DATA NOTE

The genome sequence of the Small Birch Bell, *Epinotia ramella* (Linnaeus, 1758) [version 1; peer review: awaiting peer review]

Douglas Boyes[1+], David C. Lees[2], James Hammond[3],
University of Oxford and Wytham Woods Genome Acquisition Lab,
Natural History Museum 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

[1]UK Centre for Ecology & Hydrology, Wallingford, England, UK
[2]Natural History Museum, London, England, UK
[3]University of Oxford, Oxford, England, UK

[+] Deceased author

v1 **First published:** 20 Oct 2023, **8**:479
https://doi.org/10.12688/wellcomeopenres.20150.1
Latest published: 20 Oct 2023, **8**:479
https://doi.org/10.12688/wellcomeopenres.20150.1

Abstract

We present a genome assembly from an individual female *Epinotia ramella* (the Small Birch Bell; Arthropoda; Insecta; Lepidoptera; Tortricidae). The genome sequence is 782.0 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 15.44 kilobases in length. Gene annotation of this assembly on Ensembl identified 20,893 protein coding genes.

Keywords

Epinotia ramella, Small Birch Bell, 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; **Lees DC**: Investigation, Resources; **Hammond J**: 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, Lees DC, Hammond J *et al.* **The genome sequence of the Small Birch Bell, *Epinotia ramella* (Linnaeus, 1758) [version 1; peer review: awaiting peer review]** Wellcome Open Research 2023, **8**:479 https://doi.org/10.12688/wellcomeopenres.20150.1

First published: 20 Oct 2023, **8**:479 https://doi.org/10.12688/wellcomeopenres.20150.1

Species taxonomy

Eukaryota; 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; Eucosmini; *Epinotia*; *Epinotia ramella* (Linnaeus, 1758) (NCBI:txid1594296).

Background

The Small Birch Bell *Epinotia ramella* (Linnaeus, 1758) is a moth in the Tortricidae family. The species' vernacular name is a reference to the larval foodplant, birch (*Betula*), and the species is common throughout Britain and Ireland wherever birch occurs (Bradley *et al.*, 1979). Globally the species has a Palaearctic distribution and is found across northern Eurasia to Japan (Bradley *et al.*, 1979; GBIF Secretariat, 2023).

The species overwinters as an egg or a small larva (Elliott *et al.*, 2018). The caterpillar then feeds within the birch twigs or catkins from April to May (Bradley *et al.*, 1979; Elliott *et al.*, 2018). Adult moths occur between July and September and rest by day on twigs and trunks of the foodplant (Bradley *et al.*, 1979). Adults fly by dusk and come to light (Bradley *et al.*, 1979; Elliott *et al.*, 2018). Adult moths have a pied grey-and-black appearance, with minor variation, however a form with extensive brown markings also occurs commonly (Bradley *et al.*, 1979).

The genome of *Epinotia ramella* 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 Epinotia ramella, based on one female specimen from Wytham Woods, Oxfordshire, UK.

Genome sequence report

The genome was sequenced from one female *Epinotia ramella* (Figure 1) collected from Wytham Woods, Oxfordshire, UK (51.77, –1.34). A total of 31-fold coverage in Pacific Biosciences single-molecule HiFi long reads and 64-fold coverage in 10X Genomics read clouds were generated. Primary assembly contigs were scaffolded with chromosome conformation Hi-C data. Manual assembly curation corrected 31 missing joins or mis-joins and removed 6 haplotypic duplications, reducing the assembly length by 0.21% and the scaffold number by 7.36%.

The final assembly has a total length of 782.0 Mb in 150 sequence scaffolds with a scaffold N50 of 27.5 Mb (Table 1). A summary of the assembly statistics is shown in Figure 2, 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 (95.53%) 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 specimen was identified as female based on half coverage of the Z chromosome. 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.

Figure 1. Photograph of the *Epinotia ramella* (ilEpiRame1) specimen used for genome sequencing.

The estimated Quality Value (QV) of the final assembly is 66.5 with *k*-mer completeness of 100%, and the assembly has a BUSCO v5.3.2 completeness of 97.9% (single = 96.9%, duplicated = 1.0%), using the lepidoptera_odb10 reference set (n = 5,286).

Metadata for specimens, spectral estimates, sequencing runs, contaminants and pre-curation assembly statistics can be found at https://links.tol.sanger.ac.uk/species/1594296.

