Check for updates

DATA NOTE

The genome sequence of the turban top shell, *Gibbula magus* (Linnaeus, 1758) [version 1; peer review: 2 approved]

Patrick Adkins[1], Joanna Harley[1], Teresa Darbyshire [2], Anna Holmes[2],
Marine Biological Association Genome Acquisition Lab,
Darwin Tree of Life Barcoding collective,
Wellcome Sanger Institute Tree of Life programme,
Wellcome Sanger Institute Scientific Operations: DNA Pipelines collective,
Tree of Life Core Informatics collective, Kesella Scott-Somme[1],
Nova Mieskowska [1], Darwin Tree of Life Consortium

[1]Marine Biological Association, Plymouth, UK
[2]Amgueddfa Cymru, Cardiff, Wales, UK

Abstract

We present a genome assembly from an individual *Gibbula magus* (the turban top shell; Mollusca; Gastropoda; Trochida; Trochidae). The genome sequence is 1,470 megabases in span. Most of the assembly is scaffolded into 18 chromosomal pseudomolecules. The mitochondrial genome has also been assembled and is 16.1 kilobases in length. Gene annotation of this assembly on Ensembl identified 41,167 protein coding genes.

Keywords

Gibbula magus, turban top shell, genome sequence, chromosomal, Trochida

This article is included in the Tree of Life gateway.

Open Peer Review

Approval Status ✓ ✓

	1	2
version 1 23 Jan 2023	✓ view	✓ view

1. **Iker Irisarri** , Leibniz Institute for the Analysis of Biodiversity Change (LIB), Hamburg, Germany

2. **Qinggang Xue**, Zhejiang Wanli University, Ningbo, China

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

Corresponding author: Darwin Tree of Life Consortium (mark.blaxter@sanger.ac.uk)

Author roles: Adkins P: Investigation, Resources, Writing – Original Draft Preparation, Writing – Review & Editing; **Harley J**: Investigation, Resources; **Darbyshire T**: Investigation, Resources; **Holmes A**: Investigation, Resources; **Scott-Somme K**: Writing – Original Draft Preparation; **Mieskowska N**: Writing – Original Draft Preparation;

How to cite this article: Adkins P, Harley J, Darbyshire T *et al.* **The genome sequence of the turban top shell,** *Gibbula magus* **(Linnaeus, 1758) [version 1; peer review: 2 approved]** Wellcome Open Research 2023, **8**:35 https://doi.org/10.12688/wellcomeopenres.18792.1

First published: 23 Jan 2023, **8**:35 https://doi.org/10.12688/wellcomeopenres.18792.1

Species taxonomy

Eukaryota; Metazoa; Spiralia; Lophotrochozoa; Mollusca; Gastropoda; Vetigastropoda; Trochida; Trochoidea; Trochidae; Cantharidinae; *Gibbula*; *Gibbula magus* (Linnaeus, 1758) (NCBI:txid703304).

Background

Gibbula magus is a marine gastropod mollusc in the Trochidae family, known as topshells (NBN Atlas, 2021). They are usually sublittoral, and found on muddy sandy or gravel, on algae or under stones, where they feed on microphytes (Smith, 2015). They can be found in the intertidal zone at extreme low spring tides and down to depths of 70 m (Wilson, 2017). In Great Britain, *G. magus* is seldom found on the east coast, occurring almost exclusively on the south and west coasts, and on all coasts in Ireland (de Kluijver *et al.*, 2000; NBN Atlas, 2021). The breeding times of *G. magus* varies geographically; June at Roscoff and spring and autumn at Plymouth. Fertilisation happens externally and there is a brief free-living trochophore larval stage, however, little else is known about these early life stages (Smith, 2015). The fringe present on the body of *G. magus* is thought to protect against detritus and may be able to sense poor water conditions (Smith, 2015).

The *Gibbula* genus can be difficult to identify due to high variability in the shell morphology (the main identifying feature) unspecific or missing type material and vague original descriptions (Affenzeller *et al.*, 2017). Adult *G. magus* are the largest of the trochid species found in UK waters. They have a flattened spire and a large umbilicus, and diagonal lines of pink coloured dots along the whorls of the shell. The taxonomic divides of the Trochidae family have also been investigated, with changes to the groupings and names of several species (Affenzeller *et al.*, 2017; Anistratenko, 2005). Quality genomic data can add more information and detail for how these decisions are made, and support the work of taxonomists (Coates *et al.*, 2018). Here we present a chromosomally complete genome sequence for *G. magus*, based on a specimen from Gann Bay, Pembrokeshire, UK.

