Check for updates

DATA NOTE

REVISED ERGA-BGE reference genome of *Eunicella cavolini*, an IUCN Near Threatened Gorgonian of the Mediterranean Sea

[version 2; peer review: 2 approved]

Didier Aurelle[1,2], Dorian Guillemain[3], Frédéric Zuberer[3], Denys Malengros[3], Astrid Böhne[4], Rita Monteiro[4], Thomas Marcussen[5,6], Torsten H. Struck[5], Rebekah A. Oomen[5-9], Genomescope Sequencing Team, Alice Moussy[10], Corinne Cruaud[10], Karine Labadie[10], Lola Demirdjian[11], Caroline Belser[11], Patrick Wincker[11], Pedro H. Oliveira[11], Jean-Marc Aury[11], Fergal Martin[12], Anne Lazar[12], Leanne Haggerty[12], Chiara Bortoluzzi[13]

[1]Université de Toulon, CNRS, IRD, MIO, Aix Marseille Universite, Marseille, 13009, France
[2]Institut Systématique Evolution Biodiversité (ISYEB), CNRS, Sorbonne Université, EPHE, Université des Antilles, CP 26, Muséum national d'Histoire naturelle, Paris, 75005, France
[3]CNRS, IRD, IRSTEA, OSU PYTHEAS, Aix Marseille Univer, Marseille, France
[4]Leibniz Institute for the Analysis of Biodiversity Change, Museum Koenig Bonn, Bonn, 53113, Germany
[5]University of Oslo, Natural History Museum, Oslo, Norway
[6]University of Oslo, Centre for Ecological & Evolutionary Synthesis, Oslo, Norway
[7]Department of Biological Sciences, University of New Brunswick Saint John, Saint John, Canada
[8]University of Gothenburg, Tjärnö Marine Laboratory, Gothenburg, Sweden
[9]University of Agder, Centre for Coastal Research, Kristiansand, Norway
[10]Institut François Jacob, CEA, CNRS, Univ Evry, Université Paris-Saclay, Genoscope, Evry, 91057, France
[11]Institut François Jacob, CEA, CNRS, Univ Evry, Université Paris-Saclaye, Génomique Métabolique, Genoscope, Evry, 91057, France
[12]European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, Cambridge, England, UK
[13]SIB Swiss Institute of Bioinformatics, Amphipôle, Quartier UNIL-Sorge, Lausanne, 1015, Switzerland

v2 **First published:** 22 Oct 2025, **5**:323
https://doi.org/10.12688/openreseurope.21514.1
Latest published: 27 Jul 2026, **5**:323
https://doi.org/10.12688/openreseurope.21514.2

Abstract

The *Eunicella cavolini* reference genome provides an important resource to study the adaptation of this species to different environments and anthropic pressures. This species is impacted by human activities, including climate change, and this reference genome will be useful to study the genomic evolution of this species. The entirety of the genome sequence was assembled into 17 contiguous chromosomal pseudomolecules. This chromosome-level assembly encompasses 0.49 Gb, composed of 159 contigs and 46 scaffolds, with contig and scaffold N50 values of 7.7 Mb and 51.1 Mb, respectively.

Open Peer Review

Approval Status ✓ ✓

	1	2
version 2 (revision) 27 Jul 2026		✓ view
version 1 22 Oct 2025	✓ view	✓ view

1. **Benjamin D. Young**, University of Colorado Boulder Department of Ecology and Evolutionary Biology, Boulder, USA

Keywords
Eunicella cavolini, genome assembly, European Reference Genome Atlas, Biodiversity Genomics Europe, Earth Biogenome Project, Eunicellidae, yellow gorgonian

2. **Alan Elena**, Technische Universität (TU), Dresden, Germany

Any reports and responses or comments on the article can be found at the end of the article.

This article is included in the Horizon Europe gateway.

This article is included in the Genome Reports from the Biodiversity Genomics Europe Project collection.

Corresponding author: Chiara Bortoluzzi (chiara.bortoluzzi@sib.swiss)

Author roles: Aurelle D: Investigation, Methodology, Project Administration, Resources, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing; **Guillemain D**: Investigation, Methodology, Project Administration, Resources, Supervision, Writing – Review & Editing; **Zuberer F**: Investigation, Resources, Writing – Review & Editing; **Malengros D**: Investigation, Resources, Writing – Review & Editing; **Böhne A**: Methodology, Project Administration, Supervision, Writing – Review & Editing; **Monteiro R**: Methodology, Project Administration, Supervision, Writing – Review & Editing; **Marcussen T**: Methodology, Project Administration, Supervision, Writing – Review & Editing; **Struck TH**: Methodology, Project Administration, Supervision, Writing – Review & Editing; **A. Oomen R**: Methodology, Project Administration, Supervision, Writing – Review & Editing; **Moussy A**: Investigation, Supervision, Writing – Review & Editing; **Cruaud C**: Investigation, Supervision, Writing – Review & Editing; **Labadie K**: Investigation, Writing – Original Draft Preparation, Writing – Review & Editing; **Demirdjian L**: Data Curation, Formal Analysis, Writing – Review & Editing; **Belser C**: Data Curation, Formal Analysis, Writing – Review & Editing; **Wincker P**: Investigation, Supervision, Writing – Review & Editing; **H. Oliveira P**: Investigation, Supervision, Writing – Review & Editing; **Aury JM**: Data Curation, Formal Analysis, Supervision, Writing – Review & Editing; **Martin F**: Data Curation, Formal Analysis, Writing – Review & Editing; **Lazar A**: Data Curation, Formal Analysis, Writing – Review & Editing; **Haggerty L**: Data Curation, Formal Analysis, Writing – Review & Editing; **Bortoluzzi C**: Visualization, Writing – Original Draft Preparation

Competing interests: No competing interests were disclosed.

