (Ying et al. 2018). Overall, we found a slightly higher proportion of TEs in *P. verrucosa* (30.69% of genome size) in comparison to *S. pistillata* (28.43% of genome size), despite the smaller assembly size of *P. verrucosa*. This was unexpected given the known positive correlation between TEs and genome size (Kidwell 2002; Hua-Van et al. 2005; Biscotti et al. 2015), although some margin of error exists.

At the moment we can only speculate on the significance of the increased TE content in the genome of *P. verrucosa* in comparison to *S. pistillata*. TE expansion can denote signatures of genus radiation (Wong et al. 2019), and TE activity has been associated with response to environmental stress and phenotypic plasticity in plants (Negi et al. 2016), arguably of relevance for corals in the context of climate change and ocean warming. From this perspective, an increased TE content in *P. verrucosa* fits with the species' broad geographical and depth distribution, and

arguably, broad physiological plasticity (Sawall et al. 2015; Ziegler et al. 2014; Voolstra & Ziegler 2020).

At large, however, the proportion of TEs in the genomes of both pocilloporid species were in line with expectation following Kidwell (2002), which predicted about 28.29% and 29.71% for *P. verrucosa* and *S. pistillata*, respectively. Of note, a direct comparison with *P. damicornis* was not possible, due to the lack of high-identity repetitive content, which evaded assembly (Cunning et al. 2018).

Transcriptome assembly, gene models, and protein annotation

RNA sequencing yielded 121,056,948 read pairs (i.e., 242 mio. paired-end reads) that produced 49,384 coral transcripts available for gene model prediction. Based on benchmarking universal metazoan single-copy orthologs (n = 978), we found a high number of complete single-copy orthologs (893, 91.31%), while some orthologs were complete but duplicated (33, 3.37%). Consequently, the number of fragmented (28, 2.86%) and missing (24, 2.45%) benchmarking orthologs was rather small, indicating that the reference transcriptome assembly is largely complete.

We identified a final set of 27,439 protein-coding genes with a mean CDS of 1,849.67 bp (Table 2). Of these, 96.06% (n = 26,506) have complete ORFs and 89.88% (n = 24,662) are

corroborated by RNA-Seq evidence. Despite largely corresponding gene statistics (Table 2), the proportion of intronless gene models (10.62%) in *P. verrucosa* was comparable to that of *P. damicornis* (12.26%), but three times higher than in *S. pistillata* (3.41%). Intronless genes are commonly thought to arise from horizontal gene transfer (HGT) events of sequences of bacterial origin or by reverse transcription (Baumgarten et al. 2015). As such, the higher number of intronless genes in both *Pocillopora* genomes may indicate extensive HGT or increased TE element activity, as genes encoding for reverse transcriptases are most commonly found as components of retrotransposons in eukaryotes (Lescot et al. 2016).

The majority of the 27,439 protein-coding genes were annotated (File S2): in total, 97.87% (n = 26,855) had an annotation, the large majority of which (70.80%, n = 19,427) retrieved hits from the SwissProt database, followed by 7,145 proteins (26.04%) with hits in TrEMBL, and 283 proteins (1.03%) that showed similarity to proteins from the NCBI 'nr' database. Only a small percentage of proteins (2.12%, n = 584) had no hits to either database.

Table 2. Genomic gene sets of *P. verrucosa* (present study) in comparison to the other pocilloporid genomes of *P. damicornis* and *S. pistillata*.

