



DATA NOTE

The genome sequence of the Green Oak Leaf Roller, *Tortrix viridana* (Linnaeus, 1758) [version 1; peer review: 3 approved]

Douglas Boyes¹⁺, Liam M. Crowley², Lucy M. Morley²,
 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

¹UK Centre for Ecology & Hydrology, Wallingford, England, UK

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

+ Deceased author

V1 First published: 15 May 2024, 9:260
<https://doi.org/10.12688/wellcomeopenres.21568.1>
 Latest published: 15 May 2024, 9:260
<https://doi.org/10.12688/wellcomeopenres.21568.1>

Abstract

We present a genome assembly from a female *Tortrix viridana* (the Green Oak Leaf Roller; Arthropoda; Insecta; Lepidoptera; Tortricidae). The genome sequence is 456.6 megabases in span. Most of the assembly is scaffolded into 31 chromosomal pseudomolecules, including the Z and W sex chromosomes. The mitochondrial genome has also been assembled and is 16.43 kilobases in length.

Keywords

Tortrix viridana, green oak leaf roller, genome sequence, chromosomal, Lepidoptera



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

Open Peer Review

Approval Status

	1	2	3
version 1 15 May 2024	 view	 view	 view

- Ningning Wu** , China Agricultural University College of Plant Protection, Beijing, China
- Kathy Darragh** , Indiana University Bloomington, Bloomington, USA
- Thayná Bisson Ferraz Lopes** , Universidade Estadual de Londrina, Londrina, Brazil

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: **Boyes D:** Investigation, Resources; **Crowley LM:** Investigation, Resources, Writing – Review & Editing; **Morley LM:** Writing – Original Draft Preparation;

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: © 2024 Boyes D *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: Boyes D, Crowley LM, Morley LM *et al.* **The genome sequence of the Green Oak Leaf Roller, *Tortrix viridana* (Linnaeus, 1758) [version 1; peer review: 3 approved]** Wellcome Open Research 2024, 9:260 <https://doi.org/10.12688/wellcomeopenres.21568.1>

First published: 15 May 2024, 9:260 <https://doi.org/10.12688/wellcomeopenres.21568.1>

Species taxonomy

Eukaryota; Opisthokonta; Metazoa; Eumetazoa; Bilateria; Protostomia; Ecdysozoa; Panarthropoda; Arthropoda; Mandibulata; Pancrustacea; Hexapoda; Insecta; Dicondylia; Pterygota; Neoptera; Endopterygota; Amphiesmenoptera; Lepidoptera; Glossata; Neolepidoptera; Heteroneura; Ditrysia; Apoditrysia; Tortricoidea; Tortricidae; Tortricinae; Tortricini; *Tortrix*; *Tortrix viridana* (Linnaeus, 1758) (NCBI:txid311328).

Background

The Green Oak Leaf Roller or Green Oak Tortrix, *Tortrix viridana*, is a moth of the Tortricidae family (Sterling *et al.*, 2023). It is common and widespread across much of Britain (Davis, 2012). It is widely distributed across Europe and east to Asia Minor (GBIF Secretariat, 2024). *T. viridana* has a wingspan of 17–24 mm and is a very distinctive micro-moth, being the only UK tortricid species with unmarked bright green forewings and head (Sterling *et al.*, 2023). However, it may appear similar to the macro-moth the Cream-bordered Green Pea (*Earias chlorana*; Nolidae). *Earias chlorana* may be distinguished by a broader white line along the costa and by the fact that it rests with the wings held tent-like, rather than flat as with *T. viridana* (Sterling *et al.*, 2023). *T. viridana* hindwings are light grey.

The adults flight period occurs from mid-May to July and they are attracted to light (Sterling *et al.*, 2023). The species is univoltine and females lay eggs in pairs on the bark of twigs. Larvae hatch in the following spring, between April and June, to feed on flushing buds during the earlier instars and then on the fresh expanding leaves as later instars (Du Merle, 1999). Larvae are greyish olive green in colour, with a black or dark brown head, dark pinacula and black thoracic legs. It is narrowly oligophagous, hatching and feeding predominantly on oaks (*Quercus* spp.) and *T. viridana* is an abundant Lepidopteran species in oak woodlands (Du Merle, 1999). As the common name suggests, the larvae roll or fold host leaves to form a feeding tube that they feed within (Sterling *et al.*, 2023).

T. viridana is considered a pest of oaks in Europe, as it undergoes population cycles with periodic outbreaks leading to heavy defoliation that may reduce wood and acorn production and make trees more vulnerable to further pest attacks, e.g. oak powdery mildew (*Erysiphe (Microsphaera) alphitoides*) (Schroeder & Degen, 2008a; Szabóky & Csóka, 2010). Therefore, research has aimed to untangle the influential factors on *T. viridana*'s cyclical dynamics, including the role of host tree phenology, competition (e.g. with *Operophtera brumata*), dispersal and mating behaviour, and population genetics across space, to which this full reference genome will further contribute (Hunter, 1990; Hunter, 1998; Schroeder & Degen, 2008a; Schroeder & Degen, 2008b; Simchuk *et al.*, 1999).