Genome sequence report

The genome was sequenced from a *Gibbula magus* specimen (Figure 1) collected from Gann Bay, Pembrokeshire, UK (latitude 51.71, longitude –5.17). A total of 35-fold coverage in Pacific Biosciences single-molecule HiFi long reads and 36-fold coverage in 10X Genomics read clouds was generated. Primary assembly contigs were scaffolded with chromosome conformation Hi-C data. Manual assembly curation corrected 477 missing joins or misjoins and removed 93 haplotypic duplications, reducing the assembly length by 4.9% and the scaffold number by 54.52%, and increasing the scaffold N50 by 3.38%.

The final assembly has a total length of 1,470.4 Mb in 151 sequence scaffolds with a scaffold N50 of 80.5 Mb (Table 1). Most (99.51%) of the assembly sequence was assigned to 9 chromosomal-scale scaffolds. Chromosome-scale scaffolds confirmed by the Hi-C data are named in order of size (Figure 2–Figure 5; Table 2). While not fully phased, the assembly deposited is of one haplotype. Contigs corresponding to the second haplotype have also been deposited. The mitochondrial genome was also assembled.

The assembly has a BUSCO v5.3.2 (Manni *et al.*, 2021) completeness of 84.2% (single 83.4%, duplicated 0.8%) using the OrthoDB-v10 mollusca reference set. BUSCO loci identified as fragmented accounted for a further 4.8% of loci tested. This low BUSCO score may be due to low conservation of orthologues between *G. magus* and the molluscan species in the reference set, or underperformance of the BUSCO gene finder given the particular gene structures in this species. The assembly is validated by the other assembly quality metrics (*k*-mer completeness 99.98%, consensus quality (QV) 52.2) shown in Table 1.
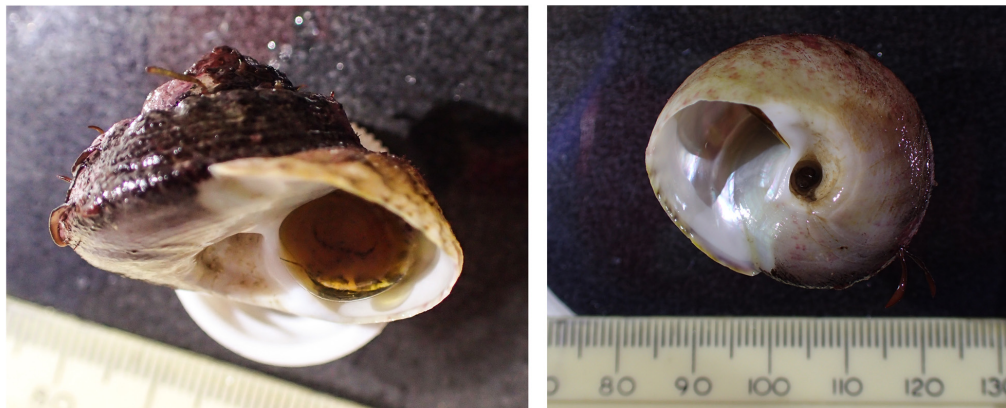

Figure 1. Photographs of the *Gibbula magus* (xgGibMagu1) specimen used for genome sequencing.

Table 1. Genome data for *Gibbula magus*, xgGibMagu1.1.

Project accession data		
Assembly identifier	xgGibMagu1.1	
Species	*Gibbula magus*	
Specimen	xgGibMagu1	
NCBI taxonomy ID	703304	
BioProject	PRJEB51161	
BioSample ID	SAMEA8717440	
Isolate information		
Assembly metrics*		***Benchmark***
Consensus quality (QV)	52.2	*≥ 50*
k-mer completeness	99.98	*≥ 95%*
BUSCO**	C:84.2%[S:83.4%,D:0.8%],F:4.8%,M:11.0%,n:5,295	*C ≥ 95%*
Percentage of assembly mapped to chromosomes	99.51%	*≥ 95%*
Organelles	mitochondrial genome assembled	*complete single alleles*
Raw data accessions		
PacificBiosciences SEQUEL II	ERR9127937–ERR9127939	
10X Genomics Illumina	ERR8974924–ERR8974927	
Hi-C Illumina	ERR8974928	
PolyA RNA-Seq Illumina	ERR10123686	
Genome assembly		
Assembly accession	GCA_936450465.1	
Accession of alternate haplotype	GCA_936439675.1	
Span (Mb)	1,470.4	
Number of contigs	1,001	
Contig N50 length (Mb)	3.4	
Number of scaffolds	151	
Scaffold N50 length (Mb)	80.5	
Longest scaffold (Mb)	117.5	

* Assembly metric benchmarks are adapted from column VGP-2020 of "Table 1: Proposed standards and metrics for defining genome assembly quality" from (Rhie *et al.*, 2021).

