Chromosome-Level Genome Assembly and Annotation of *Corallium rubrum*: A Mediterranean Coral Threatened by Overharvesting and Climate Change

Jean-Baptiste Ledoux [1,*], Jessica Gomez-Garrido [2,3], Fernando Cruz [2,3], Francisco Camara Ferreira [2,3], Ana Matos [1], Xenia Sarropoulou [4,5], Sandra Ramirez-Calero [6,7], Didier Aurelle [8,9], Paula Lopez-Sendino [6], Natalie E. Grayson [10], Bradley S. Moore [10,11], Agostinho Antunes [1,12], Laura Aguilera [2,3], Marta Gut [2,3], Judit Salces-Ortiz [13], Rosa Fernández [13], Cristina Linares [7,14], Joaquim Garrabou [6,*], Tyler Alioto [2,3]

[1]CIIMAR/CIMAR, Centro Interdisciplinar de Investigação Marinha e Ambiental, Universidade do Porto, Porto, Portugal

[2]Centro Nacional de Análisis Genómico (CNAG), Barcelona 08028, Spain

[3]Universitat de Barcelona (UB), Barcelona, Spain

[4]Department of Biology, School of Sciences and Engineering, University of Crete, Heraklion, Crete, Greece

[5]Institute of Marine Biology, Biotechnology and Aquaculture, Hellenic Centre for Marine Research (HCMR), Heraklion, Crete, Greece

[6]Departament de Biologia Marina, Institut de Ciències del Mar (CSIC), Barcelona, Spain

[7]Departament de Biologia Evolutiva, Ecologia i Ciències Ambientals, Universitat de Barcelona (UB), Barcelona 08028, Spain

[8]Aix Marseille Université, Université de Toulon, CNRS, IRD, MIO, Marseille, France

[9]Institut Systématique Evolution Biodiversité (ISYEB), Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, CP 26, 75005 Paris, France

[10]Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA

[11]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, CA, USA

[12]Department of Biology, Faculty of Sciences, University of Porto, Porto 4169-007, Portugal

[13]Metazoa Phylogenomics Lab, Institute of Evolutionary Biology (CSIC-University Pompeu Fabra), Barcelona, Spain

[14]Institut de Recerca de la Biodiversitat (IRBio), Universitat de Barcelona, Barcelona, Spain

*Corresponding authors: E-mails: jbaptiste.ledoux@gmail.com; garrabou@icm.csic.es

Accepted: November 12, 2024

Abstract

Reference genomes are key resources in biodiversity conservation. Yet, sequencing efforts are not evenly distributed across the tree of life raising concerns over our ability to enlighten conservation with genomic data. Good-quality reference genomes remain scarce in octocorals while these species are highly relevant targets for conservation. Here, we present the first annotated reference genome in the red coral, *Corallium rubrum* (Linnaeus, 1758), a habitat-forming octocoral from the Mediterranean and neighboring Atlantic, impacted by overharvesting and anthropogenic warming-induced mass mortality events. Combining long reads from Oxford Nanopore Technologies (ONT), Illumina paired-end reads for improving the base accuracy of the ONT-based genome assembly, and Arima Hi-C contact data to place the sequences into chromosomes, we assembled a genome of 532 Mb (20 chromosomes, 309 scaffolds) with contig and scaffold N50 of 1.6 and 18.5 Mb, respectively. Fifty percent of the sequence (L50) was contained in seven superscaffolds. The consensus quality value of the final assembly was 42, and the single and duplicated gene completeness reported by BUSCO was 86.4% and 1%, respectively (metazoa_odb10 database). We annotated 26,348 protein-coding genes and 34,548 noncoding transcripts. This annotated

Significance

The Mediterranean red coral, *Corallium rubrum*, is critically impacted by overharvesting and by mass mortality events linked to marine heat waves. Accordingly, *C. rubrum* is increasingly the subject of conservation efforts. Previous population genetics studies based on microsatellites contributed to improving our knowledge of the species ecology. Yet, crucial questions regarding admixture among lineages, demographic history, effective population sizes, and local adaptation are still open owing to a lack of genomic resources. Here, we present the first chromosome-level genome assembly for the species with high contiguity, good completeness, protein-coding genes, and repeat sequence annotations. This genome, one of the first in octocorals, will pave the way for the integration of population genomics data into ongoing interdisciplinary conservation efforts dedicated to *C. rubrum*.

chromosome-level genome assembly, one of the first in octocorals and the first in Scleralcyonacea order, is currently used in a project based on whole-genome resequencing dedicated to the conservation and management of *C. rubrum*.

Key words: Catalan Initiative for the Earth BioGenome Project, Biodiversity Genomics Europe, Cnidaria, Hi-C, RNA-seq, Oxford Nanopore.

