DATA NOTE

REVISED ERGA-BGE reference genome of *Holothuria (Platyperona) sanctori*: a sea cucumber from the Mediterranean Sea

[version 3; peer review: 2 approved]

Katerina Vasileiadou[1,2], Thanos Dailianis[2], Grigorios Skouradakis[2], Emmanouela Vernadou[2], Danae Karakasi[3,4], Astrid Böhne[5], Rita Monteiro[5], Rosa Fernández[6], Nuria Escudero[6], Tereza Manousaki[2], Genoscope Sequencing Team, Alice Moussy[7], Corinne Cruaud[7], Karine Labadie[7], Lola Demirdjian[8], Adama Ndar[8], Patrick Wincker[8], Pedro H Oliveira[8], Jean-Marc Aury[8], Chiara Bortoluzzi[9]

[1]Institute of Oceanography, Hellenic Centre for Marine Research, Heraklion, Crete, 70014, Greece
[2]Institute of Marine Biology Biotechnology and Aquaculture, Hellenic Centre for Marine Research, Heraklion, Crete, 70014, Greece
[3]Department of Biology, School of Sciences and Engineering, University of Crete, Vassilika Vouton, Heraklion, GR-70013, Greece
[4]Natural History Museum of Crete, School of Sciences and Engineering,, University of Crete, Knossos Avenue, Heraklion, GR-71409, Greece
[5]Leibniz Institute for the Analysis of Biodiversity Change, Adenauerallee 127, Museum Koenig Bonn, Bonn, 53113, Germany
[6]Metazoa Phylogenomics Lab, Institute for Evolutionary Biology (CSIC-UPF). Passeig marítim de la Barceloneta 37-49., Barcelona, 08003, Spain
[7]Genoscope, Institut François Jacob, CEA, CNRS, Univ Evry, Université Paris-Saclay, Evry, 91057, France
[8]Génomique Métabolique, Genoscope, Institut François Jacob, CEA, CNRS, Univ Evry, Université Paris-Saclay, Evry, 91057, France
[9]SIB Swiss Institute of Bioinformatics, Amphipôle, Quartier UNIL-Sorge, Lausanne, 1015, Switzerland

v3 **First published:** 25 Sep 2025, **5**:297
https://doi.org/10.12688/openreseurope.20841.1
Second version: 27 Jan 2026, **5**:297
https://doi.org/10.12688/openreseurope.20841.2
Latest published: 03 Feb 2026, **5**:297
https://doi.org/10.12688/openreseurope.20841.3

Abstract

Holothuria sanctori is a common species of sea cucumber found in the Mediterranean Sea and the Northeast Atlantic Ocean. It typically inhabits shallow rocky and sandy seabeds, where it plays a key ecological role as a sediment engineer processing organic matter ts and thereby contributing to nutrient cycling. As an edible species, *H. sanctori* is harvested in several countries. Although it is currently listed as a species of "Least Concern" on the IUCN Red List, the absence of a regulatory framework to prevent overexploitation poses a risk of population decline. Given its ecological significance and economic value, *H. sanctori* has become a focal point in both marine conservation and aquaculture research. The entirety of the genome sequence was assembled into 23 contiguous chromosomal

Open Peer Review

Approval Status ✓ ✓

	1	2
version 3 (revision) 03 Feb 2026		
version 2 (revision) 27 Jan 2026		
version 1 25 Sep 2025	✓ view	✓ view

1. **Poppy Hesketh-Best**, IGTP Institut Germans Trias i Pujol, Barcelona, Spain

2. **Estelle Proux-Wéra**, Stockholm

pseudomolecules. This chromosome-level assembly encompasses 1.2 Gb, composed of 135 contigs and 46 scaffolds, with contig and scaffold N50 values of 19.9 Mb and 50.7 Mb, respectively.