** BUSCO scores based on the mollusca_odb10 BUSCO set using v5.3.2. C = complete [S = single copy, D = duplicated], F = fragmented, M = missing, n = number of orthologues in comparison. A full set of BUSCO scores is available at https://blobtoolkit.genomehubs.org/view/xgGibMagu1.1/dataset/CAKZFU01/busco.

Genome annotation report

The *G. magus* (GCA_936450465.1) genome assembly was annotated using BRAKER2 in the Ensembl rapid annotation pipeline (Table 1; https://rapid.ensembl.org/Gibbula_magus_GCA_936450465.1/Info/Index). The resulting annotation includes 41,235 transcribed mRNAs from 41,167 protein-coding genes.

Methods

Sample acquisition and nucleic acid extraction

The collectors of the *G. magus* specimen (xgGibMagu1) used for genome sequencing were Patrick Adkins and Joanna Harley (Marine Biological Association) and Teresa Darbyshire and Anna Holmes (Amgueddfa Cymru), and the specimen was then identified by Patrick Adkins and Joanna Harley. The
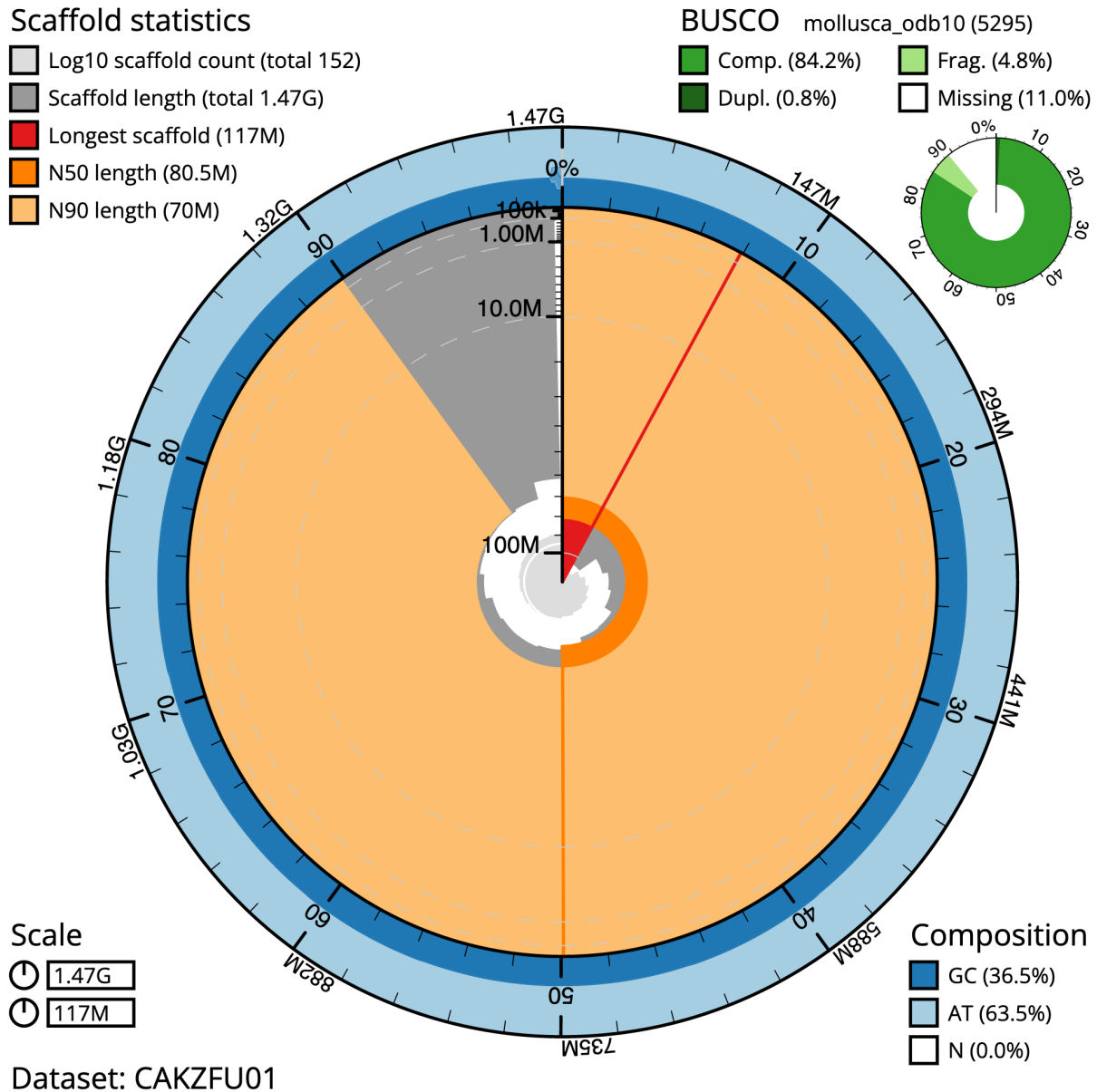
Figure 2. Genome assembly of *Gibbula magus*, xgGibMagu1.1: metrics. The BlobToolKit Snailplot shows N50 metrics and BUSCO gene completeness. The main plot is divided into 1,000 size-ordered bins around the circumference with each bin representing 0.1% of the 1,470,404,880 bp assembly. The distribution of scaffold lengths is shown in dark grey with the plot radius scaled to the longest scaffold present in the assembly (117,468,434 bp, shown in red). Orange and pale-orange arcs show the N50 and N90 scaffold lengths (80,454,948 and 69,970,916 bp), respectively. The pale grey spiral shows the cumulative scaffold 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 in the mollusca_odb10 set is shown in the top right. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/xgGibMagu1.1/dataset/CAKZFU01/snail.

