Check for updates

DATA NOTE

The genome sequence of the Olive Crescent, *Trisateles emortualis* (Denis & Schiffermüller, 1775) [version 1; peer review: awaiting peer review]

Douglas Boyes[1+], Peter W.H. Holland[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

[1]UK Centre for Ecology & Hydrology, Wallingford, England, UK
[2]University of Oxford, Oxford, England, UK

[+] Deceased author

Open Peer Review

Approval Status *AWAITING PEER REVIEW*

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

Abstract

We present a genome assembly from an individual male *Trisateles emortualis* (the Olive Crescent; Arthropoda; Insecta; Lepidoptera; Noctuidae). The genome sequence is 565.5 megabases in span. Most of the assembly is scaffolded into 31 chromosomal pseudomolecules, including the Z sex chromosome. The mitochondrial genome has also been assembled and is 16.01 kilobases in length. Gene annotation of this assembly on Ensembl identified 13,176 protein coding genes.

Keywords

Trisateles emortualis, Olive Crescent, genome sequence, chromosomal, Lepidoptera

This article is included in the Tree of Life gateway.

Corresponding author: Darwin Tree of Life Consortium (mark.blaxter@sanger.ac.uk)

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

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 Olive Crescent,** *Trisateles emortualis* **(Denis & Schiffermüller, 1775) [version 1; peer review: awaiting peer review]** Wellcome Open Research 2024, **9**:178 https://doi.org/10.12688/wellcomeopenres.20836.1

First published: 08 Apr 2024, **9**:178 https://doi.org/10.12688/wellcomeopenres.20836.1

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; Obtectomera; Noctuoidea; Noctuidae; Acontiinae; *Trisateles*; *Trisateles emortualis* (Denis & Schiffermüller, 1775) (NCBI:txid753441).

Background

The Olive Crescent *Trisateles emortualis* is a greenish-brown moth notable for local rarity, intriguing etymology and an unusual larval food source. The moth has a sparsely scattered distribution across central, eastern and northern Europe, Scandinavia, Ukraine, Russia, Korea and Japan (GBIF Secretariat, 2023). In Britain, the moth was an extreme rarity through most of the 19th and 20th centuries, with occasional individuals recorded in southern counties from the 1850s onwards and a breeding colony found in woodlands on the Chiltern Hills in the 1960s (Bretherton *et al.,* 1983). Small colonies were also found in Essex and Suffolk in the 1970s (Randle *et al.*, 2019). Since 2000, the number of records has been increasing, probably due to sporadic migration from mainland Europe plus establishment of new colonies in additional southern counties including Hampshire (HantsMoths, 2023; Randle *et al.*, 2019).

The adult moth has a very short flight period around midsummer, with the larvae feeding into autumn on dead and withered leaves of oak *Quercus* spp. and beech *Fagus sylvatica* (Bretherton *et al.,* 1983). There is an apparent preference for dry leaves attached to twigs and branches hanging underneath the tree canopy (Davis, 2013). The pupal stage overwinters.

The species has been placed in several different genera over the years and is now assigned to the genus *Trisateles* which contains very few species worldwide. The species name *emortualis*, referring to death, seems apt considering the food source of the larvae. It has been argued that this is coincidental, however, as the life cycle may not have been known when the specific name was assigned (Emmet, 1991).

A genome sequence of *Trisateles emortualis* was determined as part of the Darwin Tree of Life project. The genome sequence will facilitate research into biochemical adaptations for digesting dead organic matter and contribute to the growing set of resources for studying molecular evolution in insects.

Genome sequence report

The genome was sequenced from one male *Trisateles emortualis* (Figure 1) collected from Wytham Woods, Oxfordshire, UK (51.77, –1.34). A total of 27-fold coverage in Pacific Biosciences single-molecule HiFi long reads was generated. Primary assembly contigs were scaffolded with chromosome conformation Hi-C data. Manual assembly curation corrected 12 missing joins or mis-joins and removed 5 haplotypic duplications, reducing the assembly length by 0.65% and the scaffold number by 15.38%.

