



DATA NOTE

The genome sequence of *Eledona agricola* (Herbst, 1783)

(Coleoptera: Tenebrionidae)

[version 1; peer review: 3 approved]

Glenda M. Orledge¹, Liam M. Crowley^{id}²,
University of Oxford and Wytham Woods Genome Acquisition Lab,
Darwin Tree of Life Barcoding collective,
Wellcome Sanger Institute Tree of Life Management, Samples and Laboratory
team,
Wellcome Sanger Institute Scientific Operations: Sequencing Operations,
Wellcome Sanger Institute Tree of Life Core Informatics team,
Tree of Life Core Informatics collective, Darwin Tree of Life Consortium

¹Department of Life Sciences, University of Bath, Bath, England, UK

²Department of Biology, University of Oxford, Oxford, England, UK

V1 First published: 09 Sep 2025, 10:501
<https://doi.org/10.12688/wellcomeopenres.24809.1>

Latest published: 09 Sep 2025, 10:501
<https://doi.org/10.12688/wellcomeopenres.24809.1>

Abstract

We present a genome assembly from an individual female *Eledona agricola* (darkling beetle; Arthropoda; Insecta; Coleoptera; Tenebrionidae). The genome sequence has a total length of 358.83 megabases. Most of the assembly (75.64%) is scaffolded into 10 chromosomal pseudomolecules, including the X sex chromosome. The mitochondrial genome has also been assembled, with a length of 16.31 kilobases. This assembly was generated as part of the Darwin Tree of Life project, which produces reference genomes for eukaryotic species found in Britain and Ireland.

Keywords

Eledona agricola, darkling beetle, genome sequence, chromosomal, Coleoptera



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

Open Peer Review

Approval Status

	1	2	3
version 1			
09 Sep 2025	view	view	view

1. **João Pedro Marques** ^{id}, BIOPOLIS-CIBIO, Portugal, Portugal
2. **Matjaž Gregorič** ^{id}, Jovan Hadži Institute of Biology, Research Centre of the Slovenian Academy of Sciences and Arts, Ljubljana, Slovenia
3. **Nikoletta Andrea Nagy**, University of Debrecen, Debrecen, Hungary

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: Orledge GM: Investigation, Resources, Writing – Original Draft Preparation, Writing – Review & Editing; Crowley LM: Investigation, Resources;

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by Wellcome through core funding to the Wellcome Sanger Institute (220540) and the Darwin Tree of Life Discretionary Award [218328, <https://doi.org/10.35802/218328>].

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

Copyright: © 2025 Orledge GM *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: Orledge GM, Crowley LM, University of Oxford and Wytham Woods Genome Acquisition Lab *et al.* **The genome sequence of *Eledona agricola* (Herbst, 1783) (Coleoptera: Tenebrionidae) [version 1; peer review: 3 approved]** Wellcome Open Research 2025, 10:501 <https://doi.org/10.12688/wellcomeopenres.24809.1>

First published: 09 Sep 2025, 10:501 <https://doi.org/10.12688/wellcomeopenres.24809.1>

Species taxonomy

Eukaryota; Opisthokonta; Metazoa; Eumetazoa; Bilateria; Protostomia; Ecdysozoa; Panarthropoda; Arthropoda; Mandibulata; Pancrustacea; Hexapoda; Insecta; Dicondylia; Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia; Tenebrionoidea; Tenebrionidae; Tenebrioninae; *Eledona*; *Eledona agricola* (Herbst, 1783) (NCBI:txid347340)

Background

Initially placed in the genus *Opatrum* Fabricius, 1775, the tenebrionid beetle *Eledona agricola* (Herbst, 1783) is now one of two species placed in the genus *Eledona* Latreille, 1797, of which it is the type species (Bouchard *et al.*, 2021). The name *agricola*, used for this species by many authors, is a misspelling (Duff, 2018). Establishment of the name of the southern European *Eledona* species, *E. hellenica* Reitter, 1885, as a junior synonym of *E. agricola* (Schawaller, 1998) has been retracted (Schawaller, 2002).

Eledona agricola is a small (2.5–4.0 mm long, 1.25–1.75 mm wide), convex, dark brown or brownish-black beetle. It has a strongly sculptured, glabrous dorsal cuticle and crenulated pronotal side edges which have protruding anterior corners (see, for example, figures in Lompe, 2002; Schawaller, 2002; Sivilov & Cvetkovska-Gorgievska, 2014). These pronotal characters contrast with the simple side edges and rounded anterior corners of *E. hellenica* (Schawaller, 2002). *Eledona agricola* may be separated from other British members of the Tenebrionidae by its combination of crenulate pronotal side edges and eyes that are not divided by a canthus (Brendell, 1975). Sexually dimorphic characters include the shape of the pronotal side edges – broad and smoothly curved in the male, narrower and more or less straight in the female – and the development of the external apical angle of the hind tibia. The latter forms a stout tooth in the male and an elongate projection in the female (figured by Lompe, 2002). Male genitalia are figured by Sivilov and Cvetkovska-Gorgievska (2014) and Lompe (2002).

Eledona agricola is a widespread Palearctic species. Records are concentrated in middle and northern Europe, including southern Scandinavia, but extend several hundred kilometres to the east of Moscow, Russia, southwards into Tunisia, Turkey and Turkmenistan, and westwards into Portugal (GBIF Secretariat, 2023; Löbl *et al.*, 2008; Sayers *et al.*, 2024). The suggestion that this beetle occurs in the Nearctic (Löbl *et al.*, 2008) appears to be an error (Sivilov & Cvetkovska-Gorgievska, 2014). In the UK it is locally common across the mainland of southern England, with modern records extending into North Yorkshire. It also occurs across the south and east of Wales and there is an unverified 1892 Scottish record from Dumfriesshire (Brendell, 1975; NBN Atlas Partnership, 2024). These records justify the species' lack of conservation concern in the UK where its current status is that of 'Least Concern' (Alexander *et al.*, 2014).