| | <i>P. verrucosa</i> | <i>P. damicornis</i> | <i>S. pistillata</i> |
|--------------------------------------|---------------------|----------------------|----------------------|
| Genes | | | |
| Number of protein-encoding genes | 27,439 | 26,077 | 25,769 |
| Mean gene length [bp] | 7,566.63 | 5,859.80 | 8,378.20 |
| Genes with annotation | 97.87% | N/A | 88.91% |
| Genes with annotation to SwissProt | 70.80% | 59.70% | 67.93% |
| Exons | | | |
| Mean number of exons per gene | 7.54 | 7.17 | 7.85 |
| Mean exon length [bp] | 316.92 | 244.86 | 265.66 |
| Mean coding region (CDS) length [bp] | 1849.67 | 1365.30 | 1833.33 |
| Introns | | | |
| Percentage of genes with introns | 89.38% | 87.74% | 96.59% |
| Mean number of introns per gene | 7.37 | 7.03 | 7.09 |
| Mean intron length [bp] | 784.93 | 665.17 | 918.25 |

Towards generation of model species reference genomes

In this study, we present the first genome assembly for the coral species *P. verrucosa*. This is an important resource to advance the understanding of coral reef ecology and evolution, particularly because *P. verrucosa* is a common and globally distributed reef builder. As demonstrated above, we have assembled more than 90% of the *P. verrucosa* genome with a contig and scaffold N50 and gene completeness on par with the available genomes of *P. damicornis* and *S. pistillata*. We have accomplished this through utilization of an available genome of a close-by species used for scaffolding after initial

assembly based on the sequencing of only a single genomic library, with overlapping paired-end reads. We recognize the challenge of assembling a coral genome particularly with regard to their high heterozygosity. We hope that the *P. verrucosa* genome assembly will be useful for a number of studies ranging from comparative genomics to genetic variants identification (e.g., SNPs, microsatellites, etc.).

Data availability

The reference genome of *P. verrucosa* (Pver_genome_assembly_v1.0.fasta) together with the structural and functional annotation of genes (Pver_genome_assembly_v1.0.gff3, Pver_genes_names_v1.0.fna, Pver_proteins_names_v1.0.faa) as well as the transcriptome (Pver_transcriptome_v1.0.fasta) used for gene calling are available at <http://pver.reefgenomics.org/> (Liew et al. 2016). Detailed annotation of the gene models can be found in the supplementary information (File S2) as well as the classification of repeat elements identified in the genome (Table S4). Raw sequence data determined for the *P. verrucosa* reference genome project is available under NCBI BioProject PRJNA551401 (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA551401>). The code to reproduce the analyses described herein is available

at <https://github.com/Carol-Symbiomics/Pocillopora-verrucosa-genome>

Acknowledgements

We would like to thank the Coastal & Marine Resources Core Lab (CMOR) for the aquaria facilities as well as the Bioscience Core Lab (BCL) at KAUST for sequencing. Further, we would like to thank Larissa Morales for support with R code for the repetitive elements analysis and Yi Jin Liew for his support to analyze gff3 files of *S. pistillata*. In addition, we thank Sebastian Schmidt-Roach for confirming coral species identification based on skeletal morphology. Research reported in this publication was supported by King Abdullah University of Science and Technology (KAUST) and the University of Konstanz.

Author contributions

Conceptualization: CBL, CRV

Data Curation: CBL, KGM, AC, CRV

Formal Analysis: CBL, KGM

Funding Acquisition: CRV

Investigation: CBL, KGM, AC, CRV

Project Administration: CRV

Resources: CBL, HMG, CRV

Supervision: CRV

Validation: CBL, CRV

Writing - Original Draft Preparation: CBL, CRV

Writing - Review & Editing: CBL, AC, HMG, CRV

Competing interests

The authors declare no competing interests.