Keywords
Holothuria (Platyperona) sanctori, genome assembly, European Reference Genome Atlas, Biodiversity Genomics Europe, Earth Biogenome Project, Holothuriidae, sea cucumber

University, stockholm, Sweden

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: Vasileiadou K: Investigation, Resources, Writing – Original Draft Preparation, Writing – Review & Editing; **Dailianis T**: Investigation, Resources, Writing – Review & Editing; **Skouradakis G**: Investigation, Resources, Writing – Review & Editing; **Vernadou E**: Resources; **Karakasi D**: Investigation, Resources, Writing – Review & Editing; **Böhne A**: Methodology, Project Administration, Supervision, Writing – Review & Editing; **Monteiro R**: Methodology, Project Administration, Supervision, Writing – Review & Editing; **Fernández R**: Methodology, Project Administration, Supervision, Writing – Review & Editing; **Escudero N**: Methodology, Project Administration, Supervision, Writing – Review & Editing; **Manousaki T**: Methodology, Project Administration, Supervision, Writing – Review & Editing; **Moussy A**: Investigation, Supervision, Writing – Review & Editing; **Cruaud C**: Investigation, Supervision, Writing – Review & Editing; **Labadie K**: Investigation, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing; **Demirdjian L**: Data Curation, Formal Analysis, Writing – Review & Editing; **Ndar A**: 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; **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. 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.

How to cite this article: Vasileiadou K, Dailianis T, Skouradakis G *et al.* **ERGA-BGE reference genome of *Holothuria (Platyperona) sanctori*: a sea cucumber from the Mediterranean Sea [version 3; peer review: 2 approved]** Open Research Europe 2026, **5**:297 https://doi.org/10.12688/openreseurope.20841.3

First published: 25 Sep 2025, **5**:297 https://doi.org/10.12688/openreseurope.20841.1

REVISED **Amendments from Version 2**

In this revised version of the Genome Report of *Holothuria (Platyperona) sanctori*, we addressed all comments from the reviewers. In particular, we changed the text of the Data availability section regarding the annotation by stating that, while the RNA-seq data are available, the annotation is not yet and will be made available in the Ensembl website. We also added to the Genome assembly methods section that a bacterial contig and a substantial number of duplications were removed in the assembly step.

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

Introduction

Holothuria (Platyperona) sanctori is distributed throughout the northeastern Atlantic Ocean and is particularly common in the Mediterranean Sea. This nocturnal species inhabits shallow rocky substrates at depths of up to 70 meters (Navarro *et al.*, 2012). *Holuthuria sanctori* has a cylindrical body that is dark brown in colour, with the posterior end noticeably wider than the anterior. Its thick skin is covered with prominent dark papillae, each encircled by a whitish ring at the base. The species possesses twenty peltate tentacles and features three rows of podia and densely arranged tube feet on its ventral side (Moussa & Wirawati, 2018). While not present in all species within the Holothuriidae family (Caulier *et al.*, 2016), this sea cucumber is equipped with Cuvierian tubules, which serve as defensive structures. Phylogenetic analysis suggests that *H. sanctori* belongs to a deep evolutionary lineage dating back approximately 155 million years and it forms a distinct clade from other Mediterranean Holothuriidae species (Borrero-Pérez *et al.*, 2010).

The sea cucumber *Holothuria sanctori* is a deposit feeder that primarily consumes organic matter. It ingests large amounts of sediment, extracting nutrients from bacteria, fungi, and detritus embedded within the particles. Through this activity, it plays a key ecological role as a bioturbator, contributing to sediment oxygenation and overall seabed health. *H. sanctori* is harvested in the Mediterranean region for its saponins—bioactive compounds likely produced as a chemical defence mechanism. These saponins have attracted pharmaceutical interest due to their anti-tumoral, anti-inflammatory, and antibacterial properties (Caulier *et al.*, 2016). Additionally, the species holds significant value in aquaculture. Its ability to consume high concentrations of organic matter and filter nitrogen enables it to restore sediment quality beneath fish cages, supporting more sustainable farming practices (Moussa & Wirawati, 2018). *Holothuria sanctori* plays a crucial role in maintaining the stability and ecological integrity of marine ecosystems. Its economic importance in both pharmaceutical and aquaculture sectors highlight the need for effective conservation strategies to ensure the sustainability of its populations in the Mediterranean.

