




DATA NOTE

**REVISED** The genome sequence of the Brindled Green,***Dryobotodes eremita* (Fabricius, 1775)**

[version 2; peer review: 3 approved, 3 approved with reservations]

Douglas Boyes<sup>1+</sup>, Peter W.H. Holland <sup>2</sup>,  
 University of Oxford and Wytham Woods Genome Acquisition Lab,  
 Darwin Tree of Life Barcoding collective,  
 Wellcome Sanger Institute Tree of Life programme,  
 Wellcome Sanger Institute Scientific Operations: DNA Pipelines collective,  
 Tree of Life Core Informatics collective, Darwin Tree of Life Consortium

<sup>1</sup>UK Centre for Ecology & Hydrology, Wallingford, England, UK<sup>2</sup>University of Oxford, Oxford, England, UK

+ Deceased author

**v2** First published: 10 May 2023, 8:208  
<https://doi.org/10.12688/wellcomeopenres.19300.1>

Latest published: 18 Feb 2026, 8:208  
<https://doi.org/10.12688/wellcomeopenres.19300.2>

**Abstract**

We present a genome assembly from an individual female *Dryobotodes eremita* (the Brindled Green; Arthropoda; Insecta; Lepidoptera; Noctuidae). The genome sequence is 709.8 megabases in span. Most of the assembly is scaffolded into 32 chromosomal pseudomolecules including the Z and W sex chromosomes. The mitochondrial genome has also been assembled and is 15.5 kilobases in length. Gene annotation of this assembly on Ensembl identified 19,706 protein coding genes. 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**

*Dryobotodes eremita*, Brindled Green, genome sequence, chromosomal, Lepidoptera




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


**Open Peer Review****Approval Status** ? ✓ ? ✓ ✓ ?


	1	2	3	4	5	6
<b>version 2</b>						
(revision)		✓	?	✓	✓	?
18 Feb 2026		<a href="#">view</a>	<a href="#">view</a>	<a href="#">view</a>	<a href="#">view</a>	<a href="#">view</a>
		↑				
<b>version 1</b>						
10 May 2023	?	?				
	<a href="#">view</a>	<a href="#">view</a>				

1. **Kay Lucek** , University of Neuchâtel, Neuchâtel, Switzerland
2. **Woo Jae Kim** , Harbin Institute of Technology, Harbin, China  
**Zekun Wu**, Harbin Institute of Technology Center for Life Sciences (Ringgold ID: 598430), Harbin, China  
**Xinyue Zhou**, Harbin Institute of Technology Center for Life Sciences (Ringgold ID: 598430), Harbin, China
3. **Somjit Homchan**, Naresuan University, Tha

Pho, Thailand

4. **Wai Lok So** , The Chinese University of Hong Kong, Hong Kong, Hong Kong  
The Hang Seng University of Hong Kong  
(Ringgold ID: 272025), Hong Kong, Hong Kong

5. **Melody Clark** , Natural Environment Research Council, Cambridge, UK

6. **Noémie M.-C. Hévin** , Université de Montpellier, Montpellier, France

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](mailto:mark.blaxter@sanger.ac.uk))

**Author roles:** **Boyes D:** Investigation, Resources; **Holland PWH:** Writing – Original Draft Preparation, Writing – Review & Editing;

**Competing interests:** No competing interests were disclosed.

**Grant information:** This work was supported by Wellcome through core funding to the Wellcome Sanger Institute (206194, <https://doi.org/10.35802/206194>) 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:** © 2026 Boyes D *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**How to cite this article:** Boyes D, Holland PWH, University of Oxford and Wytham Woods Genome Acquisition Lab *et al.* **The genome sequence of the Brindled Green, *Dryobotodes eremita* (Fabricius, 1775) [version 2; peer review: 3 approved, 3 approved with reservations]** Wellcome Open Research 2026, 8:208 <https://doi.org/10.12688/wellcomeopenres.19300.2>

**First published:** 10 May 2023, 8:208 <https://doi.org/10.12688/wellcomeopenres.19300.1>

**REVISED Amendments from Version 1**

Version 2 of the data note includes a revised version of the chromosome map in Figure 5, updated links to the genome annotation on Ensembl and corrected Merqury.FK results using a new version of the software. We also link to a methods paper in which the development of wet lab protocols in the Tree of Life core laboratory is outlined.

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

**Species taxonomy**

Eukaryota; Metazoa; Ecdysozoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea; Noctuidae; Xyleninae; *Dryobotodes*; *Dryobotodes eremita* (Fabricius, 1775) (NCBI: txid988106).

**Background**

The Brindled Green, *Dryobotodes eremita*, is a small moth in the family Noctuidae (wingspan 32–39 mm) with a distinctive crinkled appearance to the forewings which are patterned with patches of green, black and cream overlain with rosy streaks. The variegated colouration extends to the thorax, head and legs such that the moth is cryptic when resting on lichen-covered tree trunks. The extent of green and red colouration is variable (Bretherton *et al.*, 1983).

*D. eremita* is found across Europe and further east into Russia (GBIF Secretariat, 2022). The species has an extensive distribution in Britain being found in woodland, parks and gardens across England, Wales, Scotland and Northern Ireland. Records range from Cornwall and the Scilly Isles in the south to Orkney and Shetland in the far north, although it is more common in the south of this range (NBN Atlas Partnership, 2021). In Ireland the moth has been recorded across much of the country (MothsIreland, 2022). The moth is on the wing in September and October in the south of England, or August and September in central and northern Scotland, and can be attracted to light or sugary substances (Bretherton *et al.*, 1983; Randle *et al.*, 2019). Eggs are laid in autumn on pedunculate oak *Quercus robur* and when larvae hatch from eggs in spring they bore into an oak leaf-bud to feed. As the larvae develop, they switch to feeding on young leaves at terminal shoots (Bretherton *et al.*, 1983; Stokoe, 1948). The species name *eremita*, meaning ‘hermit’, refers to the habit of the young larva living in a chamber formed by spinning oak leaves together with silk (Emmet, 1991); later instars feed on oak leaves openly. Pupation occurs in soil near the base of oak trees (Stokoe, 1948).