Genome annotation report

The *Epinotia ramella* genome assembly (GCA_947578815.1) was annotated using the Ensembl rapid annotation pipeline (Table 1; https://rapid.ensembl.org/Epinotia_ramella_GCA_947578815.1/Info/Index). The resulting annotation includes 21,131 transcribed mRNAs from 20,893 protein-coding genes.

Methods

Sample acquisition and nucleic acid extraction

A *Epinotia ramella* (specimen ID Ox000807, ToLID ilEpiRame1) was collected from Wytham Woods, Oxfordshire (biological vice-county Berkshire), UK (latitude 51.77, longitude –1.34) on 2020-08-01 using a light trap. The specimen was collected and identified by Douglas Boyes (University of Oxford) and preserved on dry ice. A second specimen (specimen ID NHMUK014543811, ToLID ilEpiRame4) was collected from Beinn Eighe National Nature Reserve, Scotland, UK (latitude 57.63, longitude –5.35) on 2021-09-10 using an aerial net. The specimen was collected and identified by David Lees (Natural History Museum) and preserved on dry ice. This specimen was used to produce Hi-C data.

Table 1. **Genome data for *Epinotia ramella*, ilEpiRame1.1.**

Project accession data		
Assembly identifier	ilEpiRame1.1	
Assembly release date	2022-12-18	
Species	*Epinotia ramella*	
Specimen	ilEpiRame1	
NCBI taxonomy ID	1594296	
BioProject	PRJEB55602	
BioSample ID	SAMEA7746614	
Isolate information	ilEpiRame1, female: whole organism (DNA sequencing) ilEpiRame4: head and thorax (Hi-C data)	
Assembly metrics*		***Benchmark***
Consensus quality (QV)	66.5	*≥ 50*
k-mer completeness	100%	*≥ 95%*
BUSCO**	C:97.9%[S:96.9%,D:1.0%],F:0.5%,M:1.6%,n:5,286	*C ≥ 95%*
Percentage of assembly mapped to chromosomes	95.53%	*≥ 95%*
Sex chromosomes	Z chromosome	*localised homologous pairs*
Organelles	Mitochondrial genome assembled	*complete single alleles*
Raw data accessions		
PacificBiosciences SEQUEL II	ERR10144332, ERR10144331	
10X Genomics Illumina	ERR10123717, ERR10123719, ERR10123718, ERR10123716	
Hi-C Illumina	ERR10123720	
Genome assembly		
Assembly accession	GCA_947578815.1	
Accession of alternate haplotype	GCA_947580245.1	
Span (Mb)	782.0	
Number of contigs	223	
Contig N50 length (Mb)	14.0	
Number of scaffolds	150	
Scaffold N50 length (Mb)	27.5	
Longest scaffold (Mb)	61.5	
Genome annotation		
Number of protein-coding genes	20,893	
Number of gene transcripts	21,131	

* 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/Epinotia%20ramella/dataset/CANPUT01/busco.