Eledona agricola is an obligate fungivore, with its larvae known only from the annual sporophores of the heartwood-rotting

basidiomycete *Laetiporus sulphureus* (Bull.) Murrill which have ceased sporulation (Alexander, 2012; Nikitsky & Schigel, 2004; Paviour-Smith, 1960). It is assumed to be univoltine, with each new generation of adults emerging in late summer or autumn (Alexander, 2012). Adults may be found throughout the year, sometimes in large numbers, and typically in association with *L. sulphureus*. Exceptionally, they have been recorded in association with sporophores of the wood-rotting fungi *Meripilus giganteus* (Pers.) Karst. and *Ceriporus squamosus* (Huds.) Quélet (Alexander, 2012; Brendell, 1975).

Here we present a whole genome sequence for *Eledona agricola* from an adult female (Figure 1), collected in Wytham Woods, Oxfordshire, United Kingdom on 28 May 2022. This genome has been sequenced as part of the Darwin Tree of Life Project, a collaborative effort to sequence all named eukaryotic species in the Atlantic Archipelago of Britain and Ireland.

Methods

Sample acquisition and DNA barcoding

An adult female *Eledona agricola* (specimen ID Ox002368, ToLID icEleAgri1) was collected from a sporophore of *Laetiporus sulphureus* growing on a fallen tree trunk in Wytham Woods, Oxfordshire, United Kingdom (latitude 51.77, longitude −1.34) on 28th May 2022. A second adult (specimen ID Ox002371, ToLID icEleAgri4), was collected on the same day and was used for RNA sequencing. The specimens were collected by Glenda Orledge (University of Bath) and Liam Crowley (University of Oxford), identified by Glenda Orledge and preserved on dry ice. Sample metadata were collected in line with the Darwin Tree of Life project standards described by Lawniczak *et al.* (2022)

The initial identification was verified by an additional DNA barcoding process according to the framework developed by Twyford *et al.* (2024). A small sample was dissected from each specimen and stored in ethanol, while the remaining parts were shipped on dry ice to the Wellcome Sanger Institute (WSI) (see the protocol). The tissue was lysed, the COI marker region was amplified by PCR, and amplicons were



Figure 1. Photograph of the *Eledona agricola* (icEleAgri1) specimen from which samples were taken for genome sequencing.

sequenced and compared to the BOLD database, confirming the species identification (Crowley *et al.*, 2023). Following whole genome sequence generation, the relevant DNA barcode region was also used alongside the initial barcoding data for sample tracking at the WSI (Twyford *et al.*, 2024). The standard operating procedures for Darwin Tree of Life barcoding are available on protocols.io.

Nucleic acid extraction

Protocols for high molecular weight (HMW) DNA extraction developed at the Wellcome Sanger Institute (WSI) Tree of Life Core Laboratory are available on protocols.io (Howard *et al.*, 2025). The icEleAgri1 sample was weighed and triaged to determine the appropriate extraction protocol. Tissue from the whole organism was homogenised by [powermashing](https://protocols.io) using a PowerMasher II tissue disruptor.

HMW DNA was extracted in the WSI Scientific Operations core using the [Automated MagAttract v2](https://protocols.io) protocol. DNA was sheared into an average fragment size of 12–20 kb following the [Megaruptor®3 for LI PacBio](https://protocols.io) protocol. Sheared DNA was purified by [manual SPRI](https://protocols.io) (solid-phase reversible immobilisation). The concentration of the sheared and purified DNA was assessed using a Nanodrop spectrophotometer and Qubit Fluorometer using the Qubit dsDNA High Sensitivity Assay kit. Fragment size distribution was evaluated by running the sample on the FemtoPulse system. For this sample, the final post-shearing DNA had a Qubit concentration of 9.74 ng/μL and a yield of 438.30 ng, with a fragment size of 10.9 kb. The 260/280 spectrophotometric ratio was 1.96, and the 260/230 ratio was 1.23.

RNA was extracted from whole organism tissue of icEleAgri4 in the Tree of Life Laboratory at the WSI using the [RNA Extraction: Automated MagMax™ mirVana](https://protocols.io) protocol. The RNA concentration was assessed using a Nanodrop spectrophotometer and a Qubit Fluorometer using the Qubit RNA Broad-Range Assay kit. Analysis of the integrity of the RNA was done using the Agilent RNA 6000 Pico Kit and Eukaryotic Total RNA assay.

PacBio HiFi library preparation and sequencing

Library preparation and sequencing were performed at the WSI Scientific Operations core. Libraries were prepared using the SMRTbell Prep Kit 3.0 (Pacific Biosciences, California, USA), following the manufacturer's instructions. The kit includes reagents for end repair/A-tailing, adapter ligation, post-ligation SMRTbell bead clean-up, and nuclease treatment. Size selection and clean-up were performed using diluted AMPure PB beads (Pacific Biosciences). DNA concentration was quantified using a Qubit Fluorometer v4.0 (ThermoFisher Scientific) and the Qubit IX dsDNA HS assay kit. Final library fragment size was assessed with the Agilent Femto Pulse Automated Pulsed Field CE Instrument (Agilent Technologies) using the gDNA 55 kb BAC analysis kit.

The sample was sequenced using the Sequel IIe system (Pacific Biosciences, California, USA). The concentration of the

library loaded onto the Sequel IIe was in the range 40–135 pM. The SMRT link software, a PacBio web-based end-to-end workflow manager, was used to set-up and monitor the run, and to perform primary and secondary analysis of the data upon completion.

