DATA NOTE

The genome sequence of the Pine Beauty, *Panolis flammea* (Denis & Schiffermüller), 1775 (Lepidoptera: Noctuidae)

[version 1; peer review: 2 approved]

Yannick Chittaro[1], Camille Cornet[2], Kay Lucek[2], Charlotte J. Wright[3], Joana I. Meier[3], Mark L. Blaxter[3],
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, Project Psyche Community

[1]Info fauna, Neuchâtel, Switzerland
[2]University of Neuchâtel, Neuchâtel, Switzerland
[3]Tree of Life, Wellcome Sanger Institute, Hinxton, England, UK

v1 **First published:** 15 Aug 2025, **10**:444
https://doi.org/10.12688/wellcomeopenres.24674.1
Latest published: 15 Aug 2025, **10**:444
https://doi.org/10.12688/wellcomeopenres.24674.1

Abstract

We present a genome assembly from a male specimen of *Panolis flammea* (Pine Beauty; Arthropoda; Insecta; Lepidoptera; Noctuidae). The assembly contains two haplotypes with total lengths of 709.26 megabases and 713.60 megabases. Most of haplotype 1 (99.83%) is scaffolded into 31 chromosomal pseudomolecules, including the Z sex chromosome. Haplotype 2 was assembled to scaffold level. The mitochondrial genome has also been assembled, with a length of 15.46 kilobases.

Keywords

Panolis flammea, Pine Beauty, genome sequence, chromosomal, Lepidoptera

This article is included in the Tree of Life gateway.

Open Peer Review

Approval Status ✓ ✓

	1	2
version 1 15 Aug 2025	✓ view	✓ view

1. **Arun Arumugaperumal**, Rajalakshmi Engineering College, Chennai, India

2. **Sivasankaran Kuppusamy**, Loyola College, Chennai, India

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

Corresponding author: Charlotte J. Wright (cw22@sanger.ac.uk)

Author roles: Chittaro Y: Investigation, Resources, Writing – Original Draft Preparation; **Cornet C**: Writing – Review & Editing; **Lucek K**: Resources, Supervision; **Wright CJ**: Funding Acquisition, Project Administration, Supervision; **Meier JI**: Funding Acquisition, Project Administration, Supervision; **Blaxter ML**: Funding Acquisition, Project Administration, Supervision;

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by Wellcome through core funding to the Wellcome Sanger Institute (220540). CC and KL were supported by the Swiss National Science Foundation Grant PCEFP3_202869.
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: Chittaro Y, Cornet C, Lucek K *et al.* **The genome sequence of the Pine Beauty, *Panolis flammea* (Denis & Schiffermüller), 1775 (Lepidoptera: Noctuidae) [version 1; peer review: 2 approved]** Wellcome Open Research 2025, **10**:444 https://doi.org/10.12688/wellcomeopenres.24674.1

First published: 15 Aug 2025, **10**:444 https://doi.org/10.12688/wellcomeopenres.24674.1

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; Hadeninae; *Panolis*; *Panolis flammea* (Denis & Schiffermüller), 1775 (NCBI:txid988141)

Background

The Pine Beauty moth *Panolis flammea* (Denis & Schiffermüller, 1775) is a medium-sized (wingspan 32–40 mm) moth of the family Noctuidae. Although its background colour can vary from reddish to greenish, its patterning with two large pale stigmata makes it unmistakable throughout its range.

The species is widely distributed in the western Palaearctic, from southwestern Spain (Monasterio León *et al.*, 2020) throughout Europe, the Caucasus to the temperate regions of Central Asia and the western Siberian taiga (GBIF Secretariat, 2023; Ronkay *et al.*, 2001), with the easternmost distribution from the Krasnoyarsk area in Russia (Kononenko & Mikkola, 1989). Further east and as far as Japan, the allopatric species *Panolis japonica* Draudt, 1935 is present, morphologically very similar but distinguished in particular by the simple antennae of the males and differences in the genitalia of the two sexes (Kononenko & Mikkola, 1989; Wang *et al.*, 2014), while in southern and central China there are other related species (Wang *et al.*, 2014). The species is expanding in the northern part of its range in relation to climate change, as shown in Finland (Pulgarin Díaz *et al.*, 2022).

