DATA NOTE

The genome sequence of the Sallow Marble, *Apotomis capreana* (Hübner, 1817) [version 1; peer review: awaiting peer review]

Douglas Boyes[1+], Clare Boyes[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]Independent researcher, Welshpool, Wales, UK

[+] Deceased author

v1 **First published:** 12 Apr 2024, **9**:184
https://doi.org/10.12688/wellcomeopenres.20875.1
Latest published: 12 Apr 2024, **9**:184
https://doi.org/10.12688/wellcomeopenres.20875.1

Abstract

We present a genome assembly from an individual male *Apotomis capreana* (the Sallow Marble; Arthropoda; Insecta; Lepidoptera; Tortricidae). The genome sequence is 743.2 megabases in span. Most of the assembly is scaffolded into 28 chromosomal pseudomolecules, including the Z sex chromosome. The mitochondrial genome has also been assembled and is 17.12 kilobases in length.

Keywords

Apotomis capreana, sallow marble, genome sequence, chromosomal, Lepidoptera

This article is included in the Tree of Life gateway.

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.

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

Author roles: Boyes D: Investigation, Resources; **Boyes C**: 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) and the Darwin Tree of Life Discretionary Award (218328).
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

How to cite this article: Boyes D, Boyes C, University of Oxford and Wytham Woods Genome Acquisition Lab *et al.* **The genome sequence of the Sallow Marble, *Apotomis capreana* (Hübner, 1817) [version 1; peer review: awaiting peer review]** Wellcome Open Research 2024, **9**:184 https://doi.org/10.12688/wellcomeopenres.20875.1

First published: 12 Apr 2024, **9**:184 https://doi.org/10.12688/wellcomeopenres.20875.1

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; Olethreutinae; Olethreutini; *Apotomis*; *Apotomis capreana* (Hübner, 1817) (NCBI:txid989741).

Background

Apotomis capreana (Sallow Marble) is a micro-moth in the family Tortricidae with cryptic black and white markings said to resemble a bird-dropping (Clifton & Wheeler, 2011). It is locally uncommon in the UK, largely restricted to the south of England and Wales, and rare in the north of England and Scotland. In Europe is found in northern and central Europe with scattered records as far east as Japan. It is also found in North America (GBIF Secretariat, 2023).

The adult (forewing length 8–10 mm) has dark brown forewings interspersed with darker markings, for the basal two thirds. The apical third of the forewing is white, and there is a white hook-shaped indentation extending into the dark markings in the middle of the wing. The adult moth is single-brooded and flies from June to August in wetter habitats such as woodland, marshland and stream sides where its foodplants, Goat Willow and Grey Willow, occur. (Sterling *et al.*, 2012). In Europe, the moth larvae have also been recorded feeding on Poplar, Birch and Elm. The eggs hatch in April-May of the following spring and feed in a tightly spun web on the terminal shoots (Langmaid *et al.*, 2018). The larvae pupate either in situ, or amongst leaf litter, during May and June (Hancock *et al.*, 2015).

The genome of *A. capreana* was sequenced as part of the Darwin Tree of Life Project, a collaborative effort to sequence all named eukaryotic species in the Atlantic Archipelago of Britain and Ireland. Here we present a chromosomally complete genome sequence for *Apotomis capreana* based on one male specimen from Wytham Woods, Oxfordshire, UK.

Genome sequence report

The genome was sequenced from one male *Apotomis capreana* (Figure 1) collected from Wytham Woods, Oxfordshire, UK (51.77, –1.31). A total of 31-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 15 missing joins or mis-joins and removed 15 haplotypic duplications, reducing the assembly length by 1.01% and the scaffold number by 20.00%, and increasing the scaffold N50 by 3.81%.

The final assembly has a total length of 743.2 Mb in 67 sequence scaffolds with a scaffold N50 of 27.9 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

Figure 1. Photograph of the *Apotomis capreana* (ilApoCapr1) specimen used for genome sequencing.

cumulative assembly plot in Figure 4 shows curves for subsets of scaffolds assigned to different phyla. Most (99.78%) of the assembly sequence was assigned to 28 chromosomal-level scaffolds, representing 27 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). The Z chromosome was identified based on synteny with *Apotomis betuletana* (GCA_932273695.1; ilApoBetu1) (Boyes *et al.*, 2023). 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 62.5 with *k*-mer completeness of 100.0%, and the assembly has a BUSCO v5.3.2 completeness of 98.3% (single = 97.4%, duplicated = 0.9%), 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/989741.