References

- Allen MR et al. 2018. *Framing and Context. In: Global warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the .*
https://www.ipcc.ch/site/assets/uploads/sites/2/2018/11/SR15_Chapter1_Low_Res.pdf.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403-410. doi: [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
- Baumgarten S et al. 2015. The genome of *Aiptasia*, a sea anemone model for coral symbiosis. *Proc. Natl. Acad. Sci.* 112:11893 LP - 11898. doi: [10.1073/pnas.1513318112](https://doi.org/10.1073/pnas.1513318112).
- Bellis ES, Howe DK, Denver DR. 2016. Genome-wide polymorphism and signatures of selection in the symbiotic sea anemone *Aiptasia*. *BMC Genomics.* 17:160. doi: [10.1186/s12864-016-2488-6](https://doi.org/10.1186/s12864-016-2488-6).
- Bhattacharya D et al. 2016. Comparative genomics explains the evolutionary success of reef-forming corals. *Elife.* 5:9983. doi: [10.7554/eLife.13288](https://doi.org/10.7554/eLife.13288).
- Bianchi CN et al. 2006. Dynamics and pattern of coral recolonization following the 1998 bleaching event in the reefs of the Maldives. *Proc. 10th Int. Coral Reef Symp.* 1:30-37.
- Biscotti MA, Olmo E, Heslop-Harrison JS (Pat). 2015.

Repetitive DNA in eukaryotic genomes. *Chromosom. Res.* 23:415-420. doi: 10.1007/s10577-015-9499-z.

Boetzer M, Pirovano W. 2012. Toward almost closed genomes with GapFiller. *Genome Biol.* 13:R56-R56. doi: 10.1186/gb-2012-13-6-r56.

Camacho C et al. 2009. BLAST+: architecture and applications. *BMC Bioinformatics.* 10:421. doi: 10.1186/1471-2105-10-421.

Cesar H, Burke L, Pet-soede L. 2003. The Economics of Worldwide Coral Reef Degradation. *Cesar Environ. Econ. Consult.* Arnhem, WWF-Netherlands. 14:23. www.ceec.nl (Accessed April 3, 2017).

Chen K-T et al. 2017. CSAR: a contig scaffolding tool using algebraic rearrangements. *Bioinformatics.* 34:109-111. doi: 10.1093/bioinformatics/btx543.

Combosch DJ, Vollmer S V. 2015. Trans-Pacific RAD-Seq population genomics confirms introgressive hybridization in Eastern Pacific *Pocillopora* corals. *Mol. Phylogenet. Evol.* 88:154-162. doi: 10.1016/j.ympev.2015.03.022.

Cunning R, Bay RA, Gillette P, Baker AC, Traylor-Knowles N. 2018. Comparative analysis of the *Pocillopora damicornis* genome highlights role of immune system in coral evolution. *Sci. Rep.* 8:1-10. doi: 10.1038/s41598-018-34459-8.

Doležel J, Bartoš J, Voglmayr H, Greilhuber J. 2003. Nuclear DNA content and genome size of trout and human. *Cytom. Part A.* 51A:127-128. doi: 10.1002/cyto.a.10013.

Edmunds PJ, Burgess SC. 2016. Size-dependent physiological responses of the branching coral *Pocillopora verrucosa* to elevated temperature and PCO₂. *J. Exp. Biol.* 219:3896–3906. doi: 10.1242/jeb.146381.

Flot J-F et al. 2011. Incongruence between morphotypes and genetically delimited species in the coral genus *Stylophora*: phenotypic plasticity, morphological convergence, morphological stasis or interspecific hybridization? *BMC Ecol.* 11:22. doi: 10.1186/1472-6785-11-22.

Flot JF, Magalon H, Cruaud C, Couloux A, Tillier S. 2008. Patterns of genetic structure among Hawaiian corals of the genus *Pocillopora* yield clusters of individuals that are compatible with morphology. *Comptes Rendus - Biol.* 331:239–247. doi: 10.1016/j.crvi.2007.12.003.

Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics.* 29:1072–1075. doi: 10.1093/bioinformatics/btt086.

Haas BJ et al. 2008. Automated eukaryotic gene structure annotation using EVIDENCEModeler and the Program to Assemble Spliced Alignments. *Genome Biol.* 9:R7–R7. doi: 10.1186/gb-2008-9-1-r7.

Helmkamp M, Bellinger MR, Geib SM, Sim SB, Takabayashi M. 2019. Draft Genome of the Rice Coral *Montipora capitata* Obtained from Linked-Read Sequencing. *Genome Biol. Evol.* 11:2045–2054. doi: 10.1093/gbe/evz135.