Hi-C

Sample preparation and crosslinking

The Hi-C sample was prepared from 20–50 mg of frozen tissue from the icEleAgri1 sample using the Arima-HiC v2 kit (Arima Genomics). Following the manufacturer's instructions, tissue was fixed and DNA crosslinked using TC buffer to a final formaldehyde concentration of 2%. The tissue was homogenised using the Diagnocine Power Masher-II. Crosslinked DNA was digested with a restriction enzyme master mix, biotinylated, and ligated. Clean-up was performed with SPRISelect beads before library preparation. DNA concentration was measured with the Qubit Fluorometer (Thermo Fisher Scientific) and Qubit HS Assay Kit. The biotinylation percentage was estimated using the Arima-HiC v2 QC beads.

Hi-C library preparation and sequencing

Biotinylated DNA constructs were fragmented using a Covaris E220 sonicator and size selected to 400–600 bp using SPRISelect beads. DNA was enriched with Arima-HiC v2 kit Enrichment beads. End repair, A-tailing, and adapter ligation were carried out with the NEBNext Ultra II DNA Library Prep Kit (New England Biolabs), following a modified protocol where library preparation occurs while DNA remains bound to the Enrichment beads. Library amplification was performed using KAPA HiFi HotStart mix and a custom Unique Dual Index (UDI) barcode set (Integrated DNA Technologies). Depending on sample concentration and biotinylation percentage determined at the crosslinking stage, libraries were amplified with 10 to 16 PCR cycles. Post-PCR clean-up was performed with SPRISelect beads. Libraries were quantified using the AccuClear Ultra High Sensitivity dsDNA Standards Assay Kit (Biotium) and a FLUOstar Omega plate reader (BMG Labtech).

Prior to sequencing, libraries were normalised to 10 ng/μL. Normalised libraries were quantified again and equimolar and/or weighted 2.8 nM pools. Pool concentrations were checked using the Agilent 4200 TapeStation (Agilent) with High Sensitivity D500 reagents before sequencing. Sequencing was performed using paired-end 150 bp reads on the Illumina NovaSeq 6000.

RNA library preparation and sequencing

Libraries were prepared using the NEBNext® Ultra™ II Directional RNA Library Prep Kit for Illumina (New England Biolabs), following the manufacturer's instructions. Poly(A) mRNA in the total RNA solution was isolated using oligo(dT) beads, converted to cDNA, and uniquely indexed; 14 PCR cycles were performed. Libraries were size-selected to produce fragments between 100–300 bp. Libraries were quantified, normalised, pooled to a final concentration of 2.8 nM, and diluted to 150 pM for loading. Sequencing was

carried out on the Illumina NovaSeq 6000 to generate 150-bp paired-end reads.

Genome assembly

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

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

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

Assembly curation

The assembly was decontaminated using the Assembly Screen for Cebionts and Contaminants (ASCC) pipeline. TreeVal was used to generate the flat files and maps for use in curation. Manual curation was conducted primarily in PretextView and HiGlass (Kerpedjiev *et al.*, 2018). Scaffolds were visually inspected and corrected as described by Howe *et al.* (2021). Manual corrections included 28 breaks, 45 joins, and removal of 24 haplotypic duplications. The curation process is documented at <https://gitlab.com/wtsi-grit/rapid-curation>. PretextViewSnapshot was used to generate a Hi-C contact map of the final assembly.

Assembly quality assessment

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

The genome was analysed using the BlobToolKit pipeline, a Nextflow implementation of the earlier Snakemake version (Challis *et al.*, 2020). The pipeline aligns PacBio reads using minimap2 (Li, 2018) and SAMtools (Danecek *et al.*, 2021) to generate coverage tracks. It runs BUSCO (Manni *et al.*, 2021) using lineages identified from NCBI Taxonomy (Schoch *et al.*, 2020). For the three domain-level lineages, BUSCO genes are aligned to the UniProt Reference Proteomes database (Bateman *et al.*, 2023) using DIAMOND blastp (Buchfink *et al.*, 2021). The genome is divided into chunks based on the density of BUSCO genes from the closest taxonomic lineage,

and each chunk is aligned to the UniProt Reference Proteomes database with DIAMOND blastx. Sequences without hits are chunked using seqtk and aligned to the NT database with blastn (Altschul *et al.*, 1990). The BlobToolKit suite consolidates all outputs into a blobdir for visualisation. The BlobToolKit pipeline was developed using nf-core tooling (Ewels *et al.*, 2020) and MultiQC (Ewels *et al.*, 2016), with package management via Conda and Bioconda (Grünig *et al.*, 2018), and containerisation through Docker (Merkel, 2014) and Singularity (Kurtzer *et al.*, 2017).

Genome sequence report

Sequence data

PacBio sequencing of the *Eledona agricola* specimen generated 21.63 Gb (gigabases) from 2.32 million reads, which were used to assemble the genome. GenomeScope2.0 analysis estimated the haploid genome size at 315.60 Mb, with a heterozygosity of 0.49% and repeat content of 55.85% (Figure 2). These estimates guided expectations for the assembly. Based on the estimated genome size, the sequencing data provided approximately 66x coverage. Hi-C sequencing produced 101.57 Gb from 672.62 million reads, which were used to scaffold the assembly. RNA sequencing data were also generated and are available in public sequence repositories. Table 1 summarises the specimen and sequencing details.

Assembly statistics

The primary haplotype was assembled, and contigs corresponding to an alternate haplotype were also deposited in INSDC databases. The final assembly has a total length of 358.83 Mb in 145 scaffolds, with 100 gaps, and a scaffold N50 of 22.63 Mb (Table 2).