Grant information: Biodiversity Genomics Europe (Grant no. 101059492) is funded by Horizon Europe under the Biodiversity, Circular Economy and Environment call (REA.B.3); co-funded by the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract numbers 22.00173 and 24.00054; and by the UK Research and Innovation (UKRI) under the Department for Business, Energy and Industrial Strategy's Horizon Europe Guarantee Scheme. The project leading to this publication has received funding from European FEDER Fund under project 1166-39417 and the Excellence Initiative of Aix-Marseille University - A*MIDEX, a French 'Investissements d'Avenir' programme. This work was supported by the Genoscope, the Commissariat à l'Énergie Atomique et aux Énergies Alternatives (CEA), France Génomique (ANR-10-INBS-09-08), and the exploratory research programme 'ATLASea: Atlas of marine genomes and its targeted project SEQ-Sea (ANR-22-EXAT-0003-SEQ-Sea).
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2026 Aurelle D *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License , which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Aurelle D, Guillemain D, Zuberer F *et al.* **ERGA-BGE reference genome of *Eunicella cavolini,* an IUCN Near Threatened Gorgonian of the Mediterranean Sea [version 2; peer review: 2 approved]** Open Research Europe 2026, **5**:323 https://doi.org/10.12688/openreseurope.21514.2

First published: 22 Oct 2025, **5**:323 https://doi.org/10.12688/openreseurope.21514.1

REVISED **Amendments from Version 1**

In this revised manuscript, we have address all reviewers comments. Furthermore, we provide a new set of results concerning the genome annotation that has become available for Eunicella cavolini. Therefore, this revised manuscript presents a new section in the materials and methods section describing the generation of this genome annotation. As a result, three new co-authors were added to this revised manuscript.

Any further responses from the reviewers can be found at the end of the article

Introduction

Eunicella cavolini, also known by the French local name of Gorgone jaune, is a cnidaria member of the sub-class Octocorallia and of the Eunicellidae family (McFadden *et al.*, 2022) and is one of the most common gorgonians in the Mediterranean Sea. This species can be found in the Western and Eastern Mediterranean Sea, at depths of less than 10 meters and more than 150 meters (Carugati *et al.*, 2022; Sini *et al.*, 2015). *Eunicella cavolini* is currently classified as 'Near Threatened' on the IUCN Red List (Otero *et al.*, 2017). As other Mediterranean octocorals, *Eunicella* species are impacted by mass mortality events linked with marine heat waves, fishing activities, or pollution (Garrabou *et al.*, 2022; Sini *et al.*, 2015; Topçu & Öztürk, 2015). A high-quality reference genome for *E. cavolini* is useful to study the genomic diversity of this species and its adaptive potential facing contrasted and changing environments (Aurelle *et al.*, 2022). Such genomic resources are pivotal for species management, allowing, for example, the identification of key populations to be conserved. The genome will also be used to study species limits and hybridization with other *Eunicella* species (Aurelle *et al.*, 2024).

The generation of this reference resource was coordinated by the European Reference Genome Atlas (ERGA) initiative's Biodiversity Genomics Europe (BGE) project, supporting ERGA's aims of promoting transnational cooperation to promote advances in the application of genomics technologies to protect and restore biodiversity (Mazzoni *et al.*, 2023).

Materials & methods

ERGA's sequencing strategy includes Oxford Nanopore Technology (ONT) and/or Pacific Biosciences (PacBio) for long-read sequencing, along with Hi-C sequencing for chromosomal architecture, Illumina Paired-End (PE) for polishing (i.e. recommended for ONT-only assemblies), and RNA sequencing for transcriptomic profiling, to facilitate genome assembly and annotation.

Sample and sampling information

Dorian Guillemain, Frédéric Zuberer and Denys Malengros, all from the Pythéas Institute, sampled one specimen of *Eunicella cavolini* (sex unknown), which was identified based on macro-morphology and colour. The identification was performed by Dorian Guillemain, Frédéric Zuberer and Didier Aurelle. The sample was collected on the Frioul archipelago, Marseille, France, on the 19th of July 2023. Sampling was performed under permission Arrêté n°107 issued by the Direction Interrégionale de la mer Méditerranée. Sampling was performed through scuba diving. The specimen was euthanized by being flash frozen at −80°C. The sample was preserved at this temperature until DNA and RNA extraction.