Methods

Sample acquisition and nucleic acid extraction

A male *Apotomis capreana* (specimen ID Ox001675, ToLID ilApoCapr1) was collected from Wytham Woods, Oxfordshire (biological vice-county Berkshire), UK (latitude 51.77, longitude –1.31) on 2021-07-17 using a light trap. The specimen was collected and identified by Douglas Boyes (University of Oxford) and preserved on dry ice.

Protocols developed by the Wellcome Sanger Institute (WSI) Tree of Life core laboratory have been deposited on protocols.io (Denton *et al.*, 2023b). The workflow for high molecular weight (HMW) DNA extraction at the WSI includes a sequence of core procedures: sample preparation; sample homogenisation, DNA extraction, fragmentation, and clean-up. In sample preparation, the ilApoCapr1 sample was weighed

Table 1. Genome data for *Apotomis caprean a*, ilApoCapr1.1.

Project accession data		
Assembly identifier	ilApoCapr1.1	
Species	*Apotomis capreana*	
Specimen	ilApoCapr1	
NCBI taxonomy ID	989741	
BioProject	PRJEB56802	
BioSample ID	SAMEA10978943	
Isolate information	ilApoCapr1, male: whole organism (DNA and Hi-C sequencing)	
Assembly metrics*		***Benchmark***
Consensus quality (QV)	62.5	*≥ 50*
k-mer completeness	100.0%	*≥ 95%*
BUSCO**	C:98.3%[S:97.4%,D:0.9%], F:0.4%,M:1.3%,n:5,286	*C ≥ 95%*
Percentage of assembly mapped to chromosomes	99.78%	*≥ 95%*
Sex chromosomes	Z	*localised homologous pairs*
Organelles	Mitochondrial genome: 17.12 kb	*complete single alleles*
Raw data accessions		
PacificBiosciences SEQUEL II	ERR10395978	
Hi-C Illumina	ERR10395984	
Genome assembly		
Assembly accession	GCA_947623375.1	
Accession of alternate haplotype	GCA_947623195.1	
Span (Mb)	743.2	
Number of contigs	263	
Contig N50 length (Mb)	5.0	
Number of scaffolds	67	
Scaffold N50 length (Mb)	27.9	
Longest scaffold (Mb)	62.88	

* 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/CANQKS01/dataset/CANQKS01/busco.

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 in the WSI Scientific Operations core using the Automated MagAttract v2 protocol (Oatley *et al.*, 2023). HMW DNA was sheared into an average fragment size of 12–20 kb in a Megaruptor 3 system with speed setting 31 (Bates *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