Genome sequence report

The genome was sequenced from a female *Tortrix viridana* (Figure 1) collected from Wytham Woods, Wytham Woods, UK (51.77, -1.34). A total of 60-fold coverage in Pacific



Figure 1. Photograph of the *Tortrix viridana* (ilTorViri5) specimen used for genome sequencing.

Biosciences single-molecule HiFi long reads was generated. Primary assembly contigs were scaffolded with chromosome conformation Hi-C data. Manual assembly curation corrected 3 missing joins or mis-joins and removed 1 haplotypic duplication, reducing the scaffold number by 4.44%.

The final assembly has a total length of 456.6 Mb in 42 sequence scaffolds with a scaffold N50 of 14.8 Mb (Table 1). The snail plot in Figure 2 provides a summary of the assembly statistics, while the distribution of assembly scaffolds on GC proportion and coverage is shown in Figure 3. The cumulative assembly plot in Figure 4 shows curves for subsets of scaffolds assigned to different phyla. Most (99.88%) of the assembly sequence was assigned to 31 chromosomal-level scaffolds, representing 29 autosomes and the Z and W sex chromosomes. Chromosome-scale scaffolds confirmed by the Hi-C data are named in order of size (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 and can be found as a contig within the multifasta file of the genome submission.

The estimated Quality Value (QV) of the final assembly is 64.4 with *k*-mer completeness of 100.0%, and the assembly has a BUSCO v5.3.2 completeness of 98.4% (single = 97.7%, duplicated = 0.7%), using the lepidoptera_odb10 reference set (*n* = 5,286).

Metadata for specimens, barcode results, spectra estimates, sequencing runs, contaminants and pre-curation assembly statistics are given at <https://links.tol.sanger.ac.uk/species/311328>.

Methods

Sample acquisition and nucleic acid extraction

A female *Tortrix viridana* (specimen ID Ox001663, ToLID ilTorViri5) was netted in Wytham Woods, Oxfordshire (biological vice-county Berkshire), UK (latitude 51.77, longitude -1.34) on 2021-07-17. The specimen was collected by Douglas Boyes (University of Oxford) and identified by Liam Crowley (University of Oxford) and preserved on dry ice.

Table 1. Genome data for *Tortrix viridana*, iTorViri5.1.

Project accession data		
Assembly identifier	iTorViri5.1	
Species	<i>Tortrix viridana</i>	
Specimen	iTorViri5	
NCBI taxonomy ID	311328	
BioProject	PRJEB64712	
BioSample ID	SAMEA10978932	
Isolate information	iTorViri5, female: whole organism (DNA and Hi-C sequencing)	
Assembly metrics*		Benchmark
Consensus quality (QV)	64.4	≥ 50
k-mer completeness	100.0%	≥ 95%
BUSCO**	C:98.4%,S:97.7%,D:0.7%, F:0.4%,M:1.3%,n:5,286	C ≥ 95%
Percentage of assembly mapped to chromosomes	99.88%	≥ 95%
Sex chromosomes	ZW	localised homologous pairs
Organelles	Mitochondrial genome: 16.43 kb	complete single alleles
Raw data accessions		
PacificBiosciences SEQUEL II	ERR11809137	
Hi-C Illumina	ERR11814102	
Genome assembly		
Assembly accession	GCA_963241965.1	
Accession of alternate haplotype	GCA_963243755.1	
Span (Mb)	456.6	
Number of contigs	88	
Contig N50 length (Mb)	10.1	
Number of scaffolds	42	
Scaffold N50 length (Mb)	14.8	
Longest scaffold (Mb)	62.23	

* 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 version 5.3.2. 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/CAUJKY01/dataset/CAUJKY01/busco>.

The workflow for high molecular weight (HMW) DNA extraction at the Wellcome Sanger Institute (WSI) includes a sequence of core procedures: sample preparation; sample homogenisation, DNA extraction, fragmentation, and clean-up. In sample preparation, the iTorViri5 sample was weighed and dissected on dry ice ([Jay et al., 2023](#)). Tissue from the whole

organism was homogenised using a PowerMasher II tissue disruptor ([Denton et al., 2023a](#)). HMW DNA was extracted using the Automated MagAttract v2 protocol ([Oatley et al., 2023a](#)). DNA was sheared into an average fragment size of 12–20 kb in a Megaruptor 3 system ([Bates et al., 2023](#)). Sheared DNA was purified by solid-phase reversible immobilisation

