Check for updates

DATA NOTE

REVISED ERGA-BGE Chromosome-Level Genome of *Patella rustica* Linnaeus, 1758: a Resource to Investigate Responses to Global Warming in the intertidal

[version 2; peer review: 2 approved, 1 approved with reservations]

Rui Faria[1-3], Rocío Nieto Vilela[1,2], Fernando P Lima[1,2], Rita Monteiro[4], Astrid Böhne[4], Thomas Marcussen[5,6], Torsten H Struck[5], Rebekah A Oomen[5-9], Marta Gut[10,11], Laura Aguilera[10,11], Tyler Alioto[10,11], Francisco Câmara Ferreira[10,11], Jèssica Gómez-Garrido[10,11], Fernando Cruz[10,11], Anna Lazar[12], Leanne Haggerty[12], Fergal Martin[12], Tom Brown[13,14]

[1]CIBIO, InBIO Laboratório Associado, Universidade do Porto Centro de Investigacao em Biodiversidade e Recursos Geneticos, Vairão, Porto, Portugal
[2]BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Universidade do Porto, Vairão, Portugal
[3]Departamento de Biologia, Faculdade de Ciências, Universidade do Porto, Porto, Portugal
[4]Leibniz Institute for the Analysis of Biodiversity Change - Museum Koenig Bonn, Bonn, Germany
[5]Natural History Museum, University of Oslo, Oslo, Norway
[6]University of Oslo Centre for Ecological and Evolutionary Synthesis, Oslo, Norway
[7]Department of Biological Sciences, University of New Brunswick Saint John, Saint John, New Brunswick, Canada
[8]Tjärnö Marine Laboratory, University of Gothenburg, Gothenburg, Sweden
[9]Centre for Coastal Research, University of Agder, Kristiansand, Norway
[10]Centro Nacional de Análisis Genómico (CNAG), Barcelona, Spain
[11]Universitat de Barcelona (UB), Barcelona, Spain
[12]European Molecular Biology Laboratory, European Bioinformatics Institute, Hinxton, UK
[13]Leibniz Institute for Zoo and Wildlife Research, Berlin, Germany
[14]BeGenDiv, Berlin Center for Genomics in Biodiversity Research, Berlin, Germany

v2 **First published:** 11 Nov 2025, **5**:350
https://doi.org/10.12688/openreseurope.21608.1
Latest published: 09 Mar 2026, **5**:350
https://doi.org/10.12688/openreseurope.21608.2

Abstract

The reference genome of *Patella rustica* Linnaeus, 1758 (Gastropoda, Patellidae) will offer a unique opportunity to understand how intertidal organisms respond to the effects of global warming, contributing to improve our predictions about the impact of climate change on the maintenance and diversification of marine biodiversity. A total of 9 contiguous chromosomal pseudomolecules were assembled from the genome sequence, made up from Oxford Nanopore Technologies (ONT) long-reads, WGS and Chromatin-

Open Peer Review

Approval Status ✓ ? ✓

	1	2	3
version 2 (revision) 09 Mar 2026			✓ view
version 1 11 Nov 2025	✓ view	? view	

1. **Jacob A Tennessen**, Harvard T. H. Chan

Conformation-Capture (Hi-C) short-read sequences. This chromosome-level assembly encompasses 719 Mb, composed of 157 contigs and 24 scaffolds, with contig and scaffold N50 values of 8.52 Mb and 83.6 Mb, respectively, and assembly and annotation BUSCO completeness of 97.7% and 96.5%, respectively. The production of this genome assembly and annotation was coordinated and supported by the European Reference Genome Atlas (ERGA) and Biodiversity Genomics Europe project (BGE).

Keywords
Patella rustica, genome assembly, European Reference Genome Atlas, Biodiversity Genomics Europe, Earth Biogenome Project, Lusitanian limpet, North-eastern Atlantic, Mediterranean Sea, Sentinels, Thermal stress, Evolutionary studies

This article is included in the Horizon Europe gateway.

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

School of Public Health, Boston, USA

2. **Yasuto ISHII**, Tohoku University, Sendai, Japan

3. **Craig Wilding**, University College Dublin, Dublin, Ireland

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

Corresponding author: Tom Brown (brown@izw-berlin.de)

Author roles: Faria R: Investigation, Resources, Writing – Original Draft Preparation, Writing – Review & Editing; **Nieto Vilela R**: Investigation, Resources, Writing – Review & Editing; **Lima FP**: Investigation, Resources, Writing – Review & Editing; **Monteiro R**: Methodology, Project Administration, Supervision, Writing – Review & Editing; **Böhne A**: Funding Acquisition, Methodology, Project Administration, Supervision, Writing – Review & Editing; **Marcussen T**: Methodology, Project Administration, Supervision, Writing – Review & Editing; **Struck TH**: Funding Acquisition, Methodology, Project Administration, Supervision, Writing – Review & Editing; **Oomen RA**: Methodology, Project Administration, Supervision, Writing – Review & Editing; **Gut M**: Investigation, Supervision, Writing – Review & Editing; **Aguilera L**: Investigation, Writing – Original Draft Preparation, Writing – Review & Editing; **Alioto T**: Data Curation, Formal Analysis, Supervision, Writing – Review & Editing; **Câmara Ferreira F**: Data Curation, Formal Analysis, Writing – Review & Editing; **Gómez-Garrido J**: Data Curation, Formal Analysis, Writing – Review & Editing; **Cruz F**: Data Curation, Formal Analysis, Writing – Review & Editing; **Lazar A**: Data Curation, Formal Analysis, Writing – Review & Editing; **Haggerty L**: Data Curation, Formal Analysis, Writing – Original Draft Preparation, Writing – Review & Editing; **Martin F**: Data Curation, Formal Analysis, Supervision, Writing – Review & Editing; **Brown T**: Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

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. RF and FPL received support from Fundação para a Ciência e a Tecnologia through a CEECInd contracts (https://doi.org/10.54499/2020.00275.CEECIND/CP1601/CP1649/CT0001, https://doi.org/10.54499/2020.00275.CEECIND/CP1601/CT0002 and https://doi.org/10.54499/CEECIND/03185/2018/CP1546/CP1648/CT0004) and through the project INVCONTINUUM (https://doi.org/10.54499/PTDC/BIA-EVL/1614/2021) and OceanLog (PTDC/BIA-BMA/4848/2021). RNV received support from the EU under the Horizon 2020 program (project ANERIS, ref. 101094924).
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2026 Faria R *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: Faria R, Nieto Vilela R, Lima FP *et al.* **ERGA-BGE Chromosome-Level Genome of *Patella rustica* Linnaeus, 1758: a Resource to Investigate Responses to Global Warming in the intertidal [version 2; peer review: 2 approved, 1 approved with reservations]** Open Research Europe 2026, **5**:350 https://doi.org/10.12688/openreseurope.21608.2

First published: 11 Nov 2025, **5**:350 https://doi.org/10.12688/openreseurope.21608.1

REVISED **Amendments from Version 1**

Following suggestions from the reviewers, we have added more details regarding the statistics from the intermediate genome assemblies, as well as further information on the project leading the generation of the genome. We also include more taxonomic-specific details on the sexing of the sampled molluscs.