Introduction

Octocorallia is a diverse clade of cnidarian composed of more than 3,500 species (gorgonians and soft corals) shared between two orders: Scleralcyonacea and Malacalocyonacea. This clade is characterized by an interesting phylogenetic position within the class Anthozoa as the sister group of Hexacorallia. Octocorals and hexacorals, in particular stony corals (order Scleractinia), shared various ecological features. For instance, they are characterized by a key ecological role as habitat-forming species in benthic habitats from shallow tropical to deep and polar seas (e.g. Gomez-Gras et al. 2021). They are also under strong conservation concerns owing to the impacts of global change, including extreme climatic events (e.g. Estaque et al. 2023). In spite of these similarities, genomic resources remain scarce in octocorals compared with stony corals. The few genomes available in octocorals (e.g. Ledoux et al. 2020) represent <1% of species diversity (see Ahuja et al. 2024) and target exclusively species from the Malacalocyonacea order. Besides this biodiversity genomics gap, the lack of genomic resources limits the integration of genomics and population genetics data into ongoing conservation efforts (Formenti et al. 2022).

The red coral, *Corallium rubrum*, is a habitat-forming octocoral (Fig. 1) with a central structural role in benthic communities from the Mediterranean and the neighboring Atlantic (Laborel and Vacelet 1961, Zibrowius et al. 1984). This iconic species with high cultural and economic value is critically impacted by two anthropogenic pressures. First, as a "precious coral," it has been harvested for jewelry since ancient times and owing to its market value (>1,000€/kg), the species has been overharvested and intensively poached (Ledoux et al. 2016). Second, *C. rubrum* has been recurrently impacted in the last 20 years by mass mortalities, linked to marine heatwaves, across thousands of kilometers of coastal habitats (Garrabou et al. 2022). The species with slow population dynamics (Montero-Serra et al. 2018) and restricted connectivity (Ledoux et al. 2010; Horaud et al. 2024) is characterized by a low resilience capacity (Linares et al. 2012). The combination of overharvesting and mass mortality events is driving steep demographic declines, raising concerns over the evolutionary trajectory of the species (Montero-Serra et al. 2019).

In this context, *C. rubrum* is receiving conservation attention from scientists and biodiversity managers (included in Barcelona Convention, EU Habitat Directive and listed as "endangered" by IUCN [Otero et al. 2017]). Yet, major knowledge gaps in relation to genome diversity, effective population size, and adaptation to the local environment remain and should be filled to improve existing conservation policies. As a part of the Catalan Initiative for the Earth BioGenome Project (Corominas et al. 2024), we assembled and annotated the first chromosome-level reference genome for *C. rubrum* and for the Scleralcyonacea order. This reference genome will support a conservation genomics project funded by the Biodiversity Genomics Europe (https://biodiversitygenomics.eu), which is based on whole-genome resequencing. This project will infer demographic history and contemporary processes shaping the intraspecific genetic patterns with direct applications for red coral conservation and management.

Results and Discussion

Genome Assembly

The reference genome of *C. rubrum* was assembled based on ONT long reads, Illumina paired-end reads, and Arima Hi-C contact data (supplementary table S1, Supplementary Material online) analyzed with the pipeline CLAWS v2.1 (Gomez-Garrido 2023) following the flowchart shown in supplementary fig. S1, Supplementary Material online. Results obtained with Genomescope2 (supplementary fig. S2, Supplementary Material online) suggest a genome size