specimen was collected in Gann Bay, Pembrokeshire, UK (latitude 51.71, longitude –5.17). The sample was taken from sediment by hand and placed into a container, using and then snap-frozen in liquid nitrogen.

DNA was extracted at the Tree of Life laboratory, Wellcome Sanger Institute (WSI). The xgGibMagu1 sample was

weighed and dissected on dry ice with tissue set aside for Hi-C sequencing. Muscle tissue was cryogenically disrupted to a fine powder using a Covaris cryoPREP Automated Dry Pulveriser, receiving multiple impacts. High molecular weight (HMW) DNA was extracted using the Qiagen MagAttract HMW DNA extraction kit. Low molecular weight DNA was removed from a 200 ng aliquot of extracted DNA using 0.8X
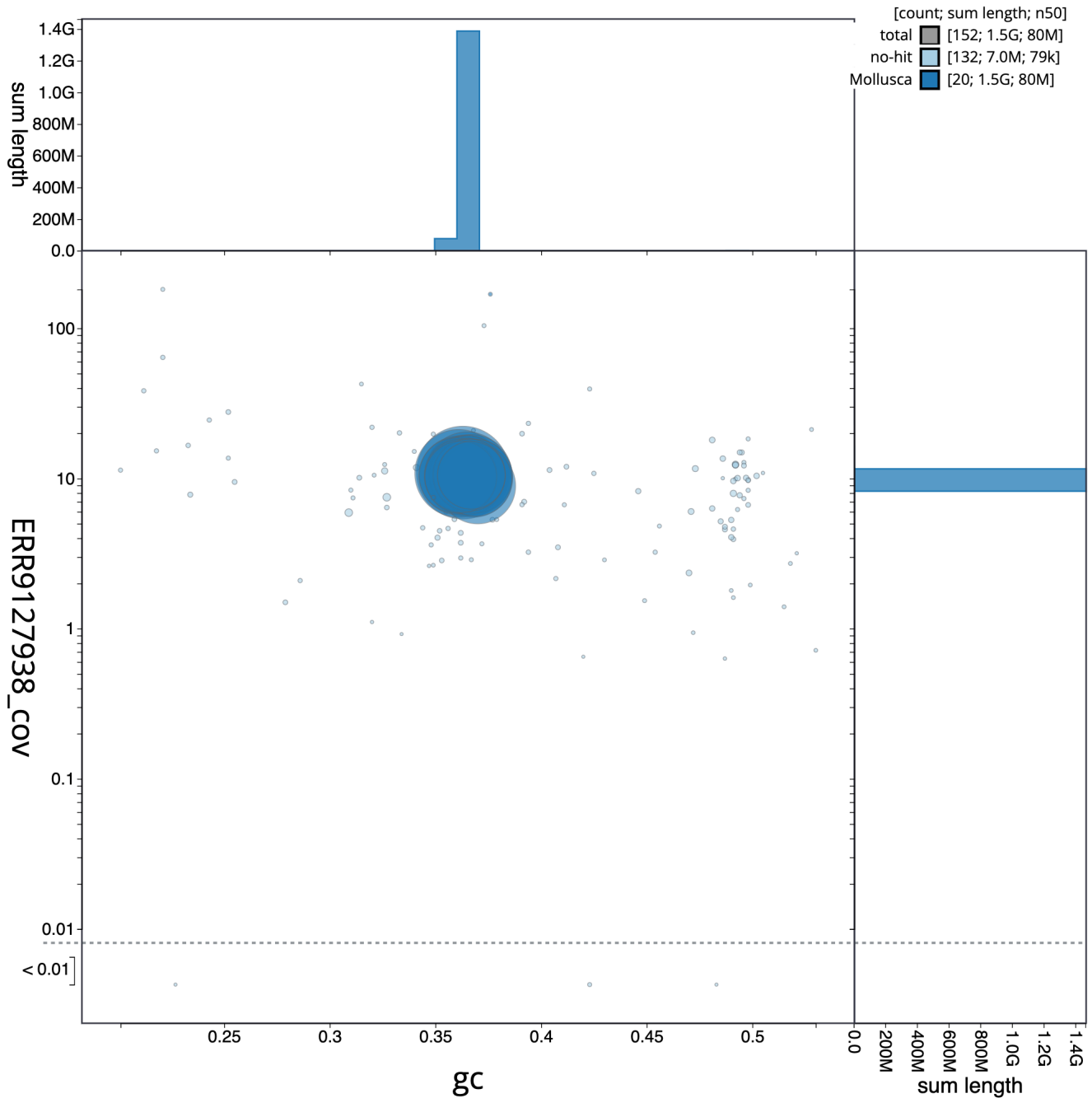
Figure 3. Genome assembly of *Gibbula magus*, xgGibMagu1.1: GC coverage. BlobToolKit GC-coverage plot. Sequences are coloured by phylum. Circles are sized in proportion to scaffold length. Histograms show the distribution of scaffold length sum along each axis. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/xgGibMagu1.1/dataset/CAKZFU01/blob.