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

Introduction

A constantly changing environment challenges organisms to adapt and survive (Hoffman and Sgrò, 2011). This is particularly true for marine organisms living in the intertidal zone, where they must cope with regular environmental fluctuations caused by tides, as well as the increasing pressure from human-caused global warming (Helmuth et al., 2006; Somero, 2010). Gaining knowledge on how intertidal organisms cope with environmental stress, adapt, and continue to diversify is therefore urgent in the context of the current climate crisis (Trégarot et al., 2024) and the pressing need for effective conservation efforts.

Limpets are among the most common marine gastropods inhabiting the rocky intertidal zones worldwide (Firth, 2021). Of these, species of the genus *Patella*, which are distributed along the Northeastern Atlantic and Mediterranean coasts, have been extensively studied in relation to environmental stress, among other species (*Siphonaria* sp., Teske et al., 2011; barnacles and topshells, Broitman et al., 2008; for more invertebrates, see Sokolova et al., 2012). These species can serve as sentinels, providing insights into how marine biodiversity responds to climate change and the health of coastal ecosystems (Hawkins et al., 2009). Consequently, they can help improve predictions of the impacts of global warming on marine communities and ecosystems.

This is especially true for *Patella rustica*, also known as the Lusitanian limpet, which is found throughout the Mediterranean basin and along the northeastern Atlantic coast from Southern France (Biarritz) to Mauritania, as well as in the Macaronesian archipelagos, excluding the Azores (Fischer-Piètte, 1955; Crisp and Fischer-Piètte, 1959; Christiaens, 1973; Weber and Hawkins, 2005). It can be easily distinguished from other *Patella* species by the presence of brown spots on the upper part of the shell (Figure 1). Several studies have documented recent changes in its



Figure 1. Example image of *Patella rustica*. Image taken by Fernando P. Lima.

distribution, particularly the bridging of a historical distribution gap in northwestern Iberia (Fischer-Piètte and Gaillard, 1959; Lima et al., 2006). This distribution change was probably driven by rising sea temperatures, increased river runoff, and decreased upwelling intensity off the Portuguese coast (Lima et al., 2006, 2007; Sousa et al., 2012). While genetic diversity has remained stable in colonizing populations (Ribeiro et al., 2010), reliance solely on conventional markers (mtDNA and microsatellites) has limited understanding of both demographic history and adaptive genomic features. Addressing this gap is crucial for predicting further northward expansions and assessing the species' impact on local communities and ecosystems.

It is now increasingly recognized that public genomic resources, specifically high-quality reference genomes, can be highly valuable for biodiversity monitoring and can support conservation efforts by providing essential information to guide management decisions (Theissinger et al., 2023). Additionally, they offer a crucial foundation for assessing how species adapt to drastic environmental changes and further diversify amid global warming (Formenti et al., 2022). While we are beginning to better understand these evolutionary processes in certain marine species with low dispersal (e.g., *Littorina* marine snails, Faria et al., 2019, 2021; Morales et al., 2019; Westram et al., 2018), we are still far from understanding how speciation occurs in most marine species, particularly those with high dispersal, where gene flow can oppose divergent selection and reproductive isolation (Potkamp and Fransen, 2019; Barry et al., 2024).

Here, we present a chromosome-level genome of *Patella rustica*, to support intertidal biodiversity monitoring amid climate change. Along with recent chromosome-level reference genomes of five other *Patella* species (GCA_964341155.2, Hawkins et al., 2023, 2024, Mieszkowska et al., 2025; Lawniczak, 2022), this will help uncover insights into their evolutionary history and investigate the development of reproductive barriers within this genus (Sá Pinto et al., 2010), including the role of structural variants in diversification (Petraccioli et al., 2010).

The creation of this reference genome results from a collaborative effort coordinated by the European Reference Genome Atlas (ERGA) within the Biodiversity Genomics Europe (BGE) project. It aligns with ERGA's objectives of fostering cooperation among European countries to develop genomic resources for biodiversity protection (Mazzoni et al., 2023), as well as with the Portuguese coalition for biodiversity genomics initiative (Marques et al., 2024).

Materials & Methods

Sample and sampling information

On 29th January 2024, an adult, female sample of *Patella rustica* was sampled by Rui Faria and Fernando P. Lima. The species was identified by Fernando Lima via morphology and color pattern. The specimen was dislodged with a knife from an intertidal rock and collected by hand in Porto, Portugal. Ethical sampling permits are not required for sampling of molluscs at this location. After dissection and sexing (based on gonad colouration - see Zegaoula *et al.* 2016 and Prusina *et al.* 2014) by Rocío Nieto Vilela, the specimen's tissues (mollusc foot, gonad, head, hepatopancreas and other somatic tissue) were snap-frozen immediately in liquid nitrogen and stored at -80°C until they were shipped to CNAG, Barcelona in dry ice for subsequent laboratory procedures.

Vouchering information

Shells for the here sequenced specimen and proxy specimen have been deposited in the Natural History and Science Museum of the University of Porto (MHNC-UP https://mhnc.up.pt/), under the accession numbers MHNCUP-MOL-24647 and MHNCUP-MOL-24646.

Frozen reference tissue material of muscle is available from the same and proxy individuals at the Museum Koenig Bonn (https://bonn.leibniz-lib.de/de/) under the voucher IDs ZFMK-TIS:89067 and ZFMK-TIS-89068.

Data availability

Patella rustica and the related genomic study were assigned to Tree of Life ID (ToLID) 'xgPatRust1' and all sample, sequence, and assembly information are available under the umbrella BioProject PRJEB84018. The sample information is available at the following BioSample accessions: SAMEA117648698, SAMEA115406481, SAMEA115406496 and SAMEA117648698. The genome assembly is accessible from ENA under accession number GCA_965111895.1 and the annotated genome is available through the Ensembl Beta site (https://beta.ensembl.org/) and BGE Project Page (https://projects.ensembl.org/erga-bge/). Sequencing data produced as part of this project are available from ENA at the following accessions: ERX13508601, ERX13629331, ERX13629332 and ERX13629333. 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/Patella_rustica/xgPatRust1. Further details and data about the project are hosted on the ERGA portal at https://portal.erga-biodiversity.eu/data_portal/87964.