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 transcriptome profiling, to facilitate genome assembly and annotation.

Sample and sampling information

On 13 July 2023, Katerina Vasileiadou sampled one specimen of *Holothuria sanctori* (sex unknown). The specimen was identified by Katerina Vasileiadou in Lasithi, Elounda bay, Crete, Greece based on the characteristics described by Moussa & Wirawati (2018). Sampling was performed under permission ΥΠΕΝ/ΔΔΔ/34284/1131 issued by the Ministry for Environment and Energy, Secretariat General for Natural Environment & Water, Directorate General for Forests & Forest Environment, Directorate for Forest Management, Greece.

Tissues were removed from the living specimen following the legal framework for scientific specimens, flash frozen in liquid nitrogen, and stored at -80°C until DNA extraction. The specimen was euthanized following the Greek and EU legal framework.

Vouchering information

Physical reference materials for the here sequenced specimen have been deposited in the Natural History Museum of Crete https://www.nhmc.uoc.gr/en/ under accession number NHMC.65.25.

Frozen reference tissue material of the here sequenced specimen is available from the same individual at the Biobank of the Natural History Museum of Crete https://www.nhmc.uoc.gr/en/ under the voucher ID NHMC.65.25.

Genetic information

The estimated genome size, based on ancestral taxa, is 2.25 Gb. This is a diploid genome with a haploid number of 21 chromosomes (2n=42). All information for this species was retrieved from Genomes on a Tree (Challis *et al.*, 2023).

DNA/RNA processing

DNA was extracted from 300 mg of intestinal tissue using a Genomic-tip 100/G Kit (QIAGEN, MD, USA) following manufacturer instructions. DNA fragment size selection was performed using Short Read Eliminator (PacBio, CA, USA). 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. RNA was extracted from intestinal tissue (50 mg) using the RNeasy Plus Universal kit (Qiagen) following manufacturer instructions. Residual genomic DNA was removed with 6U of TURBO DNase (2 U/μL) (Thermo Fisher Scientific). Quantification was performed using a Qubit RNA HS Assay kit and integrity was assessed in a Bioanalyzer system (Agilent). RNA was stored at -80 °C.

Library preparation and sequencing

Long-read DNA libraries were prepared with the SMRTbell prep kit 3.0 following manufacturers' instructions and sequenced on a Revio system (PacBio). Hi-C libraries were generated from intestinal tissue (50 mg) using the Arima High Coverage HiC kit (following the Animal Tissues low input protocol v01) and sequenced on a NovaSeq X Plus instrument (Illumina) with 2x150 bp read length. Poly(A) RNA-Seq libraries were constructed using the Illumina Stranded mRNA Prep, Ligation Prep kit (Illumina) and sequenced on an Illumina NovaSeq X Plus instrument (Illumina) with 2x150 bp read length.

In total 90x PacBio HiFi and 47x HiC data were sequenced to generate the assembly.

Genome assembly methods

The genome of *Dendarus foraminosus* 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, as well as a bacterial contig. The purged 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 to remove false joins and incorporate sequences not automatically scaffolded into their respective locations within the chromosomal pseudomolecules. This step included the removal of a bacterial contig and a substantial number of duplications. 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.*, 2024) and included in the released assembly. Summary analysis of the released assembly was performed using the ERGA-BGE Genome Report ASM Galaxy workflow (https://doi.org/10.48546/workflowhub.workflow.1104.1).