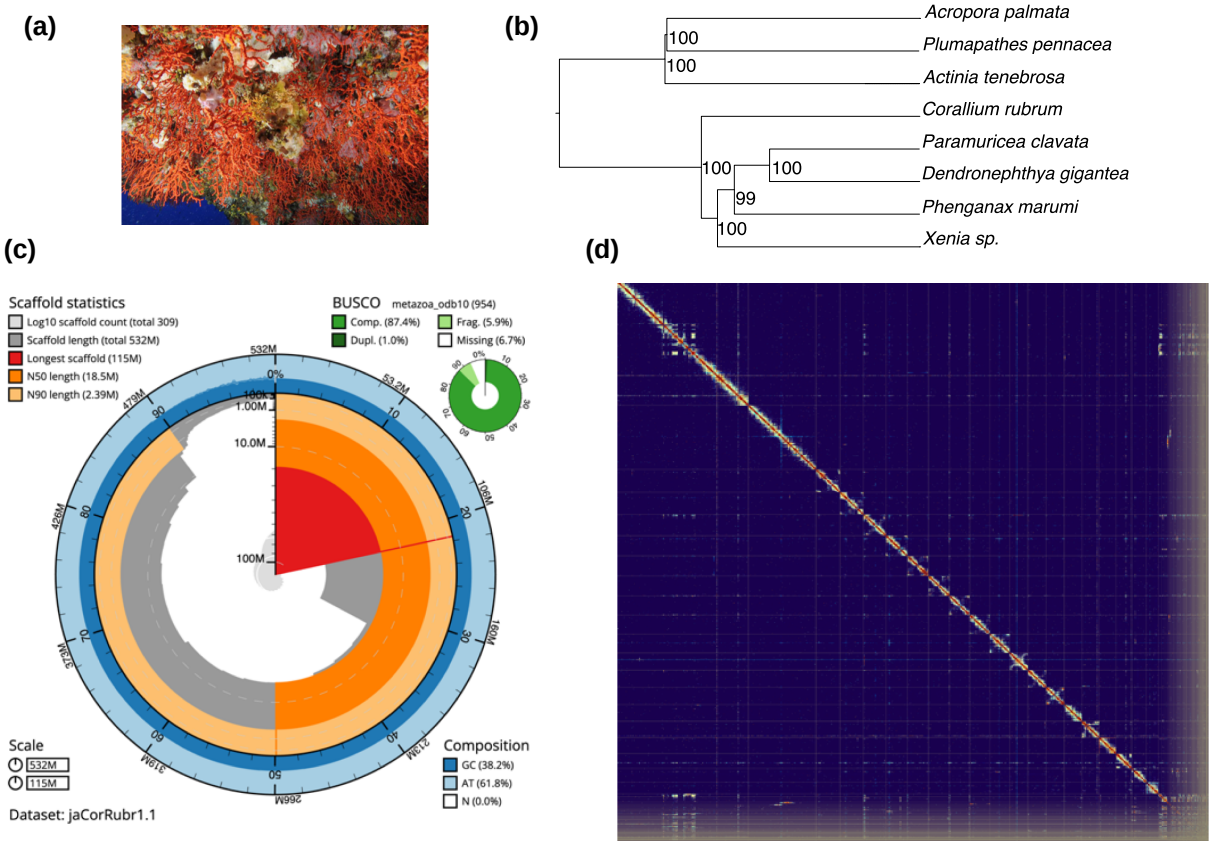

Fig. 1. a) Coralligenous habitat dominated by the red coral, *C. rubrum (picture by J. Garrabou)*. b) Phylogenetic relationships among different anthozoans species including five octocorals (*Dendronephthya gigantea*, *Paramuricea clavata*, *Phenganax marumi*, *Xenia* sp., and *C. rubrum*) and three hexacorals (*Actinia tenebrosa*, *Plumapathes pennacea*, and *Acropora palmata*) for which good-quality assemblies are available. The tree is based on 298 single-copy orthologous genes identified with BUSCO. c) BlobToolKit Snailplot showing different assembly metrics. The main plot is divided into 1,000 size-ordered bins around the circumference with each bin representing 0.1% of the 532 Mb assembly. The distribution of scaffold lengths is shown in dark gray with the plot radius scaled to the longest scaffold present in the assembly (115,004,408 bp, shown in red). Orange and pale-orange arcs show the N50 and N90 scaffold lengths (18,521,360 and 2,388,869 bp), respectively. The pale gray spiral shows the cumulative scaffold count on a log scale with white scale lines showing successive orders of magnitude. The blue and pale-blue areas around the outside of the plot show the distribution of GC, AT, and N percentages in the same bins as the inner plot. A summary of complete, fragmented, duplicated, and missing BUSCO genes in the metazoa_odb10 set is shown in the top right. d) Chromatin contact map generated from Arima2 Hi-C data shows the 20 chromosomes ($2n = 40$) that represent 89.3% of the assembled *C. rubrum* genome.

of around 500 Mb and 1.2% heterozygosity rate. The base assembly obtained with NextDenovo (ND) v2.4.1 comprised a total assembly span of 568 Mb (876 contigs, Table 1). The manual curation following the scaffolding with the Hi-C data resulted in a total of 8 cuts in contigs, 15 breaks at gaps, and 31 joins. The remaining edits corresponded to four unlocalized sequences and one haplotig. A total of 20 autosomes were assembled, and no sex chromosomes were identified. A total of 87 unplaced scaffolds (corresponding to 36 Mb of sequences) belonging to non-Cnidaria phyla were removed from the assembly (see blobplot supplementary fig. S3, Supplementary Material online). The final chromosome-level assembly comprised 532 Mb (20 chromosomes, 309 scaffolds; Table 1). The contig and scaffold N50 of the final assembly are 1.6 and 18.5 Mb, respectively, and 50% of the sequence (L50) is placed in seven superscaffolds. Merqury (Rhie et al. 2020) and BUSCO (Manni et al. 2021) were run to estimate the accuracy and completeness of the genome assembly. The consensus Phred-scaled base quality (quality value [QV] = $-10\log_{10}P$ where P is the probability of an incorrect base) of the final assembly was estimated by Merqury as 42 (>99.99% accurate). The gene completeness reported by BUSCO v5 was 87.4% (86.4% single and 1% duplicated BUSCOs) using the *metazoa_odb10* database (Fig. 1; Table 1), which is similar to values reported in other octocorals (e.g. 90.1% in Xenia sp.; see Hu et al. 2020).