Genetic information

The estimated genome size, based on measurements across the Lottioidea superfamily ranges from 421Mb - 919Mb (Hinegardner, 1974), while the estimation of the haploid genome size based on reads kmer profiling is 726 Mb. This is a diploid genome with a haploid number of 9 chromosomes (2n=18). Information for this species was retrieved from Genomes on a Tree (Challis et al., 2023).

DNA/RNA processing

DNA was extracted from the mollusc foot using the Blood & Cell Culture DNA Mini Kit (Qiagen) following the manufacturer's instructions. DNA quantification was performed using a Qubit dsDNA BR Assay Kit (Thermo Fisher Scientific), and DNA integrity was assessed using a Femtopulse system (Genomic DNA 165 Kb Kit, Agilent). DNA was stored at 4°C until use.

RNA was extracted using an RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. RNA was extracted from two different specimen body parts: mollusc foot, gonad and other somatic tissue. RNA quantification was performed using the Qubit RNA BR Kit and RNA integrity was assessed using a Fragment Analyzer system (RNA 15nt Kit, Agilent). RNA was stored at -80°C until use.

Library preparation and sequencing

A long-read whole genome library was prepared using the SQK-LSK114 kit and sequenced on a PromethION P24 A series instrument (Oxford Nanopore Technologies). For short-read whole genome sequencing (WGS), a library was prepared using the KAPA Hyper Prep Kit (Roche). A Hi-C library preparation, using mollusc foot tissue, was conducted with the Dovetail Omni-C Kit (Cantata Bio) and further processed with the KAPA Hyper Prep Kit for Illumina sequencing (Roche). The RNA library, generated from mollusc foot, gonad and other somatic tissue, was prepared with the KAPA mRNA Hyper Prep Kit (Roche). All the short-read libraries were sequenced on the Illumina NovaSeq 6000 instrument. In total, 59x Oxford Nanopore, 65x Illumina WGS shotgun, and 147x Hi-C data were sequenced to generate the assembly.

Genome assembly methods

The genome was assembled using the CNAG CLAWS pipeline (Gomez-Garrido, 2024). Briefly, reads were preprocessed for quality and length using Trim Galore v0.6.7 and Filtlong v0.2.1, and initial contigs were assembled using NextDenovo v2.5.0, followed by polishing of the assembled contigs using HyPo v1.0.3, removal of retained haplotigs using purge-dups v1.2.6 and scaffolding with YaHS v1.2a. Finally, assembled scaffolds were curated via manual inspection using Pretext v0.2.5 with the Rapid Curation Toolkit (https://gitlab.com/wtsi-grit/rapid-curation) to remove any false joins and incorporate any sequences not automatically scaffolded into their respective locations in the chromosomal pseudomolecules (or super-scaffolds). Summary analysis of the released assembly was performed using the ERGA-BGE Genome Report ASM Galaxy workflow (De Panis, 2024a), incorporating tools such as BUSCO v5.5 and Merqury v1.3. NextDenovo was chosen over the assembler Flye as it generated much better contiguity, with a contig N50 of 8.1Mb after the initial assembly, vs 301kb contig N50 generated by Flye. Flye also generated a contig assembly with a high proportion of detected duplicated BUSCO genes (29.1%), compared to the 1.8% found in the initial NextDenovo contig assembly.

Genome annotation methods

A gene set was generated using the Ensembl Gene Annotation system (Aken et al., 2016), primarily by aligning publicly available short-read RNA-seq data from BioSample: SAMEA117648698 to the genome. Gaps in the annotation were filled via protein-to-genome alignments of a select set of vertebrate proteins from UniProt (The UniProt Consortium, 2019), which had experimental evidence at the protein or transcript level. At each locus, data were aggregated and consolidated, prioritising models derived from RNA-seq data, resulting in a final set of gene models and associated non-redundant transcript sets. To distinguish true isoforms from fragments, the likelihood of each open reading frame (ORF) was evaluated against known vertebrate proteins. Low-quality transcript models, such as those showing evidence of fragmented ORFs, were removed. In cases where RNA-seq data were fragmented or absent, homology data were prioritised, favouring longer transcripts with strong intron support from short-read data. The resulting gene models were classified into three categories: protein-coding, pseudogene, and long non-coding. Models with hits to known proteins and few structural abnormalities were classified as protein-coding. Models with hits to known proteins but displaying abnormalities, such as the absence of a start codon, non-canonical splicing, unusually small intron structures (<75 bp), or excessive repeat coverage, were reclassified as pseudogenes. Single-exon models with a corresponding multi-exon copy elsewhere in the genome were classified as processed (retrotransposed) pseudogenes. Models that did not fit any of the previously described categories did not overlap protein-coding genes, and were constructed from transcriptomic data were considered potential lncRNAs. Potential lncRNAs were further filtered to remove single-exon loci due to their

unreliability. Putative miRNAs were predicted by performing a BLAST search of miRBase (Kozomara et al., 2019) against the genome, followed by RNAfold analysis (Gruber et al., 2008). Other small non-coding loci were identified by scanning the genome with Rfam (Kalvari et al., 2018) and passing the results through Infernal (Nawrocki & Eddy, 2013). Summary analysis of the released annotation was performed using the ERGA-BGE Genome Report ANNOT Galaxy workflow (De Panis, 2024b), incorporating tools such as AGAT v1.2, BUSCO v5.5 and OMArk v0.3.

