GBE

Nuclear and Mitochondrial Genome Assemblies of *Indrella ampulla*, a Terrestrial Gastropod Endemic to the Western Ghats

Gopi Krishnan [1,2], Maitreya Sil [3,4,5,*], Neelavar Anantharam Aravind [4,6,7], Govindhaswamy Umapathy [1,2], Aniruddha Datta-Roy [3]

[1]CSIR- Centre for Cellular and Molecular Biology (CSIR-CCMB), Hyderabad, India

[2]Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, India

[3]School of Biological Sciences, National Institute of Science Education and Research (NISER), An OCC of Homi Bhabha National Institute, Khorda, Odisha, India

[4]S.M. Sehgal Foundation Centre for Biodiversity and Conservation, Ashoka Trust for Research in Ecology and the Environment (ATREE), Bangalore 560064, India

[5]Center for Urban Ecology, Biodiversity, Evolution, and Climate Change, JAIN (Deemed-to-be University), Bengaluru, Karnataka 560078, India

[6]Yenepoya Research Centre, Yenepoya (Deemed to be University), Derlakatte, Mangalore 575018, India

[7]Manipal Academy of Higher Education (MAHE), Manipal 576 104, India

*Corresponding author: E-mail: maitreya.sil@gmail.com.

Accepted: October 09, 2025

Abstract

Indrella ampulla is a polymorphic, brightly colored, forest-dwelling land snail endemic to the Western Ghats of India. Studying and understanding evolutionary processes occurring within this species has remained a challenge largely due to a paucity of genomic resources. We present high-quality annotated nuclear and mitochondrial genome assemblies of *I. ampulla*. The nuclear genome is assembled through a hybrid approach using Illumina short reads and Oxford Nanopore long reads, with an N50 value of 632 kb and 93.5% BUSCO genome completeness. The mitogenome is 13,887 bp long. The demographic history reconstruction based on the genomic data exhibits signatures of population decline during the last 100,000 years. This genome will aid in deciphering the color polymorphism in *I. ampulla* and augmenting the general understanding of the evolution of gastropods.

Key words: hybrid genome assembly, mitogenome, land snail, Western Ghats, demographic history.

Significance

Indrella ampulla is a monotypic terrestrial gastropod endemic to the Western Ghats biodiversity hotspot and exhibits color polymorphism. The genome of *I. ampulla* provides a valuable resource for understanding molluscan biology, coloration genetics, adaptation, and other evolutionary processes. Our historical demographic analysis offers insights into past population dynamics of this species endemic to the Western Ghats of India, contributing to broader research on biodiversity patterns in tropical ecosystems. The bioinformatics pipeline used in this study provides a reliable framework for assembling high-quality genomes with long-read coverage as low as 10×, making it especially useful for gastropods and other species where obtaining high coverage data is a challenge. Overall, this genome will complement future studies in phylogenomics, comparative genomics, and conservation of mollusks.

© The Author(s) 2025. Published by Oxford University Press on behalf of Society for Molecular Biology and Evolution.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Downloaded from https://academic.oup.com/gbe/article/17/11/evaf199/8303493 by guest on 03 December 2025

Introduction

Gastropod mollusks are important model systems in evolutionary biology and are the only molluscan group that occupies a wide range of habitats, including marine, estuarine, freshwater, and terrestrial environments (Hayes et al. 2009; Ponder et al. 2019). They exhibit unique traits such as torsion (where internal organs twist along an axis), the presence of toxins, polychromatism in both their shells and their bodies, and a coiled shell—a feature that has been reduced or lost in certain lineages. These adaptations, along with diverse feeding strategies, including herbivory, predation, and parasitism, make gastropods an ecologically and evolutionarily significant group.