Genome Annotation

Using RNA-seq data produced for this study (supplementary table S3, Supplementary Material online), we annotated a

Table 1 Statistics of the different versions of the genome assembly

Assembly	ND	ND + hypo	ND + hypo + purged	jaCorRubr1.2
Contig N50	1,993,440 bp	1,992,814 bp	2,029,805 bp	1,625,182 bp
Scaffold N50	1,993,440 bp	1,992,814 bp	2,029,805 bp	18,521,360 bp
Scaffold L50	84	84	79	7
Total sequences	876	876	784	309
Assembly span	567,713,602 bp	567,661,090 bp	545,517,441 bp	532,310,562 bp
BUSCO[a] single complete	84.3%	88.1%	88.2%	86.4%
BUSCO[a] duplicated complete	2.3%	2.5%	1.2%	1.0%
QV	33	42	42	42
K-mer completeness	83.8%	85.8%	84.9%	83.4%

[a]BUSCO v5 metazoa_odb10 database.

total of 26,348 protein-coding genes that produce 32,180 transcripts (1.22 transcripts per gene) and encode for 30,774 unique protein products. We were able to assign functional labels to 36% of the annotated proteins. The annotated transcripts contain 7.13 exons on average, with 79% of them being multiexonic (supplementary table S2, Supplementary Material online). In addition, 35,300 noncoding transcripts were annotated, of which 31,357 and 3,943 are long and short noncoding RNA genes, respectively. A total 64.4% of the assembly was identified as repetitive. The BUSCO single and duplicated completeness on this predicted protein set are 78.9% and 1.3%, respectively (complete: 80.2%, fragmented: 1.9%, missing: 17.9% with $n = 954$).

The reference genome presented here is the backbone of an ongoing population genomics project dedicated to the conservation and management of *C. rubrum*. This chromosome-level assembly, one of the first in octocorals and the first in Scleralcyonacea order, contributes to reduce the current taxonomic bias in the generation of high-quality genome resources.

Materials and Methods

Collection and Preparation of Biological Material

The apical tip (5 cm) of one colony from the Cap Castell (42.082610; 3.201981) population in Catalunya (Spain) was sampled at 18 m depth and immediately transported in coolers to the Aquarium Experimental Zone of the Institut de Ciències del Mar (ICM-CSIC, Barcelona, Spain) in November 2021. The sample was flash frozen using liquid nitrogen and stored at −80 °C until DNA extractions. The same individual was used for short- (Illumina) and long-read (ONT) sequencing. For Hi-C sequencing, one individual colony was sampled from Meda Petita population at 12 m depth (42.043652; 3.226719), Medes Islands, Spain, in May 2023.

DNA Extraction and Illumina Whole-Genome Sequencing (WGS)

High-molecular-weight gDNA was extracted from the coenenchyme (external tissue containing the polyps) using the MagAttract HMW DNA kit (Qiagen) at the Centre Nacional d'Analisi Genomica (https://www.cnag.eu). The HMW gDNA eluate was quantified using the Qubit DNA BR Assay kit (Thermo Fisher Scientific), and its purity was assessed using Nanodrop 2000 (Thermo Fisher Scientific). The extractions integrity was analyzed in an agarose gel (1%) in a pulsed field gel electrophoresis system (Sage Science). The HMW gDNA sample was stored at 4 °C. Whole-genome sequencing (WGS) library preparation was performed using the KAPA HyperPrep kit (Roche), following the manufacturer's instructions. The libraries were sequenced on the NovaSeq 6000 (Illumina) with a read length of 2× 151 bp, following the manufacturer's protocol for dual indexing. Image analysis, base calling, and quality scoring of the run were executed using the manufacturer's Real Time Analysis (RTA 3.4.4) software.

Long-Read Whole-Genome Library Preparation and Sequencing

The sequencing libraries were prepared using the 1D Sequencing kit SQK-LSK110 from ONT. Briefly, 4.0 µg of the DNA was DNA-repaired and DNA-end-repaired using NEBNext FFPE DNA Repair Mix (NEB) and the NEBNext UltraII End Repair/dA-Tailing Module (NEB) followed by the sequencing adaptors ligation. The ligation product was purified by 0.4× AMPure XP beads (Agencourt, Beckman Coulter) and eluted in elution buffer.

The sequencing runs were performed on PromethION 24 (ONT) using a flow cell R9.4.1 FLO-PRO 002 (ONT), and the sequencing data were collected for 110 h. The quality parameters of the sequencing runs were monitored by the MinKNOW platform version 21.11.7 in real time and base called with Guppy version 5.1.13.