A genome sequence for *Dryobotodes eremita* will facilitate studies investigating molecular adaptations to oak feeding and will contribute to the growing set of genomic resources for Lepidoptera.

**Genome sequence report**

The genome was sequenced from one female *Dryobotodes eremita* (Figure 1) collected from Wytham Woods, UK (latitude 51.77, longitude –1.34). A total of 28-fold coverage in Pacific Biosciences single-molecule HiFi long reads and 55-fold coverage in 10X Genomics read clouds were generated. Primary assembly contigs were scaffolded with chromosome conformation Hi-C data. Manual assembly curation corrected 40 missing joins or mis-joins and removed 10 haplotypic duplications, reducing the assembly length by 0.83% and the scaffold number by 33.66%, and increasing the scaffold N50 by 0.78%.

The final assembly has a total length of 709.8 Mb in 67 sequence scaffolds with a scaffold N50 of 23.3 Mb (Table 1). Most (99.8%) of the assembly sequence was assigned to 32 chromosomal-level scaffolds, representing 30 autosomes, and the W and Z sex chromosome. Chromosome-scale scaffolds confirmed by the Hi-C data are named in order of size (Figure 2–Figure 5; Table 2). While not fully phased, the assembly deposited is of one haplotype. Contigs corresponding to the second haplotype have also been deposited.

The mitochondrial genome was also assembled (length 15.45 kb, OU823273.1). 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 56.9. The *k*-mer completeness is 71.73% for the primary assembly, 67.65% for the alternate haplotype, and 98.99% for the combined assemblies. The primary assembly has a BUSCO v5.3.2 completeness of 99.0% (single = 98.7%, duplicated = 0.4%), using the lepidoptera\_odb10 reference set (*n* = 5,286).



**Figure 1.** Photograph of the *Dryobotodes eremita* (ilDryErem1) specimen used for genome sequencing.

**Table 1.** Genome data for *Dryobotodes eremita*, ilDryErem1.1.

Project accession data		
Assembly identifier	ilDryErem1.1	
Species	<i>Dryobotodes eremita</i>	
Specimen	ilDryErem1	
NCBI taxonomy ID	988106	
BioProject	PRJEB46319	
BioSample ID	SAMEA8603190	
Isolate information	ilDryErem1, female: thorax (genome sequencing), head (Hi-C scaffolding)	
Assembly metrics*		Benchmark
Consensus quality (QV)	Primary: 56.7; alternate: 57.0; combined: 56.9	≥ 40
k-mer completeness	Primary: 71.73%; alternate: 67.65%; combined: 98.99%	≥ 95%
BUSCO**	C:99.0%[S:98.7%,D:0.4%], F:0.2%,M:0.8%,n:5,286	C ≥ 95%
Percentage of assembly mapped to chromosomes	99.8%	≥ 90%
Sex chromosomes	Z and W chromosomes	localised homologous pairs
Organelles	Mitochondrial genome assembled	complete single alleles

**Table 1.** Continued

<b>Raw data accessions</b>	
PacificBiosciences SEQUEL II	ERR6808003
10X Genomics Illumina	ERR6688525–ERR6688528
Hi-C Illumina	ERR6688524
PolyA RNA-Seq Illumina	ERR9435006
<b>Genome assembly</b>	
Assembly accession	GCA_917490735.1
Accession of alternate haplotype	GCA_917490515.1
Span (Mb)	709.8
Number of contigs	104
Contig N50 length (Mb)	22.6
Number of scaffolds	67
Scaffold N50 length (Mb)	23.3
Longest scaffold (Mb)	35.1
<b>Genome annotation</b>	
Number of protein-coding genes	19,706
Number of transcripts	19,901

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

\*\*BUSCO scores based on the lepidoptera\_odb10 BUSCO set using v5.3.2. C = complete [S = single copy, D = duplicated], F = fragmented, M = missing, n = number of orthologues in comparison. A full set of BUSCO scores is available at <https://blobtoolkit.genomehubs.org/view/iIDryErem1.1/dataset/CAKJSZ01.1/busco>.

### Genome annotation report

The *Dryobotodes eremita* genome assembly (GCA\_917490735.1) was annotated by Ensembl at the European Bioinformatics Institute (EBI). This annotation includes 19,901 transcribed mRNAs from 19,706 protein-coding genes. The average transcript length is 7,516.09 bp, with an average of 5.15 exons per transcript. Further details of this annotation are available from the [Ensembl annotation page](#).