Despite their considerable diversity (Aravind et al. 2005), genomic resources for gastropods, particularly those from South Asia, remain limited. This lack of genomic resources hinders the ability to explore key evolutionary and ecological questions within this group. Furthermore, invasive species pose a significant threat to native biodiversity, and gastropods are one of the most prominent invasive species worldwide. To address this gap, we present the first high-quality hybrid de novo genome assembly of a monotypic terrestrial gastropod species, *Indrella ampulla* (Ariophantidae), a large, semislug endemic to the central and southern Western Ghats mountain ranges of India (Surya Narayanan and Aravind 2021; Chakraborty et al. 2024).

The species, *I. ampulla*, exhibits three brightly colored morphs (orange, red, and yellow), a trait of significant biological interest (Chakraborty et al. 2024). Coloration plays a crucial role in the animal kingdom, serving various ecological and evolutionary functions. While cryptic coloration helps animals evade predation by blending into their surroundings, bright coloration often signals sexual maturity and plays a key role in sexual selection. Additionally, it can function as a warning signal indicating unpalatability or toxicity (Cuthill et al. 2017). The genetic basis of complex traits such as coloration often involves multiple, sometimes previously unknown, genes (San-Jose and Roulin 2017).

Here, we report high-quality nuclear and mitochondrial genome assemblies of *I. ampulla* using a hybrid approach that integrates Illumina short and Oxford Nanopore long reads. This genome provides a foundational resource for studying gastropod evolution, genetics of coloration, and adaptation in terrestrial environments.

Methods

Sample Collection and Sequencing

The focal species, *I. ampulla*, is distributed in the Southern and Central Western Ghats, which spans across parts of Karnataka, Tamil Nadu, and Kerala states in Southern India (Surya Narayanan and Aravind 2021). For genome assembly, we selected an individual belonging to the red color morph. The specimen was collected from a cardamom plantation in Nedumkandam, Kerala, and preserved in absolute ethanol. Genomic DNA was extracted from the foot muscle tissue using a modified cetyltrimethylammonium bromide) extraction protocol (Chakraborty et al. 2020). DNA concentration was assessed both visually on an agarose gel and quantitatively using a Qubit fluorometer. To generate short-read data, libraries were prepared using Illumina's TruSeq DNA PCR-Free kit, and sequencing was performed on an Illumina NovaSeq 6000 device at Novelgene Inc., Hyderabad. Libraries for long-read sequencing were prepared using Oxford Nanopore Technology's Ligation Sequencing Kit V14 (SQK-LSK114), and sequencing was carried out on a PromethION device at CSIR-Centre for Cellular and Molecular Biology, Hyderabad.

Nuclear Genome Assembly

We generated 21.3 Gb of long-read data (3,393,199 reads) and 136 Gb of short-read data (446 million paired-end reads). Short reads were quality-trimmed using fastp v0.23.2 (Chen 2023), while the long reads were base-called with Guppy Basecaller and error-corrected using the trimmed short reads with fmlrc2 (Mak et al. 2023). Both corrected long-read and short-read datasets were then used for genome assembly. First, we performed a long-read-only assembly using Flye v2.9.4-b1799 (Kolmogorov et al. 2019), resulting in Assembly 1. Next, we generated a hybrid genome assembly incorporating both short- and long-read data using MaSuRCA v 4.1.0 (Zimin et al. 2013), producing Assembly 2. Both preliminary assemblies were further polished using the corrected short-read data with Polca pipeline of MaSuRCA. To improve contiguity, we performed two rounds of QuickMerge v0.3 (Chakraborty et al. 2016), combining Assemblies 1 and 2 in different configurations. In the first round, each assembly was used as both reference and query in separate runs. The resulting assemblies were then merged again in a second round of QuickMerge. The specific assembly combinations and their metrics are detailed in Table S1.