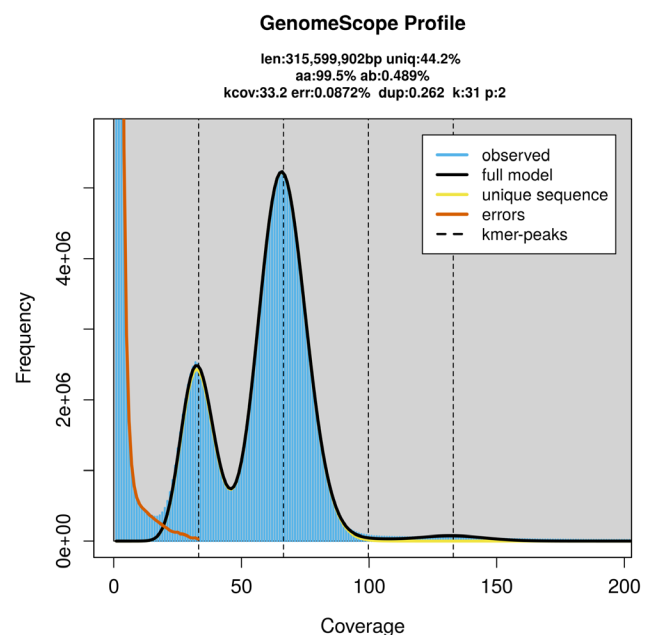


Figure 2. Frequency distribution of k -mers generated using GenomeScope2. The plot shows observed and modelled k -mer spectra, providing estimates of genome size, heterozygosity, and repeat content based on unassembled sequencing reads.

Table 1. Specimen and sequencing data for BioProject PRJEB71655.

Platform	PacBio HiFi	Hi-C	RNA-seq
ToLID	icEleAgri1	icEleAgri1	icEleAgri4
Specimen ID	Ox002368	Ox002368	Ox002371
BioSample (source individual)	SAMEA112232587	SAMEA112232587	SAMEA112232590
BioSample (tissue)	SAMEA112233052	SAMEA112233052	SAMEA112233055
Tissue	whole organism	whole organism	whole organism
Instrument	Sequel IIe	Illumina NovaSeq 6000	Illumina NovaSeq 6000
Run accessions	ERR12408808	ERR12411013	ERR12411014
Read count total	2.32 million	672.62 million	66.59 million
Base count total	21.63 Gb	101.57 Gb	10.06 Gb

Table 2. Genome assembly statistics.

Assembly name	icEleAgri1.2
Assembly accession	GCA_964023215.2
Alternate haplotype accession	GCA_964023225.2
Assembly level	chromosome
Span (Mb)	358.83
Number of chromosomes	10
Number of contigs	245
Contig N50	4.22 Mb
Number of scaffolds	145
Scaffold N50	22.63 Mb
Sex chromosomes	X
Organelles	Mitochondrion: 16.31 kb

Most of the assembly sequence (75.64%) was assigned to 10 chromosomal-level scaffolds, representing 9 autosomes and the X sex chromosome. These chromosome-level scaffolds, confirmed by Hi-C data, are named according to size (Figure 3; Table 3). the X chromosome was assigned by homology to the genome of *Tenebrio molitor* (GCA_963966145.1). During curation we noted that the exact order and orientation of the contigs in the telocentric regions is unknown. Additional centromeric repeat sequence can be found in the unassigned scaffolds. It is possible that some telocentric chromosomes are arms of acrocentric ones, but due to lack of unique HiC signal, they were not joined up.

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

The combined primary and alternate assemblies achieve an estimated QV of 63.4. The *k*-mer completeness is 87.13% for

the primary assembly, 86.13% for the alternate haplotype, and 99.07% for the combined assemblies (Figure 4).

BUSCO v.5.7.1 analysis using the endopterygota_odb10 reference set (*n* = 2 124) identified 99.5% of the expected gene set (single = 98.9%, duplicated = 0.7%). The snail plot in Figure 5 summarises the scaffold length distribution and other assembly statistics for the primary assembly. The blob plot in Figure 6 shows the distribution of scaffolds by GC proportion and coverage.

Table 4 lists the assembly metric benchmarks adapted from Rhie *et al.* (2021) the Earth BioGenome Project Report on Assembly Standards September 2024. The EBP metric, calculated for the primary assembly, is **6.7.Q63**.

Wellcome Sanger Institute – Legal and Governance
The materials that have contributed to this genome note have been supplied by a Darwin Tree of Life Partner. The submission

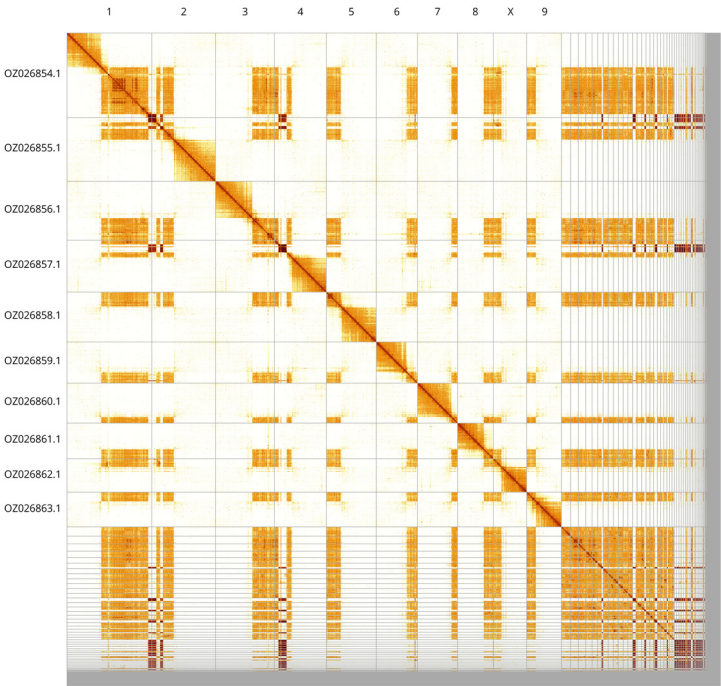


Figure 3. Hi-C contact map of the *Eledona agricola* genome assembly. Assembled chromosomes are shown in order of size and labelled along the axes. The plot was generated using PretextSnapshot.

Table 3. Chromosomal pseudomolecules in the primary genome assembly of *Eledona agricola* icEleAgri1.