Results

Genome assembly

The genome assembly has a total length of 719,386,390 bp in 24 scaffolds including the mitogenome (Figures 2 and 3), with a GC content of 36.4%. It features a contig N50 of 8,519,583 bp (L50=22) and a scaffold N50 of 83,582,575 bp (L50=4). Purge-dups removed 133 sequences identified as retained haplotigs, error or repeat sequences totalling 70Mb. This further reduced the proportion of duplicated BUSCO genes detected from 1.8% to 0.3%. There are 133 gaps, totaling

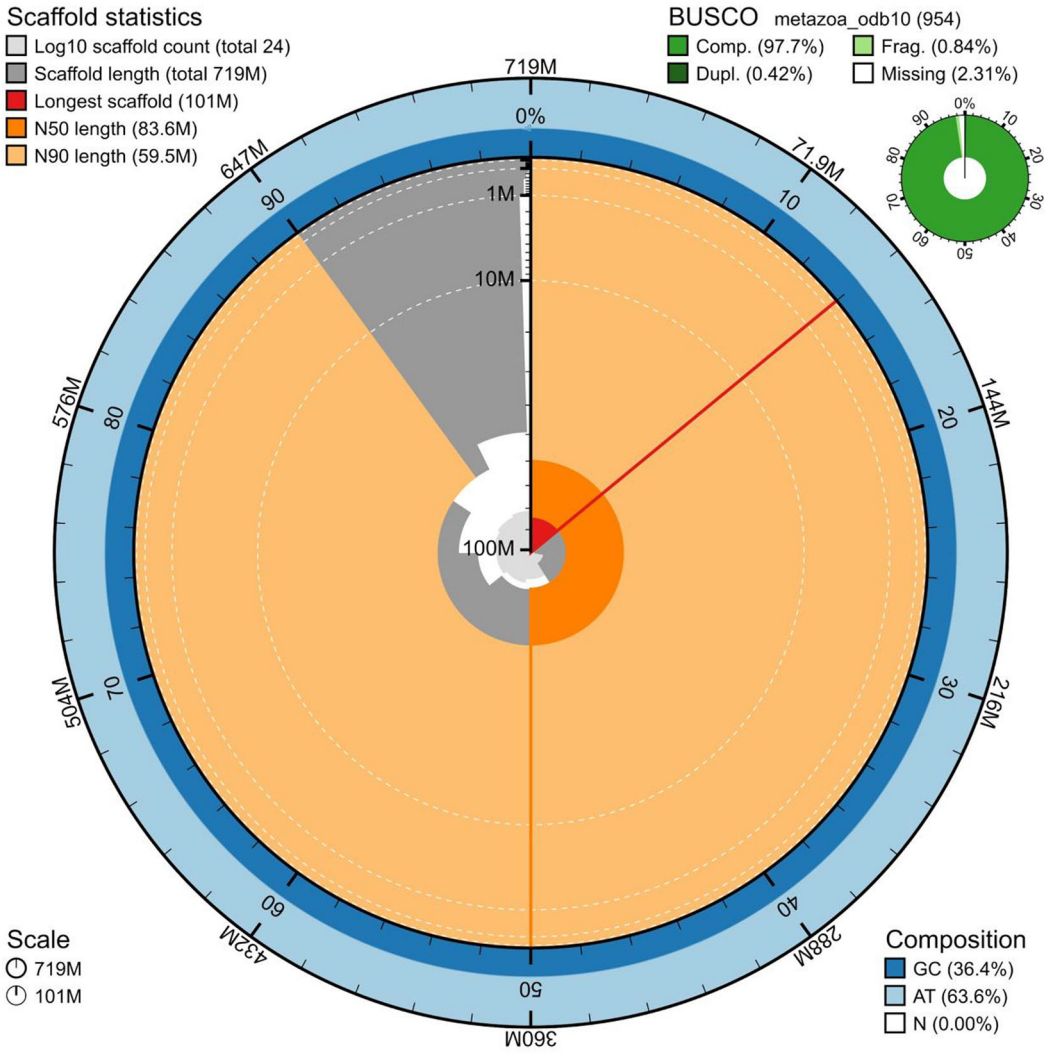

Figure 2. Snail plot summary of assembly statistics. The main plot is divided into 1,000 size-ordered bins around the circumference, with each bin representing 0.1% of the 719,386,390 bp assembly including the mitochondrial genome. The distribution of sequence lengths is shown in dark grey, with the plot radius scaled to the longest sequence present in the assembly (719,386,390 bp, shown in red). Orange and pale-orange arcs show the scaffold N50 and N90 sequence lengths 83,582,575 and 59,523,780 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 metazoa database (odb10) is shown on the top right.

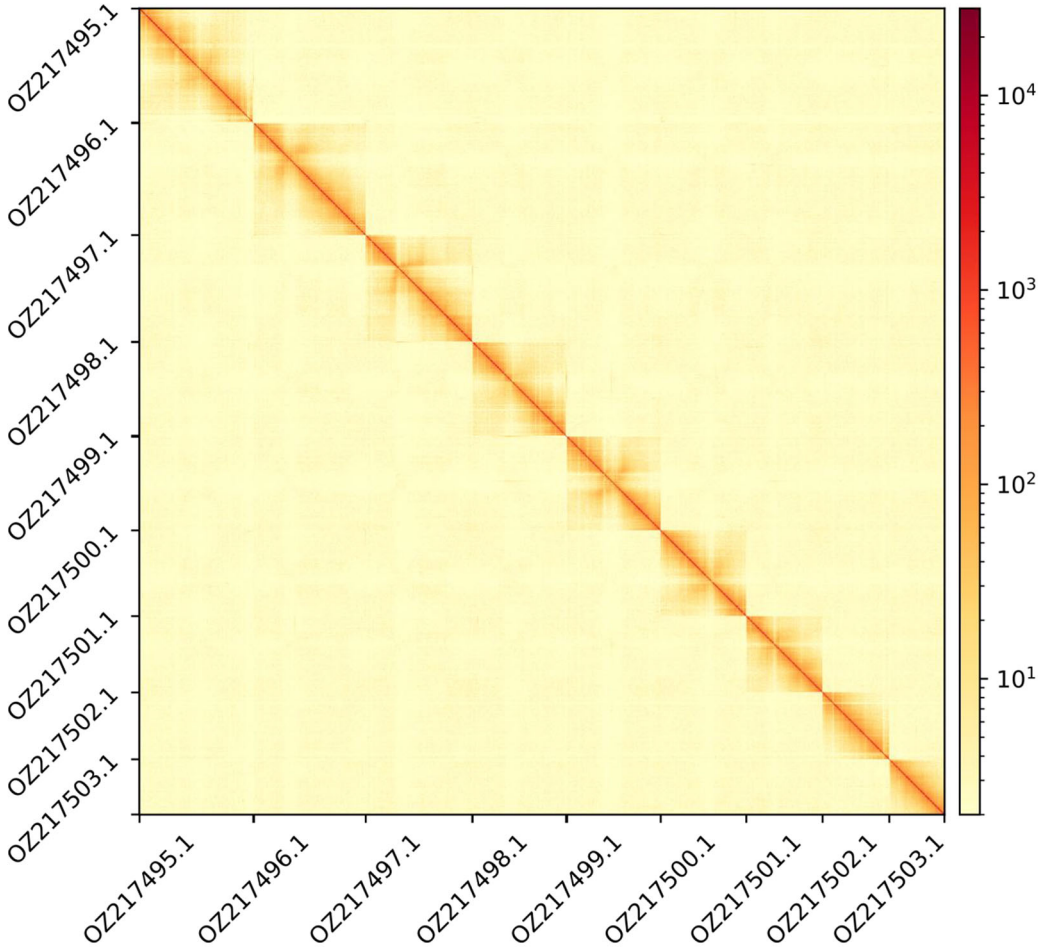

Figure 3. Hi-C contact map showing spatial interactions between regions of the genome. The diagonal corresponds to intra-chromosomal contacts, depicting chromosome boundaries. The frequency of contacts is shown on a logarithmic heatmap scale. Hi-C matrix bins were merged into a 400 kb bin size for plotting.