Figure 1. Photograph of the *Trisateles emortualis* (ilTriEmor1) specimen used for genome sequencing.

The final assembly has a total length of 565.5 Mb in 32 sequence scaffolds with a scaffold N50 of 20.1 Mb (Table 1). The snailplot 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 (99.99%) of the assembly sequence was assigned to 31 chromosomal-level scaffolds, representing 30 autosomes and the Z sex chromosome. 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 67.4 with *k*-mer completeness of 100.0%, and the assembly has a BUSCO v5.3.2 completeness of 98.7% (single = 97.9%, duplicated = 0.8%), 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/753441.

Genome annotation report

The *Trisateles emortualis* genome assembly (GCA_947095525.1) was annotated using the Ensembl rapid annotation pipeline (Table 1; https://rapid.ensembl.org/Trisateles_emortualis_GCA_947095525.1/Info/Index). The resulting annotation includes 23,580 transcribed mRNAs from 13,176 protein-coding and 1,827 non-coding genes.

Methods

Sample acquisition and nucleic acid extraction

A male *Trisateles emortualis* (specimen ID Ox001828, ToLID ilTriEmor1) was collected from Wytham Woods, Oxfordshire (biological vice-county Berkshire), UK (latitude 51.77, longitude –1.34) on 2021-07-24 using a light trap.

Table 1. Genome data for *Trisateles emortualis*, ilTriEmor1.1.

Project accession data	
Assembly identifier	ilTriEmor1.1
Species	*Trisateles emortualis*
Specimen	ilTriEmor1
NCBI taxonomy ID	753441
BioProject	PRJEB55139
BioSample ID	SAMEA10979087
Isolate information	ilTriEmor1, male: head and thorax (DNA and Hi-C sequencing); abdomen (RNA sequencing)

Assembly metrics*		Benchmark
Consensus quality (QV)	67.4	≥ 50
k-mer completeness	100.0%	$\geq 95\%$
BUSCO**	C:98.7%[S:97.9%,D:0.8%], F:0.2%,M:1.1%,n:5,286	$C \geq 95\%$
Percentage of assembly mapped to chromosomes	99.99%	$\geq 95\%$
Sex chromosomes	Z	*localised homologous pairs*
Organelles	Mitochondrial genome: 16.01 kb	*complete single alleles*

Raw data accessions	
PacificBiosciences SEQUEL II	ERR10034128
Hi-C Illumina	ERR10038433
PolyA RNA-Seq Illumina	ERR10890702

Genome assembly	
Assembly accession	GCA_947095525.1
Accession of alternate haplotype	GCA_947095545.1
Span (Mb)	565.5
Number of contigs	54
Contig N50 length (Mb)	19.0
Number of scaffolds	32
Scaffold N50 length (Mb)	20.1
Longest scaffold (Mb)	35.35

Genome annotation	
Number of protein-coding genes	13,176
Number of non-coding genes	1,827
Number of gene transcripts	23,580