Results

Genome assembly

The genome assembly has a total length of 1,171,681,863 bp in 46 scaffolds including the mitogenome (Figure 1 & Figure 2), with a GC content of 39.4%. The assembly has a contig N50 of 19,932,592 bp and L50 of 21 and a scaffold N50 of 50,699,391 bp and L50 of 10. The assembly has a total of 89

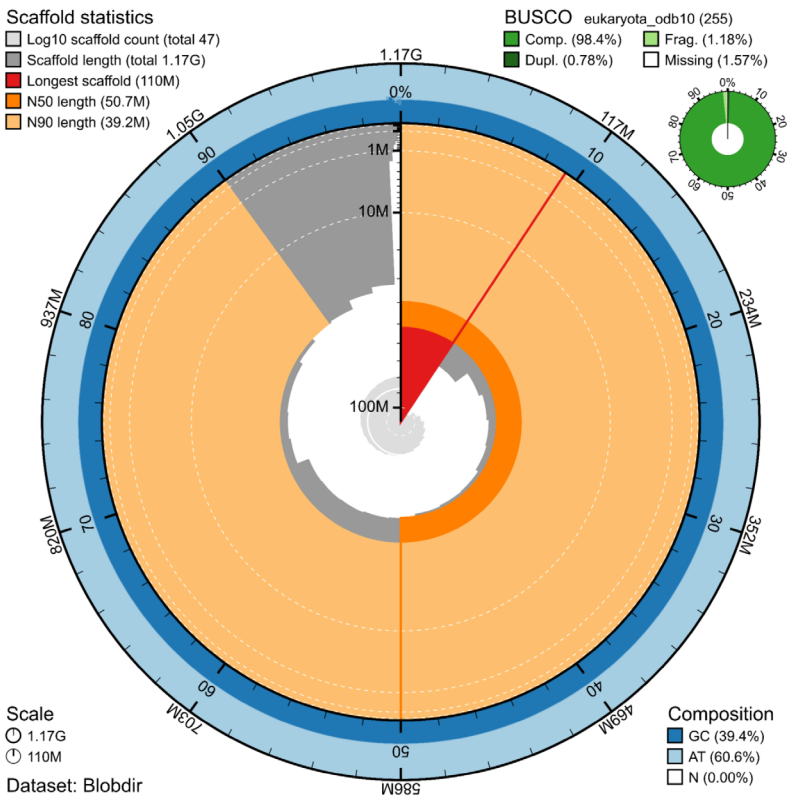

Figure 1. Snail plot summary of assembly statistics. The main plot is divided into 1,000 size-ordered bins around the circumference, with each bin representing 0.1% of the 1,171,681,863 bp assembly including the mitochondrial genome. The distribution of sequence lengths is shown in dark grey, with the plot radius scaled to the longest sequence present in the assembly (110 Mb, shown in red). Orange and pale-orange arcs show the scaffold N50 and N90 sequence lengths (50,699,391 and 39,181,847 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.