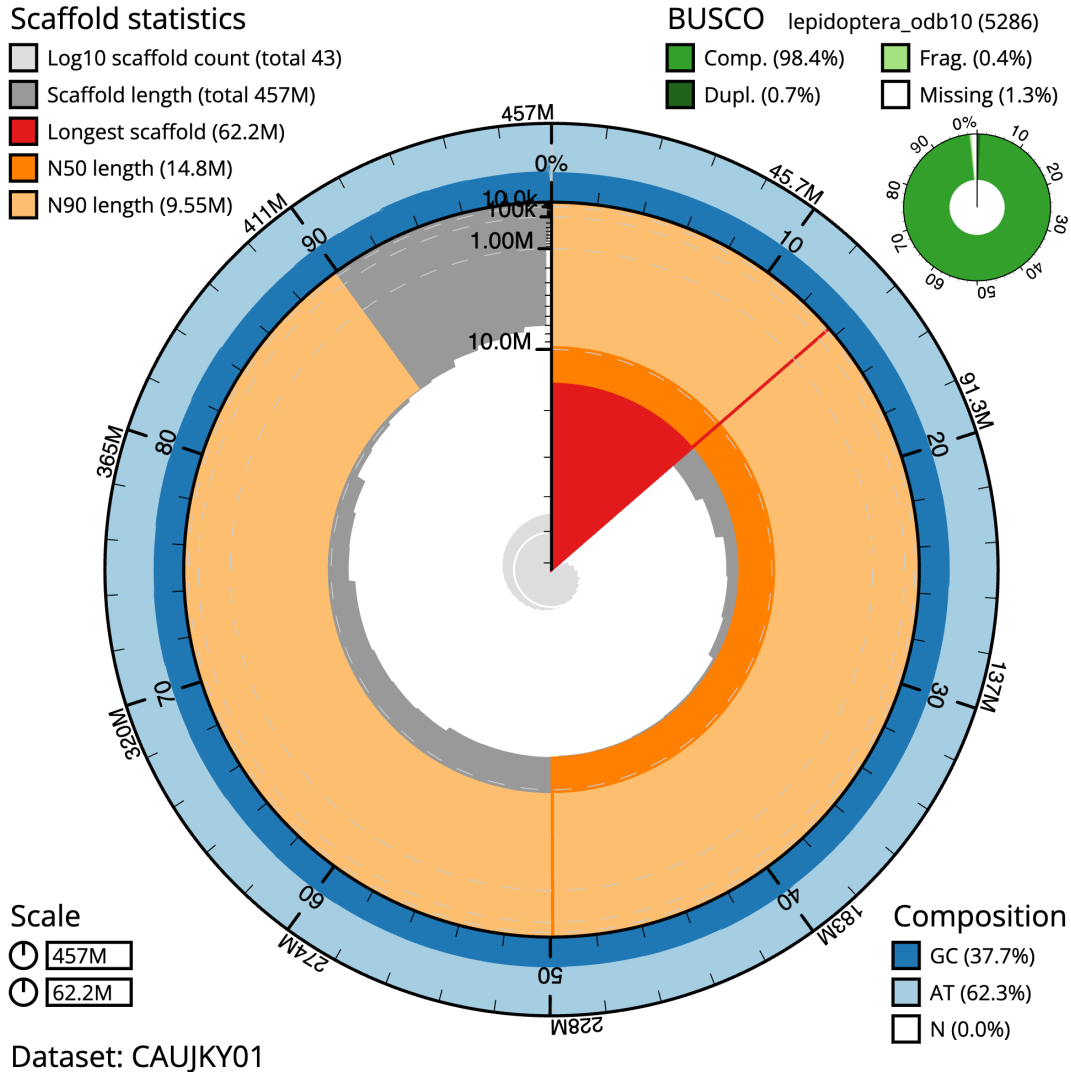


Figure 2. Genome assembly of *Tortrix viridana*, ilTorViri5.1: metrics. The BlobToolKit snail plot shows N50 metrics and BUSCO gene completeness. The main plot is divided into 1,000 size-ordered bins around the circumference with each bin representing 0.1% of the 456,586,594 bp assembly. The distribution of scaffold lengths is shown in dark grey with the plot radius scaled to the longest scaffold present in the assembly (62,234,340 bp, shown in red). Orange and pale-orange arcs show the N50 and N90 scaffold lengths (14,768,249 and 9,547,436 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/CAUJKY01/dataset/CAUJKY01/snail>.

(Oatley *et al.*, 2023b): in brief, the method employs a 1.8X ratio of AMPure PB beads to sample to eliminate shorter fragments and concentrate the DNA. 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.

Protocols developed by the WSI Tree of Life laboratory are publicly available on protocols.io (Denton *et al.*, 2023b).

