



DATA NOTE

The chromosomal genome sequence of the giant barrel sponge, *Xestospongia muta* Schmidt 1870 and its associated microbial metagenome sequences

[version 1; peer review: 2 approved]

Jose Victor Lopez ¹, Shirley A. Pomponi², Ute Hentschel ³, Dirk Erpenbeck ⁴, Nina Pruzinsky ⁵, Cara Fiore ⁶, Rebecca Mulheron², Graeme Oatley ⁷, Elizabeth Sinclair ⁷, Eerik Aunin⁷, Noah Gettle⁷, Camilla Santos⁷, Michael Paulini⁷, Haoyu Niu⁷, Victoria McKenna⁷, Rebecca O'Brien ⁷, Wellcome Sanger Institute Tree of Life Management, Samples and Laboratory Team, Wellcome Sanger Institute Scientific Operations: Sequencing Operations, Wellcome Sanger Institute Tree of Life Core Informatics Team, EBI Aquatic Symbiosis Genomics Data Portal Team, Aquatic Symbiosis Genomics Project Leadership

¹National Coral Reef Institute, Department of Biological Sciences, Nova Southeastern University, Dania Beach, Florida, USA

²Florida Atlantic University, Boca Raton, Florida, USA

³GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany

⁴Department of Earth and Environmental Sciences, Ludwig-Maximilian University of Munich, Munich, Germany

⁵Science & Technology Division and Cooperative Institute Support, NOAA Ocean Exploration (UCAR), Silver Spring, Maryland, USA

⁶Appalachian State University, Boone, North Carolina, USA

⁷Tree of Life, Wellcome Sanger Institute, Hinxton, England, UK

V1 First published: 08 Jul 2025, 10:336
<https://doi.org/10.12688/wellcomeopenres.24173.1>
Latest published: 08 Jul 2025, 10:336
<https://doi.org/10.12688/wellcomeopenres.24173.1>

Abstract

We present a genome assembly from a specimen of *Xestospongia muta* (Caribbean barrel sponge; Porifera; Demospongiae; Haplosclerida; Petrosiidae). The genome sequence has a total length of 158.52 megabases. Most of the assembly (99.56%) is scaffolded into 15 chromosomal pseudomolecules. The mitochondrial genome has also been assembled and is 18.99 kilobases in length. Several symbiotic bacterial genomes were assembled as MAGs, including *Candidatus* Poribacteria species, *Candidatus* Latescibacteria, Acidobacteriota, Actinomycetota Gemmatimonadota, multiple Chloroflexota and the archaeon Nitrosopumilus. Gene annotation of

Open Peer Review

Approval Status

	1	2
version 1 08 Jul 2025	 view	 view
1. Maximiliano M. Maronna , Federal University of Rio Grande, Rio Grande, Brazil		
2. Arzu Karahan , Middle East Technical University, Erdemli-Mersin, Turkey		
Any reports and responses or comments on the		

this assembly on Ensembl identified 20,220 protein-coding genes.

article can be found at the end of the article.

Keywords

Xestospongia muta, Caribbean barrel sponge, genome sequence, chromosomal, Haplosclerida; microbial metagenome



This article is included in the [Tree of Life gateway](#).

Corresponding author: Aquatic Symbiosis Genomics Project Leadership (Mark.Blaxter@sanger.ac.uk)

Author roles: **Lopez JV:** Investigation, Resources, Supervision, Writing – Original Draft Preparation; **Pomponi SA:** Resources, Writing – Review & Editing; **Hentschel U:** Writing – Review & Editing; **Erpenbeck D:** Investigation; **Pruzinsky N:** Resources; **Fiore C:** Writing – Review & Editing; **Mulheron R:** Resources; **Oatley G:** Investigation; **Sinclair E:** Investigation; **Aunin E:** Formal Analysis; **Gettle N:** Data Curation, Formal Analysis, Methodology; **Santos C:** Formal Analysis; **Paulini M:** Formal Analysis; **Niu H:** Project Administration; **McKenna V:** Project Administration; **O'Brien R:** Project Administration;

Competing interests: No competing interests were disclosed.

Grant information: This work was funded by the Gordon and Betty Moore Foundation through a grant (GBMF8897) to the Wellcome Sanger Institute to support the Aquatic Symbiosis Genomics Project, and by Wellcome through core funding to the Wellcome Sanger Institute (220540).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2025 Lopez JV *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: Lopez JV, Pomponi SA, Hentschel U *et al.* **The chromosomal genome sequence of the giant barrel sponge, *Xestospongia muta* Schmidt 1870 and its associated microbial metagenome sequences [version 1; peer review: 2 approved]** Wellcome Open Research 2025, 10:336 <https://doi.org/10.12688/wellcomeopenres.24173.1>

First published: 08 Jul 2025, 10:336 <https://doi.org/10.12688/wellcomeopenres.24173.1>

Species taxonomy

Eukaryota; Opisthokonta; Metazoa; Porifera; Demospongiae; Heteroscleromorpha; Haplosclerida; Petrosiidae; *Xestospongia*; *Xestospongia muta* (Schmidt, 1870) (NCBI:txid178552)

Background

For a sessile marine invertebrate, the giant barrel sponge species, *Xestospongia muta* Schmidt, 1870, can loom large as one of the most iconic species on Caribbean tropical reefs. These sponges have been nicknamed “redwoods of the reef” because of the massive sizes and their long lifespans, dated to live between hundreds of years to up 2,300 years, comparable to California redwoods (McMurray *et al.*, 2008). Individual barrel sponges can exceed 2 m in height and 3 m in diameter. The sponge mesohyl is rigid in texture and the outer, light-exposed colour ranges from reddish-brown to purple and grey (Figure 1). The external surface is irregular with jagged ridges and the spaces between the ridges have a smooth texture. Ostia and oscula (incurrent and excurrent pores) are not visible to the naked eye.

Xestospongia muta occurs on many reefs throughout the Caribbean, from the Florida Keys and northern South America to the Lesser Antilles and Trinidad and Tobago. Population densities may reach 0.28 sponge/m², with demographics on an upward trend in the Florida Keys (McMurray *et al.*, 2011). With the large biomass and frequent abundance, this sponge species provides key ecosystem services. For example, at reefs where stony coral populations have severely deteriorated, the



Figure 1. This photograph of a representative *Xestospongia muta* (odXesMuta1) was taken by Brian K Walker on 22 May 2024 at 25 m depth off of Dr. Von D. Mizell-Eula Johnson State Park in Dania Beach, FL.

barrel sponge provides shelter, rugosity, and overall topography. *X. muta* also contributes bioavailable inorganic nitrogen as nitrate to the water column (Southwell *et al.*, 2008) and may serve as harbingers for reef health, having experienced their own specific marine disease outbreaks (Angermeier *et al.*, 2011; McMurray *et al.*, 2011). This disease, termed sponge orange band (SOB) disease, on *X. muta* appeared in south Florida in 2012 (Mulheron, 2014). This event preceded by about two years the more devastating stony coral loss tissue disease (SCTLD) which struck the same geographic region (Hayes *et al.*, 2022; Lopez, 2023; Rosales *et al.*, 2023).

With the sturdy skeleton of *X. muta* and the presence of spicules as physical defenses, the species may be less reliant on chemical defences. Still, characteristic compounds, including unsaturated polyacetylenic brominated acids (Patil *et al.*, 1992), the sterols xestosterol and mutasterol (Li *et al.*, 1981), the chiral tetrahydrofurans, and the mutafurans A–G with antifungal activities were isolated from *X. muta* (Morinaka *et al.*, 2007). In terms of reproduction, the giant barrel sponge is dioecious, but not sexually dimorphic in the sense that the separate male and female sponges cannot be distinguished from one another. *X. muta* spawns its egg or sperm directly into the water column. This conspicuous event catches divers’ eyes and has been appropriately dubbed “smoking sponges”. Annual, lunar and diel cycles all play a role in the heavily synchronized spawning events (Neely & Butler, 2020).