* 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/CAMUPS01/dataset/CAMUPS01/busco.
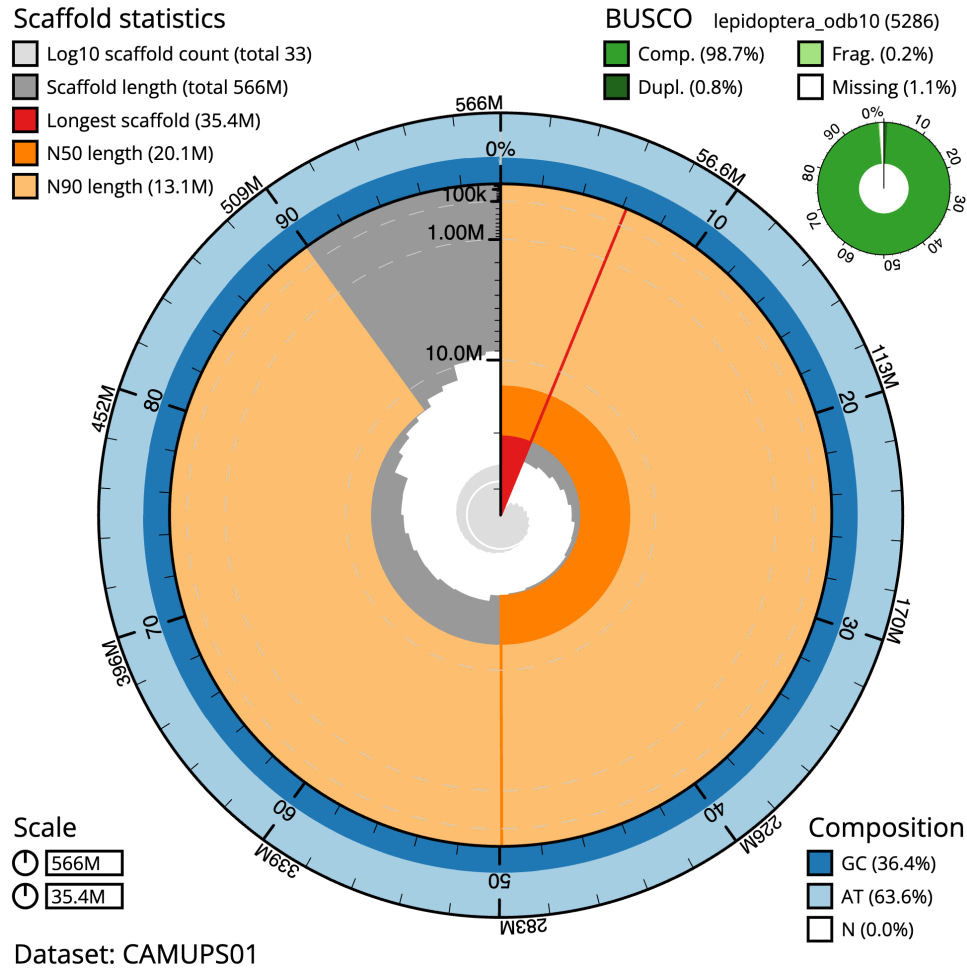
Figure 2. Genome assembly of *Trisateles emortualis*, ilTriEmor1.1: metrics. The BlobToolKit Snailplot shows N50 metrics and BUSCO gene completeness. The main plot is divided into 1,000 size-ordered bins around the circumference with each bin representing 0.1% of the 565,515,330 bp assembly. The distribution of scaffold lengths is shown in dark grey with the plot radius scaled to the longest scaffold present in the assembly (35,350,337 bp, shown in red). Orange and pale-orange arcs show the N50 and N90 scaffold lengths (20,149,241 and 13,090,150 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/CAMUPS01/dataset/CAMUPS01/snail.

The specimen was collected and identified by Douglas Boyes (University of Oxford) and preserved on dry ice.

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 ilTriEmor1 sample was weighed and dissected on dry ice (Jay *et al.*, 2023). Tissue from the head and thorax was homogenised using a PowerMasher II tissue disruptor (Denton *et al.*, 2023a).

HMW DNA was extracted using the Automated MagAttract v1 protocol (Sheerin *et al.*, 2023). DNA was sheared into an average fragment size of 12–20 kb in a Megaruptor 3 system with speed setting 30 (Todorovic *et al.*, 2023). Sheared DNA was purified by solid-phase reversible immobilisation (Strickland *et al.*, 2023): 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.