Assembly quality was assessed using Quast v5.2.0 (Gurevich et al. 2013), and completeness was evaluated with BUSCO v6.0.0 (Manni et al. 2021; Tegenfeldt et al. 2025), against the Mollusca database (mollusca_odb12). The final assembly was selected based on an optimal balance of BUSCO completeness and contiguity. The assembly was screened for contamination using NCBI Foreign Contamination Screen (FCS) (Astashyn et al. 2024), and the identified prokaryotic sequences were subsequently removed. The assembly summary was visualized using BlobToolKit v4.4.0 (Challis et al. 2020). Short reads were mapped to the final assembly using BWA-MEM2 (v2.2.1) (Vasimuddin et al. 2019) with default parameters, and alignment files were processed and summarized with SAMtools (v1.18) (Danecek et al. 2021) to evaluate the mapping rate. Assembly quality was further assessed using

Downloaded from https://academic.oup.com/gbe/article/17/11/evaf199/8303493 by guest on 03 December 2025

Table 1 Summary statistics of the final genome and mitogenome assembly and annotation

Category	Feature	Value
Sequencing data	Nanopore reads (no. of reads/Gb/coverage)	3.39 M/21.3 Gb/11×
	Illumina reads (no. of PE reads/Gb/coverage)	449 M/136 Gb/69×
Assembly statistics	Assembly size (Gb)	1.98
	Number of contigs	8782
	N50/N90 (bp)	632,279/95,350
	L50/L90	923/3810
BUSCO (mollusca_odb12)	Assembly BUSCO (%)	C: 93.5 (S: 72.0, D: 21.5); F: 3.8; M: 2.7
	Predicted gene set BUSCO (%)	C: 90.1 (S: 73.2, D: 16.9); F: 6.1; M: 3.8
Annotation	Protein-coding genes	20,418
	Functionally annotated proteins (%)	87.9
Repeat elements	Total repeat content (%)	53.69
	SINEs (%)	1.01
	LINEs (%)	25.94
	LTR elements (%)	0.28
	DNA transposons (%)	4.42
	Rolling circles (%)	0.08
	Unclassified (%)	18.58
	Small RNA (%)	0.93
	Satellites (%)	0.08
	Simple repeats (%)	2.10
	Low complexity (%)	0.27
Mitogenome statistics	Mitogenome size (bp)	13,887
	Number of protein-coding genes	13
	Number of rRNA genes	2
	Number of tRNA genes	22

Statistics of the final genome and mitogenome assembly and annotation.

Merqury (v1.3) and Meryl (v1.3) (Rhie et al. 2020) to estimate consensus accuracy (QV) and base-level error rate.

Repeat Masking and Genome Annotation

We first generated a de novo repeat library for the genome using RepeatModeler v2.0.5 (Flynn et al. 2020). The genome was then soft-masked with this custom library using RepeatMasker v4.1.4 (Smit et al. 2013–2015) in conjunction with the LTRstruct pipeline. For gene prediction, we used GALBA v1.0.11 (Bruna et al. 2023), incorporating 1.18 million molluscan protein sequences from the UniProt database as external evidence. The resulting annotation was refined by removing unsupported genes and retaining only the longest isoforms. Proteome completeness was assessed using BUSCO in protein mode against the Mollusca dataset, while the proteome quality was evaluated with OMArk v0.3.0 (Nevers et al. 2025) against the Luca.h5 dataset. Functional annotation of the proteome was performed using EggNOG-mapper (emapper v2.1.12; Cantalapiedra et al. 2021), utilizing EggNOG orthology data (eggNOG 5.0; Huerta-Cepas et al. 2019) for comprehensive annotation.

Mitogenome Assembly

We assembled the mitochondrial genome using MitoHiFi v3.2.2 (Uliano-Silva et al. 2023) with error-corrected long-read data. We annotated it de novo using MitoZ v3.6 (Meng et al. 2019) and further refined the tRNA prediction with GeSeq (Tillich et al. 2017) and ARWEN (Laslett and Canbäck 2008). We removed false-positive tRNAs predicted within protein-coding sequences and combined the results from both pipelines. Finally, we visualized the mitogenome using OGDRAW (Greiner et al. 2019).