Sequencing

Pacific Biosciences HiFi circular consensus DNA sequencing libraries were constructed according to the manufacturers' instructions. DNA sequencing was performed by the Scientific Operations core at the WSI on a Pacific Biosciences

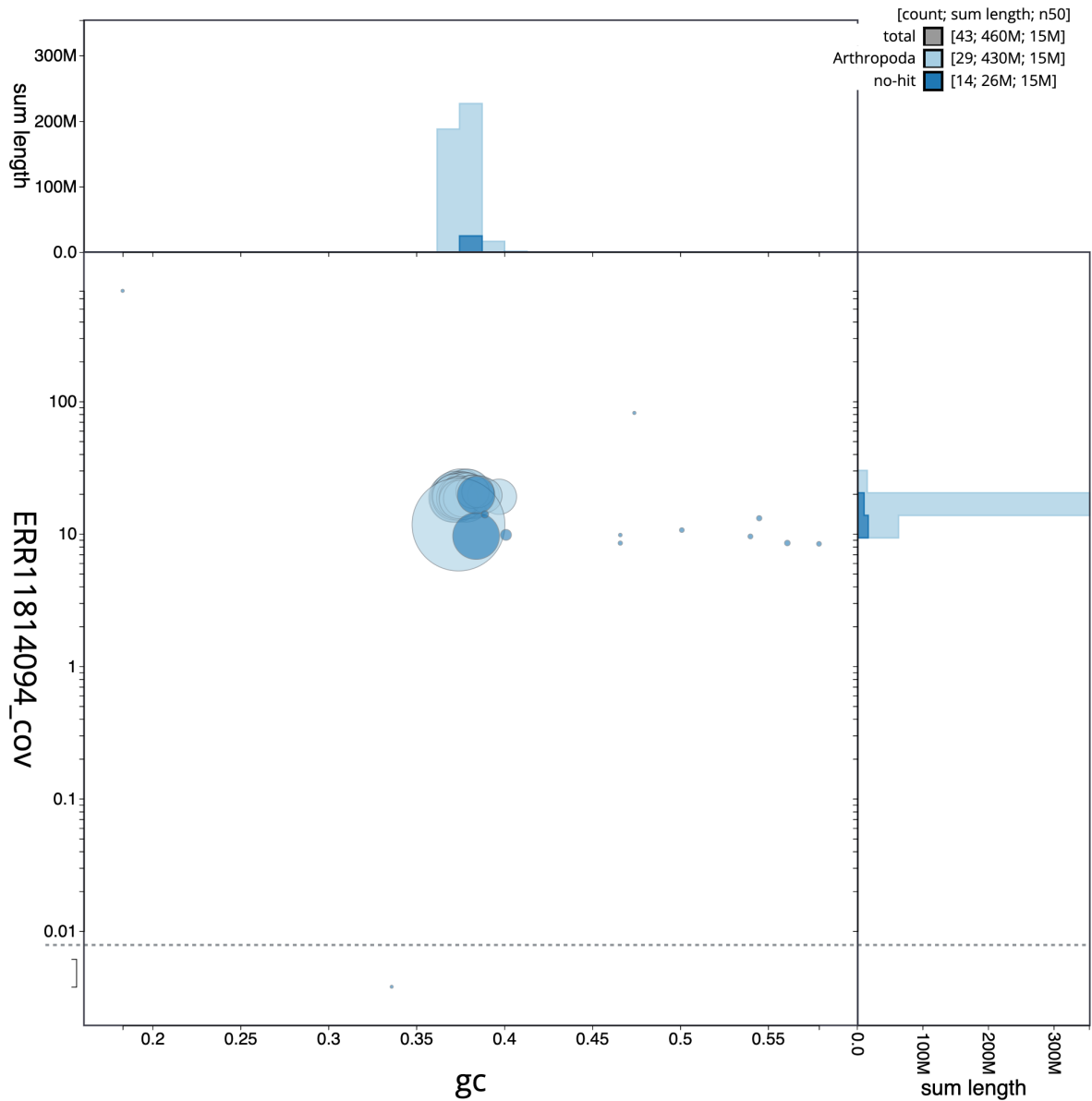


Figure 3. Genome assembly of *Tortrix viridana*, iTorViri5.1: BlobToolKit GC-coverage plot. Sequences are coloured by phylum. Circles are sized in proportion to sequence length. Histograms show the distribution of sequence length sum along each axis. An interactive version of this figure is available at <https://blobtoolkit.genomehubs.org/view/CAUJKY01/dataset/CAUJKY01/blob>.

SEQUEL II instrument. Hi-C data were also generated from remaining tissue of iTorViri5 using the Arima2 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). The assembly was then scaffolded with Hi-C data (Rao *et al.*, 2014) using YaHS (Zhou *et al.*, 2023). The assembly was checked for contamination and corrected using the gEVAL system (Chow *et al.*, 2016) as

described previously (Howe *et al.*, 2021). Manual curation was performed using gEVAL, HiGlass (Kerpedjiev *et al.*, 2018) and PretextView (Harry, 2022). The mitochondrial genome was assembled using MitoHiFi (Uliano-Silva *et al.*, 2023), which runs MitoFinder (Allio *et al.*, 2020) or MITOS (Bernt *et al.*, 2013) and uses these annotations to select the final mitochondrial contig and to ensure the general quality of the sequence.