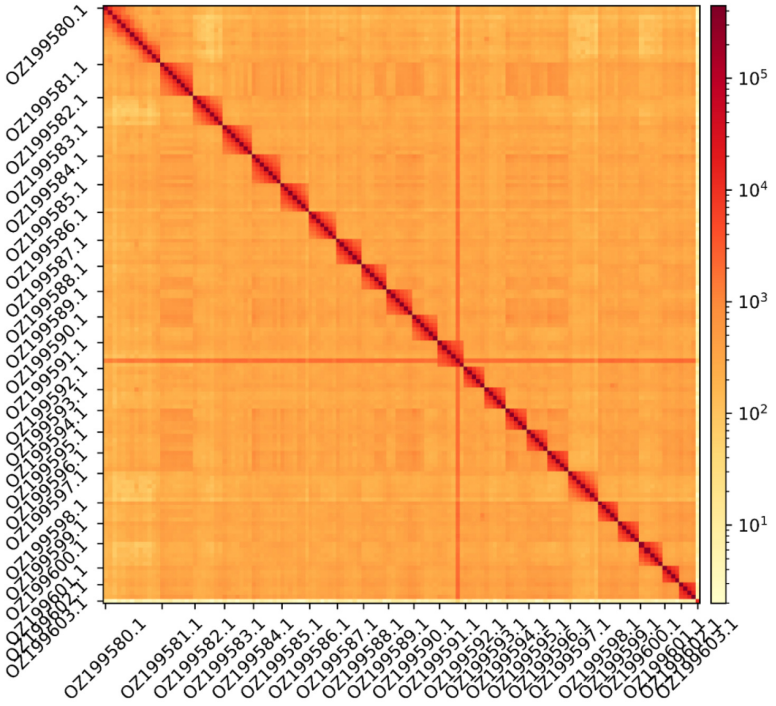

Figure 2. Hi-C contact map showing spatial interactions between regions of the genome. The diagonal corresponds to intra-chromosomal contacts, depicting chromosome boundaries. The frequency of contacts is shown on a logarithmic heatmap scale. Hi-C matrix bins were merged into a 150 kb bin size for plotting. The Hi-C contact map shows the 23 autosomes and the mitochondrial genome (GenBank accession: OZ199603.1).

gaps, totalling 15.0 kb in cumulative size. The single-copy gene content analysis using the Eukaryota database with BUSCO (Manni *et al.*, 2021) resulted in 98.4% completeness (97.6% single and 0.8% duplicated). 73.3% of reads k-mers were present in the assembly and the assembly has a base accuracy Quality Value (QV) of 62.3 as calculated by Merqury (Rhie *et al.*, 2020).

Data availability

H. sanctori and the related genomic study were assigned to Tree of Life ID (ToLID) 'ehHolSanc2' and all sample, sequence, and assembly information are available under the umbrella BioProject PRJEB77253. The sample information is available at the following BioSample accessions: SAMEA115125951 and SAMEA115125945. The genome assembly is accessible from ENA under accession number GCA_964304565.2, while the annotated genome will be made available through the Ensembl website (https://projects.ensembl.org/erga-bge/). The genome will be annotated using the RNA-Seq data generated alongside the genome assembly (see DNA/RNA processing)

Sequencing data produced as part of this project are available from ENA at the following accessions: ERX14170116, ERX14170117, ERX13549354, ERX13549355, ERX14096391, ERX14096392, and ERX14096393. 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/Holothuria_sanctori/ehHolSanc2. Further details and data about the project are hosted on the ERGA portal at https://portal.erga-biodiversity.eu/data_portal/491806.

Author contributions

KV and TM coordinated the project; TD, GS and EV collected the species; TD, GS and EV identified the species; KV, TM, and DK sampled and preserved biological material and provided metadata; AsB, RM, RF, and NE 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 and AN performed genome assembly and curation under the supervision of JMA; 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.

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, Fernando Cruz from the Centro Nacional de Análisis Genómic (CNAG).

References

Borrero-Pérez GH, Gómez-Zurita J, González-Wangüemert M, *et al.*: **Molecular systematics of the genus *Holothuria* in the Mediterranean and Northeastern Atlantic and a molecular clock for the diversification of the Holothuriidae (Echinodermata: Holothuroidea).** *Mol Phylogenet Evol.* 2010; **57**(2): 899–906.
PubMed Abstract | **Publisher Full Text**

Caulier G, Mezali K, Soualili DL, *et al.*: **Chemical characterization of saponins contained in the body wall and the Cuvierian tubules of the sea cucumber *Holothuria (Platyperona) sanctori* (Delle Chiaje, 1823).** *Biochem Syst Ecol.* 2016; **68**: 119–127.
Publisher 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**

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**

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**(7969): 252.
PubMed Abstract | **Publisher Full Text**

Moussa R, Wirawati I: **Observations on some biological characteristics of *Holothuria polii* and *Holothuria sanctori* from Mediterranean Egypt.** *Int J Fish Aquat Stud.* 2018; **6**(3): 351–357.
Reference Source

Navarro PG, Garcia-Sanz S, Tuya F: **Reproductive biology of the sea cucumber *Holothuria sanctori* (Echinodermata: Holothuroidea).** *Sci Mar.* 2012; **76**(4): 741–752.
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**