Demographic History Reconstruction

We used pairwise sequentially Markovian coalescent (PSMC) analysis (Li and Durbin 2011) using psmc v0.6.5-r67 to reconstruct the historical demography of *I. ampulla*. First, we generated a diploid consensus sequence by aligning short-read data to the reference genome (Assembly 6) using BCFtools (Danecek et al. 2021). We filtered variants with bcftools view -c - and converted the output to FASTQ format using vcfutils.pl, setting depth thresholds to a minimum of 6 and a maximum of 48. We then performed PSMC analysis with parameters -N25 -t15 -r5 -p "4 + 25*2 + 4 + 6 and with 100 bootstrap replicates.

Results and Discussion

Nuclear Genome Assembly Statistics

We present the first genome assembly of *I. ampulla*, constructed using 21 GB of error-corrected long-read data

Downloaded from https://academic.oup.com/gbe/article/17/11/evaf199/8303493 by guest on 03 December 2025

and 118 GB of filtered short-read data. We generated initial assemblies and refined them through two rounds of QuickMerge with different assembly combinations, resulting in a total of eight assemblies (see Table S1 and Fig. S1 for detailed statistics). Among these, we selected Assembly 6 as the final assembly based on its optimal balance of contiguity and completeness. NCBI's FCS had removed ten contigs, and the final assembly spans 1.98 GB, consisting of 8,782 contigs with an N50 value of 632 kb and 36.35% GC content (see Table 1 and Fig. S2). The genome size of *I. ampulla* falls well within the known range for mollusks, which exhibit considerable variation, from less than 1 GB in *Biomphalaria glabrata* (850.6 Mb; accession no. GCA_947242115.1) and *Elysia timida* (754.4 Mb; accession no. GCA_043644045.1) to over 5 GB in *Oreohelix idahoensis* (5.5 Gb; accession no. GCA_024509875.1). This diversity highlights the broad range of genome sizes within the parent subclass Heterobranchia. BUSCO analysis indicated that the genome is 93.5% complete, with 21.5% of genes found to be duplicated. Such high levels of duplication have also been observed in other species of the order Stylommatophora, such as 12.9% in *Cepaea nemoralis* (Saenko et al. 2021), 15.90% in *Meghimatium bilineatum* (Sun et al. 2024), 16.3% in *Deroceras laeve* (Miranda-Rodríguez et al. 2025), and 7.1% in *Bradybaena similaris* (Ishii et al. 2025), among others. Saenko et al. (2021) attribute such duplication to the high heterozygosity of the genome, whereas Chen et al. (2022) suggested that a whole genome duplication event occurred at the base of Stylommatophora between 93 and 109 Mya, potentially facilitating their adaptive radiation to terrestrial conditions. These explanations closely align with the high duplication of BUSCOs that we have observed in the *I. ampulla* genome. Additionally, generation of greater amounts of long-read data in the future would help in producing a more contiguous and complete assembly. Also, the high mapping rate of 99.49% indicates that the majority of short reads are well represented in the assembly. Merqury analysis showed a high consensus accuracy (QV = 42.8), corresponding to a low base-level error rate of approximately 1 in 18,910 bases.