A Hi-C map for the final assembly was produced using bwa-mem2 (Vasimuddin *et al.*, 2019) in the Cooler file

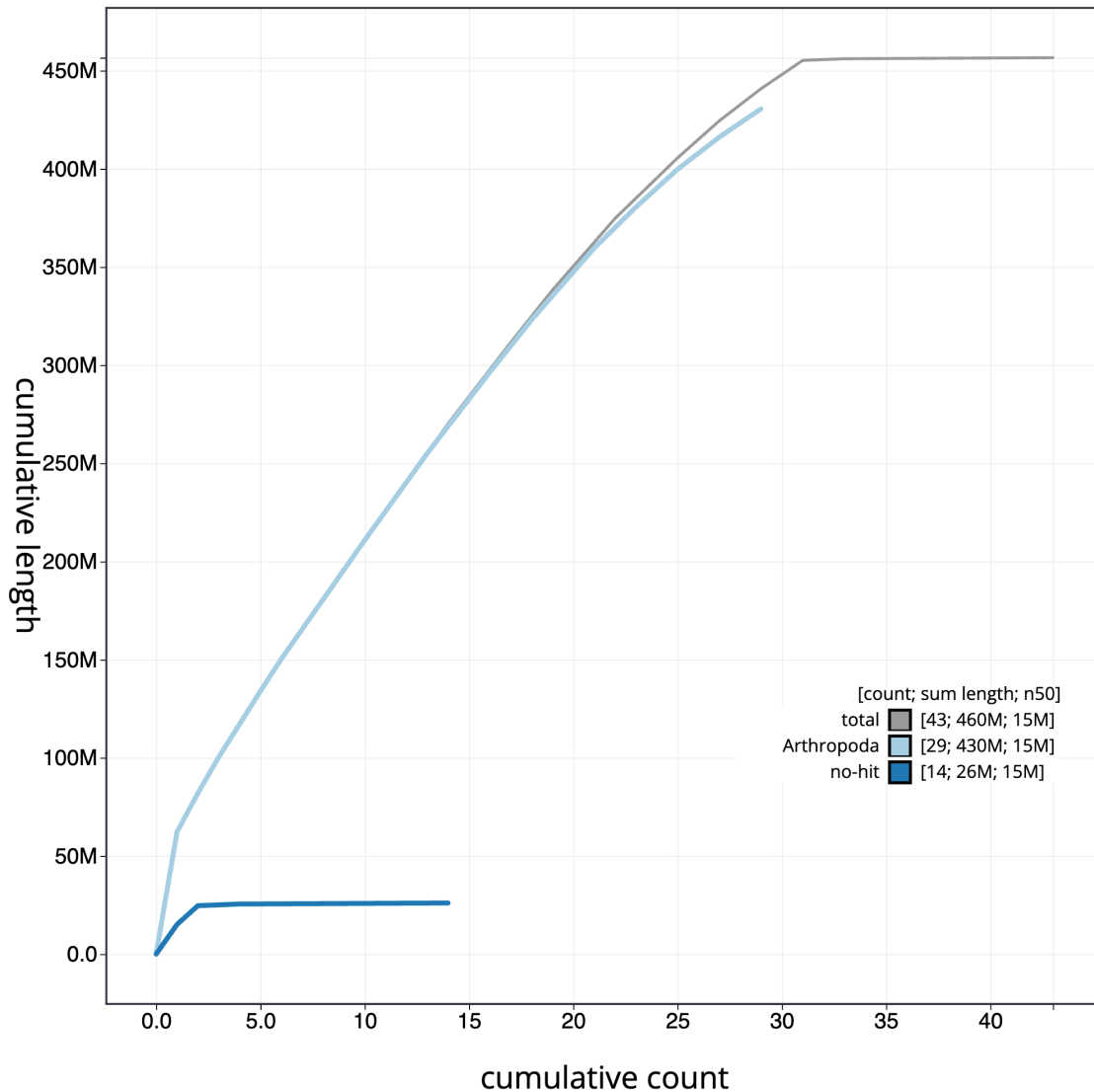


Figure 4. Genome assembly of *Tortrix viridana*, iTorViri5.1: BlobToolKit cumulative sequence plot. The grey line shows cumulative length for all sequences. Coloured lines show cumulative lengths of sequences assigned to each phylum using the buscogenes taxrule. An interactive version of this figure is available at <https://blobtoolkit.genomehubs.org/view/CAUJKY01/dataset/CAUJKY01/cumulative>.

format (Abdennur & Mirny, 2020). To assess the assembly metrics, the k -mer completeness and QV consensus quality values were calculated in Merqury (Rhie *et al.*, 2020). This work was done using Nextflow (Di Tommaso *et al.*, 2017) DSL2 pipelines “sanger-tol/readmapping” (Surana *et al.*, 2023a) and “sanger-tol/genomenote” (Surana *et al.*, 2023b). 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 relevant software tool versions and sources.