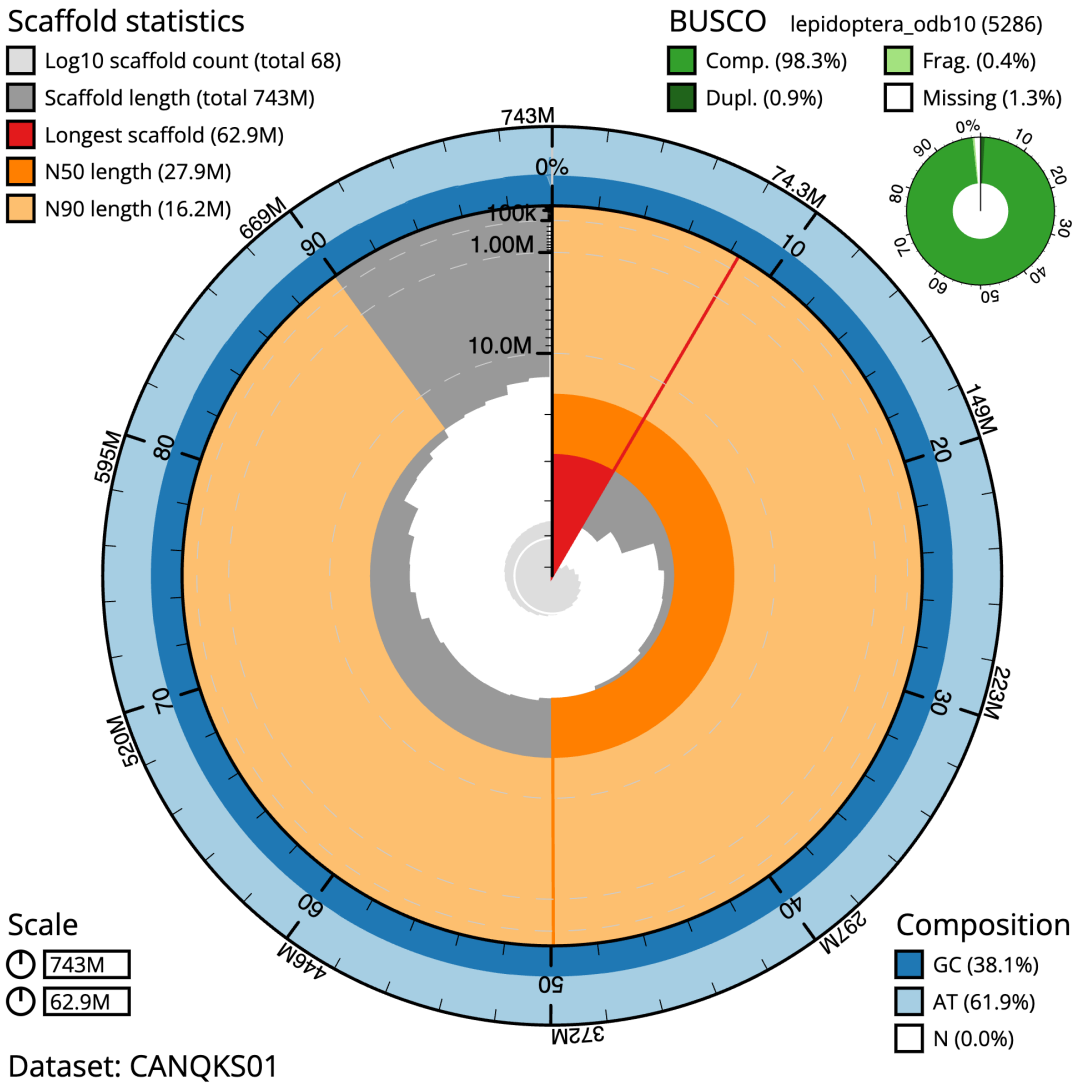

Figure 2. Genome assembly of *Apotomis capreana*, ilApoCapr1.1: metrics. The BlobToolKit Snailplot shows N50 metrics and BUSCO gene completeness. The main plot is divided into 1,000 size-ordered bins around the circumference with each bin representing 0.1% of the 743,232,780 bp assembly. The distribution of scaffold lengths is shown in dark grey with the plot radius scaled to the longest scaffold present in the assembly (62,880,285 bp, shown in red). Orange and pale-orange arcs show the N50 and N90 scaffold lengths (27,898,245 and 16,210,158 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/CANQKS01/dataset/CANQKS01/snail.

distribution was evaluated by running the sample on the FemtoPulse system.

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 SEQUEL II instrument. Hi-C data were also generated from remaining tissue of ilApoCapr1 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