Hirose M, Kinzie R a, Hidaka M. 2000. Early Development of Zooxanthella-Containing Eggs of the Corals *Pocillopora verrucosa* and *P. Eydouxi* With Special Reference to the Distribution of Zooxanthellae. Biol. Bull. 199:68-75. <http://www.ncbi.nlm.nih.gov/pubmed/10975644>.

Hua-Van A, Le Rouzic A, Maisonhaute C, Capy P. 2005. Abundance, distribution and dynamics of retrotransposable elements and transposons: similarities and differences. Cytogenet. Genome Res. 110:426-440. doi: 10.1159/000084975.

Huang S, Kang M, Xu A. 2017. HaploMerger2: rebuilding both haploid sub-assemblies from high-heterozygosity diploid genome assembly. Bioinformatics. 33:2577-2579. doi: 10.1093/bioinformatics/btx220.

Kidwell MG. 2002. Transposable elements and the evolution of genome size in eukaryotes. Genetica. 115:49-63. doi: 10.1023/A:1016072014259.

Kolde R. 2019. pheatmap: Pretty Heatmaps. R package version. <https://cran.r-project.org/package=pheatmap>.

Kruger A, Schleyer MH. 1998. Sexual reproduction in the coral *Pocillopora verrucosa* (Cnidaria: Scleractinia) in KwaZulu-Natal, South Africa. Mar. Biol. 132:703-710. doi: 10.1007/s002270050434.

Lescot M et al. 2016. Reverse transcriptase genes are highly abundant and transcriptionally active in marine plankton assemblages. ISME J. 10:1134-1146. doi:

10.1038/ismej.2015.192.

Liew YJ, Aranda M, Voolstra CR. 2016. Reefgenomics.Org - a repository for marine genomics data. Database. 2016. doi: 10.1093/database/baw152.

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

Madin JS et al. 2016. The Coral Trait Database, a curated database of trait information for coral species from the global oceans. Sci. Data. 3:160017. doi: 10.1038/sdata.2016.17.

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

Mistry J, Finn RD, Eddy SR, Bateman A, Punta M. 2013. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. Nucleic Acids Res. 41:e121-e121. doi: 10.1093/nar/gkt263.

Moberg FF, Folke C. 1999. Ecological goods and services of coral reef ecosystems. Ecol. Econ. 29:215-233. doi: 10.1016/S0921-8009(99)00009-9.

Negi P, Rai AN, Suprasanna P. 2016. Moving through the Stressed Genome: Emerging Regulatory Roles for Transposons in

Plant Stress Response. *Front. Plant Sci.* 7:1448. doi: 10.3389/fpls.2016.01448.

Ocean Biogeographic Information System. 2020. *Pocillopora verrucosa* (Ellis & Solander, 1786). Intergov. Oceanogr. Comm. UNESCO.

De Palmas S, Soto D, Denis V, Ho MJ, Chen CA. 2018. Molecular assessment of *Pocillopora verrucosa* (Scleractinia; Pocilloporidae) distribution along a depth gradient in Ludaο, Taiwan. *PeerJ.* 2018:1-14. doi: 10.7717/peerj.5797.

Pearson RG. 1981. Recovery and Recolonization of Coral Reefs. *Mar. Ecol. Prog. Ser.* 4:105-122. <http://www.jstor.org/stable/24812966>.

Pinzón JH et al. 2013. Blind to morphology: Genetics identifies several widespread ecologically common species and few endemics among Indo-Pacific cauliflower corals (*Pocillopora*, Scleractinia) Maggs, C, editor. *J. Biogeogr.* 40:1595-1608. doi: 10.1111/jbi.12110.

Plaisance L, Caley MJ, Brainard RE, Knowlton N. 2011. The diversity of coral reefs: What are we missing? *PLoS One.* 6. doi: 10.1371/journal.pone.0025026.