Chromatin Conformation Capture Sample Preparation and Sequencing

Tissue was carefully scraped from a living individual collected at Medas Petit. Chromatin conformation capture sequencing (Hi-C) libraries were prepared using the Hi-C High-Coverage kit (Arima Genomics) in the Metazoa

Phylogenomics Lab (Institute of Evolutionary Biology [CSIC-UPF]). Sample concentration was assessed by Qubit DNA HS Assay kit (Thermo Fisher Scientific), and library preparation was carried out using the ACCEL-NGS 2S PLUS DNA LIBRARY KIT (Swift Bioscience) and using the 2S Set A single indexes (Swift Bioscience). Library amplification was carried out with the KAPA HiFi DNA polymerase (Roche). The amplified libraries were sequenced on the NovaSeq 6000 (Illumina) at CNAG.

RNA Extraction and RNA Sequencing

RNA sequencing data were obtained from a parallel project characterizing the transcriptomic response of *C. rubrum* to heat stress (Ramirez-Calero et al. in preparation). RNA was extracted from 36 different samples in 2021 combining TRIzol reagent (Invitrogen) for tissue lysis and homogenization and RNA easy kit (Qiagen) for RNA isolation and purification. Eluted RNA was stored at −80 °C until shipment to CNAG. Total RNA quantification was assessed using the Qubit RNA BR Assay kit (Thermo Fisher Scientific), and the RNA integrity was estimated using the RNA 6000 Nano Bioanalyzer 2100 Assay (Agilent). To prepare the RNA-Seq libraries, the KAPA Stranded mRNA-Seq Illumina Platforms Kit (Roche) was used with 500 ng of total RNA. Library quality was assessed on an Agilent 2100 Bioanalyzer using the DNA 7500 assay. The libraries were sequenced on the NovaSeq 6000 (Illumina) as above for the WGS library.

Genome Assembly

We used the pipeline CLAWS v2.1 (Gomez-Garrido 2023) to perform this genome assembly combining ONT long reads, Illumina paired-end reads, and Arima Hi-C contact data. A flowchart with the genome assembly process is shown in supplementary fig. S1, Supplementary Material online.

Prior to assembly, adaptors present in the Illumina data were trimmed with TrimGalore (https://github.com/FelixKrueger/TrimGalore). A k-mer database was subsequently built with Meryl (https://github.com/marbl/meryl). The k-mer histogram generated by Meryl was used as input to Genomescope2 (Ranallo-Benavidez et al. 2020) to estimate haploid genome size, heterozygosity, and repeat content (supplementary fig. S2, Supplementary Material online). The ONT data were filtered with Filtlong (Wick, https://github.com/rrwick/Filtlong; *–minlen* 1000 *–min_mean_q* 80 *–target_bases* 25000000000) prior to the assembly to remove short and low-quality reads.

The filtered ONT data were assembled with ND v2.4.0 (Hu et al. 2024). To improve the base accuracy, the assembly was polished with HyPo (Kundu et al. 2019) using both Illumina and ONT data. Finally, the polished assembly was purged with *purge_dups* (Guan et al. 2020) to remove alternate haplotypes and other artificially duplicated repetitive regions.

The assembly was scaffolded using the Hi-C data with YAHS (Zhou et al. 2023). Manual curation of the resulting assembly was performed with PretextView (https://github.com/wtsi-hpag/PretextView). The Blobtoolkit (Challis et al. 2020) pipeline was run on the curated assembly, using the NCBI nucleotide database (updated in February 2023) and several BUSCO odb10 databases (metazoa, eukaryota, fungi, and bacteria).

The decontaminated assembly was scaffolded using the Hi-C data with YAHS (Zhou et al. 2023). Manual curation of the resulting assembly was performed with PretextView (https://github.com/wtsi-hpag/PretextView).

A snailplot was produced on the final assembly with Blobtoolkit (Fig. 1).

Genome Annotation

The genome annotation was obtained by running the CNAG structural genome annotation pipeline (https://github.com/cnag-aat/Annotation_AAT) that uses a combination of transcript alignments, protein alignments, and ab initio gene predictions (supplementary fig. S4, Supplementary Material online). Repeats present in the genome assembly were annotated with RedMask.