Xestospongia muta possesses a dense and diverse symbiotic community that is characteristic of high microbial abundance (HMA) sponges (Fiore *et al.*, 2013; Slaby *et al.*, 2019; Thomas *et al.*, 2016). Fiore *et al.* (2015) carried out one of the first transcriptome studies of *X. muta* that revealed active metabolic expression of multiple nitrogen transformation genes (e.g. archaeal *amoA*) and further identified many active bacterial lineages of the HMA characteristic prokaryotic phyla Gammaproteobacteria, Chloroflexi, Poribacteria, Actinobacteria (in particular, Acidimicrobiales) and Archaea (Nitrosopumilaceae). Cyanobacteria of the *Synechococcus* clade are present inter- and intracellularly in the peripheral sponge tissues, giving the animal its characteristic reddish-brown colouration (Erwin & Thacker, 2007). Sponge microbiomes may also serve as sensitive indicators for marine environmental stressors such as ocean acidification and climate change (Lesser *et al.*, 2016; Morrow *et al.*, 2016). Thus, additional genomic information for the microbiomes as well as the host are critical for improving our understanding of the ecology and evolution of sponges. The high-quality chromosomal genome sequence of the Florida *X. muta* and its Australian sister species *X. berquistia* (GCA_963965975.1) now forms a stable foundation for further studies on this charismatic sponge species’ evolution, ecology and underpinnings of sponge-microbe symbiosis.

Genome sequence report

Sequencing data

The genome of a specimen of *Xestospongia muta* (Figure 1) was sequenced using Pacific Biosciences single-molecule HiFi long reads, generating 83.29 Gb (gigabases) from 6.72 million

reads. Based on the estimated genome size, the sequencing data provided approximately 72.0x coverage of the genome. Chromosome conformation Hi-C data produced 103.39 Gb from 684.73 million reads. [Table 1](#) summarises the specimen and sequencing information.

Assembly statistics

The primary haplotype was assembled, and contigs corresponding to an alternate haplotype were also deposited in INSDC databases. The assembly was improved by manual curation, which corrected 84 misjoins or missing joins and removed 28 haplotypic duplications. These interventions reduced the total assembly length by 4.11% and decreased the scaffold count by 15.3%. The final assembly has a total length of 158.52 Mb in 298 scaffolds, with 186 gaps, and a scaffold N50 of 10.61 Mb ([Table 2](#)).

The snail plot in [Figure 2](#) provides a summary of the assembly statistics, indicating the distribution of scaffold lengths and other assembly metrics. [Figure 3](#) shows the distribution of scaffolds by GC proportion and coverage. [Figure 4](#) presents a cumulative assembly plot, with separate curves representing different scaffold subsets assigned to various phyla, illustrating the completeness of the assembly.

Most of the assembly sequence (99.57%) was assigned to 15 chromosomal-level scaffolds. These chromosome-level scaffolds, confirmed by Hi-C data, are named according to size ([Figure 5](#);

[Table 3](#)). During curation, it was noted that the exact order of the contigs in the repetitive region of chromosome 2 (16.1–18.9 Mb) is unknown.

The mitochondrial genome was also assembled. This sequence is included as a contig in the multifasta file of the genome submission and as a standalone record in GenBank.

Assembly quality metrics

The primary haplotype has a QV of 56.3, and the combined primary and alternate assemblies achieve an estimated QV of 49.7. The *k*-mer completeness for the primary haplotype is 58.30%, and for the alternate haplotype it is 85.40%. The combined primary and alternate assemblies achieve a *k*-mer completeness of 90.37%. BUSCO v.5.5.0 analysis using the metazoa_odb10 reference set (*n* = 954) indicated a completeness score of 83.7% (single = 82.4 %, duplicated = 1.4%).

Metagenome report

Fifty-five binned genomes were recovered from the metagenome assembly ([Figure 6](#)), of which 17 were classified as high-quality metagenome-assembled genomes (MAGs) (see Methods). Completeness of the bins ranged from approximately 40% to 99%, with contamination levels below 10%. Detailed information on all binned genomes is available via FigShare (<https://doi.org/10.6084/m9.figshare.28750223>). A cladogram is shown in [Figure 7](#).

Table 1. Specimen and sequencing data for *Xestospongia muta*.

Project information			
Study title	Xestospongia muta (giant barrel sponge)		
Umbrella BioProject	PRJEB63657		
Species	<i>Xestospongia muta</i>		
BioSpecimen	SAMEA9267624		
NCBI taxonomy ID	178552		
Specimen information			
Technology	ToLID	BioSample accession	Organism part
PacBio long read sequencing	odXesMuta1	SAMEA9267634	somatic animal tissue
Hi-C sequencing	odXesMuta1	SAMEA9267633	somatic animal tissue
RNA sequencing	odXesMuta3	SAMEA111491447	somatic animal tissue
Sequencing information			
Platform	Run accession	Read count	Base count (Gb)
Hi-C Illumina NovaSeq 6000	ERR12321227	6.85e+08	103.39
PacBio Sequel IIe	ERR11641085	1.85e+06	14.43
PacBio Sequel IIe	ERR11641084	4.87e+06	68.86
RNA Illumina NovaSeq X	ERR13669939	5.76e+06	0.87

Table 2. Genome assembly data for *Xestospongia muta*.

Genome assembly	
Assembly name	odXesMuta1.1
Assembly accession	GCA_963693285.1
Alternate haplotype accession	GCA_963693275.1
Assembly level for primary assembly	chromosome
Span (Mb)	158.52
Number of contigs	484
Number of scaffolds	298
Longest scaffold (Mb)	35.36
Assembly metric	Measure
Contig N50 length	1.87 Mb
Scaffold N50 length	10.61 Mb
Consensus quality (QV)	Primary: 56.3; alternate: 49.1; combined 49.7
BUSCO*	C:83.7%[S:82.4%,D:1.4%],F:5.7%,M:10.6%,n:954
Percentage of assembly mapped to chromosomes	99.57%
Organelles	Mitochondrial genome: 18.99 kb

* BUSCO scores based on the metazoa_odb10 BUSCO set using version 5.5.0. C = complete [S = single copy, D = duplicated], F = fragmented, M = missing, n = number of orthologues in comparison.

Genome annotation report

The *Xestospongia muta* genome assembly (GCA_963693285.1) was annotated at the European Bioinformatics Institute (EBI) on Ensembl Rapid Release. The resulting annotation includes 20,364 transcribed mRNAs from 20,220 protein-coding genes (Table 2; <https://beta.ensembl.org/species/24c58870-89c9-4aa6-84f1-7cae84c57ec>). The average transcript length is 2,765.34, with an average of 6.89 exons per transcript. An alternative annotation is available here:

An alternative annotation for the genome of *Xestospongia muta* performed at the WSI using BRAKER3 produced 24,448 protein-coding genes. This annotation is provided as UCSC assembly hubs and raw downloads at https://github.com/Aquatic-Symbiosis-Genomics-Project/sponge_annotations/tree/main/results/odXesMuta1.

Methods

Sample acquisition

Sponge samples were collected by SCUBA at Gulfstream Reef, Boynton Beach FL (latitude 26.51, longitude -80.03) under Florida Fish and Wildlife Conservation Commission fishing licenses and Special Activity License-12-1372-37a. Samples were immediately snap frozen in dry ice baths or stored in TRIZOL preservative before transport to the laboratory at Nova Southeastern University (Mulheron, 2014). One of the *X. muta* samples for genome sequencing stemmed from the time frame

of the SOB outbreak mentioned in the background section, although it was a healthy individual with no signs of disease (Lopez, 2023). The specimen with ID NSU0023301 (ToLID odXesMuta1) was used for PacBio HiFi sequencing.

The specimen used for RNA sequencing (specimen ID NSU0023303, ToLID odXesMuta3) was collected from the mesophotic reef east of South Carysfort Reef, Key Largo, Florida (latitude 23.25, longitude -80.19) on 2019-08-13 under Florida Keys National Marine Sanctuary Permit FKNMS-2018-175 (to SP). The specimen was collected and identified by Shirley Pomponi (Florida Atlantic University). A fragment of the sponge was dissociated into single cells (containing both sponge cells and symbionts) and cryopreserved using established methods (Munroe *et al.*, 2018). Cryopreserved cells were thawed rapidly in a 50°C water bath to minimize ice crystal damage to the cells. The cell suspension was rinsed twice by centrifugation at 4000 × rpm for 5 minutes and resuspended in filtered artificial seawater (Conkling *et al.*, 2019). After the final rinse, the supernatant was decanted and the cell pellet was submitted for sequencing.