Vouchering information

Physical reference materials for the here sequenced specimen have been deposited in the Museum National d'Histoire Naturelle, Paris (https://www.mnhn.fr/fr) under the accession number MNHN-IK-2019-2701.

Frozen reference tissue material from the same individual is being deposited at the Museum National d'Histoire Naturelle, Paris https://www.mnhn.fr/fr.

Genetic information

The estimated genome size, estimated by Genomes on a Tree (GoaT) (Challis *et al.*, 2023) by ancestral state reconstruction, is 0.62 Gb. This is a diploid genome with a haploid number of 6 chromosomes (2n = 12). All information for this species was retrieved from GoaT.

DNA/RNA processing

DNA extraction was performed by first grinding 400 mg of coenenchyme in liquid nitrogen, followed by digestion in 5 mL of buffer containing 30 mM Tris-HCl, 10 mM EDTA, 1% SDS, and proteinase K (20 μL/mL) at 53°C for 4 h. The lysate was centrifuged at 500 × g for 5 min at room temperature to remove debris, including sclerites. To the resulting

supernatant, 10 mL of absolute ethanol were added, and the mixture was incubated at −20°C for 1 h. DNA was pelleted by centrifugation at 8500 × g for 15 min at 4°C, washed once with 70% ethanol, and centrifuged again under the same conditions. The pellet was then resuspended in 9.5 mL of G2 buffer from Genomic Tip 100/G kit (QIAGEN, MD, USA) and incubated overnight at 4°C. The following day, RNase A (19 μL; 100 mg/mL) was added and the sample was incubated for 1 h at 50°C. DNA was subsequently purified using the Genomic Tip 100/G kit according to the manufacturer's protocol. DNA fragment size selection was performed using Short Read Eliminator (PacBio). Quantification was performed using a Qubit dsDNA HS Assay kit (Thermo Fisher Scientific) and integrity was assessed in a FemtoPulse system (Agilent). DNA was stored at 4°C until usage.

RNA was extracted from coenenchyme (70 mg) using RNeasy Plus Universal Kit (Qiagen) following manufacturer instructions. Residual genomic DNA was removed with 6 U of TURBO DNase (2 U/μL) (Thermo Fisher Scientific). Quantification was performed using a Qubit RNA HS Assay and integrity was assessed in a Bioanalyzer system (Agilent). RNA was stored at −80°.

Library preparation and sequencing

Long-read DNA library was prepared with SMRTbell prep kit 3.0 following manufacturers' instructions and sequenced on a Revio system (PacBio). Hi-C library was generated from reproductive tissue using the Arima High Coverage HiC Kit (following the Animal Tissues low input protocol v01) and sequenced on a NovaSeq6000 instrument (Illumina) with 2 × 150 read length. Poly(A) RNA-Seq libraries were constructed using the Illumina Stranded mRNA Prep, Ligation Prep kit (Illumina) and sequenced on a Illumina NovaSeq X Plus instrument.

In total 53× PacBio and 130× HiC data were sequenced to generate the assembly.

Genome assembly methods

The genome of *Eunicella cavolini* was assembled using the Genoscope GALOP pipeline (https://workflowhub.eu/workflows/1200). Briefly, raw PacBio HiFi reads were assembled using Hifiasm v0.19.8-r603 (Cheng *et al.*, 2021). Remaining allelic duplications were removed using purge_dups v1.2.5 (Guan *et al.*, 2020) with default parameters and the proposed cutoffs, but all contigs larger than 1 Mb were kept in the purged assembly. This assembly was scaffolded using YaHS v1.2.2 (Zhou *et al.*, 2023) and assembled scaffolds were then curated through manual inspection using PretextView v0.2.5 (Harry, 2022) to remove false joins and incorporate sequences not automatically scaffolded into their respective locations within the chromosomal pseudomolecules. Chromosome-scale scaffolds confirmed by Hi-C data were named in order of size. The mitochondrial genome was assembled as one circular contig using Oatk v1.0 (Zhou *et al.*, 2025) and included in the released assembly. Summary analysis of the released assembly was performed using the ERGA-BGE Genome Report ASM Galaxy workflow (10.48546/workflowhub.workflow.1104.1).

Genome annotation methods

A gene set was generated using the Ensembl Gene Annotation system (Aken *et al.*, 2016), primarily by aligning publicly available short-read RNA-seq data from BioSamples SAMEA115358980, SAMN38196247, SAMN38196248, SAMN38196249, SAMN38196250, SAMN38196251, and SAMN02800120 to the genome. Gaps in the annotation were filled via protein-to-genome alignments of a select set of clade-specific proteins from UniProt (Consortium, 2019) which had experimental evidence at the protein or transcript level. At each locus, data were aggregated and consolidated, prioritising models derived from RNA-seq data, resulting in a final set of gene models and associated non-redundant transcript sets. To distinguish true isoforms from fragments, the likelihood of each open reading frame (ORF) was evaluated against known metazoan proteins. Low-quality transcript models, such as those showing evidence of fragmented ORFs, were removed. In cases where RNA-seq data were fragmented or absent, homology data were prioritised, favouring longer transcripts with strong intron support from short-read data. The resulting gene models were classified into two categories: protein-coding, and long non-coding. Models that did not overlap protein-coding genes and were constructed from transcriptomic data were considered potential lncRNAs. Potential lncRNAs were further filtered to remove single-exon loci due to their unreliability. Putative miRNAs were predicted by performing a BLAST search of miRBase (Kozomara *et al.*, 2019) against the genome, followed by RNAfold analysis (Gruber *et al.*, 2008). Other small non-coding loci were identified by scanning the genome with Rfam (Kalvari *et al.*, 2018) and passing the results through Infernal (Nawrocki & Eddy, 2013). Summary analysis of the released annotation was carried out using the ERGA-BGE Genome Report ANNOT Galaxy workflow (10.48546/workflowhub.workflow.1096.1).