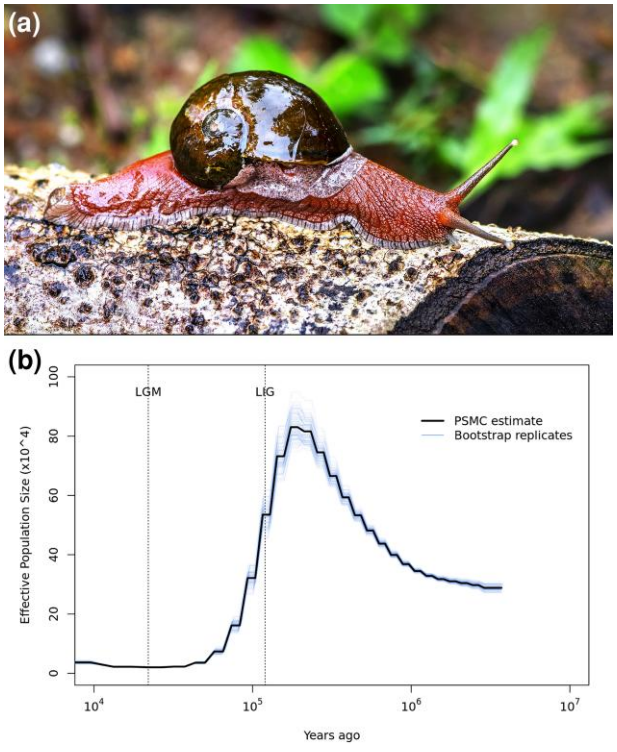

Fig. 1. a) A red morph individual of *I. ampulla*. (b) PSMC plot illustrating historical effective population size trends (dark line) of *I. ampulla* over time. The *x* axis represents time in years before the present, scaled by the generation time (g) and mutation rate (μ), and the *y* axis shows the effective population size (N_e). LGM, Last Glacial Maximum; LIG, Last Interglacial period.

Repeat Masking and Annotation

Repeat annotation revealed that approximately 54% of the genome consists of repeat elements. The major categories of repeats include SINEs, LINEs, LTR elements, DNA transposons, rolling circles, and simple repeats. LINEs constituted the largest proportion, accounting for 25.94% of the genome, nearly half of all repeat content (see Table S2 for complete repeat statistics). Gene prediction using GALBA identified a total of 20,418 protein-coding genes. BUSCO analysis of the predicted proteome indicated 90.1% completeness with a duplication rate of 16.9%, consistent with the high duplication levels reported in the proteomes of other Stylommatophora species, such as 17.7% in *M. bilineatum* (Sun et al. 2024) and 8.5% in *Candidula unifasciata* (Chueca et al. 2021). A more contiguous assembly in the future will help in predicting the proteome with more completeness. Consistency analysis using OMArk showed that 68.03% of the proteins were consistently placed within the clade Mollusca, while 20.24% were classified as unknown (Fig. S3). Importantly, OMArk predicted no contaminants in the genome. Out of the 20,418 predicted protein-coding genes, 17,952 (87.9%) were functionally annotated using EggNOG-mapper.

Mitogenome Assembly Statistics

The mitochondrial genome assembly is 13,887 bp long and is one circular contig (Fig. S4). It contains 13 protein-coding genes, 22 tRNA genes, and two ribosomal RNA genes. Four protein-coding genes (ATP6, ATP8, ND3, and COX3) are present in the minor strand, while the rest are in the major strand. Among the tRNAs, 14 are present in the plus/major strand, while eight are present in the minor strand. The 13 protein-coding genes include three Cytochrome oxidase genes, one Cytochrome b, seven NADH genes, and two

Downloaded from https://academic.oup.com/gbe/article/17/11/evaf199/8303493 by guest on 03 December 2025

ATP synthase genes. The size of the mitochondrial genome of *I. ampulla* is comparable to other mollusks; for example, 13,670 bp in *B. glabrata* (DeJong et al. 2004), 14,100 bp in *C. nemoralis* (Terrett et al. 1996), and 15,057 bp in *Lissachatina fulica* (He et al. 2016).

Demographic History

We observed a sudden drop in effective population size (N_e), concurrent with the beginning of the last glacial period (Fig. 1b). The Pleistocene glacial cycles are major drivers of various intricate phylogeographic patterns across the globe. However, its role in shaping the genetic diversity of flora and fauna of the Western Ghats is unclear. Some studies suggest a change in rainfall patterns and an increase in aridification before and during the last glacial maxima (Srivastava et al. 2016; Raja et al. 2019), which can potentially lead to population decline in forest-dwelling species. Several codistributed species, such as passerine birds, exhibit signatures of population decline around the same time (Robin et al. 2010). Interestingly, in the case of the study species, the population decline began much earlier. A striking parallel has been observed in the Asian elephant populations in the Western Ghats, based on genomic data (Khan et al. 2024). Previous attempts to assess the demographic history using genetic data were ambiguous (Chakraborthy et al. 2024), underlining the importance of incorporating genomic data to assess biodiversity patterns in tropical Asia.