Nucleic acid extraction

The workflow for high molecular weight (HMW) DNA extraction at the Wellcome Sanger Institute (WSI) Tree of Life Core Laboratory includes a sequence of procedures: sample preparation and homogenisation, DNA extraction, fragmentation and

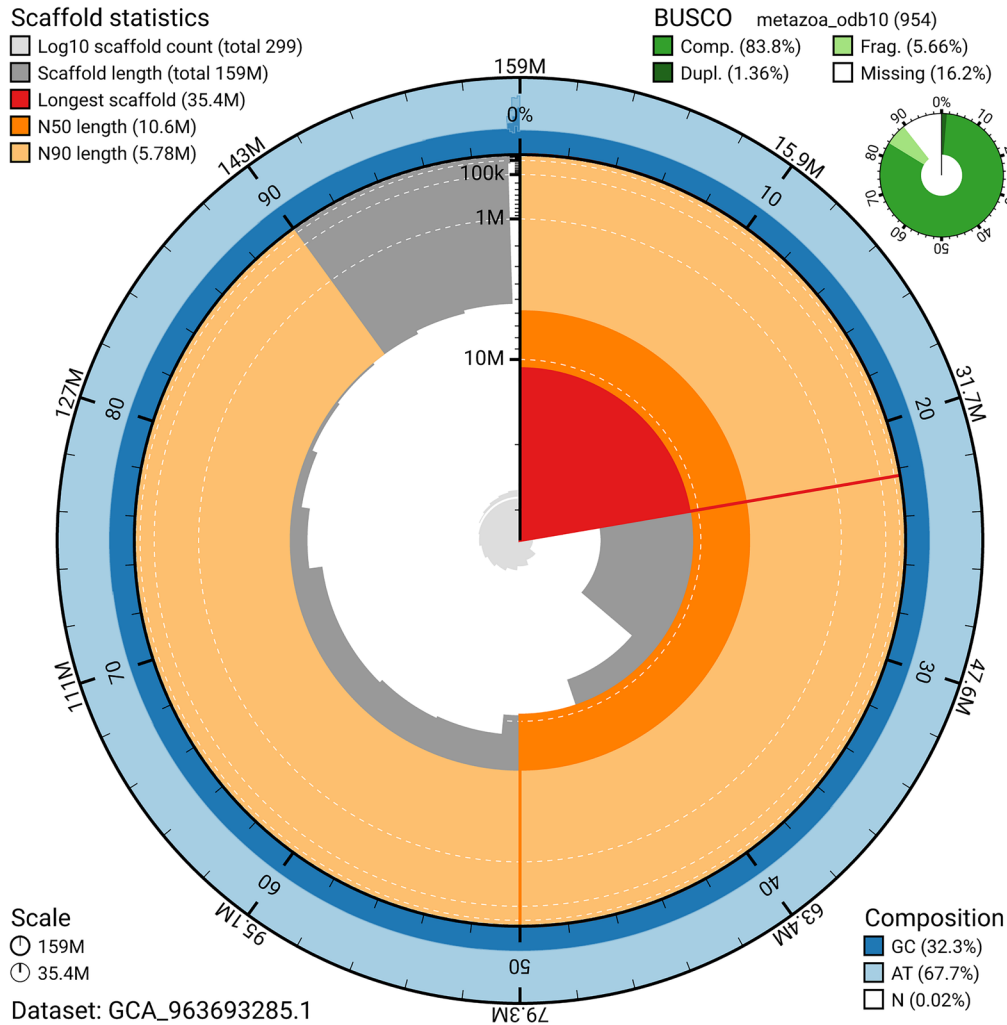


Figure 2. Genome assembly of *Xestospongia muta*, odXesMuta1.1: metrics. The BlobToolKit snail plot provides an overview of assembly metrics and BUSCO gene completeness. The circumference represents the length of the whole genome sequence, and the main plot is divided into 1,000 bins around the circumference. The outermost blue tracks display the distribution of GC, AT, and N percentages across the bins. Scaffolds are arranged clockwise from longest to shortest and are depicted in dark grey. The longest scaffold is indicated by the red arc, and the deeper orange and pale orange arcs represent the N50 and N90 lengths. A light grey spiral at the centre shows the cumulative scaffold count on a logarithmic scale. A summary of complete, fragmented, duplicated, and missing BUSCO genes in the metazoa_odb10 set is presented at the top right. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/GCA_963693285.1/dataset/GCA_963693285.1/snail.

purification. Detailed protocols are available on protocols.io (Denton *et al.*, 2023). The odXesMuta1 sample was prepared for DNA extraction by weighing and dissecting it on dry ice (Jay *et al.*, 2023). The tissue was cryogenically disrupted using the Covaris cryoPREP® Automated Dry Pulverizer (Narváez-Gómez *et al.*, 2023). Prior to DNA extraction, the sponge sample was bathed in “L buffer” (10 mM Tris, pH 7.6, 100 mM EDTA, 20 mM NaCl), minced into small pieces using a scalpel and the cellular interior separated from the mesohyl using forceps (Lopez, 2022). HMW DNA was extracted using the Automated MagAttract v2 protocol (Oatley *et al.*, 2023a). For ULI PacBio sequencing, DNA was fragmented using the Covaris g-TUBE method (Oatley *et al.*, 2023c). Sheared DNA

was purified by solid-phase reversible immobilisation, using AMPure PB beads to eliminate shorter fragments and concentrate the DNA (Oatley *et al.*, 2023b). The concentration of the sheared and purified DNA was assessed using a Nanodrop spectrophotometer and Qubit Fluorometer using the Qubit dsDNA High Sensitivity Assay kit. Fragment size distribution was evaluated by running the sample on the FemtoPulse system.

RNA was extracted from tissue of odXesMuta3 in the Tree of Life Laboratory at the WSI using the RNA Extraction: Automated MagMax™ mirVana protocol (do Amaral *et al.*, 2023). The RNA concentration was assessed using a Nanodrop spectrophotometer and a Qubit Fluorometer using the

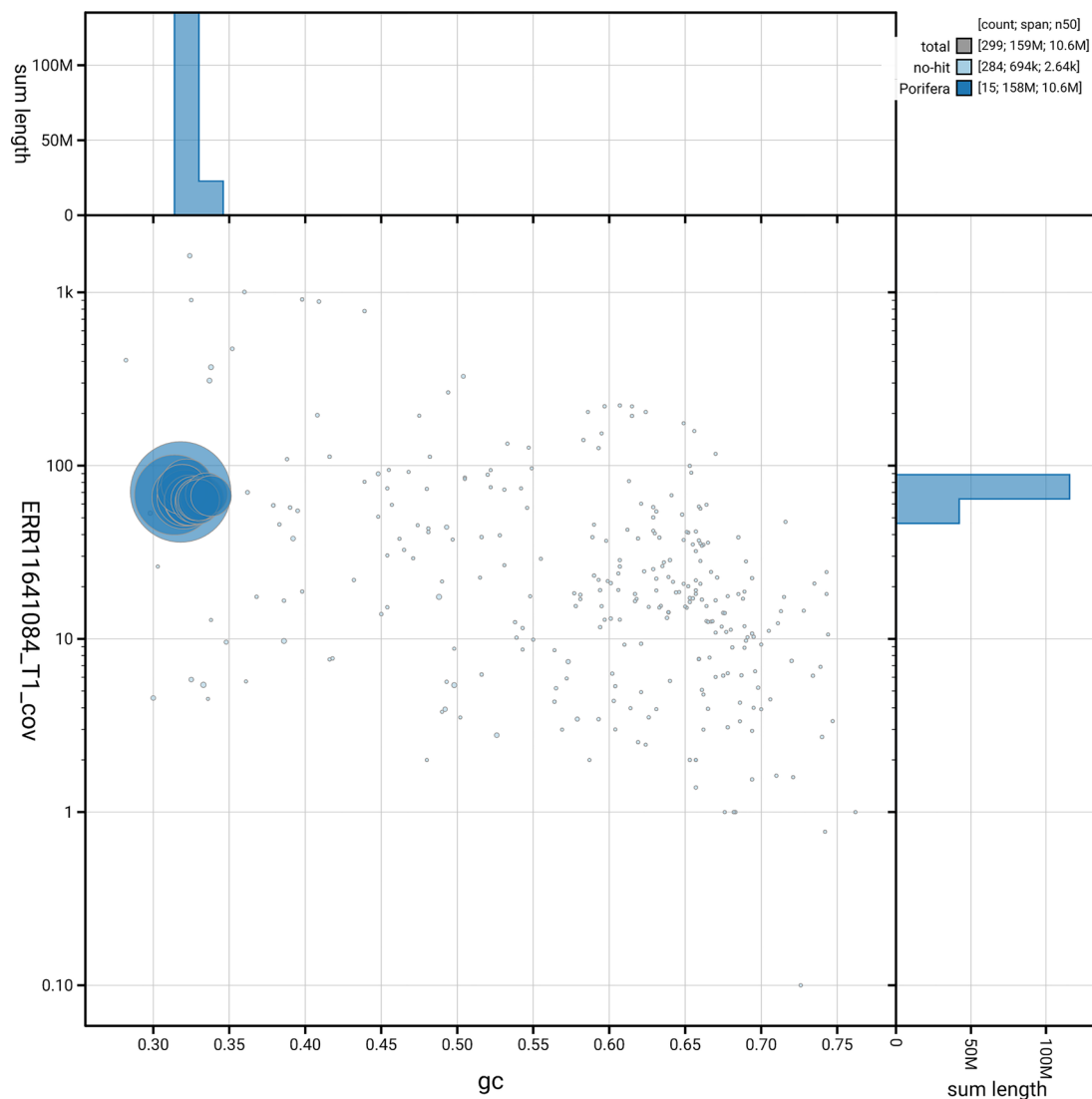


Figure 3. Genome assembly of *Xestospongia muta*, odXesMuta1.1: BlobToolKit GC-coverage plot. Blob plot showing sequence coverage (vertical axis) and GC content (horizontal axis). The circles represent scaffolds, with the size proportional to scaffold length and the colour representing phylum membership. The histograms along the axes display the total length of sequences distributed across different levels of coverage and GC content. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/GCA_963693285.1/dataset/GCA_963693285.1/blob.