After sequencing, adaptors were removed from the reads corresponding to the 36 samples used for the RNA sequencing with TrimGalore. Reads were aligned to the genome with STAR v-2.7.2a (Dobin et al. 2013). Transcript models were subsequently generated using Stringtie v2.2.1 (Pertea et al. 2015) on each BAM file and then all the models produced were combined using TACO v0.7.3 (Niknafs et al. 2017). High-quality junctions used during the annotation process were obtained by running ESPRESSO v1.3.0 (Gao et al. 2023) after mapping with STAR. Finally, PASA assemblies were produced with PASA v2.5.2 (Haas et al. 2008). The *TransDecoder* program was run on the PASA assemblies to detect the presence of coding regions in the transcripts. Additionally, the complete proteomes of *Stylophora pistillata*, *Pocillopora damicornis*, and *Paramuricea clavata* were downloaded from Swissprot/Uniprot (February 2023) and aligned to the *C. rubrum* genome using Miniprot v0.6 (Li 2023). Ab initio gene predictions were performed on the repeat-masked assembly with three different programs: GeneID v1.4 (Alioto et al. 2018), Augustus v3.5.0 (Stanke et al. 2006), and Genemark-ET v7.71 (Lomsadze et al. 2014) with and without incorporating evidence from the RNA-seq data. Geneid and Augustus were specifically trained for this species using a set of 1,000 gene candidates obtained from the longest Transdecoder complete models that had a significant (*e*value $<10^{-6}$) BLAST (Altschul et al. 1990) hit against Swissprot/Uniprot. Genemark was run in a self-training mode, and it was not specifically trained with this set of gene candidates.

Finally, all the data were combined into consensus coding sequence models using EvidenceModeler-2.1 (Haas et al. 2008). Additionally, untranslated regions and alternative splicing forms were annotated via two rounds of PASA annotation updates. To functionally annotate the proteins of the annotation, we run the Pannzer's online server (Törönen and Holm 2022). Orthofinder (Emms and Kelly 2019) was run to obtain the orthologs between *C. rubrum* and the previously downloaded proteins for *P. clavata*, *Po. damicornis*, and *S. pistillata*. The proteins that had not originally been annotated by Pannzer but for which an ortholog was found inherited the functional tags of their other paralogs in the *C. rubrum* annotation, or, if absent, they hierarchically obtained the annotation of their orthologs in *P. clavata*, *Po. Damicornis*, or *S. pistillata*.

The annotation of ncRNAs was obtained by running the following steps on the repeat-masked version of the genome assembly. First, cmsearch v1.1 (Cui et al. 2016) that is part of the Infernal package (Nawrocki and Eddy 2013) was run against the RFAM database of RNA families v12.0. Additionally, tRNAscan-SE v2.11 (Chan and Lowe 2019) was run to identify the transfer RNA genes present in the genome assembly. Identification of lncRNAs was done by first filtering the set of PASA assemblies that had not been included in the annotation of protein-coding genes to retain those longer than 200 bp and not covered more than 80% by a small ncRNA. The resulting transcripts were clustered into genes using shared splice sites or significant sequence overlap as criteria for designation as the same gene.

Supplementary Material

Supplementary material is available at *Genome Biology and Evolution* online.

Funding

This project was supported by the first call for sequencing reference genomes from the Catalan Initiative for the Earth BioGenome Project. J.-B.L. was supported by the strategic funding UIDB/04423/2020, UIDP/04423/2020, and 2021.00855.CEECIND provided by FCT—Fundação para a Ciência e Tecnologia. Institutional support to CNAG was from the Spanish Government, Ministry of Science, Innovation and Universities and Generalitat de Catalunya through the Departament de Recerca i Universitats and Departament de Salut. The project leading to this publication has received funding from European FEDER Fund under project 1166-39417. The project leading to this publication has received funding from Excellence Initiative of Aix-Marseille University—A*MIDEX, a French "Investissements d'Avenir" program. R.F. acknowledges support from the following sources of funding: Ramón y Cajal fellowship (grant agreement no. RYC2017-22492 funded by MCIN/AEI/10.13039/501100011033 and ESF "Investing in your future"), the Agencia Estatal de Investigación (project PID2019-108824GA-I00 funded by MCIN/AEI/10.13039/501100011033), the European Research Council (this project has received funding from the European Research Council [ERC] under the European's Union's Horizon 2020 research and innovation program [grant agreement no. 948281]), and the Secretaria d'Universitats i Recerca del Departament d'Economia i Coneixement de la Generalitat de Catalunya (AGAUR 2021-SGR00420). J.G. acknowledges the funding of the Spanish government through the "Severo Ochoa Centre of Excellence" accreditation (CEX2019-000928-S). C.L. gratefully acknowledges the financial support by ICREA under the ICREA Academia program. N.E.G. was supported by a Margaret A. Davidson Graduate Fellowship (NERRS NA22NOS4200050). J.G., C.L., P.L.-S., S.R.-C., and J.-B.L. are part of the Marine Conservation research group—MedRecover (2021 SGR 01073) from the "Generalitat de Catalunya".

Data Availability

Data and genome assembly presented in this article are available from CNAG (https://denovo.cnag.cat/) and ENA (Project GCA_964035015.1; https://www.ebi.ac.uk/ena/browser/view/GCA_964035015.1).

Literature Cited

Ahuja M, Cao X, Schultz DT, Picciani N, Lord A, Shao S, Jia K, Burdick DR, Haddock SHD, Li Y, et al. Giants among Cnidaria: large nuclear genomes and rearranged mitochondrial genomes in siphonophores. Genome Biol Evol. 2024:16(3):evae048. https://doi.org/10.1093/gbe/evae048.