AMpure XP purification kit prior to 10X Chromium sequencing; a minimum of 50 ng DNA was submitted for 10X sequencing. HMW DNA was sheared into an average fragment size between 12–20 kb in a Megaruptor 3 system with speed setting 30. Sheared DNA was purified by solid-phase reversible immobilisation using AMPure PB beads with a 1.8X ratio of beads to sample to remove the shorter fragments and concentrate the DNA sample. The concentration of the sheared and purified DNA was assessed using a Nanodrop spectrophotometer and Qubit Fluorometer and Qubit dsDNA High Sensitivity Assay kit. Fragment size distribution was evaluated by running the sample on the FemtoPulse system.
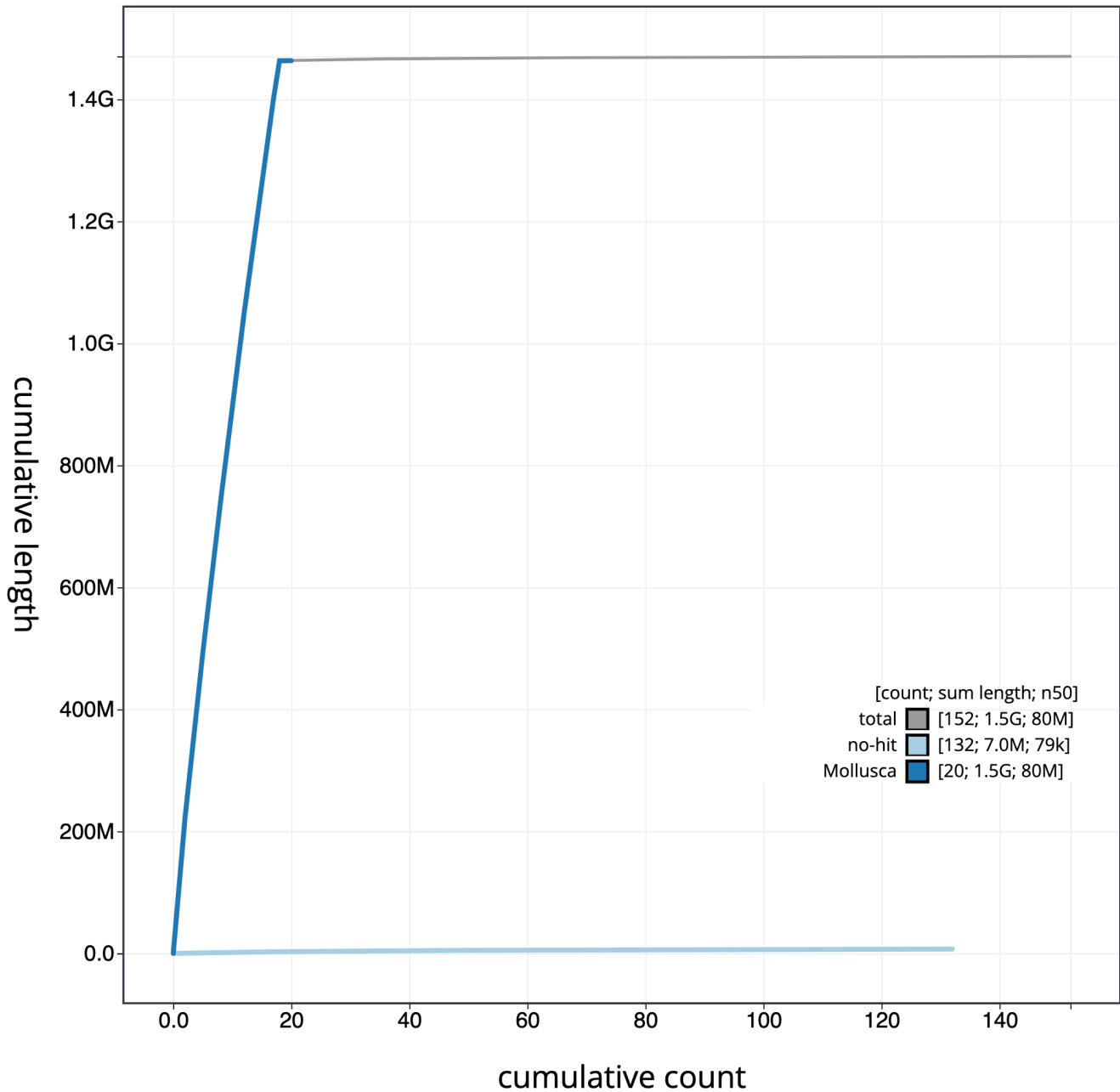
Figure 4. Genome assembly of *Gibbula magus*, xgGibMagu1.1: cumulative sequence. BlobToolKit cumulative sequence plot. The grey line shows cumulative length for all scaffolds. Coloured lines show cumulative lengths of scaffolds assigned to each phylum using the buscogenes taxrule. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/xgGibMagu1.1/dataset/CAKZFU01/cumulative.