Qubit RNA Broad-Range Assay kit. Analysis of the integrity of the RNA was done using the Agilent RNA 6000 Pico Kit and Eukaryotic Total RNA assay.

Sequencing

Library preparation and sequencing was performed at the WSI Scientific Operations core. Pacific Biosciences HiFi circular consensus DNA sequencing libraries were constructed according to the manufacturers' instructions for ULI sequencing. Samples were sequenced on a Sequel IIe instrument (Pacific Biosciences, California, USA). Hi-C data were also generated from tissue of odXesMuta1 using the Arima2 kit and sequenced

on the Illumina NovaSeq 6000 instrument. Poly(A) RNA-Seq libraries were constructed using the NEB Ultra II RNA Library Prep kit. RNA sequencing was performed on an Illumina NovaSeq X instrument.

Genome assembly, curation and evaluation

Assembly

Prior to assembly of the PacBio HiFi reads, a database of k -mer counts ($k = 31$) was generated from the filtered reads using FastK. GenomeScope2 (Ranallo-Benavidez *et al.*, 2020) was used to analyse the k -mer frequency distributions, providing estimates of genome size, heterozygosity, and repeat content.

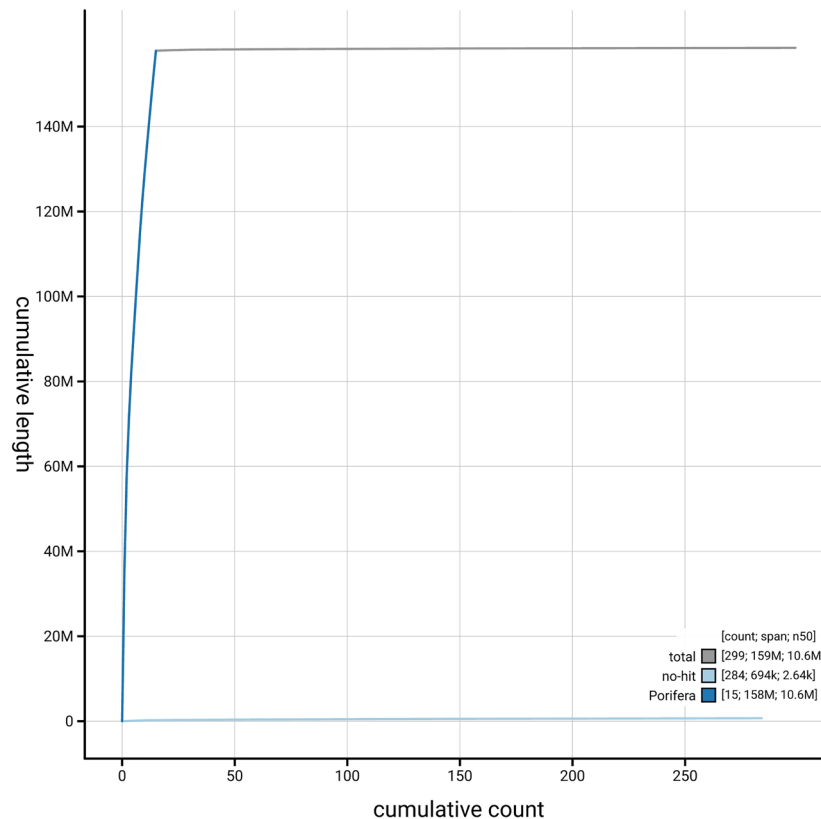


Figure 4. Genome assembly of *Xestospongia muta*, odXesMuta1.1: BlobToolKit cumulative sequence plot. The grey line shows cumulative length for all scaffolds. Coloured lines show cumulative lengths of scaffolds assigned to each phylum using the buscogenes taxrule. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/GCA_963693285.1/dataset/GCA_963693285.1/cumulative.

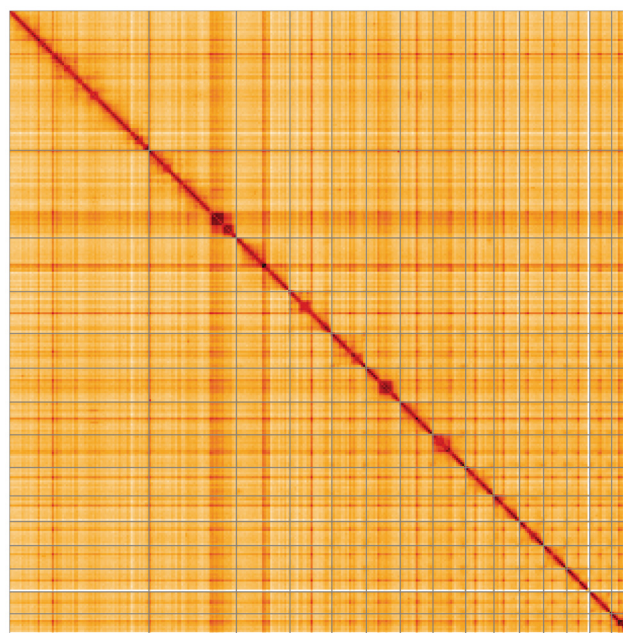


Figure 5. Genome assembly of *Xestospongia muta*: Hi-C contact map of the odXesMuta1.1 assembly, visualised using HiGlass. Chromosomes are shown in order of size from left to right and top to bottom. An interactive version of this figure may be viewed at <https://genome-note-higlass.tol.sanger.ac.uk/l/?d=Kkd4NOCzRpiniFavd6eOEg>.

Table 3. Chromosomal pseudomolecules in the genome assembly of *Xestospongia muta*, odXesMuta1.

INSDC accession	Name	Length (Mb)	GC%
OY848917.1	1	35.36	32
OY848918.1	2	22.15	31.5
OY848919.1	3	13.58	32
OY848920.1	4	10.61	32
OY848921.1	5	8.75	32.5
OY848922.1	6	8.58	32
OY848923.1	7	8.28	32.5

INSDC accession	Name	Length (Mb)	GC%
OY848924.1	8	8.28	32
OY848925.1	9	7.16	32.5
OY848926.1	10	6.49	33
OY848927.1	11	6.08	33
OY848928.1	12	5.92	33
OY848929.1	13	5.78	33.5
OY848930.1	14	5.53	33
OY848931.1	15	5.31	34
OY848932.1	MT	0.02	34