RNA was extracted from abdomen tissue of ilTriEmor1 in the Tree of Life Laboratory at the WSI using the RNA Extraction: Automated MagMax™ *mir*Vana protocol (do Amaral *et al.*, 2023). The RNA concentration was assessed using a Nanodrop spectrophotometer and a Qubit Fluorometer using the Qubit
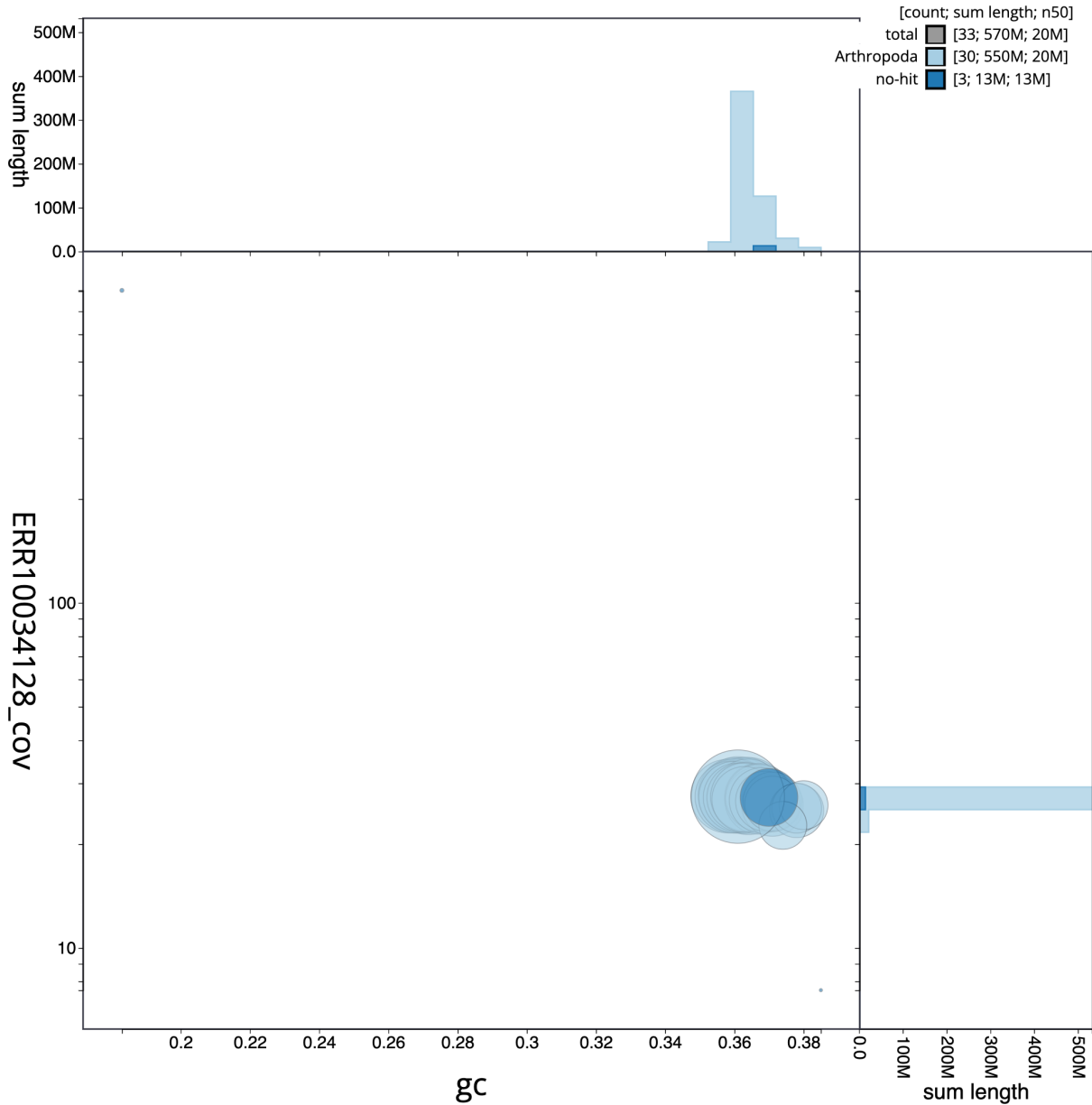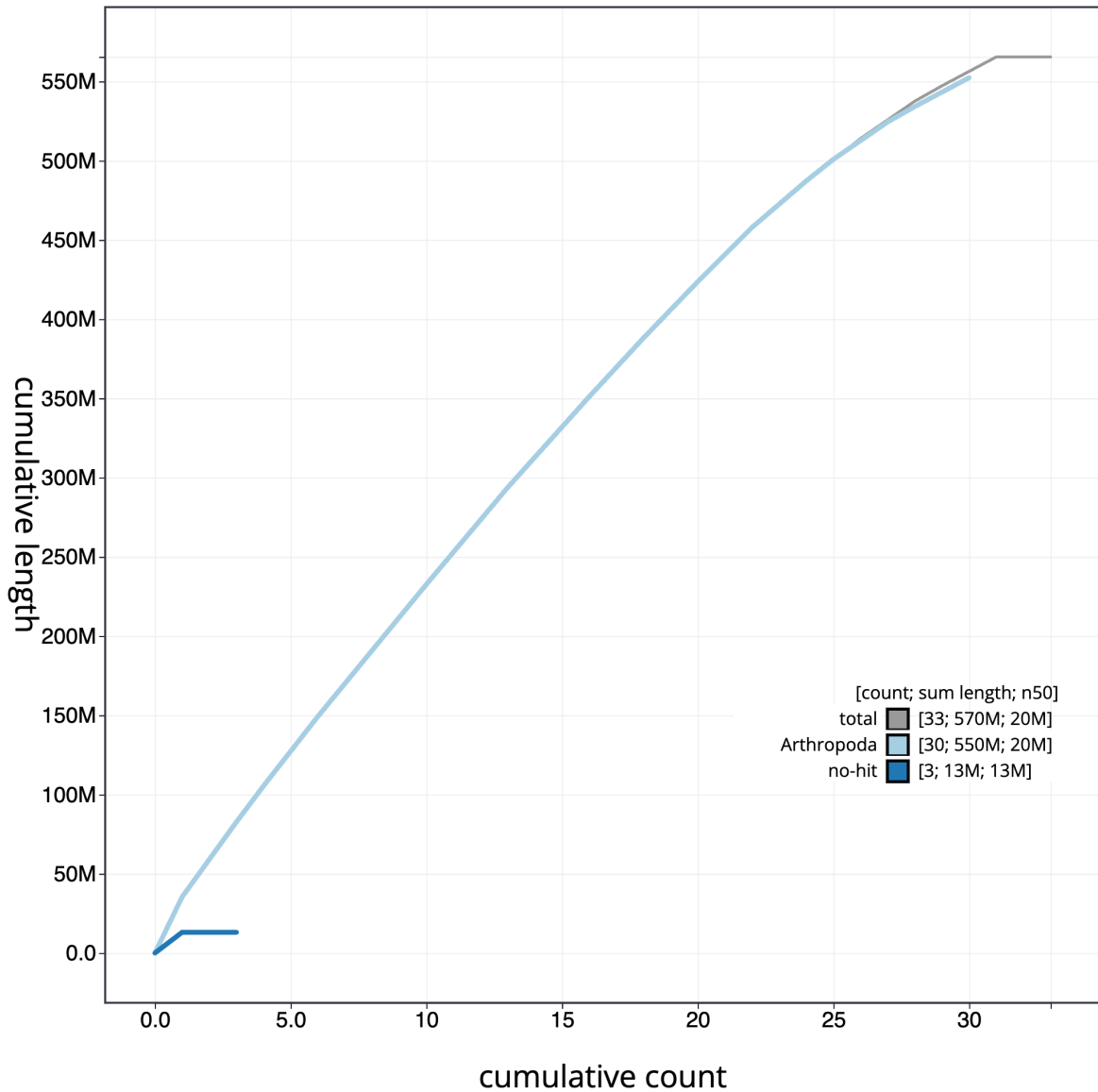
Figure 3. Genome assembly of *Trisateles emortualis*, ilTriEmor1.1: BlobToolKit GC-coverage plot. Scaffolds are coloured by phylum. Circles are sized in proportion to scaffold length. Histograms show the distribution of scaffold length sum along each axis. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/CAMUPS01/dataset/CAMUPS01/blob.