RNA was extracted from muscle tissue of xgGibMagu1 in the Tree of Life Laboratory at the WSI using TRIzol, according to the manufacturer's instructions. RNA was then eluted in 50 μl RNAse-free water and its concentration RNA assessed using a Nanodrop spectrophotometer and Qubit Fluorometer using the Qubit RNA Broad-Range (BR) Assay kit. Analysis of the integrity of the RNA was done using Agilent RNA 6000 Pico Kit and Eukaryotic Total RNA assay.

Sequencing

Pacific Biosciences HiFi circular consensus and 10X Genomics read cloud DNA sequencing libraries were constructed
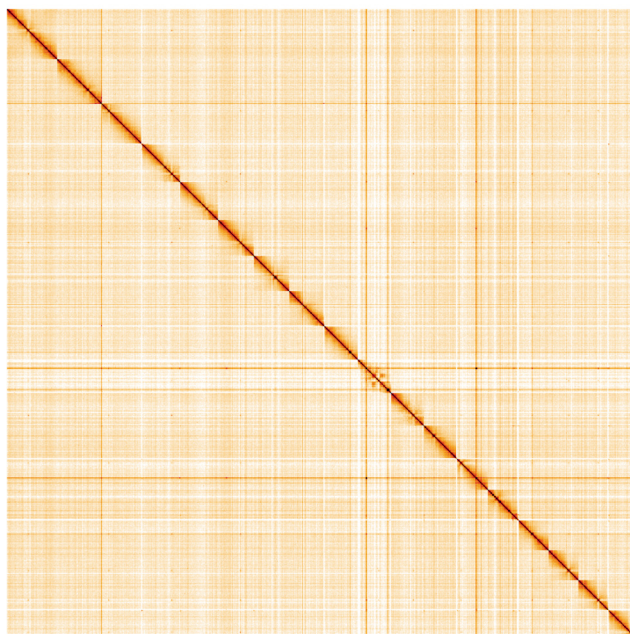
Figure 5. Genome assembly of *Gibbula magus*, xgGibMagu1.1: Hi-C contact map. Hi-C contact map of the xgGibMagu1.1 assembly, visualised using HiGlass. Chromosomes are given in order of size from left to right and top to bottom. An interactive version of this plot is at: https://genome-note-higlass.tol.sanger.ac.uk/l/?d=N9lkINhLTXmGpbn8oj1olQ.

Table 2. Chromosomal pseudomolecules in the genome assembly of *Gibbula magus*, xgGibMagu1.