Supplementary Material

Supplementary material is available at *Genome Biology and Evolution* online.

Acknowledgments

The authors acknowledge the Kerala Forest Department for permission and the CCMB sequencing facility as well as Novelgene Inc. G.K. was supported by a PhD fellowship from the Council of Scientific and Industrial Research, Govt. of India. The authors would also like to thank Dr. Kushankur Bhattacharyya for capturing the photograph of *I. ampulla*.

Funding

1. Internal funding from the National Institute of Science Education and Research (under Dept. of Atomic Energy) and SERB CRG Grant (CRG_2021_000261) awarded to A.D.R. 2. Science and Engineering Research Board (SERB), Govt. of India, for funding (EMR/2017/000199) to N. A. A.

Conflict of Interest

We declare that we have no conflict of interests.

Data Availability

The sequences and draft genome generated for this study can be found under NCBI BioProject PRJNA1227711. The short-read and long-read data are archived under SRA IDs SRR32476888 and SRR32476889, respectively. The reference genome assembly can be accessed via GenBank accession ID GCA_052818635.1 or WGS accession ID JBMUJE01. The mitogenome can be accessed through the GenBank accession ID PV650439. The gene prediction and functional annotation files are available at https://doi.org/10.5281/zenodo.17119288.

Literature Cited

Aravind NA, Rajashekhar KP, Madhyastha NA. Species diversity, endemism and distribution of land snails of the Western Ghats, India. Rec West Aust Mus Suppl. 2005:68:31–38. https://doi.org/10.18195/issn.0313-122x.68.2005.031-038.

Astashyn A, et al. Rapid and sensitive detection of genome contamination at scale with FCS-GX. Genome Biol. 2024:25:60. https://doi.org/10.1186/s13059-024-03198-7.

Brůna T, et al. Galba: genome annotation with miniprot and AUGUSTUS. BMC bioinformatics. 2023:24:327. https://doi.org/10.1186/s12859-023-05449-z.

Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. Mol Biol Evol. 2021:38:5825–5829. https://doi.org/10.1093/molbev/msab293.

Chakraborty M, Baldwin-Brown JG, Long AD, Emerson JJ. Contiguous and accurate de novo assembly of metazoan genomes with modest long read coverage. Nucleic Acids Res. 2016:44:e147. https://doi.org/10.1093/nar/gk.

Chakraborty S, et al. Phylogeographical patterns are governed by geography in endemic polymorphic snail Indrella ampulla (Gastropoda: Ariophantidae). Biol J Linn Soc Lond. 2024: 142: 44–57. https://doi.org/10.1093/biolinnean/blad148w654.

Chakraborty S, Saha A, Aravind NA. Comparison of DNA extraction methods for non-marine molluscs: is the modified CTAB DNA extraction method more efficient than DNA extraction kits? 3 Biotech. 2020:10:69. https://doi.org/10.1007/s13205-020-2051-7.

Challis R, Richards E, Rajan J, Cochrane G, Blaxter M. BlobToolKit–interactive quality assessment of genome assemblies. G3 (Bethesda). 2020:10:1361–1374. https://doi.org/10.1534/g3.119.400908.

Chen S. Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. Imeta. 2023:2:e10. https://doi.org/10.1002/imt2.107.

Chen Z, Doğan Ö, Guiglielmoni N, Guichard A, Schrödl M. Pulmonate slug evolution is reflected in the de novo genome of Arion vulgaris Moquin-Tandon, 1855. Sci Rep. 2022:12:14226. https://doi.org/10.1038/s41598-022-18099-7.