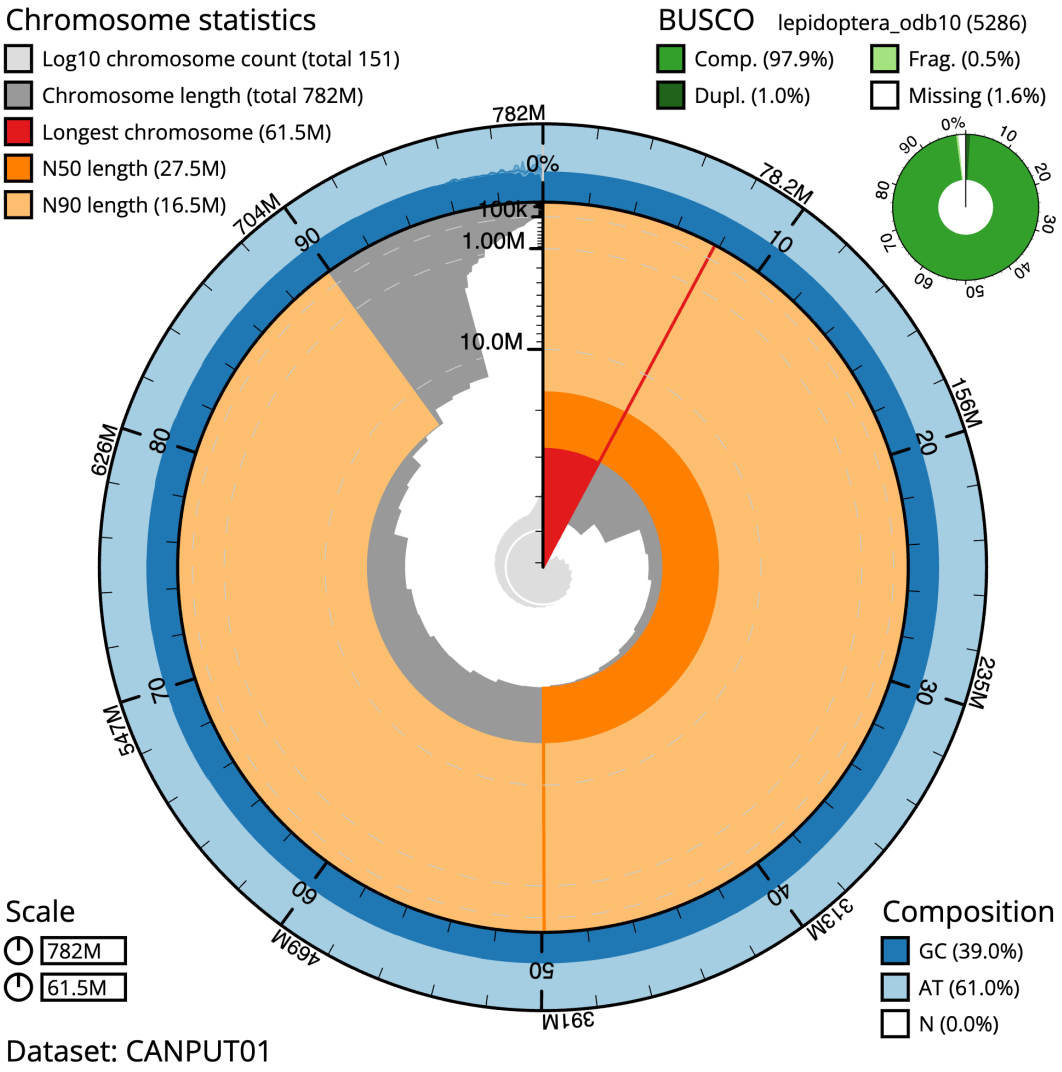

Figure 2. Genome assembly of *Epinotia ramella*, ilEpiRame1.1: metrics. The BlobToolKit Snailplot shows N50 metrics and BUSCO gene completeness. The main plot is divided into 1,000 size-ordered bins around the circumference with each bin representing 0.1% of the 781,992,136 bp assembly. The distribution of scaffold lengths is shown in dark grey with the plot radius scaled to the longest scaffold present in the assembly (61,498,029 bp, shown in red). Orange and pale-orange arcs show the N50 and N90 scaffold lengths (27,488,655 and 16,504,716 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/Epinotia%20ramella/dataset/CANPUT01/snail.

DNA was extracted at the Tree of Life laboratory, Wellcome Sanger Institute (WSI). The ilEpiRame1 sample was weighed and dissected on dry ice with tissue set aside for Hi-C sequencing. Tissue from the whole organism was disrupted using a Nippi Powermasher fitted with a BioMasher pestle. 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.