INSDC accession	Chromosome	Size (Mb)	GC%
OW388289.1	1	117.47	36.3
OW388290.1	2	104.47	36.1
OW388291.1	3	92.77	36.5
OW388292.1	4	89.61	36.7
OW388293.1	5	88	36.4
OW388294.1	6	83.61	36.1
OW388295.1	7	83.18	36.8
OW388296.1	8	80.45	36.3
OW388297.1	9	78.28	36.5
OW388298.1	10	77.47	37
OW388299.1	11	77.31	36.5
OW388300.1	12	76.35	36
OW388301.1	13	73.01	36.9
OW388302.1	14	70.77	36.9
OW388303.1	15	70.18	36.3
OW388304.1	16	69.97	36.6
OW388305.1	17	69.2	36.2
OW388306.1	18	61.23	36.7
OW388307.1	MT	0.02	37.9

according to the manufacturers' instructions. Poly(A) RNA-Seq libraries were constructed using the NEB Ultra II RNA Library Prep kit. DNA and RNA sequencing was performed by the Scientific Operations core at the WSI on Pacific Biosciences SEQUEL II (HiFi), Illumina NovaSeq 6000 (RNA-Seq and 10X) instruments. Hi-C data were also generated from muscle tissue of xgGibMagu1 using the Arima v2 kit and sequenced on the Illumina NovaSeq 6000 instrument.

Genome assembly

Assembly was carried out with Hifiasm (Cheng *et al.*, 2021) and haplotypic duplication was identified and removed with purge_dups (Guan *et al.*, 2020). One round of polishing was performed by aligning 10X Genomics read data to the assembly with Long Ranger ALIGN, calling variants with freebayes (Garrison & Marth, 2012). The assembly was then scaffolded with Hi-C data (Rao *et al.*, 2014) using YaHS (Zhou *et al.*, 2022). The assembly was checked for contamination and corrected as described previously (Howe *et al.*, 2021). Manual curation was performed using HiGlass (Kerpedjiev *et al.*, 2018) and Pretext (Harry, 2022). The mitochondrial genome was assembled using MitoHiFi (Uliano-Silva *et al.*, 2022), which performed annotation using MitoFinder (Allio *et al.*, 2020). The genome was analysed and BUSCO scores were generated within the BlobToolKit environment (Challis *et al.*, 2020). Table 3 contains a list of all software tool versions used, where appropriate.

Genome annotation

The Ensembl gene annotation system (Aken *et al.*, 2016) was used to generate annotation for the *G. magus* assembly (GCA_936450465.1). Annotation was created primarily through alignment of transcriptomic data to the genome, with

Table 3. Software tools used.

Software tool	Version	Source
BlobToolKit	3.3.10	Challis *et al.*, 2020
freebayes	1.3.1-17- gaa2ace8	Garrison & Marth, 2012
Hifiasm	0.15.3-r339	Cheng *et al.*, 2021
HiGlass	1.11.6	Kerpedjiev *et al.*, 2018
Long Ranger ALIGN	2.2.2	https://support.10xgenomics.com/genome-exome/software/pipelines/latest/advanced/other-pipelines
MitoHiFi	2.0	Uliano-Silva *et al.*, 2022
PretextView	0.2.x	Harry, 2022
purge_dups	1.2.3	Guan *et al.*, 2020
YaHS	yahs-1.1.91eebc2	Zhou *et al.*, 2022

gap filling via protein to-genome alignments of a select set of proteins from UniProt (UniProt Consortium, 2019).

Ethics/compliance issues

The materials that have contributed to this genome note have been supplied by a Darwin Tree of Life Partner. The submission of materials by a Darwin Tree of Life Partner is subject to the Darwin Tree of Life Project Sampling Code of Practice. By agreeing with and signing up to the Sampling Code of Practice, the Darwin Tree of Life Partner agrees they will meet the legal and ethical requirements and standards set out within this document in respect of all samples acquired for, and supplied to, the Darwin Tree of Life Project. Each transfer of samples is further undertaken according to a Research Collaboration Agreement or Material Transfer Agreement entered into by the Darwin Tree of Life Partner, Genome Research Limited (operating as the Wellcome Sanger Institute), and in some circumstances other Darwin Tree of Life collaborators.

Data availability

European Nucleotide Archive: *Gibbula magus* (turban top shell). Accession number PRJEB51161; https://www.ebi.ac.uk/ena/browser/view/PRJEB51161 (Wellcome Sanger Institute, 2022).

The genome sequence is released openly for reuse. The *Gibbula magus* genome sequencing initiative is part of the Darwin Tree of Life (DToL) project. All raw sequence data and the assembly have been deposited in INSDC databases. Raw data and assembly accession identifiers are reported in Table 1.

Author information

Members of the Marine Biological Association Genome Acquisition Lab are listed here: https://doi.org/10.5281/zenodo.4783604.

Members of the Darwin Tree of Life Barcoding collective are listed here: https://doi.org/10.5281/zenodo.5744972.

Members of the Wellcome Sanger Institute Tree of Life programme are listed here: https://doi.org/10.5281/zenodo.6125027.

Members of Wellcome Sanger Institute Scientific Operations: DNA Pipelines collective are listed here: https://doi.org/10.5281/zenodo.5746904.

Members of the Tree of Life Core Informatics collective are listed here: https://doi.org10.5281/zenodo.6125046.