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 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 [here](#). 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.

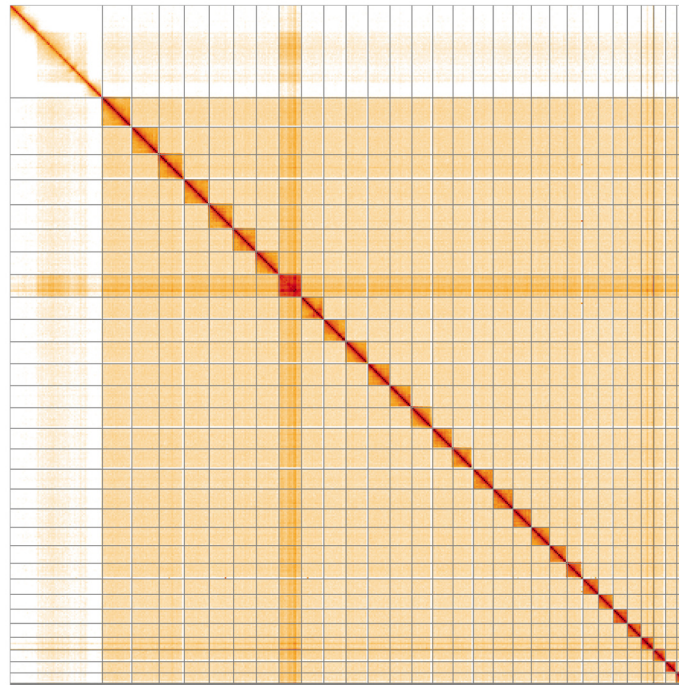


Figure 5. Genome assembly of *Tortrix viridana*, iTorViri5.1: Hi-C contact map of the iTorViri5.1 assembly, visualised using HiGlass. Chromosomes are shown in order of size from left to right and top to bottom. An interactive version of this figure may be viewed at <https://genome-note-higlass.tol.sanger.ac.uk/I/?d=YTZURUETToenilpT27Mfeg>.

Table 2. Chromosomal pseudomolecules in the genome assembly of *Tortrix viridana*, iTorViri5.

INSDC accession	Chromosome	Length (Mb)	GC%
OY725172.1	1	19.7	38.0
OY725173.1	2	18.37	37.5
OY725174.1	3	16.95	38.0
OY725175.1	4	16.72	37.0
OY725176.1	5	16.34	37.5
OY725178.1	6	15.34	37.0
OY725179.1	7	15.23	37.5
OY725180.1	8	15.0	37.5
OY725181.1	9	14.96	37.0
OY725182.1	10	14.77	37.5
OY725183.1	11	14.77	37.5
OY725184.1	12	14.73	37.0
OY725185.1	13	13.96	37.5
OY725186.1	14	13.73	37.0
OY725187.1	15	13.61	38.0

INSDC accession	Chromosome	Length (Mb)	GC%
OY725188.1	16	13.34	37.5
OY725189.1	17	13.33	38.0
OY725190.1	18	12.36	37.5
OY725191.1	19	12.25	38.5
OY725192.1	20	11.87	37.5
OY725193.1	21	10.56	38.5
OY725194.1	22	10.35	37.5
OY725195.1	23	9.95	37.5
OY725196.1	24	9.55	38.5
OY725197.1	25	9.34	37.5
OY725198.1	26	8.45	39.5
OY725199.1	27	7.9	39.0
OY725200.1	28	7.5	38.0
OY725201.1	29	6.93	38.5
OY725177.1	W	15.12	38.5
OY725171.1	Z	62.23	37.5
OY725202.1	MT	0.02	18.5

Table 3. Software tools: versions and sources.

Software tool	Version	Source
BlobToolKit	4.2.1	https://github.com/blobtoolkit/blobtoolkit
BUSCO	5.3.2	https://gitlab.com/ezlab/busco
gEVAL	N/A	https://geval.org.uk/
Hifiasm	0.19.5-r587	https://github.com/chhylp123/hifiasm
HiGlass	1.11.6	https://github.com/higlass/higlass
Merqury	MerquryFK	https://github.com/thegenemyers/MERQURY.FK
MitoHiFi	3	https://github.com/marcelauliano/MitoHiFi
PretextView	0.2	https://github.com/wtsi-hpag/PretextView
purge_dups	1.2.5	https://github.com/dfguan/purge_dups
sanger-tol/genomenote	v1.0	https://github.com/sanger-tol/genomenote
sanger-tol/readmapping	1.1.0	https://github.com/sanger-tol/readmapping/tree/1.1.0
YaHS	1.2a.2	https://github.com/c-zhou/yahs

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 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: *Tortrix viridana*. Accession number PRJEB64712; <https://identifiers.org/ena.embl/PRJEB64712> (Wellcome Sanger Institute, 2023). The genome sequence is released openly for reuse. The *Tortrix viridana* 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. 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](#).