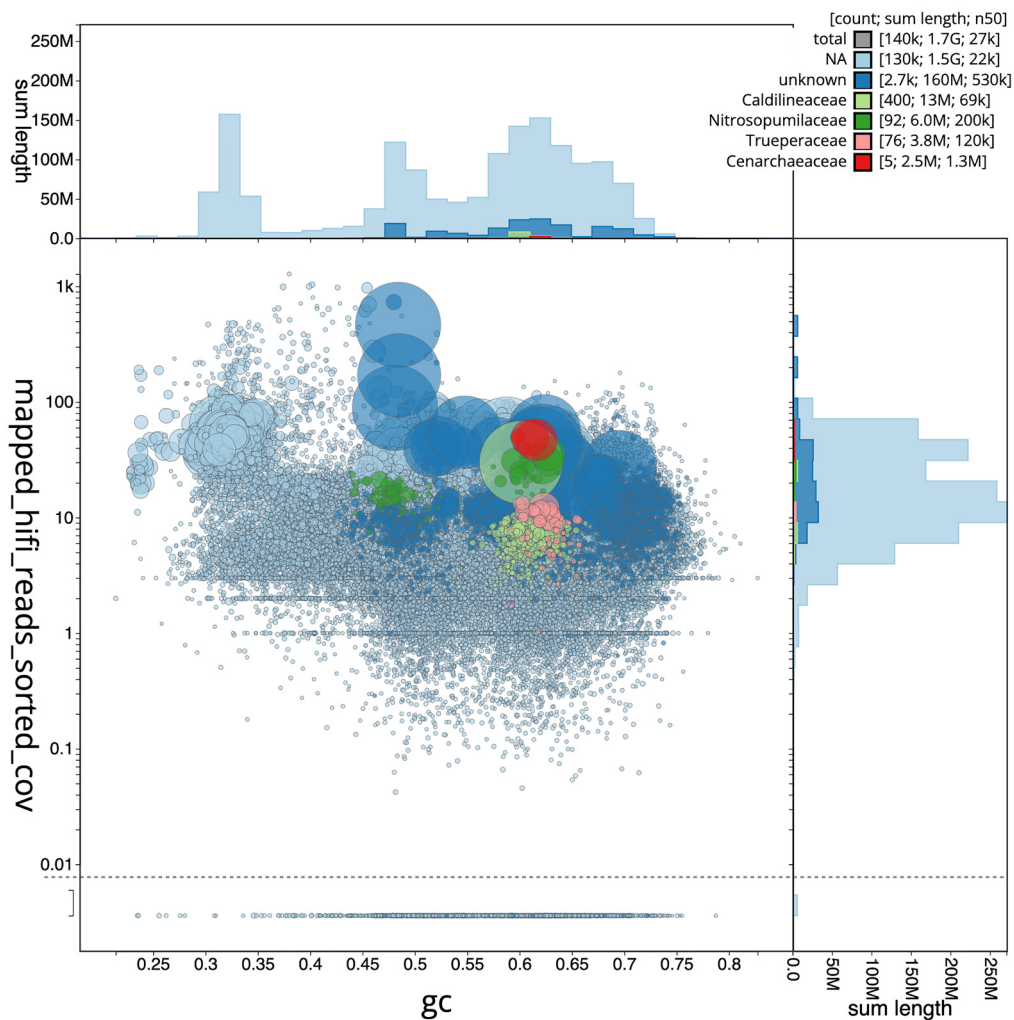


Figure 6. Blob plot of base coverage in mapped against GC proportion for sequences in the *Xestospongia muta* metagenome. Binned contigs are coloured by family. Circles are sized in proportion to sequence length on a square root scale, ranging from 505 to 5,542,969. Histograms show the distribution of sequence length sum along each axis. An interactive version of this figure may be viewed [here](#).

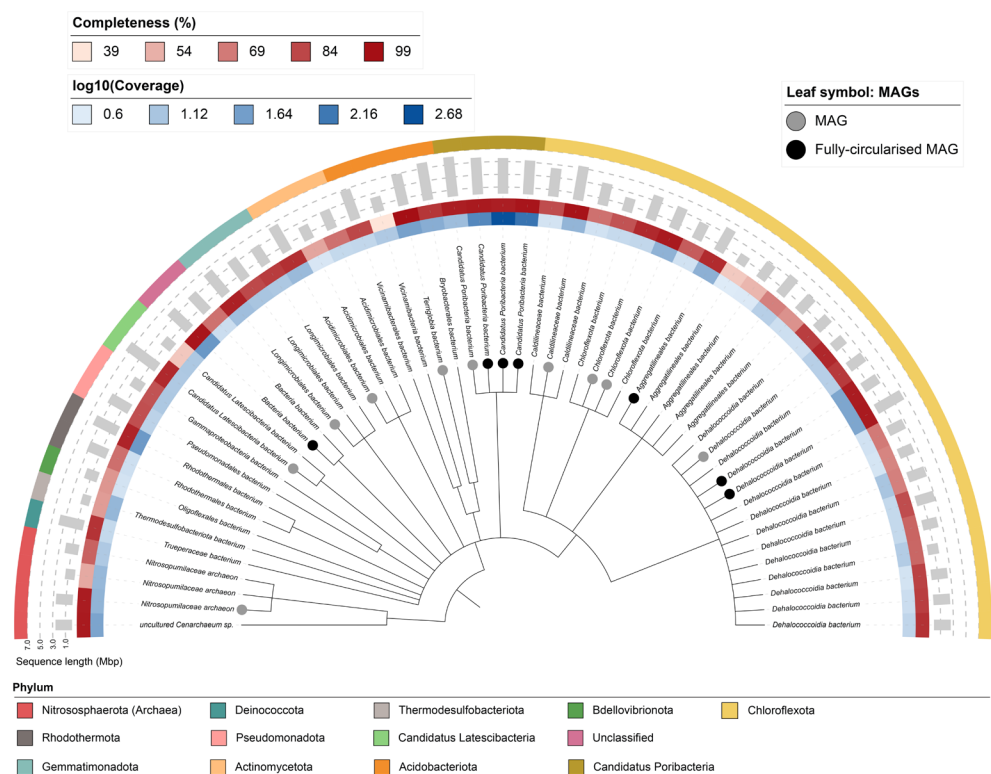


Figure 7. Cladogram showing the taxonomic placement of metagenome bins, constructed using NCBI taxonomic identifiers with *taxonomizr* and annotated in *iTOL*. Colours indicate phylum-level taxonomy. Additional tracks show sequencing coverage ($\log_{10}$), estimated genome size (Mbp), and completeness. Bins that meet the criteria for MAGs are marked with a grey circle; those that are fully circularised MAGs are marked in black.

The HiFi reads were assembled using Hifiasm (Cheng *et al.*, 2021) with the `--primary` option. Haplotypic duplications were identified and removed using `purge_dups` (Guan *et al.*, 2020). The Hi-C reads were mapped to the primary contigs using `bwa-mem2` (Vasimuddin *et al.*, 2019). The contigs were further scaffolded using the provided Hi-C data (Rao *et al.*, 2014) in YaHS (Zhou *et al.*, 2023) using the `--break` option for handling potential misassemblies. The scaffolded assemblies were evaluated using Gfastats (Formenti *et al.*, 2022), BUSCO (Manni *et al.*, 2021) and MERQURY.FK (Rhie *et al.*, 2020).

The mitochondrial genome was assembled using MitoHiFi (Uliano-Silva *et al.*, 2023), which runs MitoFinder (Allio *et al.*, 2020) and uses these annotations to select the final mitochondrial contig and to ensure the general quality of the sequence.

Taxonomic verification

Barrel sponges consist of intertwined species complexes in different ocean basins (see e.g. Swierts *et al.*, 2017). *Xestospongia muta* therefore comprise the barrel sponges in the tropical Atlantic, whose monophyletic relationship has not been shown yet. Barrel sponges are currently divided into mitochondrial haplogroups, based on CO1 (I3-M11 fragment) and ATP6. We created separate alignment for both fragments using all published *Xestospongia* spp. sequences of Swierts *et al.*

(Swierts *et al.*, 2017) with using MAFFT v7.450 (Katoh & Standley, 2013). The current specimen falls into the haplogroup C9A5 sensu, which is currently known from tropical Atlantic barrel sponge specimens only.

Assembly curation

The assembly was decontaminated using the Assembly Screen for Cobionts and Contaminants (ASCC) pipeline. Flat files and maps used in curation were generated via the TreeVal pipeline (Pointon *et al.*, 2023). Manual curation was conducted primarily in PretextView (Harry, 2022) and HiGlass (Kerpedjiev *et al.*, 2018), with additional insights provided by JBrowse2 (Diesh *et al.*, 2023). Scaffolds were visually inspected and corrected as described by Howe *et al.* (2021). Any identified contamination, missed joins, and mis-joins were amended, and duplicate sequences were tagged and removed. The curation process is documented at <https://gitlab.com/wtsi-grit/rapid-curation>.

Assembly quality assessment

The Merqury.FK tool (Rhie *et al.*, 2020), run in a Singularity container (Kurtzer *et al.*, 2017), was used to evaluate *k*-mer completeness and assembly quality for the primary and alternate haplotypes using the *k*-mer databases ($k = 31$) that were computed prior to genome assembly. The analysis outputs included assembly QV scores and completeness statistics.

A Hi-C contact map was produced for the final version of the assembly. The Hi-C reads were aligned using bwa-mem2 (Vasimuddin *et al.*, 2019) and the alignment files were combined using SAMtools (Danecek *et al.*, 2021). The Hi-C alignments were converted into a contact map using BEDTools (Quinlan & Hall, 2010) and the Cooler tool suite (Abdennur & Mirny, 2020). The contact map is visualised in HiGlass (Kerpedjiev *et al.*, 2018).