Alioto T, Blanco E, Parra G, Guigó R. Using geneid to identify genes. Curr Protoc Bioinformatics. 2018:64(1):e56. https://doi.org/10.1002/cpbi.56.

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

Challis R, Richards E, Rajan J, Cochrane G, Blaxter M. BlobToolKit—interactive quality assessment of genome assemblies. G3 (Bethesda). 2020:10(4):1361–1374. https://doi.org/10.1534/g3.119.400908.

Chan PP, Lowe TM. tRNAscan-SE: searching for tRNA genes in genomic sequences. Methods Mol Biol. 2019:1962:1–14. https://doi.org/10.1007/978-1-4939-9173-0_1.

Corominas M, Marquès-Bonet T, Arnedo MA, Bayés M, Belmonte J, Escrivà H, Fernández R, Gabaldón T, Garnatje T, Germain J, et al. The Catalan initiative for the Earth BioGenome Project: contributing local data to global biodiversity genomics. NAR Genom Bioinform. 2024:6(3):lqae075. https://doi.org/10.1093/nargab/lqae075.

Cui X, Lu Z, Wang S, Jing-Yan Wang J, Gao X. CMsearch: simultaneous exploration of protein sequence space and structure space improves not only protein homology detection but also protein structure prediction. Bioinformatics. 2016:32(12):i332–i340. https://doi.org/10.1093/bioinformatics/btw271.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013:29(1):15–21. https://doi.org/10.1093/bioinformatics/bts635.

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

Estaque T, Richaume J, Bianchimani O, Schull Q, Mérigot B, Bensoussan N, Bonhomme P, Vouriot P, Sartoretto S, Monfort T, et al. Marine heatwaves on the rise: one of the strongest ever observed mass mortality event in temperate gorgonians. Glob Change Biol. 2023:29:6159–6162. https://doi.org/10.1111/gcb.16931.

Formenti G, Theissinger K, Fernandes C, Bista I, Bombarely A, Bleidorn C, Ciofi C, Crottini A, Godoy JA, Höglund J, et al. The era of reference genomes in conservation genomics. Trends Ecol Evol. 2022:37(3):197–202. https://doi.org/10.1016/j.tree.2021.11.008.

Gao Y, Wang F, Wang R, Kutschera E, Xu Y, Xie S, Wang Y, Kadash-Edmondson KE, Lin L, Xing Y. ESPRESSO: robust discovery and quantification of transcript isoforms from error-prone long-read RNA-seq data. Sci Adv. 2023:9(3):eabq5072. https://doi.org/10.1126/sciadv.abq5072.

Garrabou J, Gómez-Gras D, Medrano A, Cerrano C, Ponti M, Schlegel R, Bensoussan N, Turicchia E, Sini M, Gerovasileiou V, et al. Marine heatwaves drive recurrent mass mortalities in the Mediterranean Sea. Glob Chang Biol. 2022:28(19):5708–5725. https://doi.org/10.1111/gcb.16301.

Gomez-Garrido J. CLAWS CNAG's long-read assembly workflow for Snakemake. 2023. https://doi.org/10.48546/WORKFLOWHUB.WORKFLOW.567.2.

Gómez-Gras D, Linares C, Dornelas M, Madin JS, Brambilla V, Ledoux J-B, López-Sendino P, Bensoussan N, Garrabou J. Climate change transforms the functional identity of Mediterranean coralligenous assemblages. Ecol Lett. 2021:24(5):1038–1051. https://doi.org/10.1111/ele.13718.

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

Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, White O, Buell CR, Wortman JR. Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. Genome Biol. 2008:9(1):R7. https://doi.org/10.1186/gb-2008-9-1-r7.

Horaud M, Arizmendi-Meija R, Nebot-Colomer E, López-Sendino P, Antunes A, Dellicour S, Viard F, Leblois R, Linares C, Garrabou J, et al. Comparative population genetics of habitat-forming octocorals in two marine protected areas: eco-evolutionary and management implications. Conserv Genet. 2024:25:319–334. https://doi.org/10.1007/s10592-023-01573-8.

Hu J, Wang Z, Sun Z, Hu B, Ayoola AO, Liang F, Li J, Sandoval JR, Cooper DN, Ye K, et al. NextDenovo: an efficient error correction and accurate assembly tool for noisy long reads. Genome Biol. 2024:25(1):107. https://doi.org/10.1186/s13059-024-03252-4.

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

Kundu R, Casey J, Sung W-K. HyPo: super fast & accurate polisher for long read genome assemblies. bioRxiv 882506. https://doi.org/10.1101/2019.12.19.882506, 20 December 2019, preprint: not peer reviewed.