Author information

Members of the University of Oxford and Wytham Woods Genome Acquisition Lab are listed here: <https://doi.org/10.5281/zenodo.7125292>.

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 Management, Samples and Laboratory team are listed here: <https://doi.org/10.5281/zenodo.10066175>.

Members of Wellcome Sanger Institute Scientific Operations: Sequencing Operations are listed here: <https://doi.org/10.5281/zenodo.10043364>.

Members of the Wellcome Sanger Institute Tree of Life Core Informatics team are listed here: <https://doi.org/10.5281/zenodo.10066637>.

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

- Abdennur N, Mirny LA: **Cooler: scalable storage for Hi-C data and other genomically labeled arrays.** *Bioinformatics.* 2020; **36**(1): 311–316.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Allio R, Schomaker-Bastos A, Romiguiet 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](#)
- Bates A, Clayton-Lucey I, Howard C: **Sanger Tree of Life HMW DNA fragmentation: diagenode Megaruptor[®]3 for LI PacBio.** *protocols.io.* 2023.
[Publisher Full Text](#)
- Bernt M, Donath A, Jühling F, et al.: **MITOS: improved de novo metazoan mitochondrial genome annotation.** *Mol Phylogenet Evol.* 2013; **69**(2): 313–319.
[PubMed Abstract](#) | [Publisher 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](#)
- Chow W, Brugger K, Caccamo M, et al.: **gEVAL - a web-based browser for evaluating genome assemblies.** *Bioinformatics.* 2016; **32**(16): 2508–2510.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Davis T: **A review of the status of Microlepidoptera in Britain.** *Butterfly Conservation Report.* 2012; [Accessed 6 December 2022].
[Reference Source](#)
- Denton A, Oatley G, Cornwell C, et al.: **Sanger Tree of Life sample homogenisation: PowerMash.** *protocols.io.* 2023a.
[Publisher Full Text](#)
- Denton A, Yatsenko H, Jay J, et al.: **Sanger Tree of Life wet laboratory protocol collection v1.** *protocols.io.* 2023b.
[Publisher Full Text](#)
- Di Tommaso P, Chatzou M, Floden EW, et al.: **Nextflow enables reproducible computational workflows.** *Nat Biotechnol.* 2017; **35**(4): 316–319.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Du Merle P: **Egg development and diapause: ecophysiological and genetic basis of phenological polymorphism and adaptation to varied hosts in the green oak tortrix, *Tortrix viridana* L. (Lepidoptera: Tortricidae).** *J Insect Physiol.* 1999; **45**(6): 599–611.
[PubMed Abstract](#) | [Publisher Full Text](#)
- GBIF Secretariat: ***Tortrix viridana* (Linnaeus, 1758).** *GBIF Backbone Taxonomy.* 2024; [Accessed 20 March 2024].
[Reference Source](#)
- Guan D, McCarthy SA, Wood J, et al.: **Identifying and removing haplotypic duplication in primary genome assemblies.** *Bioinformatics.* 2020; **36**(9): 2896–2898.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Harry E: **PretextView (Paired REad TEXTure Viewer): a desktop application for viewing pretext contact maps.** 2022; [Accessed 19 October 2022].
[Reference Source](#)
- Howe K, Chow W, Collins J, et al.: **Significantly improving the quality of genome assemblies through curation.** *GigaScience.* Oxford University Press, 2021; **10**(1): gjaa153.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Hunter MD: **Differential susceptibility to variable plant phenology and its role in competition between two insect herbivores on oak.** *Ecol Entomol.* 1990; **15**(4): 401–408.
[Publisher Full Text](#)
- Hunter MD: **Interactions between *Operophtera brumata* and *Tortrix viridana* on oak: new evidence from time-series analysis.** *Ecol Entomol.* 1998; **23**(2): 168–173.
[Publisher Full Text](#)
- Jay J, Yatsenko H, Narváez-Gómez JP, et al.: **Sanger Tree of Life sample preparation: triage and dissection.** *protocols.io.* 2023.
[Publisher 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, 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](#)
- Oatley G, Denton A, Howard C: **Sanger Tree of Life HMW DNA extraction: automated magAttract v.2.** *protocols.io.* 2023a.
[Publisher Full Text](#)
- Oatley G, Sampaio F, Howard C: **Sanger Tree of Life fragmented DNA clean up: automated SPRI.** *protocols.io.* 2023b; [Accessed 21 November 2023].
[Publisher 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](#)
- Schroeder H, Degen B: **Genetic structure of the green oak leaf roller (*Tortrix viridana* L.) and one of its hosts, *Quercus robur* L.** *For Ecol Manage.* 2008a; **256**(6): 1270–1279.
[Publisher Full Text](#)
- Schroeder H, Degen B: **Spatial genetic structure in populations of the green oak leaf roller, *Tortrix viridana* L. (Lepidoptera, Tortricidae).** *Eur J For Res.* 2008b; **127**(6): 447–453.
[Publisher 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](#)
- Simchuk AP, Ivashov AV, Companyntsev VA: **Genetic patterns as possible factors causing population cycles in oak leafroller moth, *Tortrix viridana* L.** *For Ecol Manage.* 1999; **113**(1): 35–49.
[Publisher Full Text](#)
- Sterling P, Parsons M, Lewington R: **Field guide to the micro-moths of Great Britain and Ireland, second edition.** London: Bloomsbury Publishing, 2023.
[Reference Source](#)
- Surana P, Muffato M, Qi G: **sanger-tol/readmapping: sanger-tol/readmapping v1.1.0 - Hebridean Black (1.1.0).** *Zenodo.* 2023a.
[Publisher Full Text](#)
- Surana P, Muffato M, Sadasivan Baby C: **sanger-tol/genomenote (v1.0.dev).** *Zenodo.* 2023b.
[Publisher Full Text](#)
- Szabó C, Csóka G: **Sodrómolysok - Tortricids.** Miskolc, Hungary: Forest Research Institute, 2010.
- Uliano-Silva M, Ferreira JGRN, Krashennikova 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](#)
- Wellcome Sanger Institute: **The genome sequence of the Green Oak Leaf Roller, *Tortrix viridana* (Linnaeus, 1758).** European Nucleotide Archive, [dataset], accession number PRJEB64712, 2023.
- 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 10 October 2024