Chueca LJ, Schell T, Pfenninger M. De novo genome assembly of the land snail *Candidula unifasciata* (Mollusca: Gastropoda). G3 (Bethesda). 2021:11:jkab180. https://doi.org/10.1093/g3journal/jkab180.

Cuthill IC, et al. The biology of colour. Science. 2017:357:eaan0221. https://doi.org/10.1126/science.aan0221.

Danecek P, et al. Twelve years of SAMtools and BCFtools. Gigascience. 2021:10:giab008. https://doi.org/10.1093/gigascience/giab008.

Downloaded from https://academic.oup.com/gbe/article/17/11/evaf199/8303493 by guest on 03 December 2025

DeJong RJ, Emery AM, Adema CM. The mitochondrial genome of *Biomphalaria glabrata* (Gastropoda: Basommatophora), intermediate host of Schistosoma mansoni. J Parasitol. 2004:90:991–997. https://doi.org/10.1645/GE-284R.

Flynn JM, et al. RepeatModeler2 for automated genomic discovery of transposable element families. Proc Natl Acad Sci U S A. 2020:117:9451–9457. https://doi.org/10.1073/pnas.1921046117.

Greiner S, Lehwark P, Bock R. OrganellarGenomeDRAW (OGDRAW) version 1.3. 1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Res. 2019:47:W59–W64. https://doi.org/10.1093/nar/gkz238.

Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013:29:1072–1075. https://doi.org/10.1093/bioinformatics/btt086.

Hayes KA, et al. Molluscan models in evolutionary biology: apple snails (Gastropoda: Ampullariidae) as a system for addressing fundamental questions. Am Malacol Bull. 2009:27:47–58. https://doi.org/10.4003/006.027.0204.

He ZP, et al. Complete mitochondrial genome of the giant African snail, *Achatina fulica* (Mollusca: Achatinidae): a novel location of putative control regions (CR) in the mitogenome within Pulmonate species. Mitochondrial DNA A DNA Mapp Seq Anal. 2016:27:1084–1085. https://doi.org/10.3109/19401736.2014.930833.

Huerta-Cepas J, et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res. 2019:47:D309–D314. https://doi.org/10.1093/nar/gky1085.

Ishii Y, et al. Chromosome-level genome assembly of the Asian tramp snail Bradybaena similaris (Stylommatophora: Camaenidae). Genome Biol Evol. 2025:17:evaf070. https://doi.org/10.1093/gbe/evaf070.

Khan A, et al. Divergence and serial colonization shape genetic variation and define conservation units in Asian elephants. Curr Biol. 2024:34:4692–4703. https://doi.org/10.1016/j.cub.2024.08.062.

Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol. 2019:37:540–546. https://doi.org/10.1038/s41587-019-0072-8.

Laslett D, Canbäck B. ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. Bioinformatics. 2008:24:172–175. https://doi.org/10.1093/bioinformatics/btm573

Li H, Durbin R. Inference of human population history from individual whole-genome sequences. Nature. 2011:475:493–496. https://doi.org/10.1038/nature10231.

Mak QC, Wick RR, Holt JM, Wang JR. Polishing de novo nanopore assemblies of bacteria and eukaryotes with FMLRC2. Mol Biol Evol. 2023:40:msad048. https://doi.org/10.1093/molbev/msad048.

Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. Mol Biol Evol. 2021:38:4647–4654. https://doi.org/10.1093/molbev/msab199.

Meng G, Li Y, Yang C, Liu S. Mitoz: a toolkit for animal mitochondrial genome assembly, annotation and visualization. Nucleic Acids Res. 2019:47:e63. https://doi.org/10.1093/nar/gkz173.

Miranda-Rodríguez JR, et al. The chromosome-level genome assembly of the slug Deroceras laeve facilitates its use as a comparative model of regeneration. G3 (Bethesda). 2025:15:jkaf164. https://doi.org/10.1093/g3journal/jkaf164.