The blobtoolkit pipeline is a Nextflow port of the previous Snakemake Blobtoolkit pipeline (Challis *et al.*, 2020). It aligns the PacBio reads in SAMtools and minimap2 (Li, 2018) and generates coverage tracks for regions of fixed size. In parallel, it queries the GoaT database (Challis *et al.*, 2023) to identify all matching BUSCO lineages to run BUSCO (Manni *et al.*, 2021). For the three domain-level BUSCO lineages, the pipeline aligns the BUSCO genes to the UniProt Reference Proteomes database (Bateman *et al.*, 2023) with DIAMOND blastp (Buchfink *et al.*, 2021). The genome is also divided into chunks according to the density of the BUSCO genes from the closest taxonomic lineage, and each chunk is aligned to the UniProt Reference Proteomes database using DIAMOND blastx. Genome sequences without a hit are chunked using seqtk and aligned to the NT database with blastn (Altschul *et al.*, 1990). The blobtools suite combines all these outputs into a blobdir for visualisation.

The blobtoolkit pipeline was developed using nf-core tooling (Ewels *et al.*, 2020) and MultiQC (Ewels *et al.*, 2016), relying on the Conda package manager, the Bioconda initiative (Grüning *et al.*, 2018), the Biocontainers infrastructure (da Veiga Leprevost *et al.*, 2017), as well as the Docker (Merkel, 2014) and Singularity (Kurtzer *et al.*, 2017) containerisation solutions.

Table 4 contains a list of relevant software tool versions and sources.

Metagenome assembly

The metagenome assembly was generated using metaMDBG (Benoit *et al.*, 2024) and binned using MetaBAT2 (Kang *et al.*, 2019), MaxBin (Wu *et al.*, 2014), bin3C (DeMaere & Darling, 2019), and MetaTOR. The resulting bin sets of each binning algorithm were optimised and refined using MAGScoT (Rühlemann *et al.*, 2022). PROKKA (Seemann, 2014) was used to identify tRNAs and rRNAs in each bin, CheckM (Parks *et al.*, 2015) (checkM_DB release 2015-01-16) was used to assess bin completeness/contamination, and GTDB-TK (Chaumeil *et al.*, 2022) (GTDB release 214) was used to taxonomically classify bins. Taxonomic replicate bins were identified using dRep (Olm *et al.*, 2017), with default settings (95% ANI threshold). The final bin set was filtered for bacteria and archaea. All bins were assessed for quality and categorised

Table 4. Software tools: versions and sources.

Software tool	Version	Source
BEDTools	2.30.0	https://github.com/arq5x/bedtools2
bin3C	0.3.3	https://github.com/cerebis/bin3C
BLAST	2.14.0	ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/
BlobToolKit	4.3.3	https://github.com/blobtoolkit/blobtoolkit
BUSCO	5.5.0	https://gitlab.com/ezlab/busco
bwa-mem2	2.2.1	https://github.com/bwa-mem2/bwa-mem2
CheckM	1.2.1	https://github.com/Ecogenomics/CheckM
Cooler	0.8.11	https://github.com/open2c/cooler
DIAMOND	2.1.8	https://github.com/bbuchfink/diamond
dRep	3.4.0	https://github.com/MrOlm/drep
fasta_windows	0.2.4	https://github.com/tolkit/fasta_windows
FastK	427104ea91c78c3b8b8b49f1a7d6bbeaa869ba1c	https://github.com/thegenemyers/FASTK
Gfastats	1.3.6	https://github.com/vgl-hub/gfastats
GoaT CLI	0.2.5	https://github.com/genomehubs/goat-cli
GTDB-TK	2.3.2	https://github.com/Ecogenomics/GTDBTk

Software tool	Version	Source
Hifiasm	0.16.1	https://github.com/chhylp123/hifiasm
HiGlass	44086069ee7d4d3f6f3f0012569789ec138f42b84aa44357826c0b6753eb28de	https://github.com/higlass/higlass
MAGScoT	1.0.0	https://github.com/ikmb/MAGScoT
MaxBin	2.7	https://sourceforge.net/projects/maxbin/
MerquryFK	d00d98157618f4e8d1a9190026b19b471055b22e	https://github.com/thegenemyers/MERQURY.FK
MetaBat2	2.15-15-gd6ea400	https://bitbucket.org/berkeleylab/metabat/src/master/
MetaTOR	-	https://github.com/koszullab/metaTOR
Minimap2	2.24-r1122	https://github.com/lh3/minimap2
MitoHiFi	2	https://github.com/marcelauliano/MitoHiFi
MultiQC	1.14, 1.17, and 1.18	https://github.com/MultiQC/MultiQC
Nextflow	23.04.1	https://github.com/nextflow-io/nextflow
PretextView	0.2.5	https://github.com/sanger-tol/PretextView
PROKKA	1.14.5	https://github.com/vdeJager/prokka
purge_dups	1.2.3	https://github.com/dfguan/purge_dups
samtools	1.18	https://github.com/samtools/samtools
sanger-tol/ascc	-	https://github.com/sanger-tol/ascc
sanger-tol/blobtoolkit	0.3.0	https://github.com/sanger-tol/blobtoolkit
Seqtk	1.3	https://github.com/lh3/seqtk
Singularity	3.9.0	https://github.com/sylabs/singularity
TreeVal	1.2.0	https://github.com/sanger-tol/treeval
YaHS	1.1a.2	https://github.com/c-zhou/yahs

as metagenome-assembled genomes (MAGs) if they met the following criteria: contamination $\leq 5\%$, presence of 5S, 16S, and 23S rRNA genes, at least 18 unique tRNAs, and either $\geq 90\%$ completeness or $\geq 50\%$ completeness with fully circularised chromosomes. Bins that did not meet these thresholds, or were identified as taxonomic replicates of MAGs, were retained as ‘binned metagenomes’ provided they had $\geq 50\%$ completeness and $\leq 10\%$ contamination. A cladogram based on NCBI taxonomic assignments was generated using the ‘taxonomizr’ package in R. The tree was visualised and annotated using iTOL (Letunic & Bork, 2024). Software tool versions and sources are given in Table 4.

Host genome annotation

The BRAKER2 pipeline (Brůna *et al.*, 2021) was used in the default protein mode to generate annotation for the *Xestospongia muta* assembly (GCA_963693285.1) in Ensembl Rapid Release at the EBI.

Wellcome Sanger Institute – Legal and Governance

The materials that have contributed to this genome note have been supplied by a Tree of Life collaborator. The Wellcome Sanger Institute employs a process whereby due diligence is carried out proportionate to the nature of the materials themselves, and the circumstances under which they have been/are to be collected and provided for use. The purpose of this is to address and mitigate any potential legal and/or ethical implications of receipt and use of the materials as part of the research project, and to ensure that in doing so we align with best practice wherever possible. The overarching areas of consideration are:

- Ethical review of provenance and sourcing of the material
- Legality of collection, transfer and use (national and international)

Each transfer of samples is undertaken according to a Research Collaboration Agreement or Material Transfer Agreement entered into by the Tree of Life collaborator, Genome Research Limited (operating as the Wellcome Sanger Institute) and in some circumstances other Tree of Life collaborators.

Data availability

European Nucleotide Archive: *Xestospongia muta* (giant barrel sponge). Accession number PRJEB63657; <https://identifiers.org/ena.embl/PRJEB63657>. The genome sequence is released openly for reuse. The *Xestospongia muta* genome sequencing initiative is part of the Aquatics Symbiosis Genomics (ASG) project. All raw sequence data and the assembly have been deposited in INSDC databases. The genome will be annotated using available RNA-Seq data and presented through the [Ensembl](#) pipeline at the European Bioinformatics Institute. Raw data and assembly accession identifiers are reported in [Table 1](#) and [Table 2](#).

Author information

Members of the Wellcome Sanger Institute Tree of Life Management, Samples and Laboratory Team are listed here: <https://doi.org/10.5281/zenodo.10066175>.

Members of the Wellcome Sanger Institute Scientific Operations: Sequencing Operations are listed here: <https://doi.org/10.5281/zenodo.10043364>.

Members of the Wellcome Sanger Institute Tree of Life Core Informatics team are listed here: <https://doi.org/10.5281/zenodo.10066637>.

Members of the European Bioinformatics Institute ASG Data Portal team are listed here: <https://doi.org/10.5281/zenodo.10076466>.

Members of the Wellcome Sanger Institute/Aquatic Symbiosis Genomics Project Leadership are listed here: <https://doi.org/10.5281/zenodo.10184833>.