The caterpillar develops mainly on Scots pine *Pinus sylvestris* and other pine species, feeding on their needles, but other conifers and sometimes even broad-leaved trees have also been reported as larval food plants (Robinson *et al.*, 2023), particularly in the event of an outbreak. Two genetically controlled melanic forms also exist, their frequency probably being linked to availability of food plants and resting sites (Majerus, 1982).

The species is undemanding, and can be found wherever pine trees grow, from dry Mediterranean pine forests to cooler forests, as well as plantations and even isolated trees in urban areas, from the plains to around 1 600 m in the Alps (Wymann *et al.*, 2015). The species hibernates as a pupa in the litter and the moth flies early in the year, in one generation, mainly between January and May, depending on the region (Monasterio León *et al.*, 2020) and altitude. The female lays her eggs in rows on the previous year's needles (Hicks *et al.*, 2001).

The Pine Beauty moth is sometimes a forest pest (Carter, 1984) and the caterpillars can swarm and cause major damage, which has been reported in Europe for over 200 years (Klimetzek, 1979), particularly in pine plantations (Hicks *et al.*, 2001) with fewer natural enemies (Hicks *et al.*, 2008). A study of Scottish populations using nuclear and mitochondrial markers suggested relatively high population structure and low dispersal ability (Lowe *et al.*, 2005). However, pest outbreaks are likely affected by climate change (Haynes *et al.*, 2014), and reference genomes of pest species such as *P. flammea* will contribute to efficient monitoring in a changing environment.

Methods

Sample acquisition

The specimen used for genome sequencing was an adult male *Panolis flammea* (specimen ID SAN28000123, ToLID ilPanFlam1; Figure 1), collected from Essertines VD, Switzerland (latitude 46.7165, longitude 6.641; elevation 600 m) on 15/04/2023. The specimen was collected and identified by Yannick Chittaro (Info Fauna, Neuchâtel, Switzerland).

Nucleic acid extraction

Protocols for high molecular weight (HMW) DNA extraction developed at the Wellcome Sanger Institute (WSI) Tree of Life Core Laboratory are available on protocols.io (Howard *et al.*, 2025). The ilPanFlam1 sample was weighed and triaged to determine the appropriate extraction protocol. Tissue from the thorax was homogenised by powermashing using a PowerMasher II tissue disruptor.

HMW DNA was extracted in the WSI Scientific Operations core using the Automated MagAttract v2 protocol. DNA was sheared into an average fragment size of 12–20 kb following the Megaruptor®3 for LI PacBio protocol. Sheared DNA was purified by automated SPRI (solid-phase reversible immobilisation). The concentration of the sheared and purified DNA was assessed using a Nanodrop spectrophotometer and Qubit Fluorometer using the Qubit dsDNA High Sensitivity Assay kit.

Figure 1. Voucher photograph of the *Panolis flammea* (ilPanFlam1) specimen used for genome sequencing.

Fragment size distribution was evaluated by running the sample on the FemtoPulse system. For this sample, the final post-shearing DNA had a Qubit concentration of 54.85 ng/μL and a yield of 2 577.95 ng, with a fragment size of 14.1 kb. The 260/280 spectrophotometric ratio was 1.88, and the 260/230 ratio was 2.62.

PacBio HiFi library preparation and sequencing

Library preparation and sequencing were performed at the WSI Scientific Operations core. Samples with an average fragment size greater than 8 kb and total mass exceeding 400 ng were eligible for the low-input SMRTbell Prep Kit 3.0 protocol (Pacific Biosciences, California, USA), depending on genome size and required sequencing depth. Libraries were prepared using the SMRTbell Prep Kit 3.0 according to the manufacturer's instructions. The kit includes reagents for end repair/A-tailing, adapter ligation, post-ligation SMRTbell bead clean-up, and nuclease treatment. Size selection and clean-up were performed using diluted AMPure PB beads (Pacific Biosciences). DNA concentration was quantified using a Qubit Fluorometer v4.0 (ThermoFisher Scientific) and the Qubit 1X dsDNA HS assay kit. Final library fragment size was assessed with the Agilent Femto Pulse Automated Pulsed Field CE Instrument (Agilent Technologies) using the gDNA 55 kb BAC analysis kit.