Narayanan S, Aravind NA. Observations on natural diet and reproductive behaviour of an endemic snail *Indrella ampulla* (Benson 1850) (Gastropoda: Ariophantidae) from the Western Ghats, India. J Nat Hist. 2021:55:2961–2972. https://doi.org/10.1080/00222933.2022.2032857.

Nevers Y, et al. Quality assessment of gene repertoire annotations with OMArk. Nat Biotechnol. 2025:43:124–133. https://doi.org/10.1038/s41587-024-02147-w.

Ponder WF, Lindberg DR, Ponder JM. Gastropoda I: introduction and the stem groups. Biol Evol Mollusca. 2019:2:289–363. https://doi.org/10.1201/9781351115254-7.

Raja P, et al. Tropical rainforest dynamics and palaeoclimate implications since the late Pleistocene, Nilgiris, India. Quat Res. 2019:91:367–382. https://doi.org/10.1017/qua.2018.58.

Rhie A, Walenz BP, Koren S, Phillippy AM. Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies. Genome Biol. 2020:21:245. https://doi.org/10.1186/s13059-020-02134-9.

Robin VV, Sinha A, Ramakrishnan U. Ancient geographical gaps and paleo-climate shape the phylogeography of an endemic bird in the sky islands of southern India. PLoS One. 2010:5:e13321. https://doi.org/10.1371/journal.pone.0013321.

Saenko SV, Groenenberg DS, Davison A, Schilthuizen M. The draft genome sequence of the grove snail Cepaea nemoralis. G3 (Bethesda). 2021:11:jkaa071. https://doi.org/10.1093/g3journal/jkaa071.

San-Jose LM, Roulin A. Genomics of colouration in natural animal populations. Philos Trans R Soc Lond B Biol Sci. 2017:372:20160337. https://doi.org/10.1098/rstb.2016.0337.

Smit AF, Hubley R, Green P. RepeatMasker Open-4.0: 2013–2015. https://www.repeatmasker.org

Srivastava G, et al. Monsoon variability over Peninsular India during late Pleistocene: signatures of vegetation shift recorded in the terrestrial archive from the corridors of Western Ghats. Palaeogeogr Palaeoclimatol Palaeoecol. 2016:443:57–65. https://doi.org/10.1016/j.palaeo.2015.11.045.

Sun S, Han X, Han Z, Liu Q. Chromosomal-scale genome assembly and annotation of the land slug (Meghimatium bilineatum). Sci Data. 2024:11:35. https://doi.org/10.1038/s41597-023-02893-7.

Tegenfeldt F, et al. OrthoDB and BUSCO update: annotation of orthologs with wider sampling of genomes. Nucleic Acids Res. 2025:53:D516–D522. https://doi.org/10.1093/nar/gkae987.

Terrett JA, Miles S, Thomas RH. Complete DNA sequence of the mitochondrial genome of Cepaea nemoralis (Gastropoda: Pulmonata). J Mol Evol. 1996:42:160–168. https://doi.org/10.1007/BF02198842.

Tillich M, et al. Geseq–versatile and accurate annotation of organelle genomes. Nucleic Acids Res. 2017:45:W6–11. https://doi.org/10.1093/nar/gkx391.

Uliano-Silva M, et al. Mitohifi: a Python pipeline for mitochondrial genome assembly from PacBio high-fidelity reads. BMC Bioinformatics. 2023:24:288. https://doi.org/10.1186/s12859-023-05385-y.

Vasimuddin M, Misra S, Li H, Aluru S. Efficient architecture-aware acceleration of BWA-MEM for multicore systems. In: 2019 IEEE international parallel and distributed processing symposium (IPDPS). IEEE; 2019. p. 314–324.

Zimin AV, Marçais G, Puiu D, Roberts M, Salzberg SL. The MaSuRCA genome assembler. Bioinformatics. 2013:29:2669–2677. https://doi.org/10.1093/bioinformatics/btt476.

Associate editor: John Wang

Downloaded from https://academic.oup.com/gbe/article/17/11/evaf199/8303493 by guest on 03 December 2025