Price AL, Jones NC, Pevzner PA. 2005. De novo identification of repeat families in large genomes. *Bioinformatics.* 21:i351-i358. doi: 10.1093/bioinformatics/bti1018.

Robbins SJ et al. 2019. A genomic view of the reef-building coral *Porites lutea* and its microbial symbionts. *Nat.*

Microbiol. 4:2090-2100. doi: 10.1038/s41564-019-0532-4.

Robitzch V, Banguera-Hinestroza E, Sawall Y, Al-Sofyani A, Woolstra CR. 2015. Absence of genetic differentiation in the coral *Pocillopora verrucosa* along environmental gradients of the Saudi Arabian Red Sea. *Front. Mar. Sci.* 2:5. doi: 10.3389/fmars.2015.00005.

Sawall Y et al. 2015. Extensive phenotypic plasticity of a Red Sea coral over a strong latitudinal temperature gradient suggests limited acclimatization potential to warming. *Sci. Rep.* 5. doi: 10.1038/srep08940.

Schmidt-Roach S, Miller KJ, Lundgren P, Andreakis N. 2014. With eyes wide open: A revision of species within and closely related to the *Pocillopora damicornis* species complex (Scleractinia; Pocilloporidae) using morphology and genetics. *Zool. J. Linn. Soc.* 170:1-33. doi: 10.1111/zoj.12092.

Soto D, De Palmas S, Ho MJ, Denis V, Chen CA. 2018. Spatial variation in the morphological traits of *Pocillopora verrucosa* along a depth gradient in Taiwan. *PLoS One.* 13:e0202586. <https://doi.org/10.1371/journal.pone.0202586>.

Spalding D, Ravilious C. 2002. *World atlas of coral reefs*. doi: 10.5860/choice.39-2540.

Stanke M, Morgenstern B. 2005. AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Res.* 33:W465-W467. doi: 10.1093/nar/gki458.

Storey JD, Bass AJ, Dabney A, Robinson D. 2019. qvalue: Q-

value estimation for false discovery rate control.

<http://github.com/jdstorey/qvalue>.

Voolstra CR et al. 2017. Comparative analysis of the genomes of *Stylophora pistillata* and *Acropora digitifera* provides evidence for extensive differences between species of corals. *Sci. Rep.* 7:1-14. doi: 10.1038/s41598-017-17484-x.

Voolstra CR et al. 2015. The ReFuGe 2020 Consortium—using “omics” approaches to explore the adaptability and resilience of coral holobionts to environmental change. *Front. Mar. Sci.* 2:68.

<https://www.frontiersin.org/article/10.3389/fmars.2015.00068>.

Voolstra CR, Ziegler M. 2020. Adapting with microbial help: Microbiome flexibility facilitates rapid responses to environmental change. *BioEssays*. doi: 10.1002/bies.202000004.

Weisenfeld NI et al. 2014. Comprehensive variation discovery in single human genomes. *Nat. Genet.* 46:1350-1355. doi: 10.1038/ng.3121.

Wong WY et al. 2019. Expansion of a single transposable element family is associated with genome-size increase and radiation in the genus *Hydra*. *Proc. Natl. Acad. Sci.* 116:22915 LP - 22917. doi: 10.1073/pnas.1910106116.

Xia K, Fu Z, Hou L, Han J-DJ. 2008. Impacts of protein-protein interaction domains on organism and network complexity. *Genome Res.* 18:1500-1508. doi: 10.1101/gr.068130.107.

Ying H et al. 2018. Comparative genomics reveals the distinct

evolutionary trajectories of the robust and complex coral lineages. *Genome Biol.* 19:175. doi: 10.1186/s13059-018-1552-8.

Ziegler M, Roder CM, Buchel C, Voolstra CR. 2014. Limits to physiological plasticity of the coral *Pocillopora verrucosa* from the central Red Sea. *Coral Reefs.* 33:1115–1129. doi: 10.1007/s00338-014-1192-8.