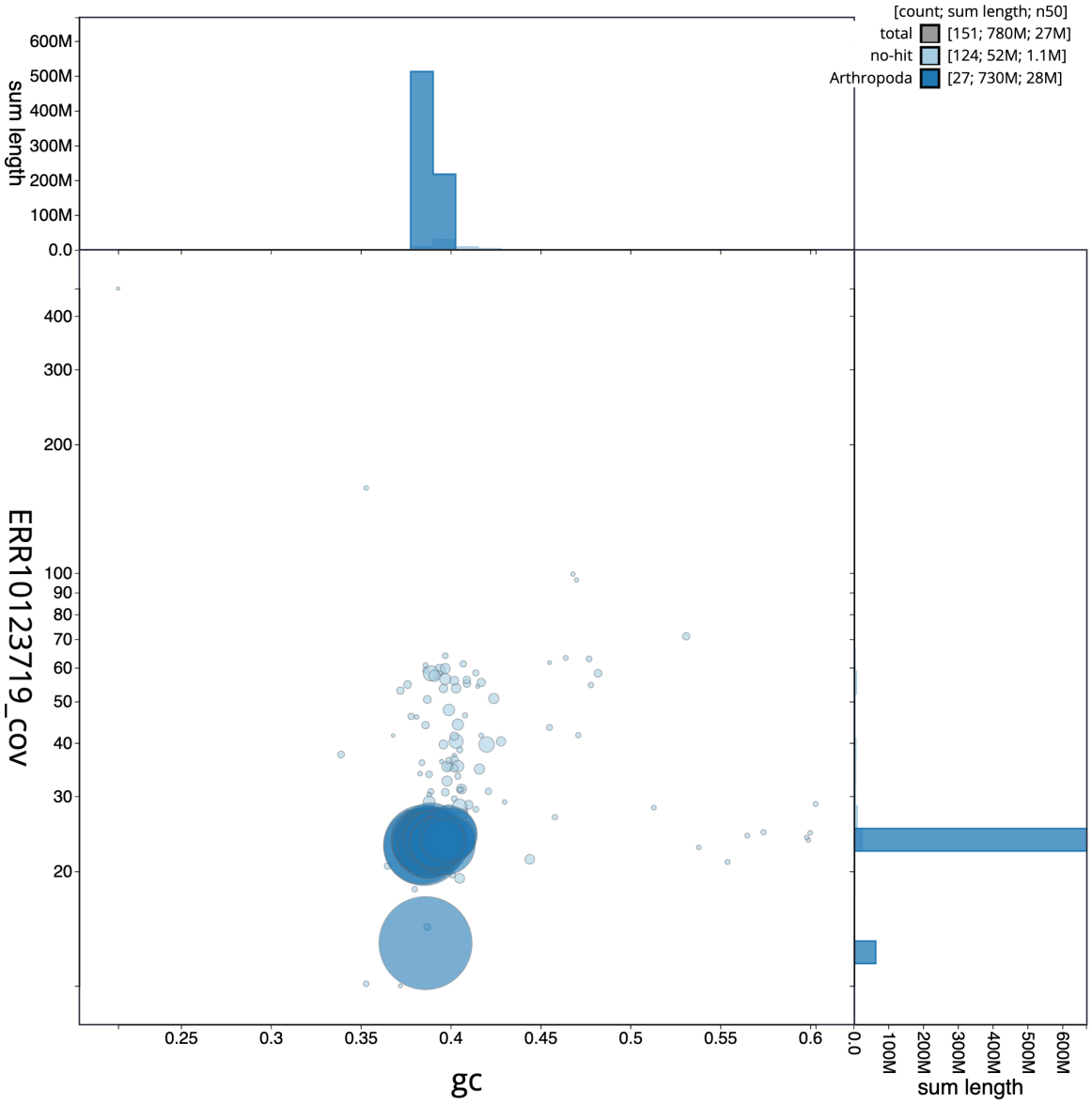

Figure 3. Genome assembly of *Epinotia ramella*, ilEpiRame1.1: BlobToolKit GC-coverage plot. Scaffolds are coloured by phylum. Circles are sized in proportion to scaffold length. Histograms show the distribution of scaffold length sum along each axis. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/Epinotia%20ramella/dataset/CANPUT01/blob.

Sequencing

Pacific Biosciences HiFi circular consensus and 10X Genomics read cloud DNA sequencing libraries were constructed according to the manufacturers' instructions. DNA sequencing was performed by the Scientific Operations core at the WSI on Pacific Biosciences SEQUEL II (HiFi) and Illumina NovaSeq 6000 (10X) instruments. Hi-C data were also generated from head and thorax tissue of ilEpiRame4 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.