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.

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. 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
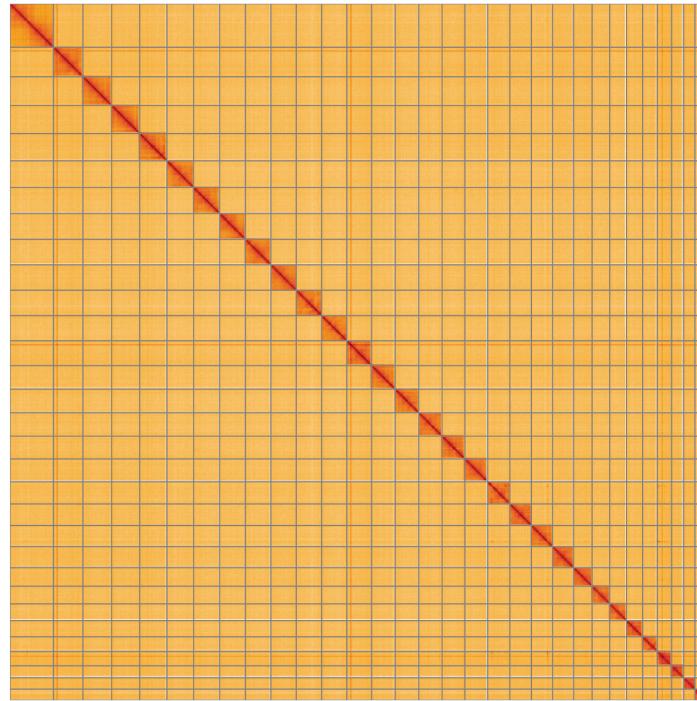
Figure 4. Genome assembly of *Trisateles emortualis*, ilTriEmor1.1: BlobToolKit cumulative sequence plot. The grey line shows cumulative length for all scaffolds. Coloured lines show cumulative lengths of scaffolds assigned to each phylum using the buscogenes taxrule. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/CAMUPS01/dataset/CAMUPS01/cumulative.

at the WSI on Pacific Biosciences SEQUEL II (HiFi) and Illumina NovaSeq 6000 (RNA-Seq) instruments. Hi-C data were also generated from remaining head and thorax tissue of ilTriEmor1 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 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.*, 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.

Figure 5. **Genome assembly of *Trisateles emortualis*, ilTriEmor1.1: Hi-C contact map of the ilTriEmor1.1 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/l/?d=G5aEfKQ7S1m40bzhBaTUVA.

Table 2. Chromosomal pseudomolecules in the genome assembly of *Trisateles emortualis*, ilTriEmor1.

INSDC accession	Chromosome	Length (Mb)	GC%
OX352730.1	1	23.9	36.5
OX352731.1	2	23.4	36.0
OX352732.1	3	22.76	36.0
OX352733.1	4	22.14	36.5
OX352734.1	5	21.66	36.0
OX352735.1	6	21.11	36.0
OX352736.1	7	20.98	36.0
OX352737.1	8	20.67	36.0
OX352738.1	9	20.66	36.5
OX352739.1	10	20.65	36.0
OX352740.1	11	20.41	36.0
OX352741.1	12	20.15	36.5
OX352742.1	13	19.14	36.5
OX352743.1	14	19.12	36.0
OX352744.1	15	18.98	36.0

INSDC accession	Chromosome	Length (Mb)	GC%
OX352745.1	16	18.56	36.5
OX352746.1	17	18.47	36.5
OX352747.1	18	18.02	36.5
OX352748.1	19	17.52	36.5
OX352749.1	20	17.17	36.5
OX352750.1	21	17.1	37.0
OX352751.1	22	15.05	37.0
OX352752.1	23	14.3	37.0
OX352753.1	24	13.64	37.0
OX352754.1	25	13.09	37.0
OX352755.1	26	12.08	37.0
OX352756.1	27	11.51	38.0
OX352757.1	28	9.7	38.0
OX352758.1	29	9.29	38.0
OX352759.1	30	8.92	37.5
OX352729.1	Z	35.35	36.0
OX352760.1	MT	0.02	18.5

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

Genome annotation
The Ensembl gene annotation system (Aken *et al.*, 2016) was used to generate annotation for the *Trisateles emortualis* assembly (GCA_947095525.1). Annotation was created primarily through alignment of transcriptomic data to the genome, with gap filling via protein-to-genome alignments of a select set of proteins from UniProt (UniProt Consortium, 2019).