26.6 kb in cumulative size. The single-copy gene content analysis using the metazoa_odb10 database with BUSCO, run in genome assembly mode, resulted in 97.7% completeness (97.3% single and 0.4% duplicated). 81.41% of reads k-mers were present in the assembly and the assembly has a base accuracy Quality Value (QV) of 43.04 as calculated by Merqury. Overall 99.6% of the assembly was scaffolded into the 9 pseudo-chromosomal molecules.

Table 1. Statistics from assembled gene models.

	No. genes	No. transcripts	Mean* gene length (bp)	No. single-exon genes	Mean* exons per transcript
Protein-coding	20,636	37,053	18,629	834	8.9
lncRNA	13,909	19,144	16,846	5	2.2
snRNA	279	279	140	43	1.0
snoRNA	43	43	142	43	1.0
rRNA	1,120	1,120	221	1,120	1.0
tRNA	2,579	2,579	78	2,579	1.9
scRNA	1	1	133	1	1.0
Other non-coding	4,303	4,303	58-77	4,303	1.0

*Combined categories show the range of the mean values.

Table 2. Annotation completeness and consistency scores calculated by BUSCO run in protein mode (metazoa_odb) and OMArk (Lophotrochozoa).

	Complete	Singular	Duplicated	Fragmented	Missing
BUSCO	921 (96.5%)	915 (95.9%)	6 (0.6%)	9 (0.9%)	24 (2.5%)
OMArk	2,065 (95.91%)	1,918 (89.08%)	147 (6.83%)	-	88 (4.09%)
	Consistent	**Inconsistent**	**Contaminants**	**Unknown**	
OMArk	13,451 (65.18%)	1,914 (9.28%)	0 (0.00%)	5,271 (25.54%)	

Genome annotation

The genome annotation consists of 20,636 protein-coding genes with an associated 37,053 transcripts, in addition to 22,234 non-coding RNA genes of various types (Table 1). Using the longest isoform per transcript, the single-copy gene content analysis using the metazoa_odb10 database with BUSCO, run in protein mode, resulted in 93.2% completeness. Using the OMAmer Lophotrochozoa database for OMArk resulted in 95.91% completeness and 65.19% consistency (Table 2).

Author contributions

RF coordinated the project, RF and FL collected the species, FL and RNV identified the species, RF, FL and RNV sampled and preserved biological material and provided metadata, RF, RM, AB, TM, TS and RO provided sampling and metadata support and management, LA and MG extracted DNA, prepared libraries, and performed sequencing, FCC, JG-G and FC performed genome assembly and curation under the supervision of TA, LH and FM performed genome annotation, TB 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.

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, Emilie Teodori from Genoscope.

References

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

Barry P, Balmori-de la Puente A, Martinez IM, *et al.*: **Evolutionary processes in marine environments.** *eLS, Encyclopedia of Life Sciences.* 2024; **3**: a0029536.
Publisher Full Text

Broitman BR, Mieszkowska N, Helmuth B, *et al.*: **Climate and recruitment of rocky shore intertidal invertebrates in the eastern North Atlantic.** *Ecology.* 2008; **89**(11, Suppl. S): S81–S90.

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.
Publisher Full Text

Christiaens J: **Révision du genre *Patella* (Mollusca, Gastropoda).** *Bulletin du Museum National d'Histoire Naturelle, Paris.* 1973; **182**: 1305–1392.

Crisp DJ, Fischer-Piètte E: **Répartition des principales espèces intercotidales de la côte atlantique française en 1954–1955.** *Annales de l'Institut Océanographique, Paris.* 1959; **36**: 275–388.

De Panis D: **ERGA-BGE Genome Report ASM analyses (one-asm WGS Illumina PE + HiC).** *WorkflowHub.* 2024a.
Publisher Full Text

De Panis D: **ERGA-BGE Genome Report ANNOT analyses.** *WorkflowHub.* 2024b.
Publisher Full Text

Faria R, Chaube P, Morales HE, *et al.*: **Multiple chromosomal rearrangements in a hybrid zone between *Littorina saxatilis* ecotypes.** *Molecular Ecology.* 2019; **28**(6): 1375–1393.
PubMed Abstract | Publisher Full Text | Free Full Text

Faria R, Johannesson K, Stankowski S: **Speciation in marine environments: Diving under the surface.** *Journal of Evolutionary Biology.* 2021; **34**(1): 4–15.
PubMed Abstract | Publisher Full Text

Fischer-Piètte E: **Répartition le long des côtes septentrionales de l'Espagne, des principales espèces peuplant les rochers intercotidaux.** *Annales de l'Institut Océanographique, Paris.* 1955; **31**: 37–124.

Fischer-Piètte E, Gaillard J-M: **Les patelles au long des côtes atlantiques Ibériques et Nord-Marocaines.** *Journal of Conchology.* 1959; **99**: 135–200.