Zhou C, Brown M, Blaxter M, *et al.*: **Oatk: a *de novo* assembly tool for complex plant organelle genomes.** *bioRxiv.* 2024; 2024–10.
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**

Open Peer Review

Current Peer Review Status: ✔ ✔

Version 1

Reviewer Report 30 December 2025

https://doi.org/10.21956/openreseurope.22549.r64808

✔ **Estelle Proux-Wéra**
[1] Stockholm University, stockholm, Sweden
[2] Stockholm University, stockholm, Sweden
[3] Stockholm University, stockholm, Sweden

The authors present the genome assembly of *Holothuria (Platyperona) sanctori*, a sea cucumber.

The materials and methods section describes in details how the specimen was sampled, how the DNA was extracted, and how the genome assembly was performed.

The raw data and the final assembly are available and find able through the links provided in the article.

The genome assembly seems to be of high quality, with an expected number of chromosomes and good metrics (N50, L50). The BUSCO score is very good.

I was wondering about contamination, as there is nothing mentioned in the article. I could see in the EAR report that 1 bacterial contig was removed. I think it would be good to write it as well in the article.

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, genome assembly, comparative genomics

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 05 Jan 2026

Chiara Bortoluzzi

We thank the reviewer for their comments. We have addressed all of them as following:
1. Following reviewer 1, we added to the Genome assembly methods section additional information about the removal of a bacterial contig and a substantial number of duplications during the assembly step.

Competing Interests: No competing interests were disclosed.

Reviewer Report 29 December 2025

https://doi.org/10.21956/openreseurope.22549.r64816

Poppy Hesketh-Best
[1] IGTP Institut Germans Trias i Pujol, Barcelona, Spain
[2] IGTP Institut Germans Trias i Pujol, Barcelona, Spain
[3] IGTP Institut Germans Trias i Pujol, Barcelona, Spain

In this data note, Vasileiadou and colleagues present a reference genome for *Holothuria (Platyperona) sanctori*, a sea cucumber species important for maintaining seabed health through its bioturbation activities. The reference genome consists of 46 scaffolds, including the mitochondrial genome, and represents a valuable resource for future research.

Is the rationale for creating the dataset clearly described?
Yes. The rationale is clearly articulated within the broader context of the European Reference Genome Atlas initiatives. In addition, the specific justification for generating a reference genome for this sea cucumber species is well described.

Are the protocols appropriate and is the work technically sound?
The protocols used for sequencing and assembling the genome are appropriate for marine invertebrates, and the technical approach for assembly is sound.

Are sufficient details of methods and materials provided to allow replication by others?
Yes. The manuscript provides sufficient detail on DNA extraction, library preparation, and genome assembly to enable replication and to support appropriate downstream use of the reference genome.

One point of confusion is the subtitle "DNA/RNA processing" in the methods. Transcriptomics is not integrated into the results describing dataset generation. The authors could clarify the availability and relevance of the transcriptome, or remove mention of the RNA storage and focus on DNA and the genome.

Additionally, the curator's notes indicate the removal of a bacterial contig and a substantial number of duplications; these steps contributed to the final reference genome and should be explicitly reported in the data note.

Are the datasets clearly presented in a usable and accessible format?
Yes. As of 2025-12-09, both the genome and transcriptome datasets were accessible through ENA and NCBI. The GitHub repository containing curation and assembly reports is also available and useful, although more of this information could be incorporated directly into the data note.

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 and (meta-)genomics

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 05 Jan 2026

Chiara Bortoluzzi

We thank the reviewer for their comments. We have addressed all of them as following:

1. We slightly changed the text of the Data availability section regarding the annotation by stating that, while the RNA-seq data are available, the annotation is not yet and will be made available in the Ensembl website.

2. We added in the Genome assembly methods section that a bacterial contig and a substantial number of duplications were removed in the assembly step.

Competing Interests: No competing interests were disclosed.