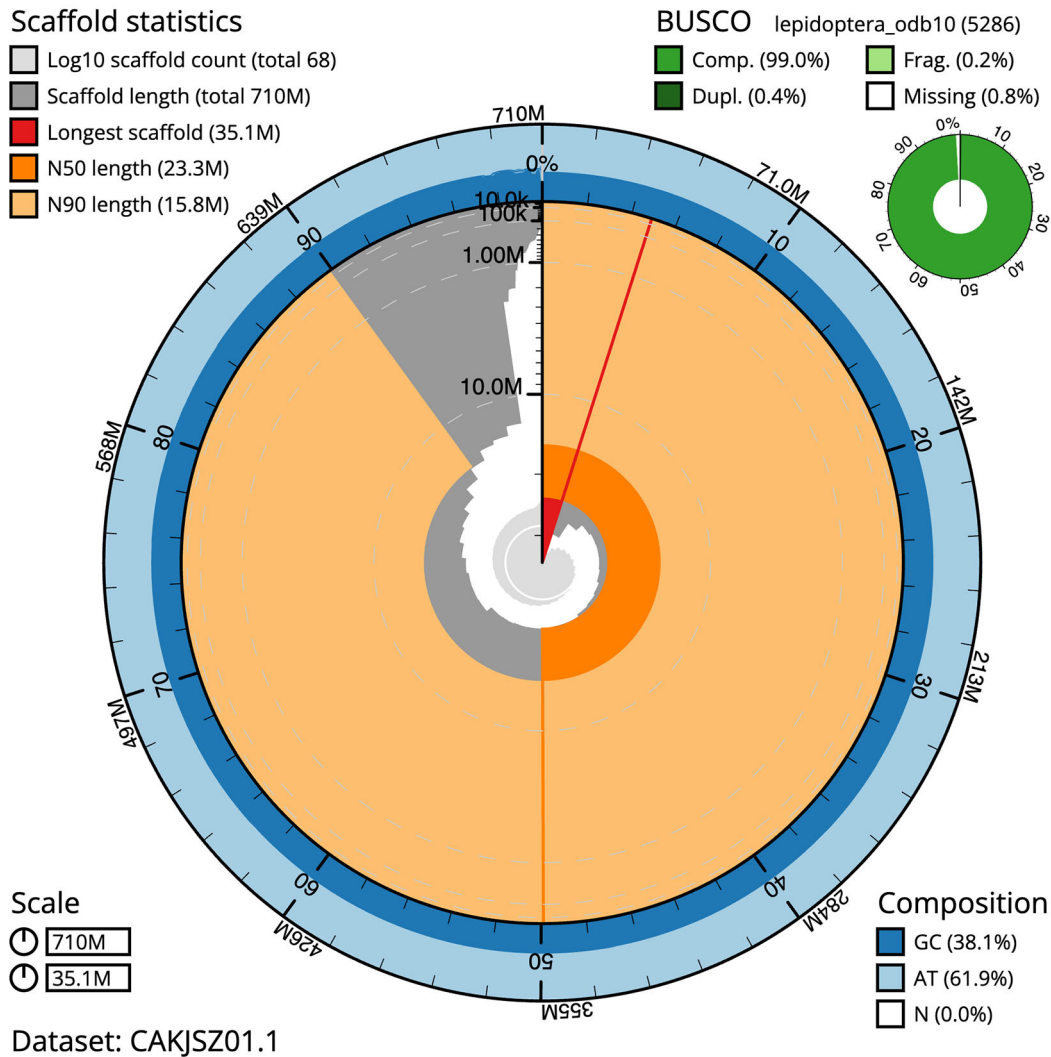
### Methods

#### Sample acquisition and nucleic acid extraction

A female *Dryobotodes eremita* (specimen number Ox000959, ToLID iIDryErem1) was collected from Wytham Woods, Oxfordshire (biological vice-county: Berkshire), UK (latitude 51.77, longitude –1.34) on 8 September 2020. The specimen was taken from woodland habitat by Douglas Boyes (University of Oxford) using a light trap. The specimen was identified by the collector and snap-frozen on dry ice.

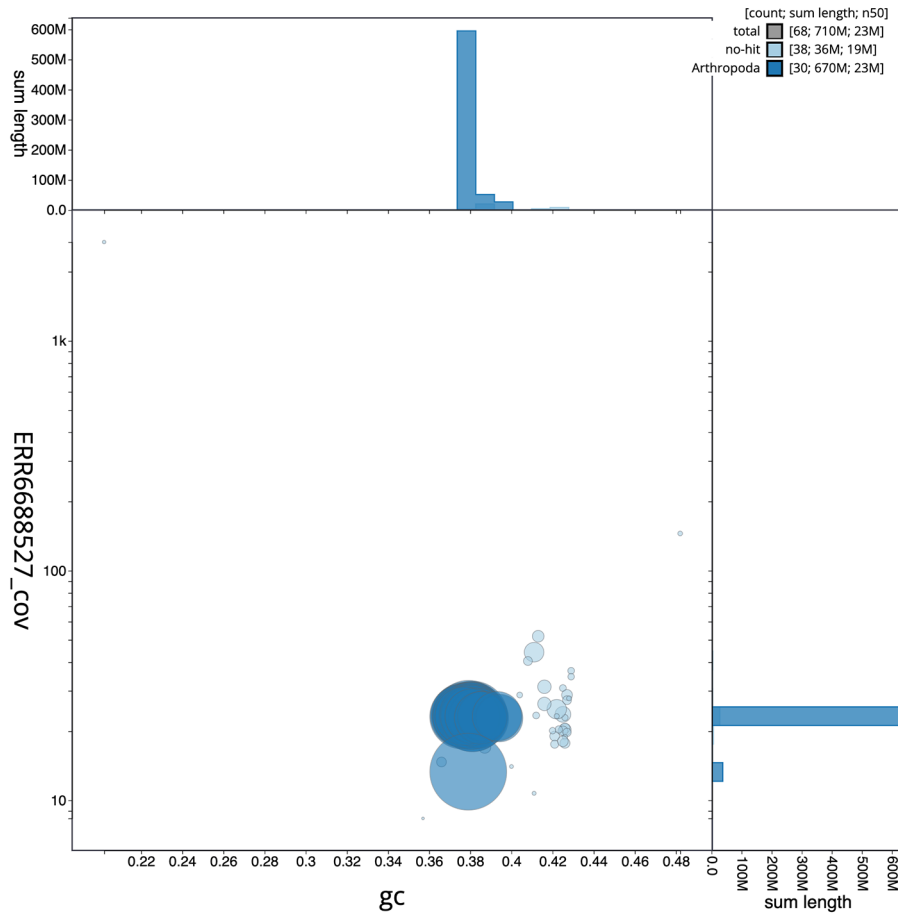
Protocols for high molecular weight (HMW) DNA extraction developed at the Wellcome Sanger Institute (WSI) Tree of Life Core Laboratory are available on [protocols.io](https://protocols.io) (Howard *et al.*, 2025).