Results

Genome assembly

Results reported here are for the 17 contiguous chromosomal pseudomolecules. The genome assembly has a total length of 489,969,558 bp in 46 scaffolds (Figure 1 & Figure 2), with a GC content of 37.3%. The assembly has a contig N50 of

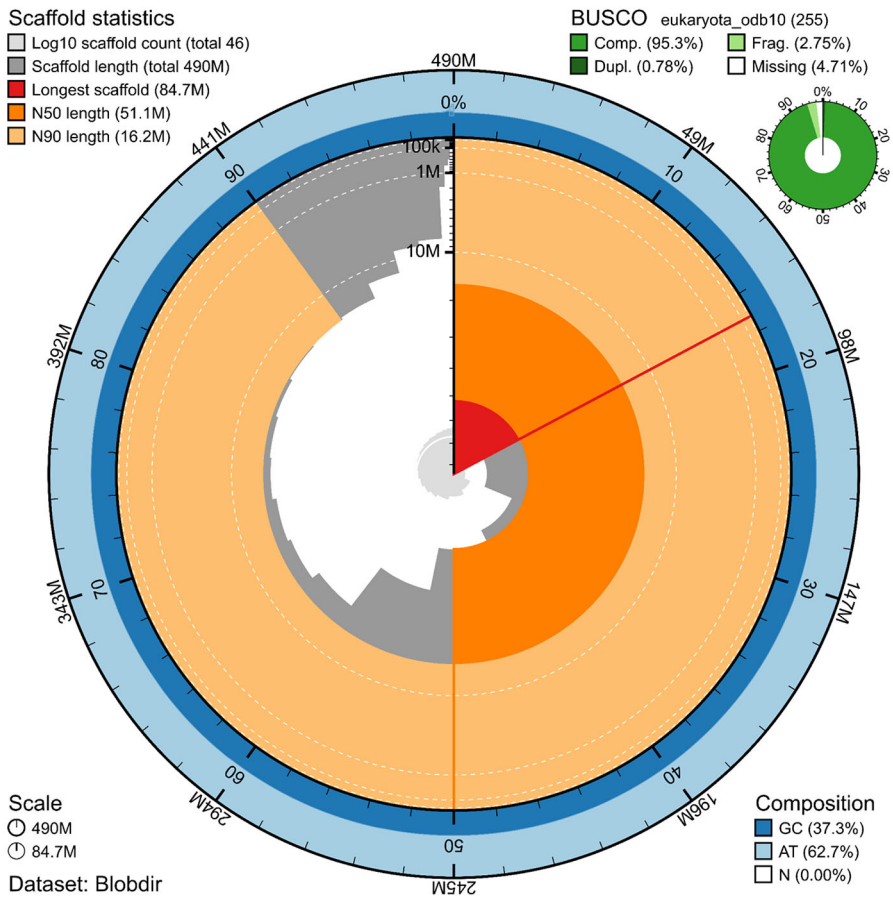


Figure 1. Snail plot summary of assembly statistics. The main plot is divided into 1,000 size-ordered bins around the circumference, with each bin representing 0.1% of the 489,969,558 bp assembly. The distribution of sequence lengths is shown in dark grey, with the plot radius scaled to the longest sequence present in the assembly (84.7 Mb, shown in red). Orange and pale-orange arcs show the scaffold N50 and N90 sequence lengths (51,102,291 and 16,173,997 bp), respectively. The pale grey spiral shows the cumulative sequence count on a log-scale, with white scale lines showing successive orders of magnitude. The blue and pale-blue area around the outside of the plot shows 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 found in the assembled genome from the Eukaryota database (odb10) is shown in the top right.

7,689,143 bp and L50 of 20 and a scaffold N50 of 51,102,291 bp and L50 of 4. The assembly has a total of 113 gaps, totalling 15.3 kb in cumulative size. The single-copy gene content analysis using the Eukaryota database (odb10) with BUSCO v5.8.0 (Manni *et al.*, 2021) resulted in 95.3% completeness (94.5% single and 0.8% duplicated). 80.6% of reads k-mers were present in the assembly and the assembly has a base accuracy Quality Value (QV) of 62.4 as calculated by Merqury (Rhie *et al.*, 2020).