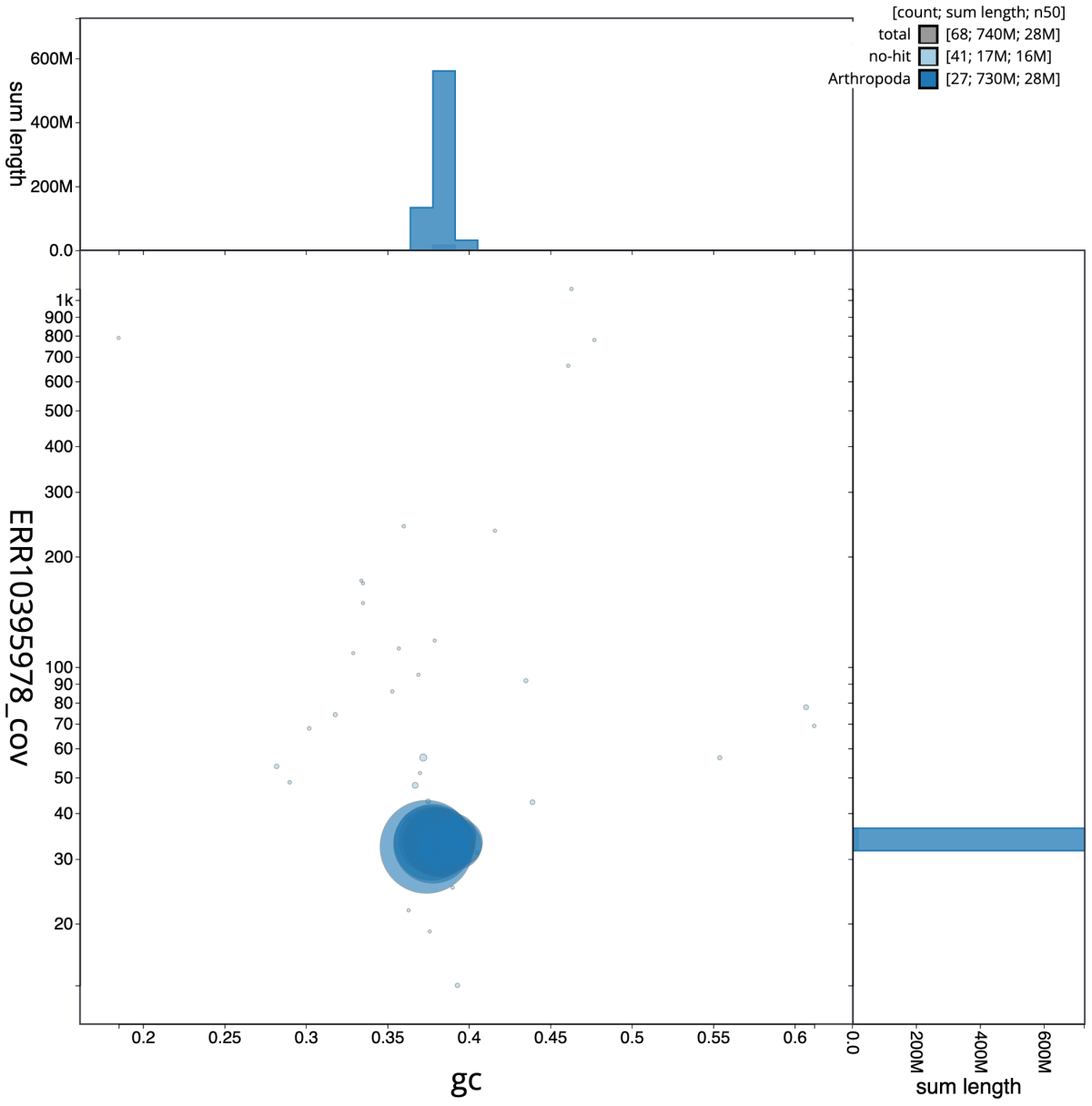

Figure 3. Genome assembly of *Apotomis capreana*, ilApoCapr1.1: BlobToolKit GC-coverage plot. Scaffolds are coloured by phylum. Circles are sized in proportion to scaffold length. Histograms show the distribution of scaffold length sum along each axis. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/CANQKS01/dataset/CANQKS01/blob.

(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 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