DNA was extracted at the Tree of Life laboratory, Wellcome Sanger Institute (WSI). The iIDryErem1 sample was weighed and dissected on dry ice with head tissue set aside for Hi-C sequencing. Thorax tissue was cryogenically disrupted to a fine powder using a Covaris cryoPREP Automated Dry Pulveriser, receiving multiple impacts. High molecular weight (HMW) DNA was extracted using the Qiagen MagAttract HMW DNA extraction kit. Low molecular weight DNA was removed from a 20 ng aliquot of extracted DNA using the 0.8X AMPure XP purification kit prior to 10X Chromium sequencing; a minimum of 50 ng DNA was submitted for 10X sequencing. HMW DNA was sheared into an average fragment size of 12–20 kb in a Megaruptor 3 system with speed setting 30. Sheared DNA was purified by solid-phase reversible immobilisation using AMPure PB beads with a 1.8X ratio of beads to sample to remove the shorter fragments and concentrate the DNA sample. The concentration of the sheared and purified DNA was assessed using a Nanodrop spectrophotometer and Qubit Fluorometer and Qubit dsDNA High Sensitivity Assay kit. Fragment size distribution was evaluated by running the sample on the FemtoPulse system.



**Figure 2. Genome assembly of *Dryobotodes eremita*, iLDryErem1.1: metrics.** The BlobToolKit Snailplot shows N50 metrics and BUSCO gene completeness. The main plot is divided into 1,000 size-ordered bins around the circumference with each bin representing 0.1% of the 709,806,353 bp assembly. The distribution of scaffold lengths is shown in dark grey with the plot radius scaled to the longest scaffold present in the assembly (35,077,990 bp, shown in red). Orange and pale-orange arcs show the N50 and N90 scaffold lengths (23,327,181 and 15,843,934 bp), respectively. The pale grey spiral shows the cumulative scaffold count on a log scale with white scale lines showing successive orders of magnitude. The blue and pale-blue area around the outside of the plot shows the distribution of GC, AT and N percentages in the same bins as the inner plot. A summary of complete, fragmented, duplicated and missing BUSCO genes in the lepidoptera\_odb10 set is shown in the top right. An interactive version of this figure is available at <https://blobtoolkit.genomehubs.org/view/iLDryErem1.1/dataset/CAKJSZ01.1/snail>.

RNA was extracted from abdomen tissue of iLDryErem1 in the Tree of Life Laboratory at the WSI using TRIzol, according to the manufacturer's instructions. RNA was then eluted in 50  $\mu$ l RNase-free water and its concentration assessed using a Nanodrop spectrophotometer and Qubit Fluorometer using the Qubit RNA Broad-Range (BR) Assay kit. Analysis of the integrity of the RNA was done using Agilent RNA 6000 Pico Kit and Eukaryotic Total RNA assay.



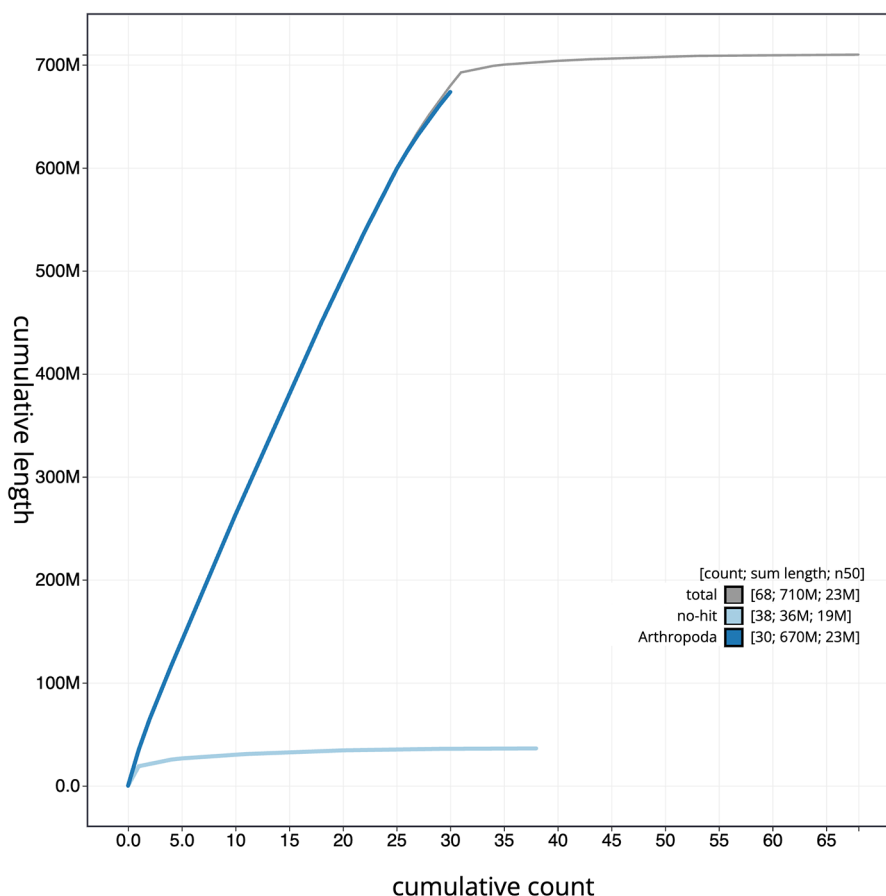
**Figure 3. Genome assembly of *Dryobotodes eremita*, iLDryErem1.1: BlobToolKit GC-coverage plot.** Scaffolds are coloured by phylum. Circles are sized in proportion to scaffold length. Histograms show the distribution of scaffold length sum along each axis. An interactive version of this figure is available at <https://blobtoolkit.genomehubs.org/view/iLDryErem1.1/dataset/CAKJSZ01.1/blob>.

### Sequencing

Pacific Biosciences HiFi circular consensus and 10X Genomics read cloud DNA sequencing libraries were constructed according to the manufacturers' instructions. Poly(A) RNA-Seq libraries were constructed using the NEB Ultra II RNA Library Prep kit. DNA and RNA sequencing was performed by the Scientific Operations core at the WSI on Pacific Biosciences SEQUEL II (HiFi), Illumina HiSeq 4000 (RNA-Seq) and Illumina NovaSeq 6000 (10X) instruments. Hi-C data were also generated from head tissue of iLDryErem1 using the Arima v2 kit and sequenced on the Illumina NovaSeq 6000 instrument.