Members of the Darwin Tree of Life Consortium are listed here: https://doi.org/10.5281/zenodo.5638618.

References

Affenzeller S, Haar N, Steiner G: **Revision of the genus complex *Gibbula*: an integrative approach to delineating the Eastern Mediterranean genera *Gibbula* Risso, 1826, *Steromphala* Gray, 1847, and *Phorcus* Risso, 1826 using DNA-barcoding and geometric morphometrics (Vetigastropoda, Trochoidea).** *Organisms Diversity & Evolution.* 2017; **17**(4): 789–812.
Publisher Full Text

Aken BL, Ayling S, Barrell D, *et al.*: **The Ensembl gene annotation system.**

Database (Oxford). 2016; **2016**: baw093.
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**

Anistratenko VV: **Lectotypes for *Tricolia pullus*, *Gibbula divaricata* and**

Theodoxus fluviatilis **(Mollusca, Gastropoda) revisited.** *Vestnik Zoologii.* 2005; **39**(6): 3–10.
Reference Source

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**

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**

Coates DJ, Byrne M, Moritz C: **Genetic Diversity and Conservation Units: Dealing With the Species-Population Continuum in the Age of Genomics.** *Front Ecol Evol.* 2018; **6**: 165.
Publisher Full Text

de Kluijver MJ, Ingalsuo SS, de Bruyne RH: **Mollusca of the North Sea: Gibbula magus**. 2000; (Accessed: 2 December 2022).
Reference Source

Garrison E, Marth G: **Haplotype-based variant detection from short-read sequencing**. 2012.
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; (Accessed: 19 October 2022).
Reference Source

Howe K, Chow W, Collins J, *et al.*: **Significantly improving the quality of genome assemblies through curation.** *Gigascience.* Oxford University Press, 2021; **10**(1): giaa153.
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**

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**

NBN Atlas: *Gibbula magus* **(Linnaeus, 1758) Turban Top Shell.** 2021; (Accessed: 2 December 2022).
Reference Source

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, McCarthy SA, Fedrigo O, *et al.*: **Towards complete and error-free genome assemblies of all vertebrate species.** *Nature.* 2021; **592**(7856): 737–746.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Smith IF: *Gibbula magus* **(Linnaeus, 1758) Identification and Biology.** 2015.
Publisher Full Text

Uliano-Silva M, Ferreira JGRN, Krasheninnikova K, *et al.*: **MitoHiFi: a python pipeline for mitochondrial genome assembly from PacBio High Fidelity reads.** *bioRxiv.* 2022.
Publisher Full Text

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

Wellcome Sanger Institute: **The genome sequence of the turban top shell, *Gibbula magus* (Linnaeus, 1758), European Nucleotide Archive [Dataset].** 2022.
http://identifiers.org/ena.embl/PRJEB51161

Wilson C: **Turban top shell (*Gibbula magus*).** 2017; (Accessed: 2 December 2022).
Reference Source

Zhou C, McCarthy SA, Durbin R: **YaHS: yet another Hi-C scaffolding tool.** *bioRxiv.* 2022.
Publisher Full Text

Open Peer Review

Current Peer Review Status: ✔ ✔

Version 1

Reviewer Report 01 August 2023

https://doi.org/10.21956/wellcomeopenres.20837.r63422

✔ **Qinggang Xue**
Zhejiang Wanli University, Ningbo, Zhejiang, China

The authors reported a high quality chromosome –level genome assembly for the *Gibbula magus*, it would be useful for supporting the work of taxonomists and evolution analysis .The manuscript is acceptable after minor revision.

Minor suggestions
1. In the abstract, the mitochondrial genome were mentioned, however, the sequence was not uploaded, at least not appeared in the data availability section.

2. In the method section, the genome prediction and annotation methods should be much clearly present.

3. Figure 1, the unit of the ruler are not marked out.

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?
Partly

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

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Gene function, genome

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.

✓   **Iker Irisarri** (iD)
Centre for Molecular Biodiversity Research, Leibniz Institute for the Analysis of Biodiversity Change (LIB), Hamburg, Germany

The rationale for this new genome is clearly explained. The genome assembly and annotation pipelines represent the current gold standards and are sufficiently explained. The resulting genome is thus of high quality and it is made fully accessible to the community.
I have only two very minor comments on the text that I feel could increase clarity:
  ○ Abstract. The sentence "Most of the assembly is scaffolded into 18 chromosomal pseudomolecules" could be understood as the assembly consisting of 18 pseudochromosomes when it is actually a phased assembly of 9 pseudochromosomes.

  ○ Blobtoolkit and Figure 3 are not mentioned in the text. Also, the caption of Figure 3 ("GC coverage") could be clarified.

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: Phylogenomics and 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.