Firth LB: **What have limpets ever done for us? On the past and present provisioning and cultural services of limpets.** *International Review of Environmental History.* 2021; **7**(2): 5–45.
Publisher Full Text

Formenti G, Theissinger K, Fernandes C, *et al.*: **The era of reference genomes in conservation genomics.** *Trends in Ecology & Evolution.* 2022; **37**(3): 197–202.
Publisher Full Text

Gomez-Garrido J: **CLAWS (CNAG's long-read assembly workflow in Snakemake).** *WorkflowHub.* 2024.
Publisher Full Text

Gruber AR, Lorenz R, Bernhart SH, *et al.*: **The Vienna RNA Websuite.** *Nucleic Acids Research.* 2008; **36**(suppl_2), W70–W74.
PubMed Abstract | Publisher Full Text | Free Full Text

Hawkins SJ, Mieszkowska N, Mrowicki R: **The genome sequence of the common limpet, Patella vulgata (Linnaeus, 1758).** *Wellcome Open*

Research. 2023; **8**: 418.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Hawkins SJ, Mieszkowska N, Mrowicki R: **The genome sequence of the black-footed limpet, Patella depressa (Pennant, 1777).** *Wellcome Open Research.* 2024; **9**: 47.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Hawkins SJ, Sugden HE, Mieszkowska N, *et al.*: **Consequences of climate-driven biodiversity changes for ecosystem functioning of north European rocky shores.** *Marine Ecology Progress Series.* 2009; **396**: 245–259.
Publisher Full Text

Helmuth B, Mieszkowska N, Moore P, *et al.*: **Living on the edge of two changing worlds: Forecasting the responses of rocky intertidal ecosystems to climate change.** *Annual Review of Ecology, Evolution, and Systematics.* 2006; **37**(1): 373–404.
Publisher Full Text

Hinegardner R: **Cellular DNA content of the Mollusca.** *Comparative Biochemistry and Physiology Part A: Physiology.* 1974; **47**(2): 447–460.
Publisher Full Text

Hoffmann A, Sgrò C: **Climate change and evolutionary adaptation.** *Nature.* 2011; **470**: 479–485.
Publisher Full Text

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

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

Lawniczak MKN: **The genome sequence of the blue-rayed limpet, Patella pellucida Linnaeus, 1758.** *Wellcome Open Research.* 2022; **7**: 126.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Lima FP, Queiroz N, Ribeiro PA, *et al.*: **Recent changes in the distribution of a marine gastropod, *Patella rustica* Linnaeus, 1758, and their relationship to unusual climatic events.** *Journal of Biogeography.* 2006; **33**(5): 812–822.
Publisher Full Text

Lima FP, Ribeiro PA, Queiroz N, *et al.*: **Modelling past and present geographical distribution of the marine gastropod *Patella rustica* as a tool for exploring responses to environmental change.** *Global Change Biology.* 2007; **13**(10): 2065–2077.
Publisher Full Text

Marques JP, Alves PC, Amorim IR, *et al.*: **Building a Portuguese coalition for biodiversity genomics.** *NPJ Biodiversity.* 2024; **3**(1): 27.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

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

Mieszkowska N, Hawkins SJ, Côrte-Real HBSM, *et al.*: **The genome sequence of the China Limpet, Patella ulyssiponensis Gmelin, 1791 (Patellidae).** *Wellcome Open Research.* 2025; **10**: 419.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Morales HE, Faria R, Johannesson K, *et al.*: **Genomic architecture of parallel ecological divergence: Beyond a single environmental contrast.** *Science Advances.* 2019; **5**(12): eaav9963.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

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

Petraccioli A, Guarino FM, Maio N, *et al.*: **Molecular cytogenetic study of three common Mediterranean limpets, *Patella caerulea*, *P. rustica* and *P. ulyssiponensis* (Archaeogastropoda, Mollusca).** *Genetica.* 2010; **138**(2): 219–225.
PubMed Abstract | **Publisher Full Text**

Potkamp G, Fransen CH: **Speciation with gene flow in marine systems.** *Contributions to Zoology.* 2019; **88**(2): 1–40.
Publisher Full Text

Prusina I, Ezgeta-Balić D, Ljubimir S, *et al.*: **On the reproduction of the Mediterranean keystone limpet Patella rustica: histological overview.** *Journal of the Marine Biological Association of the United Kingdom.* 2014; **94**(8): 1651–1660.
Publisher Full Text

Ribeiro PA, Branco M, Hawkins SJ, *et al.*: **Recent changes in the distribution of a marine gastropod, *Patella rustica*, across the Iberian Atlantic coast did not result in diminished genetic diversity or increased connectivity.** *Journal of Biogeography.* 2010; **37**(9): 1782–1796.
Publisher Full Text

Somero GN: **The physiology of climate change: How potentials for acclimatization and genetic adaptation will determine 'winners' and 'losers'.** *Journal of Experimental Biology.* 2010; **213**(6): 912–920.
PubMed Abstract | **Publisher Full Text**

Sá-Pinto A, Baird SJ, Pinho C, *et al.*: **A three-way contact zone between forms of *Patella rustica* (Mollusca: Patellidae) in the central Mediterranean Sea.** *Biological Journal of the Linnean Society.* 2010; **100**(1): 154–169.
Publisher Full Text

Sokolova IM, Frederich M, Bagwe R, *et al.*: **Energy homeostasis as an integrative tool for assessing limits of environmental stress tolerance in aquatic invertebrates.** *Marine Environmental Research.* 2012; **79**: 1–15.
PubMed Abstract | **Publisher Full Text**

Sousa LL, Seabra R, Wethey DS, *et al.*: **Fate of a climate-driven colonisation: Demography of newly established populations of the limpet *Patella rustica* Linnaeus, 1758, in northern Portugal.** *Journal of Experimental Marine Biology and Ecology.* 2012; **438**: 68–75.
Publisher Full Text

Teske PR, Papadopoulos I, Mmonwa KL, *et al.*: **Climate-driven genetic divergence of limpets with different life histories across a southeast African marine biogeographic disjunction: Different processes, same outcome.** *Molecular Ecology.* 2011; **20**(23): 5025–5041.
PubMed Abstract | **Publisher Full Text**

Theissinger K, Fernandes C, Formenti G, *et al.*: **How genomics can help biodiversity conservation.** *Trends in Genetics.* 2023; **39**(7): 545–559.
Publisher Full Text

Trégarot E, D'Olivo JP, Botelho AZ, *et al.*: **Effects of climate change on marine coastal ecosystems: A review to guide research and management.** *Biological Conservation.* 2024; **289**: 110394.
Publisher Full Text

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

Weber LI, Hawkins SJ: ***Patella aspera* and *P. ulyssiponensis*: Genetic evidence of speciation in the north-east Atlantic.** *Marine Biology.* 2005; **147**(1): 153–162.
Publisher Full Text

Westram AM, Rafajlović M, Chaube P, *et al.*: **Clines on the seashore: The genomic architecture underlying rapid divergence in the face of gene flow.** *Evolution Letters.* 2018; **2**(4): 297–309.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Zegaoula B, Beldi H, Draredja B, *et al.*: **Reproduction of Patella rustica (Mollusca, Gastropoda) in the gulf of Annaba (Algeria, Mediterranean South Western).** *Advances in Environmental Biology.* 2016; **10**(11): 42–51.