References

- Abdennur N, Mirny LA: **Cooler: scalable storage for Hi-C data and other genomically labeled arrays.** *Bioinformatics.* 2020; **36**(1): 311–316. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Allio R, Schomaker-Bastos A, Romiguier J, et al.: **MitoFinder: efficient automated large-scale extraction of mitogenomic data in target enrichment phylogenomics.** *Mol Ecol Resour.* 2020; **20**(4): 892–905. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Altschul SF, Gish W, Miller W, et al.: **Basic Local Alignment Search Tool.** *J Mol Biol.* 1990; **215**(3): 403–410. [PubMed Abstract](#) | [Publisher Full Text](#)
- Angermeier H, Kamke J, Abdelmohsen UR, et al.: **The pathology of sponge orange band disease affecting the Caribbean barrel sponge *Xestospongia muta*.** *FEMS Microbiol Ecol.* 2011; **75**(2): 218–230. [PubMed Abstract](#) | [Publisher Full Text](#)
- Bateman A, Martin MJ, Orchard S, et al.: **UniProt: the universal protein knowledgebase in 2023.** *Nucleic Acids Res.* 2023; **51**(D1): D523–D531. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Benoit G, Raguideau S, James R, et al.: **High-quality metagenome assembly from long accurate reads with metaMDBG.** *Nat Biotechnol.* 2024; **42**(9): 1378–1383. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Brúna T, Hoff KJ, Lomsadze A, et al.: **BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database.** *NAR Genom Bioinform.* 2021; **3**(1): lqaa108. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Buchfink B, Reuter K, Drost HG: **Sensitive protein alignments at Tree-of-Life scale using DIAMOND.** *Nat Methods.* 2021; **18**(4): 366–368. [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](#)
- Challis R, Richards E, Rajan J, et al.: **BlobToolKit – interactive quality assessment of genome assemblies.** *G3 (Bethesda).* 2020; **10**(4): 1361–1374. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Chaumeil PA, Mussig AJ, Hugenholtz P, et al.: **GTDB-Tk v2: memory friendly classification with the genome taxonomy database.** *Bioinformatics.* 2022; **38**(23): 5315–5316. [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](#)
- Conkling M, Hesp K, Munroe S, et al.: **Breakthrough in marine invertebrate cell culture: sponge cells divide rapidly in improved nutrient medium.** *Sci Rep.* 2019; **9**(1): 17321. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- da Veiga Leprevost F, Grüning BA, Alves Aflitos S, et al.: **BioContainers: an open-source and community-driven framework for software standardization.** *Bioinformatics.* 2017; **33**(16): 2580–2582. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Danecek P, Bonfield JK, Liddle J, et al.: **Twelve years of SAMtools and BCFtools.** *GigaScience.* 2021; **10**(2): giab008. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- DeMaere MZ, Darling AE: **bin3C: exploiting Hi-C sequencing data to accurately resolve metagenome-assembled genomes.** *Genome Biol.* 2019; **20**(1): 46. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Denton A, Yatsenko H, Jay J, et al.: **Sanger Tree of Life wet laboratory protocol collection V.1.** *protocols.io.* 2023. [Publisher Full Text](#)
- Diesh C, Stevens GJ, Xie P, et al.: **JBrowse 2: a modular genome browser with views of synteny and structural variation.** *Genome Biol.* 2023; **24**(1): 74. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- do Amaral RJV, Bates A, Denton A, et al.: **Sanger Tree of Life RNA extraction: automated MagMax™ mirVana.** *protocols.io.* 2023. [Publisher Full Text](#)
- Erwin PM, Thacker RW: **Incidence and identity of photosynthetic symbionts in Caribbean coral reef sponge assemblages.** *J Mar Biol Assoc U K.* 2007; **87**(6): 1683–1692. [Publisher Full Text](#)
- Ewels P, Magnusson M, Lundin S, et al.: **MultiQC: summarize analysis results for multiple tools and samples in a single report.** *Bioinformatics.* 2016; **32**(19): 3047–3048. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Ewels PA, Peltzer A, Fillinger S, et al.: **The nf-core framework for community-curated bioinformatics pipelines.** *Nat Biotechnol.* 2020; **38**(3): 276–278. [PubMed Abstract](#) | [Publisher Full Text](#)
- Fiore CL, Jarett JK, Lesser MP: **Symbiotic prokaryotic communities from different populations of the giant barrel sponge, *Xestospongia muta*.** *MicrobiologyOpen.* 2013; **2**(6): 938–952. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Fiore CL, Labrie M, Jarett JK, et al.: **Transcriptional activity of the giant barrel sponge, *Xestospongia muta* holobiont: molecular evidence for metabolic interchange.** *Front Microbiol.* 2015; **6**: 364. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Formenti G, Abueg L, Brajuka A, et al.: **Gfastats: conversion, evaluation and**

manipulation of genome sequences using assembly graphs. *Bioinformatics*. 2022; **38**(17): 4214–4216.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Grüning B, Dale R, Sjödin A, *et al.*: **Bioconda: sustainable and comprehensive software distribution for the life sciences.** *Nat Methods*. 2018; **15**(7): 475–476.

[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](#)

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

[Reference Source](#)

Hayes NK, Walton CJ, Gilliam DS: **Tissue loss disease outbreak significantly alters the Southeast Florida stony coral assemblage.** *Front Mar Sci*. 2022; **9**: 975894.

[Publisher Full Text](#)

Howe K, Chow W, Collins J, *et al.*: **Significantly improving the quality of genome assemblies through curation.** *GigaScience*. 2021; **10**(1): g1aa153.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Jay J, Yatsenko H, Narváez-Gómez JP, *et al.*: **Sanger Tree of Life sample preparation: triage and dissection.** *protocols.io*. 2023.

[Publisher Full Text](#)

Kang DD, Li F, Kirton E, *et al.*: **MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies.** *PeerJ*. 2019; **7**: e7359.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Katoh K, Standley DM: **MAFFT multiple sequence alignment software version 7: improvements in performance and usability.** *Mol Biol Evol*. 2013; **30**(4): 772–80.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Kerpedjiev P, Abdennur N, Lekschas F, *et al.*: **HiGlass: web-based visual exploration and analysis of genome interaction maps.** *Genome Biol*. 2018; **19**(1): 125.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Kurtzer GM, Sochat V, Bauer MW: **Singularity: scientific containers for mobility of compute.** *PLoS One*. 2017; **12**(5): e0177459.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Lesser MP, Fiore C, Slattery M, *et al.*: **Climate change stressors destabilize the microbiome of the Caribbean barrel sponge, *Xestospongia muta*.** *J Exp Mar Biol Ecol*. 2016; **475**: 11–18.

[Publisher Full Text](#)

Letunic I, Bork P: **Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool.** *Nucleic Acids Res*. 2024; **52**(W1): W78–W82.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Li H: **Minimap2: pairwise alignment for nucleotide sequences.** *Bioinformatics*. 2018; **34**(18): 3094–3100.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Li LN, Sjostrand U, Djerassi C: **Minor and trace sterols in marine invertebrates. 19. Isolation, structure elucidation, and partial synthesis of 24-methylene-25-ethylcholesterol (mutasterol): first example of sterol side-chain bioalkylation at position.** *J Am Chem Soc*. 1981; **103**(1): 115–119.

[Publisher Full Text](#)

Lopez J: **Squeeze-enrichment of intact cells (eukaryotic and prokaryotic) from marine sponge tissues prior to routine DNA extraction.** *protocols.io*. 2022.

[Publisher Full Text](#)

Lopez JV: **Assessments and conservation of biological diversity from coral reefs to the deep sea.** London: Academic Press, 2023.

[Reference Source](#)

Manni M, Berkeley MR, Seppely 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](#)

McMurray SE, Blum JE, Leichter JJ, *et al.*: **Bleaching of the giant barrel sponge *Xestospongia muta* in the Florida Keys.** *Limnol Oceanogr*. 2011; **56**(6): 2243–2250.

[Publisher Full Text](#)

McMurray SE, Blum JE, Pawlik JR: **Redwood of the reef: growth and age of the giant barrel sponge *Xestospongia muta* in the Florida Keys.** *Mar Biol*. 2008; **155**(2): 159–171.

[Publisher Full Text](#)

Merkel D: **Docker: lightweight Linux containers for consistent development and deployment.** *Linux J*. 2014; **2014**(239): 2, [Accessed 2 April 2024].