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.

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.

Table 3. Software tools: versions and sources

Software tool	Version	Source
BlobToolKit	4.0.7	https://github.com/blobtoolkit/blobtoolkit
BUSCO	5.3.2	https://gitlab.com/ezlab/busco
Hifiasm	0.16.1-r375	https://github.com/chhylp123/hifiasm
HiGlass	1.11.6	https://github.com/higlass/higlass
Merqury	MerquryFK	https://github.com/thegenemyers/MERQURY.FK
MitoHiFi	2	https://github.com/marcelauliano/MitoHiFi
PretextView	0.2	https://github.com/wtsi-hpag/PretextView
purge_dups	1.2.3	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	yahs-1.1.91eebc2	https://github.com/c-zhou/yahs

Data availability
European Nucleotide Archive: *Trisateles emortualis* (olive crescent). Accession number PRJEB55139; https://identifiers.org/ena.embl/PRJEB55139 (Wellcome Sanger Institute, 2022). The genome sequence is released openly for reuse. The *Trisateles emortualis* 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.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.

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**

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

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**

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**

Bretherton RF, Goater B, Lorimer RI: **Noctuidae: Cuculliinae to Hypeninae.** In: Heath, J. and Emmet, A. M. *The Moths and Butterflies of Great Britain and Ireland.* Colchester: Harley Books,1983.

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**

Davis T: **Rare moth on the move.** *Butterfly Conservation.* 2013; [Accessed 30 December 2023].
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.** *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**

do Amaral RJV, Bates A, Denton A, *et al.*: **Sanger Tree of Life RNA Extraction: Automated MagMax™ mirVana.** *protocols.io.* 2023.
Publisher Full Text

Emmet AM: **The Scientific Names of the British Lepidoptera - their History and Meaning.** Colchester: Harley Books, 1991.
Reference Source

GBIF Secretariat: ***Biston strataria* (Hufnagel, 1767).** *GBIF Backbone Taxonomy.* 2023; [Accessed 30 December 2023].
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**

HantsMoths: **Olive Crescent (*Trisateles emortualis*) ([Denis & Schiffermüller], 1775).** *hantsmoths.org.uk.* 2023; [Accessed 30 December 2023].
Reference Source

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): giaa153.
PubMed Abstract | **Publisher Full Text** | **Free 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, Seppey 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**

Randle Z, Evans-Hill LJ, Parsons MS: **Atlas of Britain & Ireland's Larger Moths.** Newbury: NatureBureau, 2019.

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.*: **Merqury: Reference-free quality, completeness, and phasing assessment for genome assemblies.** *Genome Biol.* 2020; **21**(1): 245.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Sheerin E, Sampaio F, Oatley G, *et al.*: **Sanger Tree of Life HMW DNA Extraction: Automated MagAttract v.1.** *protocols.io.* 2023.
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**

Strickland M, Cornwell C, Howard C: **Sanger Tree of Life Fragmented DNA clean up: Manual SPRI.** *protocols.io.* 2023.
Publisher Full Text

Surana P, Muffato M, Qi G: **sanger-tol/readmapping: sanger-tol/readmapping v1.1.0 - Hebridean Black (1.1.0).** *Zenodo.* 2023a.
Reference Source

Surana P, Muffato M, Sadasivan Baby C: **sanger-tol/genomenote (v1.0.dev).** *Zenodo.* 2023b.
Publisher Full Text

Todorovic M, Sampaio F, Howard C: **Sanger Tree of Life HMW DNA Fragmentation: Diagenode Megaruptor®3 for PacBio HiFi.** *protocols.io.* 2023.
Publisher Full Text

Uliano-Silva M, Ferreira JGRN, Krasheninnikova 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**

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

Vasimuddin Md, 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 Olive Crescent, *Trisateles emortualis* (Denis & Schiffermüller, 1775).** European Nucleotide Archive, [dataset], accession number PRJEB55139. 2022.

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**