### Genome assembly, curation and evaluation

Assembly was carried out with Hifiasm (Cheng *et al.*, 2021) and haplotypic duplication was identified and removed with purge\_dups (Guan *et al.*, 2020). One round of polishing was performed by aligning 10X Genomics read data to the assembly with Long Ranger ALIGN, calling variants with FreeBayes (Garrison & Marth, 2012). The assembly was then scaffolded with Hi-C data (Rao *et al.*, 2014) using SALSA2 (Ghurye *et al.*, 2019). The assembly was checked for contamination as described previously (Howe *et al.*, 2021). Manual curation was performed using HiGlass (Kerpedjiev *et al.*, 2018) and Pretext (Harry, 2022). The mitochondrial genome was assembled using MitoHiFi (Uliano-Silva *et al.*, 2022), which 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. To evaluate the assembly, Merqury.FK was used to estimate consensus quality

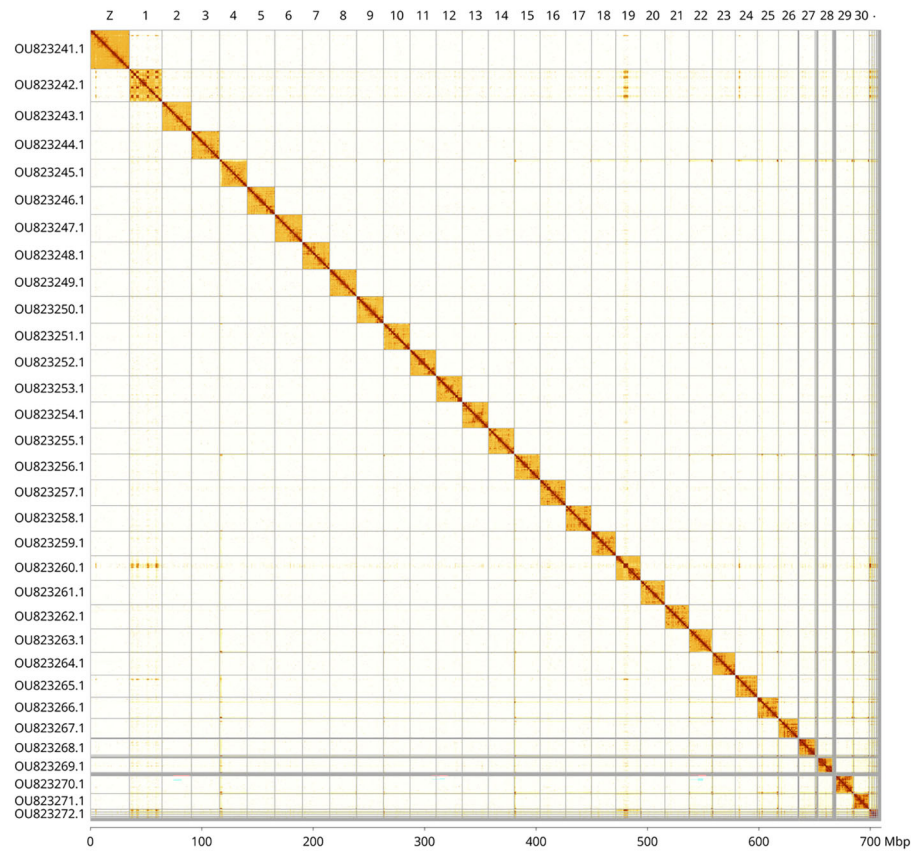


**Figure 4. Genome assembly of *Dryobotodes eremita*, iLDryErem1.1: BlobToolKit cumulative sequence plot.** The grey line shows cumulative length for all scaffolds. Coloured lines show cumulative lengths of scaffolds assigned to each phylum using the buscogenes taxrule. An interactive version of this figure is available at <https://blobtoolkit.genomehubs.org/view/iLDryErem1.1/dataset/CAKJSZ01.1/cumulative>.

(QV) scores and  $k$ -mer completeness (Rhie *et al.*, 2020). The genome was analysed within the BlobToolKit environment (Challis *et al.*, 2020) and BUSCO scores (Manni *et al.*, 2021; Simão *et al.*, 2015) were calculated. Table 3 contains a list of software tool versions and sources.

#### Ethics and compliance issues

The materials that have contributed to this genome note have been supplied by a Darwin Tree of Life Partner. The submission of materials by a Darwin Tree of Life Partner is subject to the [Darwin Tree of Life Project Sampling Code of Practice](#). By agreeing with and signing up to the Sampling Code of Practice, the Darwin Tree of Life Partner agrees they will meet the legal and ethical requirements and standards set out within this document in respect of all samples acquired for, and supplied to, the Darwin Tree of Life Project. All efforts are undertaken to minimise the suffering of animals used for sequencing. 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.



**Figure 5. Hi-C contact map of the *Dryobotodes eremita* genome assembly.** Assembled chromosomes are shown in order of size and labelled along the axes, with a megabase scale shown below. The plot was generated using PretextSnapshot. Chromosomes are shown in order of size from left to right and top to bottom. An interactive HiGlass version of this map may be viewed at <https://genome-note-higlass.tol.sanger.ac.uk/?d=a-aARzNPStu0-F6GjMD60XA>.

**Table 2. Chromosomal pseudomolecules in the genome assembly of *Dryobotodes eremita*, iIDryErem1.**

INSDC accession	Chromosome	Size (Mb)	GC%
OU823242.1	1	29.33	38.1
OU823243.1	2	26.44	37.9
OU823244.1	3	25.13	37.7
OU823245.1	4	24.83	38.1
OU823246.1	5	24.77	38
OU823247.1	6	24.73	38.1
OU823248.1	7	24.58	37.6
OU823249.1	8	24.22	37.8
OU823250.1	9	23.99	38.1
OU823251.1	10	23.83	37.7
OU823252.1	11	23.47	37.7
OU823253.1	12	23.46	37.6
OU823254.1	13	23.33	38.1
OU823255.1	14	23.26	38