[Reference Source](#)

Morinaka BI, Skepper CK, Molinski TF: **Ene-yne tetrahydrofurans from the sponge *Xestospongia muta*. Exploiting a weak CD effect for assignment of configuration.** *Org Lett*. 2007; **9**(10): 1975–1978.

[PubMed Abstract](#) | [Publisher Full Text](#)

Morrow KM, Fiore CL, Lesser MP: **Environmental drivers of microbial community shifts in the giant barrel sponge, *Xestospongia muta*, over a shallow to mesophotic depth gradient.** *Environ Microbiol*. 2016; **18**(6): 2025–2038.

[PubMed Abstract](#) | [Publisher Full Text](#)

Mulheron R: **Microbial community assembly found with sponge orange band disease in *Xestospongia muta* (Giant Barrel Sponge).** 2014.

[Reference Source](#)

Munroe S, Martens DE, Sipkema D, *et al.*: **Comparison of cryopreservation techniques for cells of the marine sponge *Dysidea etheria*.** *Cryo Letters*. 2018; **39**(4): 269–278.

[PubMed Abstract](#)

Narváez-Gómez JP, Mbye H, Oatley G, *et al.*: **Sanger Tree of Life sample homogenisation: Covaris cryoPREP® automated Dry Pulverizer V.1.** *protocols.io*. 2023.

[Publisher Full Text](#)

Neely KL, Butler CB: **Seasonal, lunar, and diel patterns in spawning by the giant barrel sponge, *Xestospongia muta*.** *Coral Reefs*. 2020; **39**(6): 1511–1515.

[Publisher Full Text](#)

Oatley G, Denton A, Howard C: **Sanger Tree of Life HMW DNA extraction: automated MagAttract v.2.** *protocols.io*. 2023a.

[Publisher Full Text](#)

Oatley G, Sampaio F, Howard C: **Sanger Tree of Life fragmented DNA clean up: automated SPRI.** *protocols.io*. 2023b.

[Publisher Full Text](#)

Oatley G, Sampaio F, Kitchin L, *et al.*: **Sanger Tree of Life HMW DNA fragmentation: Covaris g-TUBE for ULI PacBio.** *protocols.io*. 2023c; [Accessed 13 June 2024].

[Publisher Full Text](#)

Olm MR, Brown CT, Brooks B, *et al.*: **dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication.** *ISME J*. 2017; **11**(12): 2864–2868.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Parks DH, Imelfort M, Skennerton CT, *et al.*: **CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes.** *Genome Res*. 2015; **25**(7): 1043–55.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Patil AD, Kokke WC, Cochran S, *et al.*: **Brominated polyacetylenic acids from the marine sponge *Xestospongia muta*: inhibitors of HIV protease.** *J Nat Prod*. 1992; **55**(9): 1170–1177.

[PubMed Abstract](#) | [Publisher Full Text](#)

Pointon DL, Eagles W, Sims Y, *et al.*: **sanger-tol/treeval: v1.0.0 – Ancient Atlantis.** 2023.

[Publisher Full Text](#)

Quinlan AR, Hall IM: **BEDTools: a flexible suite of utilities for comparing genomic features.** *Bioinformatics*. 2010; **26**(6): 841–842.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

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.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Rao SSP, Huntley MH, Durand NC, *et al.*: **A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping.** *Cell*. 2014; **159**(7): 1665–1680.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Rhie A, Walenz BP, Koren S, *et al.*: **Merquy: reference-free quality, completeness, and phasing assessment for genome assemblies.** *Genome Biol*. 2020; **21**(1): 245.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Rosales SM, Huebner LK, Evans JS, *et al.*: **A meta-analysis of the stony coral tissue loss disease microbiome finds key bacteria in unaffected and lesion tissue in diseased colonies.** *ISME Commun*. 2023; **3**(1): 19.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Rühlemann MC, Wacker EM, Ellinghaus D, *et al.*: **MAGScoT: a fast, lightweight and accurate bin-refinement tool.** *Bioinformatics*. 2022; **38**(24): 5430–5433.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Seemann T: **Prokka: rapid prokaryotic genome annotation.** *Bioinformatics*. 2014; **30**(14): 2068–2069.

[PubMed Abstract](#) | [Publisher Full Text](#)

Slaby BM, Franke A, Rix L, *et al.*: **Marine sponge holobionts in health and disease.** In: *Symbiotic microbiomes of coral reefs sponges and corals*. 2019; 81–104.

[Publisher Full Text](#)

Southwell MW, Weisz JB, Martens CS, *et al.*: **In situ fluxes of dissolved inorganic nitrogen from the sponge community on Conch Reef, Key Largo, Florida.** *Limnol Oceanogr*. 2008; **53**(3): 986–996.

[Publisher Full Text](#)

Swierts T, Peijnenburg KTCA, de Leeuw CA, *et al.*: **Globally intertwined evolutionary history of giant barrel sponges.** *Coral Reefs*. 2017; **36**(3): 933–945.

[Publisher Full Text](#)

Thomas T, Moitinho-Silva L, Lurgi M, *et al.*: **Diversity, structure and**

convergent evolution of the global sponge microbiome. *Nat Commun.* 2016; **7**(1): 11870.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Ulliano-Silva M, Ferreira JGRN, Krasheninnikova K, *et al.*: **MitoHiFi: a python pipeline for mitochondrial genome assembly from PacBio high fidelity reads.** *BMC Bioinformatics.* 2023; **24**(1): 288.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Vasimuddin M, Misra S, Li H, *et al.*: **Efficient architecture-aware acceleration of BWA-MEM for multicore systems.** In: *2019 IEEE International Parallel and*

Distributed Processing Symposium (IPDPS). IEEE, 2019; 314–324.

[Publisher Full Text](#)

Wu YW, Tang YH, Tringe SG, *et al.*: **MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm.** *Microbiome.* 2014; **2**(1): 26.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free 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 August 2025

<https://doi.org/10.21956/wellcomeopenres.26664.r127266>

© 2025 Karahan 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.



Arzu Karahan 

Middle East Technical University, Erdemli-Mersin, Turkey

The research titled 'The chromosomal genome sequence of the giant barrel sponge, *Xestospongia muta* Schmidt 1870 and its associated microbial metagenome sequences' is a well-designed, prepared, and concluded study. It can be accepted for indexed with minor revision. I only recommend rotating the text on the Y-axis of the graphs upwards.

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: metabarcoding, metagenomics, transcriptomics on Tunicates

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.

Reviewer Report 28 August 2025

<https://doi.org/10.21956/wellcomeopenres.26664.r128020>

© 2025 Maronna M. 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.



Maximiliano M. Maronna 

Federal University of Rio Grande, Rio Grande, Brazil

The DATA NOTE submitted by Lopez et al. describes the genome assembly of one individual of *Xestospongia muta*. Main text presents proper details about general biology, biogeography and recent ecological events.

The applied methodology is a high-quality approach, including long reads, short reads for Hi-C spatial proximity method and transcriptomic sequencing. This manuscript is well-written and provides information related to materials and methods plus genome assembly, including exogenous DNA removal and metagenome assemblies.

Minor suggestions to improve this article:

- (1) To detail number of current species for *Xestospongia* in the Background section.
- (2) To better describe the “holobiont transcriptomic profile” in Background section, “carried out one of the first transcriptome studies of *X. muta* that revealed active metabolic expression of multiple nitrogen transformation genes (e.g. archaeal *amoA*)...” The way it is presented in this first version, it would be interpreted that *X. muta* is expressing archaeal genes, among others.
- (3) To describe basic genome heterozygosity, a histogram of k-mer multiplicity from Illumina reads would be easy to calculate and informative as part of main figures.

Also, I would like to remark two important points from another review I did a few months ago, in this same journal:

“(5) To create as Supplementary Material a genome report sheet as recommended in Santander et al 2022 (refer to 1). Doing so, it will be easier to learn about all inputs_outputs for each step, and to combine+compare with similar datasets for comparative analyses. Also, an additional document briefly describing workflow steps, scripts and command lines via GitHub or similar (e.g., GitLab) would be very useful. Current format at Github is quite fragmented and not very easy to understand.

(4) Busco values are not represented in the GenBank datasets section. I understand this is a Data Note, but the relatively medium-low BUSCO values for a chromosome-level assembly would be interesting to remark (maybe this is not an artifact, and this early-branching lineage present a reduced BUSCO elements).”

References

1. Santander M, Maronna M, Ryan J, Andrade S: The state of Medusozoa genomics: current evidence and future challenges. *GigaScience*. 2022; **11**. [Publisher Full Text](#)

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: Evolutionary Genetics, Marine Genetics, 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.