Genome annotation

The genome annotation consists of 34,126 protein-coding genes with associated 49,606 transcripts, in addition to 7,956 non-coding genes (Table 1). Using the longest isoform per transcript, the single-copy gene content analysis using the Eukaryota database (odb10) with BUSCO resulted in 96.9% completeness. Using the Metazoa-v2.0.0.h5 database for OMArk (Nevers *et al.*, 2025) resulted in 93.4% completeness and 53.5% consistency (Table 2).

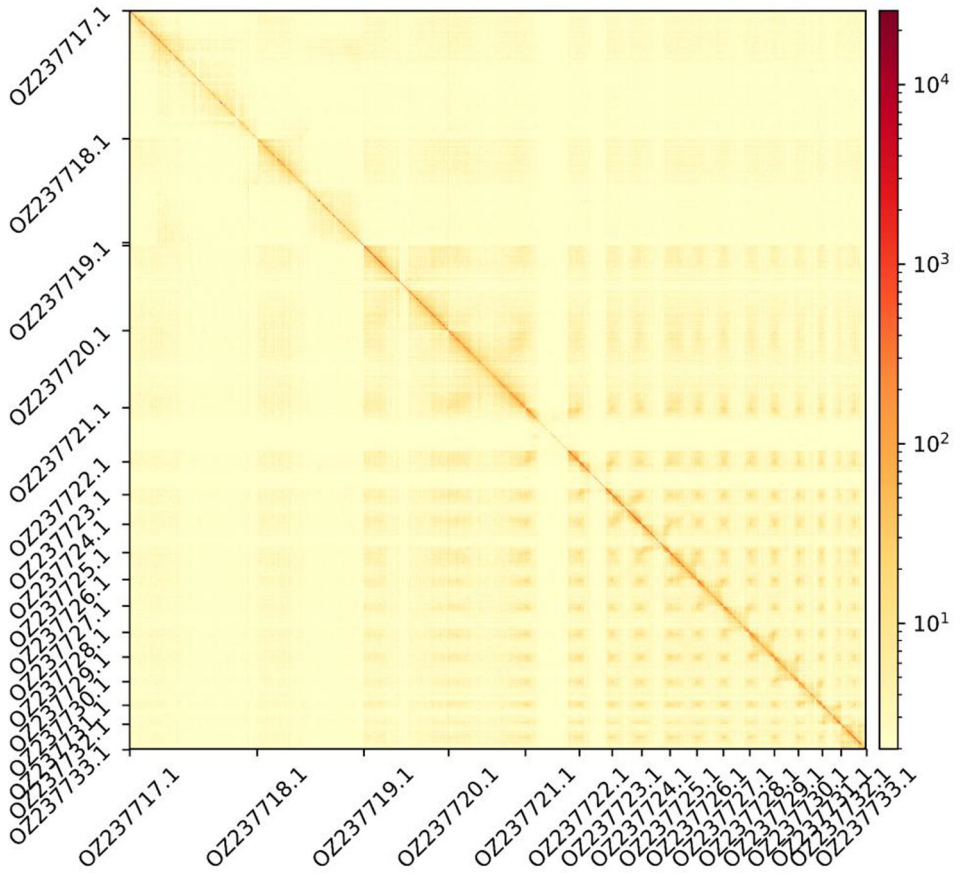


Figure 2. Hi-C contact map showing spatial interactions between regions of the genome. The diagonal corresponds to intra-chromosomal contacts, depicting chromosome boundaries. The frequency of contacts is shown on a logarithmic heatmap scale. Hi-C matrix bins were merged into a 100 kb bin size for plotting.

Table 1. Statistics from assembled gene models.

	No. genes	No. transcripts	Mean gene length (bp)	No. single-exon genes	Mean exons per transcript
mRNA	34,126	49,606	7,891	1,984	6.4
pseudogene	0.0	0.0	0.0	0.0	0.0
snoRNA	28	28	139	28	1.0
lncRNA	4,243	4,901	5,740	34	2.3
tRNA	2,914	2,914	76	2,914	1.0
snRNA	310	310	171	310	1.0
rRNA	460	460	355	460	1.0
scRNA	1	1	134	1	1.0
Other ncRNA	16,068	16,068	59–73	16,068	1.0–1.0

Table 2. Annotation completeness and consistency scores calculated by BUSCO run in protein mode (eukaryota_odb10) and OMArk (Metazoa-v2.0.0.h5).

	Complete	Singular	Duplicated	Fragmented	Missing
BUSCO	247 (96.9%)	246 (96.5%)	1 (0.4%)	5 (2.0%)	3 (1.1%)
OMArk	3,032 (93.4%)	2,761 (85.1%)	271 (8.3%)	-	212 (6.5%)
	Consistent	**Inconsistent**	**Contaminants**	**Unknown**	
OMArk	18,265 (53.5%)	0.0 (0.0%)	0.0 (0.0%)	15,861 (46.5%)	

Author contributions

DA coordinated the project; DG, FZ, and DM collected the species; DG, FZ, and DA identified the species; DA sampled and preserved biological material and provided metadata; DG, DA, AsB, RM, TM, THS, and RAM provided support in sampling, shipping of biological material, metadata collection, and management; GST extracted DNA, prepared libraries, and performed sequencing under the supervision of AM, CC, KL, PHO, and PW; LD, CB, and JMA performed genome assembly and curation under the supervision of JMA; FM, AL, and LH performed genome annotation; CB generated the analysis and report. All authors contributed to the writing, review, and editing of this genome note and read and approved the final version. This work is part of the species assigned to Genoscope, which was instrumental in the wet lab, sequencing, and assembly processes, and represents a key contribution to BGE's outputs.