Open Peer Review

Current Peer Review Status: ✔ ? ✔

Version 2

Reviewer Report 12 March 2026

https://doi.org/10.21956/openreseurope.24968.r70969

© **2026 Wilding C.** 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.

✔ **Craig Wilding**
University College Dublin, Dublin, Ireland

The authors present a high-quality chromosomally anchored reference genome for *Patella rustica* which will, as the authors state in the introduction, provide a valuable resource for both empowering population genomics in this species, but also facilitate comparative genomic analysis within the limpets.

I have only minor comments:

Materials and methods:

I thought the sub-heading 'Genetic information' to be unhelpful since this section deals only with genome size estimation, and perhaps therefore this is a more useful heading. The first sentence of this section "The estimated genome size, based on measurements across the Lottioidea superfamily ranges from 421Mb - 919Mb (Hinegardner, 1974)" could perhaps be more clearly worded since, presumably, this large range in values is not the estimated genome size of this species. Perhaps it could be worded more clearly as "There is no published estimate of genome size for *Patella rustica*, but measurements across the Lottioidea superfamily range from 421Mb - 919Mb (Hinegardner, 1974)"

DNA/RNA processing section: "RNA was extracted from two different specimen body parts: mollusc foot, gonad and other somatic tissue." Isn't this three different body parts?

Results:
Genome annotation section - needs a citation for "OMAmer Lophotrochozoa database for OMArk" both here, and in the methods.

Table 1 - please be clear on that "Mean * gene length" includes introns.

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

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

Version 1

Reviewer Report 11 December 2025

https://doi.org/10.21956/openreseurope.23376.r64049

© 2025 ISHII Y. 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.

?

Yasuto ISHII
[1] Tohoku University, Sendai, Japan
[2] Tohoku University, Sendai, Japan

GENERAL COMMENTS
The present study released a genome assembly of *Patella rustica*. The assembly have high continuity and completeness, so thereby it should be valuable genetic resource. Also, the authors provided a high-quality gene annotation which is also a promising genetic resource. The assembly and annotation are available from a public database. Collectively, the assembly and annotation should promote further genomic research in *Patella* snails and other related species. Yet, this manuscript does not contain fundamental information, such as details of pre-Hi-C scaffolding assembly and how much sequences were anchored on the Hi-C scaffolds. Also, some fruitful analyses, including repeat masking, are lack in the current version. I listed some minor comments and suggestions below.

COMMENTS
Title and abstract: ERGA-BGE is not a common abbreviation. The full name should be addressed in

the abstract. Also, I suggest that the word "chromosome-level assembly" should be included in the title to be appealing to readers. If the word limit is not strict, sequencing methods (e.g., the use of long-read, read type [ONT+Hi-C+short-read]) and BUSCO values should be noted in the abstract. This information ensures the validity of the assembly. Furthermore, the taxonomic information (e.g., family, class; I'm not familiar with limpets, so not sure what rank is useful) is helpful for readers to reach this article.

In the introduction, while the authors stated that the species is serve as a "sentinel", the reason why this species particularly has such a character.

Add appropriate reference(s) for the following sentence: "Several studies have documented recent changes in its distribution, particularly the bridging of a historical distribution gap in northwestern Iberia."

How did you detect the sex?: "female sample of *Patella rustica* was sampled"

I suggest that you should do the genome size estimation by yourself. In my opinion, articles should be self-contained if possible.: "the estimation based on reads kmer profiling is 726 Mb"

I cannot grasp the meaning of this sentence. Does it mean that the estimation is for the haploid genome?: "This is a diploid genome with a haploid number of 9 chromosomes (2n=18) "

To be clear, did you perform genome annotation without repeat masking? One intriguing feature of molluscan genomes are the evolution of repeat elements. If you can do it, it should be a valuable resource. Also, while I'm not familiar with the gene annotation pipeline, generally repeat masking improves the quality of gene annotations.

In Table 2, you reported the proportion of contaminations is 0.05 (lack of "%"?) but the number of contaminated transcripts is zero. It looks weird. The proportion should be 0.00%; otherwise, the number of contaminated transcripts is wrong.

More details of the assembly should be reported, including the statistics of the assembly before Hi-C scaffolding. Moreover, according to Fig. 3, the sequences look to be anchored on nine mega-scaffolds, but it is not clear due to the lack of the mention in the main text. If so, the number of mega-scaffolds is matched with the haploid number of chromosomes of this species (documented in the method section); this information is noteworthy. It is also unclear how much sequences are anchored on the mega-scaffolds.

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

Partly

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: Genomics on molluscs

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, however I have significant reservations, as outlined above.

Author Response 03 Feb 2026

Thomas Brown

GENERAL COMMENTS The present study released a genome assembly of Patella rustica. The assembly have high continuity and completeness, so thereby it should be valuable genetic resource. Also, the authors provided a high-quality gene annotation which is also a promising genetic resource. The assembly and annotation are available from a public database. Collectively, the assembly and annotation should promote further genomic research in Patella snails and other related species. Yet, this manuscript does not contain fundamental information, such as details of pre-Hi-C scaffolding assembly and how much sequences were anchored on the Hi-C scaffolds. Also, some fruitful analyses, including repeat masking, are lack in the current version. I listed some minor comments and suggestions below.

Thank you for the suggestions. We have added to the text following the guidelines of the reviewer:

COMMENTS
Title and abstract: ERGA-BGE is not a common abbreviation. The full name should be addressed in the abstract. Also, I suggest that the word "chromosome-level assembly" should be included in the title to be appealing to readers. If the word limit is not strict, sequencing methods (e.g., the use of long-read, read type [ONT+Hi-C+short-read]) and BUSCO values should be noted in the abstract. This information ensures the validity of the assembly. Furthermore, the taxonomic information (e.g., family, class; I'm not familiar with limpets, so not sure what rank is useful) is helpful for readers to reach this article.
Thank you for the suggestions. We have added "Chromosome-Level" to the title, further details of the class and family (Gastropoda, Patellidae), as well as the sequencing recipe and the project and consortia to the abstract.

In the introduction, while the authors stated that the species is serve as a "sentinel",

the reason why this species particularly has such a character.
We would like to clarify that this is not a particular attribute of our focal species and that the term was not coined by us. We only use it in the introduction to refer to other intertidal species, in the sense that they can inform us about environmental change, for instance, via shifts in their distribution ranges as a consequence of global warming (as mentioned in the introduction and supported by references). We hope that this is clear in the main text and no further changes are needed.