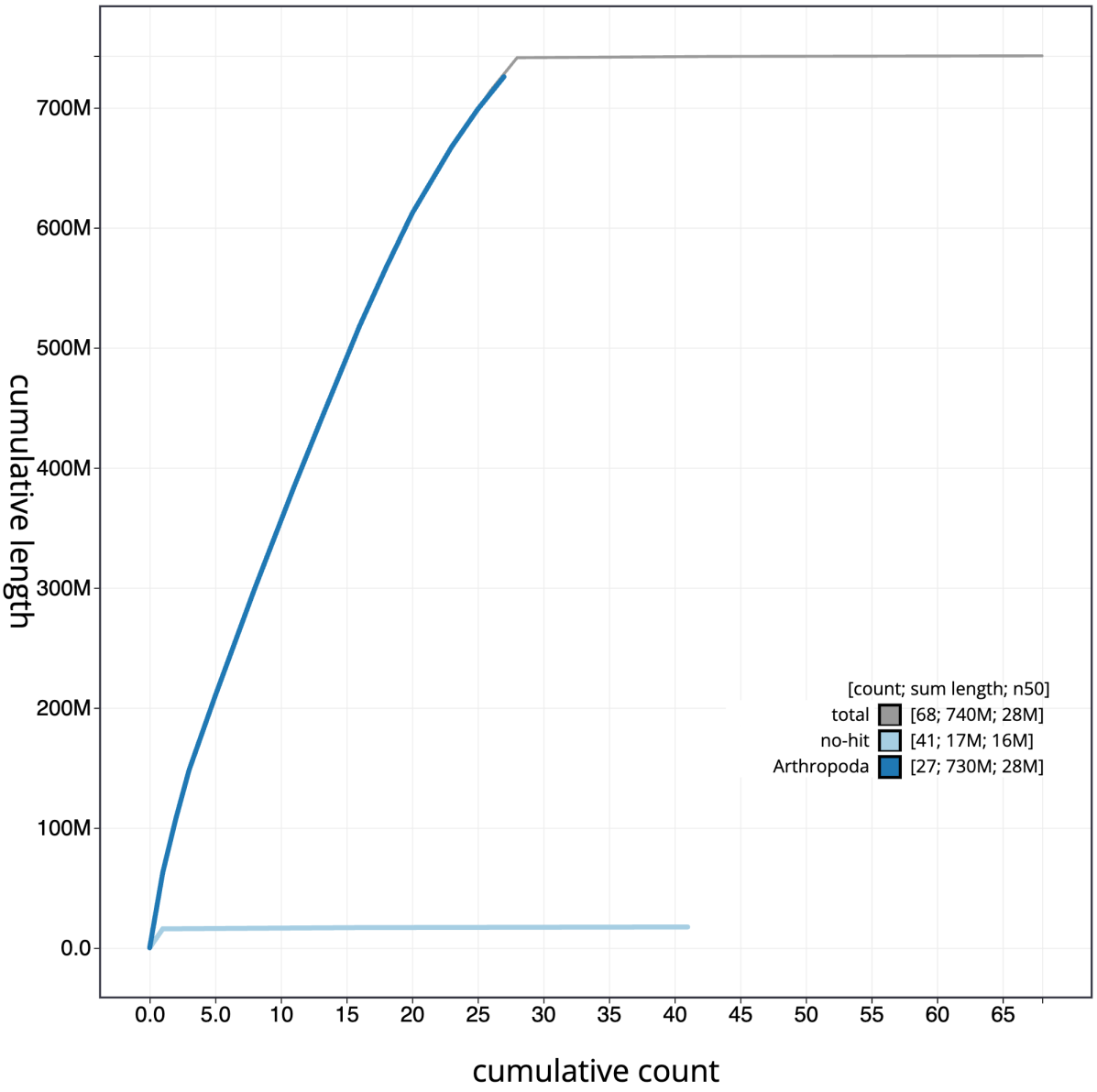

Figure 4. Genome assembly of *Apotomis capreana*, ilApoCapr1.1: BlobToolKit cumulative sequence plot. The grey line shows cumulative length for all scaffolds. Coloured lines show cumulative lengths of scaffolds assigned to each phylum using the buscogenes taxrule. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/CANQKS01/dataset/CANQKS01/cumulative.