Author information

Members of the Genoscope Sequencing Team are listed here: https://zenodo.org/records/14611490.

Data availability

Eunicella cavolini and the related genomic study were assigned to Tree of Life ID (ToLID) 'jaEunCavo1' and all sample, sequence, and assembly information are available under the umbrella BioProject PRJEB79972. The sample information is available at the following BioSample accession SAMEA115358980. The genome assembly is accessible from ENA under accession number GCA_965177985. 1. The annotated genome is available through the Ensembl website (https://projects.ensembl.org/erga-bge/). Sequencing data produced as part of this project are available from ENA at the following accessions: ERX14096372, ERX14096373, ERX14096374, ERX14096395, ERX14096397, ERX14169049, and ERX14169050. Documentation related to the genome assembly and curation can be found in the ERGA Assembly Report (EAR) document available at https://github.com/ERGA-consortium/EARs/tree/main/Assembly_Reports/Eunicella_cavolini/jaEunCavo1. Further details and data about the project are hosted on the ERGA portal at https://portal.erga-biodiversity.eu/data_portal/317547.

Acknowledgements

We acknowledge the support of the Freiburg Galaxy Team: Saim Momin and Björn Grüning, Bioinformatics, University of Freiburg (Germany), funded by the German Federal Ministry of Education and Research BMBF grant 031 A538A de. NBI-RBC and the Ministry of Science, Research and the Arts Baden-Württemberg (MWK) within the framework of LIBIS/de. NBI Freiburg. We would like to acknowledge the assembly reviewer, Michael Paulini from the Wellcome Sanger Institute. We acknowledge Magalie Castelin responsible for the cnidaria collection at MNHN.

References

Aken BL, Ayling S, Barrell D, *et al.*: **The Ensembl gene annotation system.** *Database.* 2016, 2016: baw093.
Publisher Full Text

Aurelle D, Haguenauer A, Bally M, *et al.*: **Symbiosis, hybridization, and speciation in Mediterranean octocorals (Octocorallia, Eunicellidae).** *Biol J Linn Soc.* 2024; **143**(4): blae116.
Publisher Full Text

Aurelle D, Thomas S, Albert C, *et al.*: **Biodiversity, climate change, and adaptation in the Mediterranean.** *Ecosphere.* 2022; **13**(4): e3915.
Publisher Full Text

Carugati L, Moccia D, Bramanti L, *et al.*: **Deep-dwelling populations of Mediterranean *Corallium rubrum* and *Eunicella cavolini*: distribution, demography, and co-occurrence.** *Biology (Basel).* 2022; **11**: 333.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Challis R, Kumar S, Sotero-Caio C, *et al.*: **Genomes on a Tree (GoaT): a versatile, scalable search engine for genomic and sequencing project metadata across the eukaryotic Tree of Life [version 1; peer review: 2 approved].** *Wellcome Open Res.* 2023; **8**: 24.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Cheng H, Concepcion GT, Feng X, *et al.*: **Haplotype-resolved *de novo* assembly using phased assembly graphs with hifiasm.** *Nat Methods.* 2021; **18**(2): 170–175.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Consortium, U: **UniProt: A worldwide hub of protein knowledge.** *Nucleic Acids Res.* 2019; **47**(D1): D506–D515.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Garrabou J, Gómez-Gras D, Medrano A, *et al.*: **Marine heatwaves drive recurrent mass mortalities in the Mediterranean Sea.** *Glob Chang Biol.* 2022; **28**(19): 5708–5725.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Gruber AR, Lorenz R, Bernhart SH, *et al.*: **The Vienna RNA websuite.** *Nucleic Acids Res.* 2008; **36**(suppl_2): W70–W74.
Publisher Full Text

Guan D, McCarthy SA, Wood J, *et al.*: **Identifying and removing haplotypic duplication in primary genome assemblies.** *Bioinformatics.* 2020; **36**(9): 2896–2898.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Harry E: *PretextView (Paired REad TEXTure Viewer): A desktop application for viewing pretext contact maps.* 2022.