INSDC accession	Molecule	Length (Mb)	GC%
OZ026854.1	1	46.67	35.50
OZ026855.1	2	34.99	35.50
OZ026856.1	3	32.32	36
OZ026857.1	4	28.48	35.50
OZ026858.1	5	27.36	36
OZ026859.1	6	22.63	35.50
OZ026860.1	7	21.95	36
OZ026861.1	8	19.62	35
OZ026863.1	9	18.34	35.50
OZ026862.1	X	19.05	36

of materials by a Darwin Tree of Life Partner is subject to the ‘**Darwin Tree of Life Project Sampling Code of Practice**’, which can be found in full on the [Darwin Tree of Life website](#). 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. Further, the Wellcome Sanger Institute employs a process whereby due diligence is carried out proportionate to the nature of the materials themselves, and the circumstances under which they have been/are to be collected and provided for use. The purpose of this is to address and mitigate any potential legal and/or ethical implications of receipt and use of the materials as part of

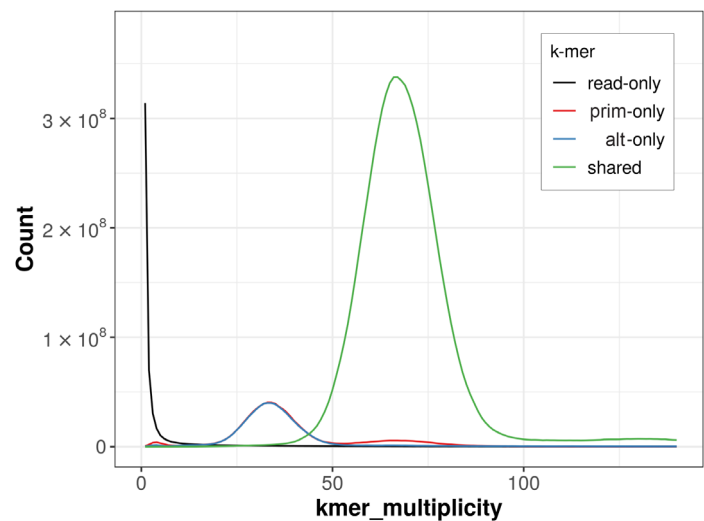


Figure 4. Evaluation of *k*-mer completeness using MerquryFK. This plot illustrates the recovery of *k*-mers from the original read data in the final assemblies. The horizontal axis represents *k*-mer multiplicity, and the vertical axis shows the number of *k*-mers. The black curve represents *k*-mers that appear in the reads but are not assembled. The green curve corresponds to *k*-mers shared by both haplotypes, and the red and blue curves show *k*-mers found only in one of the haplotypes.

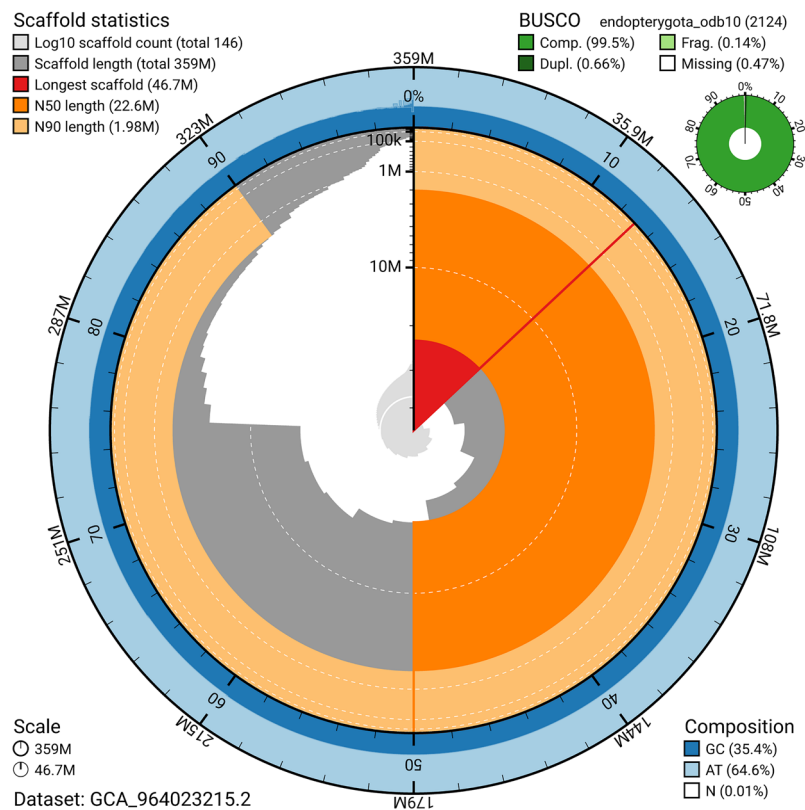


Figure 5. Assembly metrics for icEleAgri1.2. The BlobToolKit snail plot provides an overview of assembly metrics and BUSCO gene completeness. The circumference represents the length of the whole genome sequence, and the main plot is divided into 1,000 bins around the circumference. The outermost blue tracks display the distribution of GC, AT, and N percentages across the bins. Scaffolds are arranged clockwise from longest to shortest and are depicted in dark grey. The longest scaffold is indicated by the red arc, and the deeper orange and pale orange arcs represent the N50 and N90 lengths. A light grey spiral at the centre shows the cumulative scaffold count on a logarithmic scale. A summary of complete, fragmented, duplicated, and missing BUSCO genes in the endopterygota_odb10 set is presented at the top right. An interactive version of this figure can be accessed on the [BlobToolKit viewer](#).