<https://doi.org/10.21956/wellcomeopenres.23840.r102248>

© 2024 Lopes T. 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.



Thayná Bisson Ferraz Lopes 

Universidade Estadual de Londrina, Londrina, State of Paraná, Brazil

This data note presents a high-quality assembly of the genome of the Green Oak Leaf Roller, *Tortrix viridana*.

The background section encompasses all pertinent information regarding this species. The methodologies used for sample acquisition, nucleic acid extraction, sequencing, and genome assembly were adequate and well described. The figures and assembly statistics provide valuable insights into the quality and validation of the assembly.

This data will serve as an invaluable resource for researchers studying moth genomics and ecology in future investigations.

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: Bioinformatics and molecular genetics of insects

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 10 October 2024

<https://doi.org/10.21956/wellcomeopenres.23840.r101525>

© 2024 Darragh 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.



Kathy Darragh 

Indiana University Bloomington, Bloomington, Indiana, USA

The authors describe their genome assembly of a female *Tortrix viridana* moth (green oak leaf roller). Background on the ecology and natural history of the moth is provided. The data collection, sequencing, and assembly methods are all clearly described and appropriate. This provides useful data for public use through the 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?

Yes

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

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: 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 17 September 2024

<https://doi.org/10.21956/wellcomeopenres.23840.r97129>

© 2024 Wu 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.



Ningning Wu 

China Agricultural University College of Plant Protection, Beijing, Beijing, China

The rationale for generating the dataset is clearly. The genome of *Tortrix viridana* is presented as valuable for understanding the species' cyclical population dynamics and its role as a pest in European oak forests. The dataset is contextualized within research on Lepidopteran species and pest management, with potential contributions to future studies on population genetics, pest control, and forest ecosystem health.

However, the rationale could be enhanced by directly highlighting specific scientific questions or hypotheses this dataset addresses, to emphasize its broader relevance. The protocols are appropriate for genome assembly. The inclusion of multiple steps such as manual curation and contamination checks further ensures the technical rigor of the work.

Additionally, the use of standardized tools like Hifiasm, purge_dups, and BUSCO analysis provides confidence in the accuracy and completeness of the assembly. Overall, the technical work appears sound. The manuscript provides thorough descriptions of the methods, including the DNA extraction workflow, sequencing, assembly, and curation steps. Protocols for high molecular weight DNA extraction, shearing, and library preparation are sufficiently detailed, with relevant references to publicly available protocols.io entries, making replication feasible.

The datasets are presented in a clear and accessible manner. The assembly statistics, BUSCO scores, and figures are informative and provide useful insights into the quality of the assembly. The inclusion of metadata and external links to barcode results, spectra estimates, and assembly statistics further enhances the usability of the data.

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: Genome evolution

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.