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

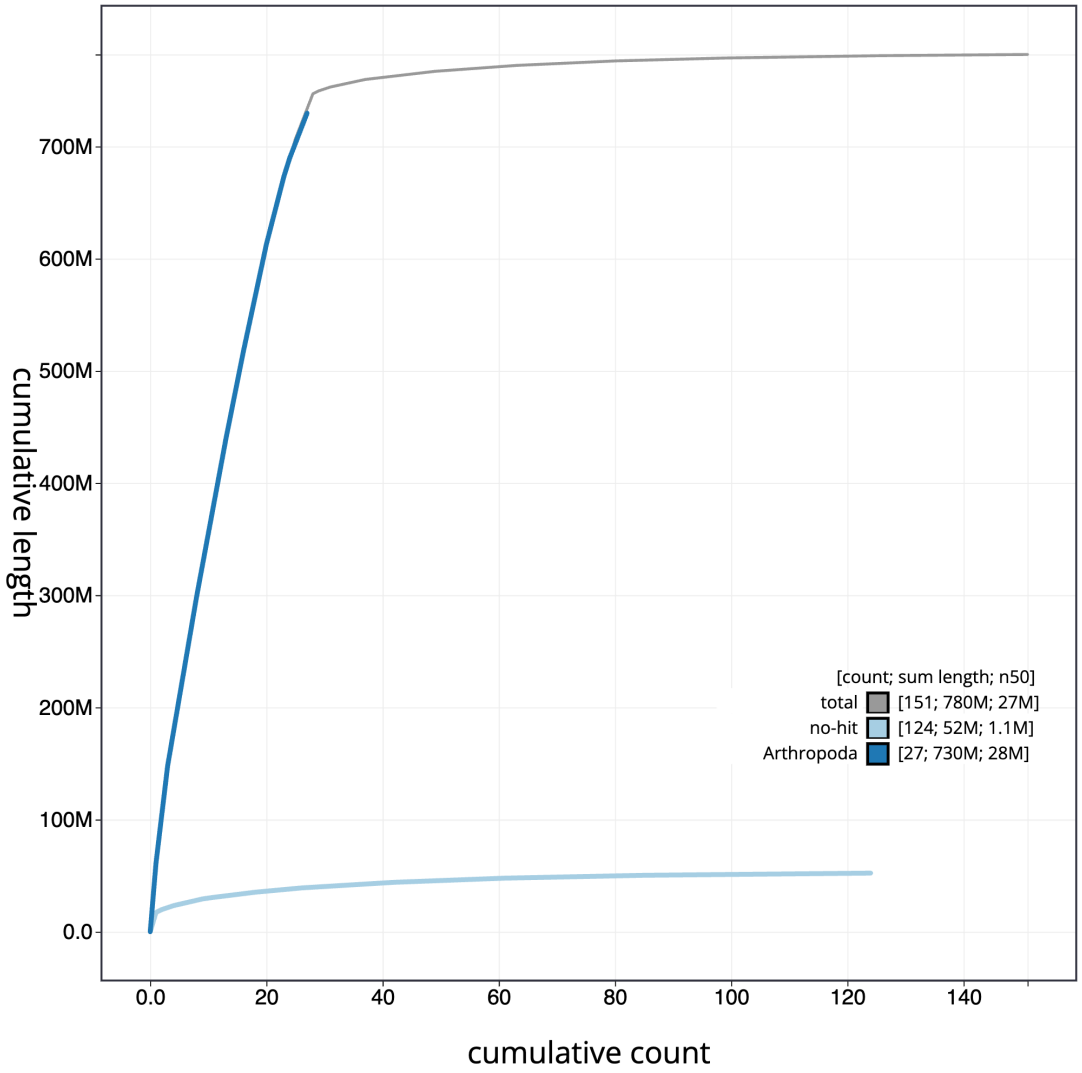

Figure 4. Genome assembly of *Epinotia ramella*, ilEpiRame1.1: BlobToolKit cumulative sequence plot. The grey line shows cumulative length for all scaffolds. Coloured lines show cumulative lengths of scaffolds assigned to each phylum using the buscogenes taxrule. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/Epinotia%20ramella/dataset/CANPUT01/cumulative.

(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 BRAKER2 pipeline (Brůna *et al.*, 2021) was used in the default protein mode to generate annotation for the *Epinotia ramella* assembly (GCA_947578815.1) in Ensembl Rapid Release.