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

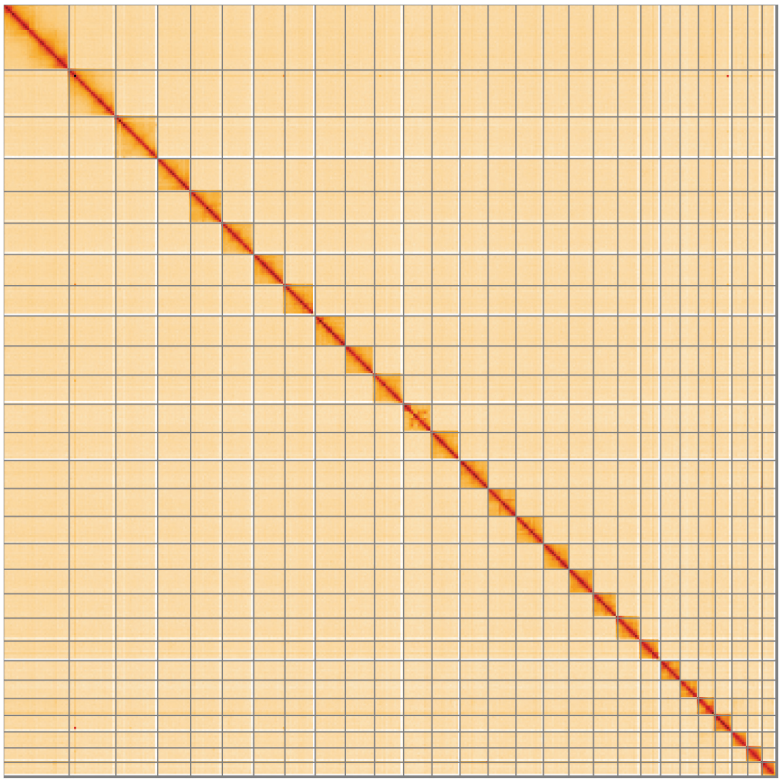

Figure 5. Genome assembly of *Apotomis capreana*, ilApoCapr1.1: Hi-C contact map of the ilApoCapr1.1 assembly, visualised using HiGlass. Chromosomes are shown in order of size from left to right and top to bottom. An interactive version of this figure may be viewed at https://genome-note-higlass.tol.sanger.ac.uk/l/?d=P7aRJjhzTVKm08qVuUbc9g.

Table 2. Chromosomal pseudomolecules in the genome assembly of *Apotomis capreana*, ilApoCapr1.

INSDC accession	Chromosome	Length (Mb)	GC%
OX392499.1	1	44.93	38.0
OX392500.1	2	40.22	37.5
OX392501.1	3	31.66	38.0
OX392502.1	4	30.37	37.5
OX392503.1	5	30.1	38.0
OX392504.1	6	29.98	38.0
OX392505.1	7	29.11	38.0
OX392506.1	8	28.9	38.0
OX392507.1	9	28.05	38.0
OX392508.1	10	27.9	38.0
OX392509.1	11	27.29	38.5
OX392510.1	12	26.95	38.5
OX392511.1	13	26.87	38.0
OX392512.1	14	26.83	38.0
OX392513.1	15	26.21	38.0
OX392514.1	16	24.61	38.0
OX392515.1	17	23.58	38.5
OX392516.1	18	23.4	38.5
OX392517.1	19	22.07	38.5
OX392518.1	20	18.94	39.0
OX392519.1	21	18.7	38.0
OX392520.1	22	17.67	38.5
OX392521.1	23	16.21	39.5
OX392522.1	24	15.87	38.5
OX392523.1	25	15.32	39.5
OX392524.1	26	13.78	38.5
OX392525.1	27	13.29	39.0
OX392498.1	Z	62.88	37.5
OX392526.1	MT	0.02	19.0

Table 3. Software tools: versions and sources.

Software tool	Version	Source
BlobToolKit	4.1.7	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

Research Limited (operating as the Wellcome Sanger Institute), and in some circumstances other Darwin Tree of Life collaborators.

Data availability

European Nucleotide Archive: *Apotomis capreana*. Accession number PRJEB56802; https://identifiers.org/ena.embl/PRJEB56802 (Wellcome Sanger Institute, 2022). The genome sequence is released openly for reuse. The *Apotomis capreana* 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, 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

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

Boyes D; University of Oxford and Wytham Woods Genome Acquisition Lab,

Darwin Tree of Life Barcoding collective, *et al.*: **The genome sequence of the Birch Marble, *Apotomis betuletana* (Haworth, 1811) [version 1; peer review: 2 approved].** *Wellcome Open Res.* 2023; **8**: 66.
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**

Clifton J, Wheeler J: **Bird-dropping Tortrix Moths of the British Isles**. Dorchester: Clifton & Wheeler, 2011.

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**

GBIF Secretariat: ***Apotomis capreana* (Hübner, 1817).** *GBIF Backbone Taxonomy.* 2023; **[Accessed 9 January 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**

Hancock EF, Bland KP, Razowski J: **The Moths and Butterflies of Great Britain and Ireland Tortricidae Olethreutinae**. Leiden: Brill. 2015.

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**

Langmaid JR, Palmer S, Young MR: **A Field Guide to the Smaller Moths of Great Britain and Ireland**. 3rd ed. British Entomological and Natural History Society, 2018.
Reference Source

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**

Oatley G, Denton A, Howard C: **Sanger Tree of Life HMW DNA Extraction: Automated MagAttract v.2.** *protocols.io.* 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.*: **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**

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**

Sterling P, Parsons M, Lewington R: **Field Guide to the Micro Moths of Great Britain and Ireland**. Gillingham, Dorset: British Wildlife Publishing, 2012.
Reference Source

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

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**

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 Sallow Marble, *Apotomis capreana* (Hübner, 1817).** European Nucleotide Archive, [dataset], accession number PRJEB56802. 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**