**Table 2.** *Continued*

INSDC accession	Chromosome	Size (Mb)	GC%
OU823256.1	15	23.21	38
OU823257.1	16	23.05	38
OU823258.1	17	22.99	37.7
OU823259.1	18	22.07	38.2
OU823260.1	19	22.01	38.1
OU823261.1	20	21.97	38
OU823262.1	21	21.65	38
OU823263.1	22	21.02	38.4
OU823264.1	23	20.32	38
OU823265.1	24	19.94	37.7
OU823266.1	25	18.97	38.3
OU823267.1	26	18.26	38.1
OU823268.1	27	17.13	38.6
OU823269.1	28	16.09	39.4
OU823270.1	29	15.84	38.5
OU823271.1	30	14.09	39.3
OU823272.1	W	9.31	38.1
OU823241.1	Z	35.08	37.9
OU823273.1	MT	0.02	20.3

**Table 3. Software tools: versions and sources.**

Software tool	Version	Source
BlobToolKit	4.0.7	<a href="https://github.com/blobtoolkit/blobtoolkit">https://github.com/blobtoolkit/blobtoolkit</a>
BUSCO	5.3.2	<a href="https://gitlab.com/ezlab/busco">https://gitlab.com/ezlab/busco</a>
FreeBayes	1.3.1-17-gaa2ace8	<a href="https://github.com/freebayes/freebayes">https://github.com/freebayes/freebayes</a>
Hifiasm	0.15.3-r339	<a href="https://github.com/chhylp123/hifiasm">https://github.com/chhylp123/hifiasm</a>
HiGlass	1.11.6	<a href="https://github.com/higlass/higlass">https://github.com/higlass/higlass</a>
Long Ranger ALIGN	2.2.2	<a href="https://support.10xgenomics.com/genome-exome/software/pipelines/latest/advanced/other-pipelines">https://support.10xgenomics.com/genome-exome/software/pipelines/latest/advanced/other-pipelines</a>
Merqury.FK	1.1.2	<a href="https://github.com/thegenemyers/MERQURY.FK">https://github.com/thegenemyers/MERQURY.FK</a>
MitoHiFi	2	<a href="https://github.com/marcelauliano/MitoHiFi">https://github.com/marcelauliano/MitoHiFi</a>
PretextView	0.2	<a href="https://github.com/wtsi-hpag/PretextView">https://github.com/wtsi-hpag/PretextView</a>
purge_dups	1.2.3	<a href="https://github.com/dfguan/purge_dups">https://github.com/dfguan/purge_dups</a>
SALSA	2.2	<a href="https://github.com/salsa-rs/salsa">https://github.com/salsa-rs/salsa</a>

**Data availability**

European Nucleotide Archive: *Dryobotodes eremita* (brindled green). Accession number [PRJEB46319](https://identifiers.org/ena.embl/PRJEB46319); <https://identifiers.org/ena.embl/PRJEB46319>. The genome sequence is released openly for reuse. The *Dryobotodes eremita* genome sequencing initiative is part of the Darwin Tree of Life (DToL) project. All raw sequence data and the assembly have been deposited in INSDC databases. Raw data and assembly accession identifiers are reported in [Table 1](#).

**Author information**

Members of the University of Oxford and Wytham Woods Genome Acquisition Lab are listed here: <https://doi.org/10.5281/zenodo.4789928>.

Members of the Darwin Tree of Life Barcoding collective are listed here: <https://doi.org/10.5281/zenodo.4893703>.

Members of the Wellcome Sanger Institute Tree of Life programme are listed here: <https://doi.org/10.5281/zenodo.4783585>.

Members of Wellcome Sanger Institute Scientific Operations: DNA Pipelines collective are listed here: <https://doi.org/10.5281/zenodo.4790455>.

Members of the Tree of Life Core Informatics collective are listed here: <https://doi.org/10.5281/zenodo.5013541>.

Members of the Darwin Tree of Life Consortium are listed here: <https://doi.org/10.5281/zenodo.4783558>.

## References

- Allio R, Schomaker-Bastos A, Romiguer 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](#)
- Bretherton RF, Goater B, Lorimer RI: **Noctuidae: Cucullinae to Hypeninae.** In: Heath J, and Emmet AM (eds) *The Moths and Butterflies of Great Britain and Ireland* Colchester: Harley Books. 1983.
- Brůna T, Hoff KJ, Lomsadze A, *et al.*: **BRAKER2: Automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database.** *NAR Genom Bioinform.* 2021; **3**(1): lqaa108.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- 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](#)
- Emmet AM: 1991.  
[Reference Source](#) ;*The Scientific Names of the British Lepidoptera - their History and Meaning* Colchester: Harley Books. | [Publisher Full Text](#)
- Garrison E, Marth G: *Haplotype-based variant detection from short-read sequencing.* 2012.  
[Publisher Full Text](#)
- GBIF Secretariat: **Dryobotodes eremita Fabricius, 1775.** *GBIF Backbone Taxonomy.* 2022; (Accessed: 23 February 2023).  
[Reference Source](#)
- Ghurye J, Rhie A, Walenz BP, *et al.*: **Integrating Hi-C links with assembly graphs for chromosome-scale assembly.** *PLoS Comput Biol.* 2019; **15**(8): e1007273.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Guan D, McCarthy SA, Wood J, *et al.*: **Identifying and removing haplotypic duplication in primary genome assemblies.** *Bioinformatics.* 2020; **36**(9): 2896–2898.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Harry E: **PretextView (Paired REad TEXTure Viewer): A desktop application for viewing pretext contact maps.** 2022; (Accessed: 19 October 2022).  
[Reference Source](#)
- 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.* Oxford University Press, 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](#)
- Manni M, Berkeley MR, Seppy 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](#)
- MothsIreland: *73.225 Brindled Green (Dryobotodes eremita).* 2022; (Accessed: 23 February 2023).  
[Reference Source](#)
- NBN Atlas Partnership: *Dryobotodes eremita (Fabricius, 1775).* Brindled Green: NBN Atlas. 2021; (Accessed: 23 February 2023).  
[Reference Source](#)
- Randle Z, Evans-Hill LJ, Parsons MS, *et al.*: 2019; *Atlas of Britain & Ireland's Larger Moths.* Newbury: NatureBureau.  
[Reference Source](#)
- Rao SSP, Huntley MH, Durand NC, *et al.*: **A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping.** *Cell.* 2014; **159**(7): 1665–1680.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Rhie A, 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](#)
- 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](#)
- Simão FA, Waterhouse RM, Ioannidis P, *et al.*: **BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs.** *Bioinformatics.* 2015; **31**(19): 3210–3212.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Stokoe WJ: 1948; *The Caterpillars of British Moths. Series 1.* London and New York: Warne and Co.  
[Reference Source](#)
- Uliano-Silva M, Ferreira JGRN, Krashenninnikova K, *et al.*: **MitoHiFi: a python pipeline for mitochondrial genome assembly from PacBio High Fidelity reads.** *bioRxiv. [Preprint]* 2022.  
[Publisher Full Text](#)

# Open Peer Review

Current Peer Review Status: ? ✓ ? ✓ ✓ ?