Laborel J, Vacelet J. Répartition bionomique du *Corallium rubrum* LMCK dans les grottes et falaises sous-marines. Rapp Comm int Mer Médit. 1961:16(2):464–469.

Ledoux J-B, Antunes A, Haguenauer A, Pratlong M, Costantini F, Abbiati M, Aurelle D. Molecular forensics into the sea: how molecular markers can help to struggle against poaching and illegal trade in precious corals? In: Stefano G, Zvy D, editors. Medusa and her children. Cnidarian evolution, through global climate change effects. Cham: Springer; 2016. p. 729–746.

Ledoux J-B, Cruz F, Gomez-Garrido J, Antoni R, Blanc J, Gómez-Gras D, López-Sendino P, Antunes A, Linares C, Gut M, et al. The genome sequence of the octocoral *Paramuricea clavata*—a key resource to study the impact of climate change in the Mediterranean. G3 (Bethesda). 2020:10(9):2941–2952. https://doi.org/10.1101/849158.

Ledoux J-B, Mokhtar-Jamaï K, Roby C, Féral J-P, Garrabou J, Aurelle D. Genetic survey of shallow populations of the Mediterranean red coral (*Corallium rubrum* (Linnaeus, 1758)): new insights into evolutionary processes shaping current nuclear diversity and implications for conservation. Mol Ecol. 2010:19:675–690. https://doi.org/10.1111/j.1365-294X.2009.04516.x.

Li H. Protein-to-genome alignment with miniprot. Bioinformatics. 2023:39(1):btad014. https://doi.org/10.1093/bioinformatics/btad014.

Linares C, Garrabou J, Hereu B, Diaz D, Marschal C, Sala E, Zabala M. Assessing the effectiveness of marine reserves on unsustainably harvested long-lived sessile invertebrates. Conserv Biol. 2012:26(1):88–96. https://doi.org/10.1111/j.1523-1739.2011.01795.x.

Lomsadze A, Burns PD, Borodovsky M. Integration of mapped RNA-Seq reads into automatic training of eukaryotic gene finding algorithm. Nucleic Acids Res. 2014:42(15):e119. https://doi.org/10.1093/nar/gku557.

Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. Mol Biol Evol. 2021:38(10):4647–4654. https://doi.org/10.1093/molbev/msab199.

Montero-Serra I, Garrabou J, Doak DF, Ledoux J-B, Linares C. Marine protected areas enhance structural complexity but do not buffer the consequences of ocean warming for an overexploited precious coral. J Appl Ecol. 2019:56(5):1063–1074. https://doi.org/10.1111/1365-2664.13321.

Montero-Serra I, Linares C, Doak DF, Ledoux J-B, Garrabou J. Strong linkages between depth, longevity and demographic stability across marine sessile species. Proc Biol Sci. 2018:285(1873):2017.2688. https://doi.org/10.1098/rspb.2017.2688.

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

Niknafs YS, Pandian B, Iyer HK, Chinnaiyan AM, Iyer MK. TACO produces robust multisample transcriptome assemblies from RNA-Seq. Nat Methods. 2017:14(1):68–70. https://doi.org/10.1038/nmeth.4078.

Otero M, Numa C, Bo M, Orejas C, Garrabou J, Cerrano C, Kružić P, Chryssanthi A, Aguilar R, Kipson S, et al. Overview of the conservation status of Mediterranean anthozoa. IUCN; 2017. https://doi.org/10.2305/IUCN.CH.2017.RA.2.en.

Pertea M, Pertea GM, Antonescu CM, Chang T-C, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol. 2015:33(3):290–295. https://doi.org/10.1038/nbt.3122.

Ranallo-Benavidez TR, Jaron KS, Schatz MC. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. Nat Commun. 2020:11(1):1432. https://doi.org/10.1038/s41467-020-14998-3.

Rhie A, Walenz BP, Koren S, Phillippy AM. Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies. Genome Biol. 2020:21(1):245. https://doi.org/10.1186/s13059-020-02134-9.

Stanke M, Schöffmann O, Morgenstern B, Waack S. Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. BMC Bioinformatics. 2006:7:62. https://doi.org/10.1186/1471-2105-7-62.
Törönen P, Holm L. PANNZER—a practical tool for protein function prediction. Protein Sci. 2022:31(1):118–128. https://doi.org/10.1002/pro.4193.
Wick R. FiltLong. https://github.com/rrwick/Filtlong.
Zhou C, McCarthy SA, Durbin R. YaHS: yet another Hi-C scaffolding tool. Bioinformatics. 2023:39(1):btac808.
Zibrowius H, Monteiro-Marques V, Grasshoff M. La répartition du *Corallium rubrum* dans l'Atlantique (Cnidaria: Anthozoa: Gorgonaria). TÉTHYS. 1984:11(2):163–170.

Associate editor: Christopher Wheat