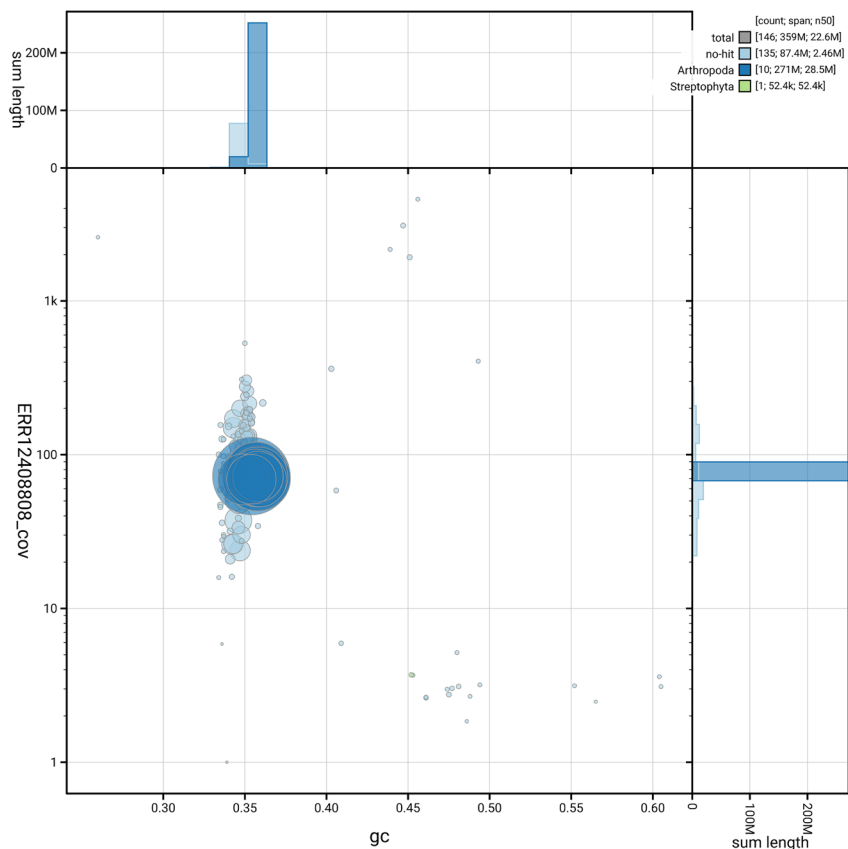


Figure 6. BlobToolKit GC-coverage plot for icEleAgri1.2. Blob plot showing sequence coverage (vertical axis) and GC content (horizontal axis). The circles represent scaffolds, with the size proportional to scaffold length and the colour representing phylum membership. The histograms along the axes display the total length of sequences distributed across different levels of coverage and GC content. An interactive version of this figure is available on the [BlobToolKit viewer](#).

Table 4. Earth Biogenome Project summary metrics for the *Eledona agricola* assembly.

Measure	Value	Benchmark
EBP summary (primary)	6.7.Q63	6.C.Q40
Contig N50 length	4.22 Mb	≥ 1 Mb
Scaffold N50 length	22.63 Mb	= chromosome N50
Consensus quality (QV)	Primary: 63.7; alternate: 63.2; combined: 63.4	≥ 40
k-mer completeness	Primary: 87.13%; alternate: 86.13%; combined: 99.07%	≥ 95%
BUSCO	C:99.5% [S:98.9%; D:0.7%]; F:0.1%; M:0.3%; n:2 124	S > 90%; D < 5%
Percentage of assembly assigned to chromosomes	75.64%	≥ 90%

the research project, and to ensure that in doing so we align with best practice wherever possible. The overarching areas of consideration are:

- Ethical review of provenance and sourcing of the material

- Legality of collection, transfer and use (national and international)

Each transfer of samples is 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: *Eledona agricola*. Accession number [PRJEB71655](#). The genome sequence is released openly for reuse. The *Eledona agricola* genome sequencing initiative is part of the Darwin Tree of Life Project (PRJEB40665)

and the Sanger Institute Tree of Life Programme (PRJEB43745). All raw sequence data and the assembly have been deposited in INSDC databases. The genome will be annotated using available RNA-Seq data and presented through the [Ensembl](#) pipeline at the European Bioinformatics Institute. Raw data and assembly accession identifiers are reported in [Table 1](#) and [Table 2](#)

Production code used in genome assembly at the WSI Tree of Life are available at <https://github.com/sanger-tol>. [Table 5](#) lists software versions used in this study.

Table 5. Software versions and sources.

Software	Version	Source
BEDTools	2.30.0	https://github.com/arq5x/bedtools2
BLAST	2.14.0	ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/-
BlobToolKit	4.4.5	https://github.com/blobtoolkit/blobtoolkit
BUSCO	5.7.1	https://gitlab.com/ezlab/busco
bwa-mem2	2.2.1	https://github.com/bwa-mem2/bwa-mem2
Cooler	0.8.11	https://github.com/open2c/cooler
DIAMOND	2.1.8	https://github.com/bbuchfink/diamond
fasta_windows	0.2.4	https://github.com/tolkita/fasta_windows
FastK	1.1	https://github.com/thegenemyers/FASTK
GenomeScope2.0	2.0.1	https://github.com/tbenavi1/genomescope2.0
Gfastats	1.3.6	https://github.com/vgl-hub/gfastats
GoaT CLI	0.2.5	https://github.com/genomehubs/goat-cli
Hifiasm	0.19.8-r603	https://github.com/chhyllp123/hifiasm
HiGlass	1.13.4	https://github.com/higlass/higlass
MerquryFK	1.1.2	https://github.com/thegenemyers/MERQURY.FK
Minimap2	2.28-r1209	https://github.com/lh3/minimap2
MitoHiFi	3	https://github.com/marcelauliano/MitoHiFi
MultiQC	1.14; 1.17 and 1.18	https://github.com/MultiQC/MultiQC
Nextflow	24.10.4	https://github.com/nextflow-io/nextflow
PretextView	N/A	https://github.com/sanger-tol/PretextView
PretextView	0.2.5	https://github.com/sanger-tol/PretextView
purge_dups	1.2.5	https://github.com/dfguan/purge_dups
samtools	1.21	https://github.com/samtools/samtools
sanger-tol/ascc	0.1.0	https://github.com/sanger-tol/ascc
sanger-tol/blobtoolkit	v0.7.1	https://github.com/sanger-tol/blobtoolkit
Seqtk	1.3	https://github.com/lh3/seqtk
Singularity	3.9.0	https://github.com/sylabs/singularity
TreeVal	1.2.0	https://github.com/sanger-tol/treeval
YaHS	1.2a.2	https://github.com/c-zhou/yahs