## Version 2

Reviewer Report 14 March 2026

<https://doi.org/10.21956/wellcomeopenres.28658.r148621>

© 2026 Hévin 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.



Noémie M.-C. Hévin 

Université de Montpellier, Montpellier, France

The authors present a high-quality genome assembly for *Dryobotodes eremita*. This is a valuable contribution: the genome is robust, and the methodological pipeline is clearly described. I particularly appreciated the informative figures and the background section, which provides useful biological context on the species. I only have a few suggestions:

1. A reference to this sentence could be added to the abstract: "A genome sequence for *Dryobotodes eremita* will facilitate studies investigating molecular adaptations to oak feeding and will contribute to the growing set of genomic resources for Lepidoptera".
2. Keywords should be added
3. Authors should detail how they identified Z and W chromosomes
4. The manuscript would benefit substantially from a concise discussion section. This should include a brief comparison of the newly generated genome with what is already known for related genera regarding chromosome number, genome size, and genome composition.

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

Yes

**Are the protocols appropriate and is the work technically sound?**

Yes

**Are sufficient details of methods and materials provided to allow replication by others?**

Partly

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

Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Genomics, macroevolution

**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.**

Reviewer Report 11 March 2026

<https://doi.org/10.21956/wellcomeopenres.28658.r148226>

© 2026 Kim W et al. 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.



**Zekun Wu**

Center for Life Sciences, Harbin Institute of Technology Center for Life Sciences (Ringgold ID: 598430), Harbin, Heilongjiang, China

**Woo Jae Kim** 

Harbin Institute of Technology, Harbin, China

**Xinyue Zhou**

Center for Life Sciences, Harbin Institute of Technology Center for Life Sciences (Ringgold ID: 598430), Harbin, Heilongjiang, China

This article presents a high-quality, chromosome-scale genome assembly of the Brindled Green moth, *Dryobotodes eremita*. The authors adopted well-established, technically sound long-read sequencing and Hi-C scaffolding workflows, and generated a highly complete, robust genomic resource for this species. All relevant sequencing data and assembly results have been deposited in public databases for open access, the dataset is clearly structured, and the manuscript fits the publishing scope of the journal's Data Note format. I have no suggestions for revision or further improvement.

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** *Drosophila*, Behavioral genetics, Neural circuits, Neuropeptide, Social behavior, Social neuroscience, fruit fly, Science and Society, Honey bee

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

Reviewer Report 11 March 2026

<https://doi.org/10.21956/wellcomeopenres.28658.r148615>

© 2026 Clark 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.

✓ **Melody Clark** 

Natural Environment Research Council, Cambridge, UK

This data note describes the genome sequence of a female Brindled Green moth *Dryobotodes eremita*. The background biology and biogeography are well described, especially the early life stage reliance on oak trees. DNA extraction techniques, sequencing, assembly and annotation are well described and use referenced methodologies and are open source, as expected by the Darwin Tree of Life. The final assembly comprises 32 chromosomal pseudochromosomes including the Z and W, plus the mitochondrial genome. This genome which was constructed using a small level of manual annotation is of very high quality.

**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:** transcriptomics and genomics of non-model species

**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 11 March 2026

<https://doi.org/10.21956/wellcomeopenres.28658.r148618>

© 2026 So W. 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.

✓ **Wai Lok So** 

<sup>1</sup> The Chinese University of Hong Kong, Hong Kong, Hong Kong

<sup>2</sup> The Hang Seng University of Hong Kong (Ringgold ID: 272025), Hong Kong, Hong Kong

The authors present a high-quality genome assembly of the Brindled Green, *Dryobotodes eremita*, as a part of Darwin Tree of Life project. The methods are standard, solid, clear and technically sound. The assembly was well assigned and annotated, including the Z and W sex chromosomes and a complete mitochondrial genome. The resultant resources are also transparent and accessible.

The genome assembly is considered to be a very useful resource for research on Lepidoptera ecology and evolution, and other related research including lepidopteran molecular adaptations as the authors have suggested.

I have no major concerns with the manuscript, except for the following two comments:

1. Can a scale bar be provided in Figure 1?
2. Can the authors state how the assembly of the sex chromosomes being validated? Are the sex chromosomes assembled by mapping to other reference genomes, or based on some marker genes with *de novo* assembly?

**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:** evolution, genomics, arthropods

**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 09 March 2026

<https://doi.org/10.21956/wellcomeopenres.28658.r148613>

© 2026 Homchan S. 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.

 **Somjit Homchan**

Naresuan University, Tha Pho, Thailand

This article presents valuable data with high-quality genome assembly of an adult female moth, *Dryobotodes eremita*. The authors also detail a chromosome-scale genome assembly achieved using various tools. However, I have a couple of suggestions for improvement:

1. It would be helpful if the authors included a scale in the specimen images so the reader can imagine how big or small of the insect specimen.
2. Additionally, more detailed information about the RNA analysis would be beneficial. Rather than just outlining the RNA extraction methods, a clearer explanation of the process would help readers understand and replicate the RNA analysis.
3. In figure 5 Hi-C contact map of the *Dryobotodes eremita* genome assembly, the chromosome W should be labeled next to the chromosome 30.