Add appropriate reference(s) for the following sentence: "Several studies have documented recent changes in its distribution, particularly the bridging of a historical distribution gap in northwestern Iberia.
Thank you for the suggestion. We added "(Fischer-Piètte and Gaillard, 1959; Lima et al., 2006)" at the end of the sentence. And added the first reference to the final list.

How did you detect the sex?: "female sample of Patella rustica was sampled"
The sex of the individual was determined macroscopically as they are easily distinguishable by gonad coloration: females have green or brown gonads, while males present a pinkish-cream color (based on previous work describing the reproductive cycles of the species, with some examples in Zegaoula 2016 and Prusina 2014). We added the following information in one of the subsequent sections of the same section with the new text underlined: After dissection and sexing (based on gonad colouration) by Rocío Nieto Vilela, the specimen's tissues (mollusc foot, gonad, head, hepatopancreas and other somatic tissue) Refs Zegaoula, B., Beldi, H., Draredja, B., & Soltani, N. (2016). Reproduction of Patella rustica (Mollusca, Gastropoda) in the gulf of Annaba (Algeria, Mediterranean South Western). Advances in Environmental Biology, 10(11), 42-51. Prusina, I., Ezgeta-Balić, D., Ljubimir, S., Dobroslavić, T., & Glamuzina, B. (2014). On the reproduction of the Mediterranean keystone limpet Patella rustica: histological overview. Journal of the Marine Biological Association of the United Kingdom, 94(8), 1651-1660.

I suggest that you should do the genome size estimation by yourself. In my opinion, articles should be self-contained if possible.: "the estimation based on reads kmer profiling is 726 Mb"
We performed genome size estimation using the sequenced genomic DNA, estimating the size based on total coverage and ratio of heterozygous - homozygous kmers. To add context in the manuscript, we have added the range of genome sizes identified by Hinegardner, however performing genome size calculation from scratch is not within the scope of this study.
I cannot grasp the meaning of this sentence. Does it mean that the estimation is for the haploid genome?: "This is a diploid genome with a haploid number of 9 chromosomes (2n=18) "
We apologise for the confusion and have explicitly stated that the size estimates correspond to the haploid genome size.

To be clear, did you perform genome annotation without repeat masking? One intriguing feature of molluscan genomes are the evolution of repeat elements. If you can do it, it should be a valuable resource. Also, while I'm not familiar with the gene annotation pipeline, generally repeat masking improves the quality of gene

annotations.
The gene annotation pipeline used did not make use of masked repetitive elements, however soft- and hard-masked versions of the genome assembly are available via the Ensembl download site alongside the annotation files.

In Table 2, you reported the proportion of contaminations is 0.05 (lack of "%"?) but the number of contaminated transcripts is zero. It looks weird. The proportion should be 0.00%; otherwise, the number of contaminated transcripts is wrong.
Thank you for noticing our typo - which came from 5 and % sharing a key - this has been corrected in the table.

More details of the assembly should be reported, including the statistics of the assembly before Hi-C scaffolding. Moreover, according to Fig. 3, the sequences look to be anchored on nine mega-scaffolds, but it is not clear due to the lack of the mention in the main text. If so, the number of mega-scaffolds is matched with the haploid number of chromosomes of this species (documented in the method section); this information is noteworthy. It is also unclear how much sequences are anchored on the mega-scaffolds.
We have added details as to the choice of assembler in the methods section and the amount of sequence removed by the purge-dups step in creating the assembly for a single haplotype. We have furthermore added the total proportion of the assembly anchored to the pseudo-chromosomal molecules to the results section.

Competing Interests: No competing interests were disclosed.

Reviewer Report 24 November 2025

https://doi.org/10.21956/openreseurope.23376.r64050

© 2025 Tennessen J. 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.

Jacob A Tennessen
[1] Harvard T. H. Chan School of Public Health, Boston, Massachusetts, USA
[2] Harvard T. H. Chan School of Public Health, Boston, Massachusetts, USA

This is a nicely assembled and annotated genome that will be a useful resource for the mollusk genomics community. Assembly and annotation methods are appropriately described, albeit with minimal details.

The genome is from an outbred diploid individual, and it would be nice to see more details on how heterozygous sites and more complex structural polymorphisms were handled. Presumably they were arbitrarily collapsed and it says haplotigs were removed. Is there an estimate of genome-wide heterozygosity (e.g. from alignment of the short reads to the final assembly)? Is it possible

that some contigs/scaffolds are allelic copies of each other?

The authors mention the genomes of five other *Patella* species. These should be cited in this paper (if no publication exists, then link to their accessions on ENA or https://beta.ensembl.org). The utility of this genome will largely be found in the context of those congener genomes: comparative work will allow errors to be identified and corrected, while also revealing informative evolutionary patterns. It would be interesting to see such comparative work included here, though I understand if such analysis is beyond the scope of the funded project.

Change (Hoffmann, 2011) to (Hoffmann & Sgrò, 2011)

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, including genomics of snails

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 03 Feb 2026

Thomas Brown

Review: This is a nicely assembled and annotated genome that will be a useful resource for the mollusk genomics community. Assembly and annotation methods are appropriately described, albeit with minimal details. The genome is from an outbred diploid individual, and it would be nice to see more details on how heterozygous sites and more complex structural polymorphisms were handled. Presumably they were arbitrarily collapsed and it says haplotigs were removed. Is there an estimate of genome-wide heterozygosity (e.g. from alignment of the short reads to the final assembly)? Is it possible that some contigs/scaffolds are allelic copies of each other? As much as possible, any retained sequences from the other haplotype were removed both during the purge-dups and manual curation steps. An analysis of any heterozygosity levels or regions is not considered part of this data note. The authors mention the genomes of five other Patella species. These should be cited in this paper (if no publication exists, then link to their accessions on ENA or https://beta.ensembl.org). The utility of this genome will largely be found in the context of those congener genomes: comparative work will allow errors to be identified and

corrected, while also revealing informative evolutionary patterns. It would be interesting to see such comparative work included here, though I understand if such analysis is beyond the scope of the funded project.

Response: Thank you for the suggestion. We have included the citations for 4 papers where the genomes have been published in a corresponding article. For the 5th genome, we have included the genbank accession. Change (Hoffmann, 2011) to (Hoffmann & Sgrò, 2011) Thank you for noticing the error - we have corrected the citation in the text

Competing Interests: No competing interests were disclosed.