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

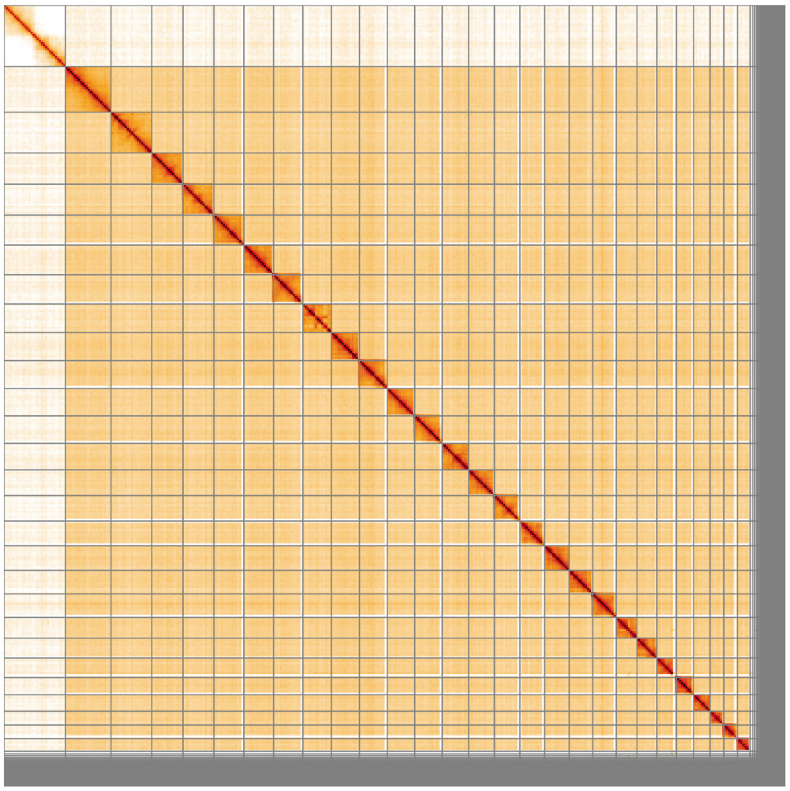

Figure 5. Genome assembly of *Epinotia ramella*, ilEpiRame1.1: Hi-C contact map of the ilEpiRame1.1 assembly, visualised using HiGlass. Chromosomes are shown in order of size from left to right and top to bottom. An interactive version of this figure may be viewed at https://genome-note-higlass.tol.sanger.ac.uk/l/?d=P3H2ujAMQH61DVtTxVKj3w.

Table 2. Chromosomal pseudomolecules in the genome assembly of *Epinotia ramella*, ilEpiRame1.

INSDC accession	Chromosome	Length (Mb)	GC%
OX388201.1	1	45.62	38.5
OX388202.1	2	40.79	38.5
OX388203.1	3	31.36	38.5
OX388204.1	4	30.83	38.5
OX388205.1	5	29.96	38.5
OX388206.1	6	29.82	39.0
OX388207.1	7	29.26	38.5
OX388208.1	8	28.6	39.0
OX388209.1	9	28.08	39.0
OX388210.1	10	27.86	39.0
OX388211.1	11	27.49	39.0
OX388212.1	12	27.4	38.5
OX388213.1	13	26.56	39.0
OX388214.1	14	25.68	38.5
OX388215.1	15	25.56	39.0
OX388216.1	16	24.76	39.5
OX388217.1	17	24.58	39.0
OX388218.1	18	23.62	39.5
OX388219.1	19	23.6	39.0
OX388220.1	20	20.69	40.0
OX388221.1	21	19.98	39.0
OX388222.1	22	19.44	40.0
OX388223.1	23	17.24	39.5
OX388224.1	24	16.5	39.0
OX388225.1	25	13.91	40.0
OX388226.1	26	13.49	39.5
OX388227.1	27	12.75	39.5
OX388200.1	Z	61.5	38.5
OX388228.1	MT	0.02	21.5

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

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.

Data availability

European Nucleotide Archive: *Epinotia ramella* (small birch bell). Accession number PRJEB55602; https://identifiers.org/ena.embl/PRJEB55602. (Wellcome Sanger Institute, 2022)

The genome sequence is released openly for reuse. The *Epinotia ramella* 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 Natural History Museum Genome Acquisition Lab are listed here: https://doi.org/10.5281/zenodo.4790042.

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

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

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

Bradley JD, Tremewan WD, Smith A: **British Tortricoid Moths – Tortricidae: Olethreutinae.** The Ray Society, 1979.

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

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

Elliott B, Langmaid JR, Palmer RM, *et al.*: **Tortricidae.** In: Langmaid, J. R., Palmer, S. M., and Young, M. R. (eds.) *A Field Guide to the Smaller Moths of Great Britain and Ireland.* The British Entomological and Natural History Society, 2018; 279.

GBIF Secretariat: ***Epinotia ramella* (Linnaeus, 1758).** *GBIF Backbone Taxonomy. Checklist dataset.* 2023; [Accessed 23 September 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

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

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

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

Surana P, Muffato M, Qi G: **sanger-tol/readmapping: sanger-tol/readmapping v1.1.0 - Hebridean Black (1.1.0).** *Zenodo.* 2023a; [Accessed 21 July 2023].
Publisher Full Text

Surana P, Muffato M, Sadasivan Baby C: **sanger-tol/genomenote (v1.0.dev).** *Zenodo.* 2023b; [Accessed 21 July 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

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 Small Birch Bell, *Epinotia ramella* (Linnaeus, 1758).** European Nucleotide Archive. [dataset], accession number PRJEB55602, 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