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

Yes

**Are the protocols appropriate and is the work technically sound?**

Yes

**Are sufficient details of methods and materials provided to allow replication by others?**

Partly

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

Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Molecular Markers, Population Genetics, Cytogenetics

**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.**

---

**Version 1**

Reviewer Report 01 October 2024

<https://doi.org/10.21956/wellcomeopenres.21383.r101162>

© 2024 Kim W et al. 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.



**Woo Jae Kim**

Harbin Institute of Technology, Harbin, China

**Xinyue Zhou**

Center for Life Sciences, Harbin Institute of Technology Center for Life Sciences (Ringgold ID: 598430), Harbin, Heilongjiang, China

**Zekun Wu**

Center for Life Sciences, Harbin Institute of Technology Center for Life Sciences (Ringgold ID: 598430), Harbin, Heilongjiang, China

The manuscript presents a high-quality de novo genome assembly of the adult female *Dryobotodes eremita* as part of the Darwin Tree of Life Project. Using good sequencing methods, including Pacific Biosciences HiFi long reads and Hi-C data, the authors assembled primary contigs into 32 chromosomal-level scaffolds, representing 30 autosomes and the W and Z sex chromosomes, with an N50 value of 23.3 Mb. The assembly achieves 99% completeness according to BUSCO, indicating a well-structured and robust genomic resource for this species.

However, there are a few areas that require further clarification:

1. While the methods are generally well-articulated, several steps, particularly related to DNA extraction and tissue selection, could be elaborated on further. For instance, the manuscript mentions that thorax tissue was cryogenically disrupted for DNA extraction. It would be beneficial to clarify the rationale behind selecting thorax tissue and how this choice may have impacted the quality or integrity of the extracted DNA.
2. The hermeneutics of the BlobToolKit snailplot remain obfuscated, rendering the comprehension of its semantic content arduous for the naive observer. The intricacies of its representational schema are not readily accessible, thereby necessitating a specialized level of expertise for the accurate interpretation of its graphic expression.
3. I note the presence of an unplaced scaffold measuring 14.73 Mbps in Table 2. Please elucidate the content of this scaffold, specifically whether it harbors any genes of functional importance. If so, please provide an explanation of their roles and potential significance. In the absence of such genes, please indicate whether further investigation or annotation of this scaffold is planned. Addressing this matter is crucial for assessing the comprehensiveness and applicability of the genome assembly.

Additionally, I have a few broader suggestions and inquiries regarding the overall context of the study:

1. I have observed that the Darwin Tree of Life Project has published several similar genome sequencing articles with consistent workflows. While I have no concerns about the quality of the genome assembly itself, as a reader, I would appreciate more context about the broader scope of the study.

For example:

How many insects have been sequenced, are plans in place to expand this dataset if necessary, and what role do Brindled Green play in the long-term insect genomics strategy of the Darwin Tree of Life Project?

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

Yes

**Are the protocols appropriate and is the work technically sound?**

Yes

**Are sufficient details of methods and materials provided to allow replication by others?**

Partly

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

Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Drosophila, Behavioral genetics, Neural circuits, Neuropeptide, Social behavior, Social neuroscience, fruit fly, Science and Society, Honey bee

**We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.**

Author Response 05 Feb 2026

#### **Tree of Life Team Sanger**

We thank the reviewer for their comments and suggestions. We address each point below.

**1. Tissue choice for DNA extraction** For Lepidoptera processed within the Darwin Tree of Life programme, a standardised tissue strategy is used where specimen size permits: thoracic tissue for high molecular weight DNA extraction, head tissue for Hi-C library preparation, and abdominal tissue for RNA sequencing. A reference to Howard *et al.*, 2025, in which the development of Tree of Life wet lab protocols is discussed has been added to the methods section. **2. Interpretation of the BlobToolKit snail plot** The BlobToolKit snail plot is a standard summary visualisation used across Darwin Tree of Life genome notes to present assembly size, contiguity, GC content, and BUSCO completeness as a high-level quality overview. **3. Unplaced sequence reported in Table 2** The value that was reported in Table 2 does not represent a single unplaced scaffold and has therefore been removed from the table. Rather, it combines the lengths of unlocalised scaffolds that could not be confidently ordered or oriented but are nonetheless assigned to chromosomes. For this assembly, 99.8% of the sequence is assigned to chromosomes. The revised Figure 5, generated using PretextSnapshot, shows these unlocalised scaffolds immediately after the chromosomes to which they belong. **Broader context of the study** This genome was generated as part of the Darwin Tree of Life Project, an ongoing initiative to produce high-quality reference genomes for all named eukaryotic species in Britain and Ireland using a consistent and scalable workflow. Genome notes in this series are intended to document individual assemblies rather than to provide a synthesis of project-wide progress or taxon-level strategy. We have added a short contextual sentence to the Background section. Reference: 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;  
<https://doi.org/10.1101/2025.04.11.648334>

**Competing Interests:** No competing interests were disclosed.

Reviewer Report 22 August 2024

<https://doi.org/10.21956/wellcomeopenres.21383.r94260>

© 2024 Lucek K. 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.



**Kay Lucek**

University of Neuchâtel, Neuchâtel, Switzerland

The authors present the chromosome level genome assembly of an apparently female specimen of the Brindled Green *Dryobotodes*. The assembly consists of 32 chromosomes. The assembly is highly complete as revealed by the high BUSCO score. Sequencing and genome assembly seem to have followed using established methods. The assembly has been generated using an established pipeline of the Darwin Tree of Life Consortium.

Overall, this genome note leaves me with a lot of questions. For instance, RNA was extracted and sequenced, but no indication is given what it was used for. Similarly, 55-fold coverage in 10X Genomics read clouds were generated, but what for? The genome note lacks the basic information about the sex of the sequenced specimen. Agreeably the presence of a W chromosome suggests that it is a female, yet this could also simply stated. More importantly, Table 2 records an unplaced scaffold of 14.73 Mbps.

No indication is given why this relatively large amount of data is unplaced. Nevertheless, the presented assembly will be of great value to study genome evolution in Lepidoptera.

**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?**

No

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

Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Speciation 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, however I have significant reservations, as outlined above.**

---