Kalvari I, Nawrocki EP, Argasinska J, *et al.*: **Non-Coding RNA Analysis Using the Rfam Database.** *Curr Protoc Bioinformatics.* 2018; **62**(1): e51.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Kozomara A, Birgaoanu M, Griffiths-Jones S: **miRBase: From microRNA sequences to function.** *Nucleic Acids Res.* 2019; **47**(D1): D155–D162.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Manni M, Berkeley MR, Seppey M, *et al.*: **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.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Mazzoni CJ, Ciofi C, Waterhouse RM: **Biodiversity: an atlas of European reference genomes.** *Nature.* 2023; **619**: 252.
PubMed Abstract | **Publisher Full Text**

McFadden CS, Van Ofwegen LP, Quattrini AM: **Revisionary systematics of Octocorallia (Cnidaria: Anthozoa) guided by phylogenomics.** *Bull Soc Syst Biol.* 2022; **1**(3).
Publisher Full Text

Nawrocki EP, Eddy SR: **Infernal 1.1: 100-fold faster RNA homology searches.** *Bioinformatics.* 2013; **29**(22): 2933–2935.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Nevers Y, Warwick Vesztrocy A, Rossier V, *et al.*: **Quality assessment of gene repertoire annotations with OMArk.** *Nat Biotechnol.* 2025; **43**(1): 124–133.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Otero MM, Numa C, Bo M, *et al.*: **Overview of the conservation status of Mediterranean Anthozoa.** *International Union for Conservation of Nature.* 2017.
Publisher Full Text

Rhie A, Walenz BP, Koren S, *et al.*: **Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies.** *Genome Biol.* 2020; **21**(1): 245.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Sini M, Kipson S, Linares C, *et al.*: **The yellow gorgonian *Eunicella cavolini*: demography and disturbance levels across the Mediterranean Sea.** *PLoS One.* 2015; **10**(5): e0126253.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Topçu EN, Öztürk B: **Composition and abundance of octocorals in the Sea of Marmara, where the Mediterranean meets the Black Sea.** *Sci Mar.* 2015; **79**(1): 125–135.
Publisher Full Text

Zhou C, McCarthy SA, Durbin R: **YaHS: Yet another Hi-C Scaffolding tool.** *Bioinformatics.* 2023; **39**(1): btac808.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Zhou C, Brown M, Blaxter M, *et al.*: **Oatk: a de novo assembly tool for complex plant organelle genomes.** *Genome Biol.* 2025; **26**: 235.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Open Peer Review

Current Peer Review Status: ✔ ✔

Version 2

Reviewer Report 28 July 2026

https://doi.org/10.21956/openreseurope.25369.r78106

© **2026 Elena A.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **Alan Elena**
Technische Universität (TU), Dresden, Germany

I thank the authors for the responses to my comments. I have nothing more to add.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Bioinformatics, genomics, microbiology, environment, clinic.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 02 January 2026

https://doi.org/10.21956/openreseurope.23272.r66318

© **2026 Elena A.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **Alan Elena**
[1] Technische Universität (TU), Dresden, Germany
[2] Technische Universität (TU), Dresden, Germany

Aurelle *et al.* describe the generation of the only available *E. cavolini* genome. The assembly quality is high enough to be classified as a reference sequence. The provided data is relevant since it provides a high quality benchmark for future comparison. Authors highlight the relevance of such

sequences in a proper way. I believe the communication provides what it is supposed to provide and there are no major comments on my side.

Just a few minor questions:

1) Was this analysis done only once on one individual? Maybe to minimise genetic variability in a reference sequence and ensure robustness of results, this protocol could be repeated on independent organisms.
2) Given the fragmentation of the genome and the status of this entry as RefSeq, was there/will there be an attempt to close the scaffolds?
3) Why was reproductive tissue selected for sequencing?
4) How did the authors deal with potential DNA contamination from colonising species?
5) Aside from the present, there are very few sequences belonging to this species and most of these focus on specific structures (e.g.: Cytochromes) rather than on broader fragments of DNA. What is the reason for this?

Is the rationale for creating the dataset(s) clearly described?
Yes

Are the protocols appropriate and is the work technically sound?
Yes

Are sufficient details of methods and materials provided to allow replication by others?
Yes

Are the datasets clearly presented in a useable and accessible format?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Bioinformatics, genomics, microbiology, environment, clinic.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 13 Mar 2026

Chiara Bortoluzzi

We thank the reviewer for their enthusiasm regarding this Genome Report and production of the E. cavolini genome. We have addressed their comments as below:

1. The analysis was done only once and it cannot be repeated as there is no additional material available for the species and funding is also finished due to the end of the BGE project that funded this genome assembly;
2. This is a chromosome-level assembly and is not that fragmented as indicated by its contig N50 and scaffold N50. The genome assembly passed the ERGA Assembly Report, which is also meant to flag any potential fragmentation issues. None were found;

3. The use of the reproductive tissue (now referred to as coenenchyme) is a standard procedure for octocorals as it is difficult to separate specific tissues;
4. There was no evidence of macro-organism at the surface of the colony. Therefore, the risk of DNA contamination was very low. This is also confirmed by the contamination analysis reported in the ERGA Assembly Report: https://github.com/ERGA-consortium/EARs/blob/main/Assembly_Reports/Eunicella_cavolini/jaEunCavo1/jaEunCavo1.1_EAR.pdf ;
5. The genetic diversity and genetic structure of this species had previously been studied with microsatellites (Masmoudi et al., 2016), and more recently with transcriptome and RAD sequencing data (Romiguier et al., 2014; Aurelle et al., 2024).