Author information

Contributors are listed at the following links:

- Members of the [University of Oxford and Wytham Woods Genome Acquisition Lab](#)
- Members of the [Darwin Tree of Life Barcoding collective](#)
- Members of the [Wellcome Sanger Institute Tree of Life Management, Samples and Laboratory team](#)
- Members of [Wellcome Sanger Institute Scientific Operations – Sequencing Operations](#)

- Members of the [Wellcome Sanger Institute Tree of Life Core Informatics team](#)
- Members of the [Tree of Life Core Informatics collective](#)
- Members of the [Darwin Tree of Life Consortium](#)

Acknowledgements

Grateful thanks to Wolfgang Schawaller for providing a copy of his 2002 paper cited here, also to the Royal Entomological Society's librarian Rose Pearson, and to Harriet Jameson for their help with literature access.

References

- Alexander KNA: **The ecology of *Eledona agricola* (Coleoptera: Tenebrionidae).** *British Journal of Entomology and Natural History*. 2012; **25**(3): 129–35.
- Alexander KNA, Dodd S, Denton JS: **A review of the beetles of Great Britain. The darkling beetles and their allies.** 2014.
[Reference Source](#)
- Allio R, Schomaker-Bastos A, Romiguier J, et al.: **MitoFinder: efficient automated large-scale extraction of mitogenomic data in target enrichment phylogenomics.** *Mol Ecol Resour*. 2020; **20**(4): 892–905.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Altschul SF, Gish W, Miller W, et al.: **Basic Local Alignment Search Tool.** *J Mol Biol*. 1990; **215**(3): 403–410.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Bateman A, Martin MJ, Orchard S, et al.: **UniProt: the Universal Protein Knowledgebase in 2023.** *Nucleic Acids Res*. 2023; **51**(D1): D523–D531.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Bouchard P, Bousquet Y, Aalbu RL, et al.: **Review of the genus-group names in the family Tenebrionidae (Insecta, Coleoptera).** *ZooKeys*. 2021; **1050**: 1–633.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Brendell MJD: **Coleoptera: Tenebrionidae.** London: Royal Entomological Society of London, 1975; 5.
[Reference Source](#)
- Buchfink B, Reuter K, Drost HG: **Sensitive protein alignments at Tree-of-Life scale using DIAMOND.** *Nat Methods*. 2021; **18**(4): 366–368.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Challis R, 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](#)
- Crowley L, Allen H, Barnes I, et al.: **A sampling strategy for genome sequencing the British terrestrial Arthropod fauna [version 1; peer review: 2 approved].** *Wellcome Open Res*. 2023; **8**: 123.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Danecek P, Bonfield JK, Liddle J, et al.: **Twelve years of SAMtools and BCFtools.** *GigaScience*. 2021; **10**(2): giab008.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Duff AG: **Checklist of beetles of the British Isles (Third Edition).** Iver: Pemberley Books, 2018.
[Reference Source](#)
- Ewels P, Magnusson M, Lundin S, et al.: **MultiQC: summarize analysis results for multiple tools and samples in a single report.** *Bioinformatics*. 2016; **32**(19): 3047–3048.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Ewels PA, Peltzer A, Fillinger S, et al.: **The nf-core framework for community-curated bioinformatics pipelines.** *Nat Biotechnol*. 2020; **38**(3): 276–278.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Formenti G, Abueg L, Brajuka A, et al.: **Gfastats: conversion, evaluation and manipulation of genome sequences using assembly graphs.** *Bioinformatics*. 2022; **38**(17): 4214–4216.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- GBIF Secretariat: ***Eledona agricola* (Herbst, 1783).** GBIF Backbone Taxonomy. 2023.
[Reference Source](#)
- Grüning B, Dale R, Sjödin A, et al.: **Bioconda: sustainable and comprehensive software distribution for the life sciences.** *Nat Methods*. 2018; **15**(7): 475–476.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Guan D, McCarthy SA, Wood J, et al.: **Identifying and removing haplotypic duplication in primary genome assemblies.** *Bioinformatics*. 2020; **36**(9): 2896–2898.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Howard C, Denton A, Jackson B, et al.: **On the path to reference genomes for all biodiversity: lessons learned and laboratory protocols created in the Sanger Tree of Life core laboratory over the first 2000 species.** *bioRxiv*. 2025.
[Publisher Full Text](#)
- Howe K, Chow W, Collins J, et al.: **Significantly improving the quality of genome assemblies through curation.** *GigaScience*. 2021; **10**(1): g1aa153.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Kerpedjiev P, Abdennur N, Lekschas F, et al.: **HiGlass: web-based visual exploration and analysis of genome interaction maps.** *Genome Biol*. 2018; **19**(1): 125.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Kurtzer GM, Sochat V, Bauer MW: **Singularity: scientific containers for mobility of compute.** *PLoS One*. 2017; **12**(5): e0177459.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Lawnczak MKN, Davey RP, Rajan J, et al.: **Specimen and sample metadata standards for biodiversity genomics: a proposal from the Darwin Tree of Life project [version 1; peer review: 2 approved with reservations].** *Wellcome Open Res*. 2022; **7**: 187.
[Publisher Full Text](#)
- Li H: **Minimap2: pairwise alignment for nucleotide sequences.** *Bioinformatics*. 2018; **34**(18): 3094–3100.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Löbl I, Merkl O, Ando K, et al.: **Tenebrionidae.** In: I. Löbl and A. Smetana (eds), *Catalogue of Palaearctic Coleoptera, Vol. 5: Tenebrionidae*. Stenstrup: Apollo Books, 2008; 105–352.
- Lompe A: **Gattung *Eledona* Latreille, 1796.** Die Käfer Europas: Ein bestimmungswerk im Internet, 2002.
[Reference Source](#)
- Manni M, Berkeley MR, Seppely M, et al.: **BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes.** *Mol Biol Evol*. 2021; **38**(10): 4647–4654.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Merkel D: **Docker: lightweight Linux containers for consistent development and deployment.** *Linux J*. 2014; **2014**(239): 2, [Accessed 2 April 2024].
[Reference Source](#)
- NBN Atlas Partnership: ***Eledona agricola* (Herbst, 1783).** NBN Atlas 2024, 2024.
[Reference Source](#)
- Nikitsky NB, Schigel DS: **Beetles in polypores of the Moscow region: checklist and ecological notes.** *Entomol Fenn*. 2004; **15**(1): 6–22.
[Publisher Full Text](#)
- Paviour-Smith K: **The fruiting bodies of macrofungi as habitats for beetles**