References: Masmoudi, M. B., Chaoui, L., Topçu, N. E., Hammami, P., Kara, M. H., & Aurelle, D. (2016). Contrasted levels of genetic diversity in a benthic Mediterranean octocoral: Consequences of different demographic histories?. Ecology and Evolution, 6(24), 8665-8678. Romiguier, J., Gayral, P., Ballenghien, M., Bernard, A., Cahais, V., Chenuil, A., ... & Galtier, N. (2014). Comparative population genomics in animals uncovers the determinants of genetic diversity. Nature, 515(7526), 261-263. Aurelle, D., Haguenauer, A., Bally, M., Zuberer, F., Guillemain, D., Ledoux, J. B., ... & Pontarotti, P. (2024). Symbiosis, hybridization, and speciation in Mediterranean octocorals (Octocorallia, Eunicellidae). Biological Journal of the Linnean Society, 143(4), blae116.

Competing Interests: No competing interests were disclosed.

Reviewer Report 05 November 2025

https://doi.org/10.21956/openreseurope.23272.r62849

© 2025 Young B. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Benjamin D. Young
[1] University of Colorado Boulder Department of Ecology and Evolutionary Biology, Boulder, Colorado, USA
[2] University of Colorado Boulder Department of Ecology and Evolutionary Biology, Boulder, Colorado, USA

This is a fantastic resource for the soft coral community and the assembly pipeline follows best practices. I have a few small comments that I believe will improve the manuscript readability and should be super easy to incorporate!

- *Along with other gorgonians, E. cavolini plays an important ecological role by creating 3D habitats supporting a diversity of species (Gibson et al., 2006), including various bacterial species depending on the environment. Nonetheless, the dominant bacterial taxa are from the genus Endozoicomonas (Gammaproteobacteri) (Bayer et al., 2013).*

This sentence reads a little weird to me, especially as the *Endozoicomonas* are within the coral microbiome form the Bayer paper and the statement you are making seems to be about the

habitat the coral generates. I would recommend rewording this section or removing the section about the bacteria and microbiome as I do not think it is needed.

- *The specimen was euthanized by being put flash frozen at -80 °C.*

Remove the “put” it makes the sentence read weird.
You say “reproductive tissue” a few times, can you be a little more specific? I understand the sex was unknown so did you just grind up several whole polyps? Or did you extract mesenteries form within the polyps?

- *In total 53x PacBio and 130x HiC data were sequenced to generate the assembly.*

This is wonderful but did you get any results from the RNA-sequencing? You do specify you did RNA-prep and sequencing in the methods but then there is no further mention anywhere in the report of why you did this or what it was used for. I assume it was/is being used for annotation so maybe stating that would be good, and potentially providing some preliminary results for the gene prediction and annotation? If not, then I would recommend removing the RNA aspects completely.

- *and assembled scaffolds were then curated through manual inspection using PretextView v0.2.5 to remove false joins and incorporate sequences not automatically scaffolded into their respective locations within the chromosomal pseudomolecules.*

Is there an article/github citation you can include for PretextView ?

- *The genome assembly has a total length of 489,969,558 bp in 46 scaffolds (Figure 1 & Figure 2), with a GC content of 37.3%. The assembly has a contig N50 of 7,689,143 bp and L50 of 20 and a scaffold N50 of 51,102,291 bp and L50 of 4.*

This is a little nit-picky, but can you add in at the beginning the results that these results you are stating are for the 17 contiguous chromosomal pseudomolecules? It just makes it instantly obvious which assembly (initial versus chromosomal) you are talking about. I had to scroll to the abstract to make sure this was the chromosomal stats.
This is not necessary but could be nice, before giving the chromosomal results could you give the initial assembly size/stats from just HiFiAsm? I think it is interesting to see how use of HiC can improve the assembly. Again, not necessary at all if this is too much extra work!

- *The single-copy gene content analysis using the Eukaryota database with BUSCO*

Can you please include the version (i.e. odb10, odb12) BUSCO database you used.

Again, this is wonderful work and a amazing resource for researchers using this species!

Is the rationale for creating the dataset(s) clearly described?

Yes

Are the protocols appropriate and is the work technically sound?

Yes

Are sufficient details of methods and materials provided to allow replication by others?

Yes

Are the datasets clearly presented in a useable and accessible format?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: I work in a number of different organisms (e.g., bacteria, fungi, coral, insects, fish) assembling genomes and undertaking genetic research.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 13 Mar 2026

Chiara Bortoluzzi

We thank the reviewer for their enthusiasm regarding this Genome Report and production of the E. cavolini genome. We have addressed their comments as below:

1. We removed the small paragraph in the introduction about the bacteria and microbiome;
2. We removed 'put' as suggested;
3. We decided to change the term 'reproductive tissue' with the term 'coenenchyme' as this is more appropriate for octocoral;
4. We did produce RNA-Seq data, which are available at the following BioProject: PRJEB79972. The RNA-Seq data were used for the annotation, which is available in Ensembl. We added a new section about the genome annotation for completeness.
5. We added the citation to PretextView as suggested.
6. We added a new sentence at the beginning of the results section as suggested;
7. We added the version of the BUSCO database used.

Competing Interests: No competing interests were disclosed.