of the family Ciidae (Coleoptera). *Oikos*. 1960; **11**(1): 43–71.

[Publisher Full Text](#)

Ranallo-Benavidez TR, Jaron KS, Schatz MC: **GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes**. *Nat Commun*. 2020; **11**(1): 1432.

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

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

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

Rhie A, 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](#)

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

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

Sayers EW, Cavanaugh M, Clark K, *et al.*: **GenBank 2024 update**. *Nucleic Acids Res*. 2024; **52**(D1): D134–D137.

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

Schawaller W: **Taxonomie und faunistik von *Eledona agricola* (Herbst 1793) (*hellenica* Reitter 1885, n. Syn.) (Coleoptera: Tenebrionidae)**. *Entomologische Zeitschrift*. 1998; **108**: 501–3.

Schawaller W: **Evidence of two species of *Eledona* Latreille and *Eledonoprius***

Reitter in Europe (Coleoptera, Tenebrionidae). *Entomologische Blätter*. 2002; **98**: 111–14.

Schoch CL, Ciufo S, Domrachev M, *et al.*: **NCBI taxonomy: a comprehensive update on curation, resources and tools**. *Database (Oxford)*. 2020; **2020**: baaa062.

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

Sivilov O, Cvetkovska-Gorgievska A: **Faunistical records of Tenebrionidae (Coleoptera) from Bulgaria and Republic of Macedonia**. *Zoonotes*. 2014; **53**: 1–7.

[Reference Source](#)

Twyford AD, Beasley J, Barnes I, *et al.*: **A DNA barcoding framework for taxonomic verification in the Darwin Tree of Life project [version 1; peer review: 2 approved]**. *Wellcome Open Res*. 2024; **9**: 339.

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

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

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

Vasimuddin M, Misra S, Li H, *et al.*: **Efficient architecture-aware acceleration of BWA-MEM for multicore systems**. In: *2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*. IEEE, 2019; 314–324.

[Publisher Full Text](#)

Zhou C, McCarthy SA, Durbin R: **YaHS: yet another Hi-C Scaffolding tool**. *Bioinformatics*. 2023; **39**(1): btac808.

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

Open Peer Review

Current Peer Review Status:   

Version 1

Reviewer Report 23 December 2025

<https://doi.org/10.21956/wellcomeopenres.27332.r137905>

© 2025 Nagy N. 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.



Nikoletta Andrea Nagy

University of Debrecen, Debrecen, Hungary

The manuscript is well-written and technically sound. I have a few comments:

- "he suggestion that this beetle occurs in the Nearctic..." - missing "T" at the beginning of the sentence
- species scientific names should be italicized in the methods
- please include the sex of the specimen used for RNA sequencing

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: arthropod 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.

Reviewer Report 08 December 2025

<https://doi.org/10.21956/wellcomeopenres.27332.r137906>

© 2025 Gregorič M. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Matjaž Gregorič 

Jovan Hadži Institute of Biology, Research Centre of the Slovenian Academy of Sciences and Arts, Ljubljana, Slovenia

This genome report on the tenebrionid beetle *Eledona agricola* (Herbst, 1783) is thoroughly written and I have no concerns regarding its indexing. There are a few typos but I can't mark them in text: "he" to "The" in the third paragraph, and *Eledona agricola* and *Laetiporus sulphureus* need to be italicized at the beginning of the Methods.

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: Systematics and taxonomy, evolutionary biology.

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

Reviewer Report 25 November 2025

<https://doi.org/10.21956/wellcomeopenres.27332.r137904>

© 2025 Marques 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.



João Pedro Marques 

BIOPOLIS-CIBIO, Portugal, Portugal

I read the manuscript "*The genome sequence of Eledona agricola*" carefully. In my opinion, the genome was produced following the current best practices in reference genome generation, fully consistent with the standards adopted by large-scale initiatives such as the Darwin Tree of Life

project.

The authors provide a clear and comprehensive description of specimen processing, barcoding verification, high-molecular-weight DNA extraction, PacBio HiFi sequencing, Hi-C scaffolding, RNA-seq generation, and extensive quality control. Assembly statistics (e.g., scaffold N50 of 22.63 Mb, BUSCO completeness of 99.5%, QV ~ 63) fall well within, and in several aspects exceed, accepted benchmarks for high-quality chromosome-level assemblies. Figures and tables are clear, and the dataset is accessible through the ENA. No methodological weaknesses or analytical shortcomings were identified.

Importantly, the authors report no steps that raise scientific concerns, and all protocols (sequencing, assembly, curation, and contamination screening) follow validated pipelines and softwares (e.g., Hifiasm, YaHS, purge_dups, BlobToolKit, Merqury). The manuscript is consistent and scientifically correct.

I do not have any requests for revision. In my opinion, the manuscript meets all criteria for publication in its current form.

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 Biology, Population Genomics and Genome Assembly

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.