The sample was sequenced on a Revio instrument (Pacific Biosciences). The prepared library was normalised to 2 nM, and 15 μL was used for making complexes. Primers were annealed and polymerases bound to generate circularised complexes, following the manufacturer's instructions. Complexes were purified using 1.2X SMRTbell beads, then diluted to the Revio loading concentration (200–300 pM) and spiked with a Revio sequencing internal control. The sample was sequenced on a Revio 25M SMRT cell. The SMRT Link software (Pacific Biosciences), a web-based workflow manager, was used to configure and monitor the run and to carry out primary and secondary data analysis.

Specimen details, sequencing platforms, and data yields are summarised in Table 1.

Table 1. Specimen and sequencing data for BioProject.

Platform	PacBio HiFi	Hi-C
ToLID	ilPanFlam1	ilPanFlam1
Specimen ID	SAN28000123	SAN28000123
BioSample (source individual)	SAMEA115110018	SAMEA115110018
BioSample (tissue)	SAMEA115110068	SAMEA115110066
Tissue	thorax	head
Sequencing platform and model	Revio	Illumina NovaSeq X
Run accessions	ERR13605500	ERR13602163
Read count total	1.86 million	931.08 million
Base count total	19.03 Gb	140.59 Gb

Hi-C

Sample preparation and crosslinking

The Hi-C sample was prepared from 20–50 mg of frozen head tissue of the ilPanFlam1 sample using the Arima-HiC v2 kit (Arima Genomics). Following the manufacturer's instructions, tissue was fixed and DNA crosslinked using TC buffer to a final formaldehyde concentration of 2%. The tissue was homogenised using the Diagnocine Power Masher-II. Crosslinked DNA was digested with a restriction enzyme master mix, biotinylated, and ligated. Clean-up was performed with SPRISelect beads before library preparation. DNA concentration was measured with the Qubit Fluorometer (Thermo Fisher Scientific) and Qubit HS Assay Kit. The biotinylation percentage was estimated using the Arima-HiC v2 QC beads.

Hi-C library preparation and sequencing

Biotinylated DNA constructs were fragmented using a Covaris E220 sonicator and size selected to 400–600 bp using SPRISelect beads. DNA was enriched with Arima-HiC v2 kit Enrichment beads. End repair, A-tailing, and adapter ligation were carried out with the NEBNext Ultra II DNA Library Prep Kit (New England Biolabs), following a modified protocol where library preparation occurs while DNA remains bound to the Enrichment beads. Library amplification was performed using KAPA HiFi HotStart mix and a custom Unique Dual Index (UDI) barcode set (Integrated DNA Technologies). Depending on sample concentration and biotinylation percentage determined at the crosslinking stage, libraries were amplified with 10–16 PCR cycles. Post-PCR clean-up was performed with SPRISelect beads. Libraries were quantified using the Accu-Clear Ultra High Sensitivity dsDNA Standards Assay Kit (Biotium) and a FLUOstar Omega plate reader (BMG Labtech).

Prior to sequencing, libraries were normalised to 10 ng/μL. Normalised libraries were quantified again and equimolar and/or weighted 2.8 nM pools. Pool concentrations were checked using the Agilent 4200 TapeStation (Agilent) with High Sensitivity D500 reagents before sequencing. Sequencing was performed using paired-end 150 bp reads on the Illumina NovaSeq X.

Specimen details, sequencing platforms, and data yields are summarised in Table 1.

Genome assembly

Prior to assembly of the PacBio HiFi reads, a database of *k*-mer counts ($k = 31$) was generated from the filtered reads using FastK. GenomeScope2 (Ranallo-Benavidez *et al.*, 2020) was used to analyse the *k*-mer frequency distributions, providing estimates of genome size, heterozygosity, and repeat content.

The HiFi reads were assembled using Hifiasm in Hi-C phasing mode (Cheng *et al.*, 2021; Cheng *et al.*, 2022), producing two haplotypes. Hi-C reads (Rao *et al.*, 2014) were mapped to the primary contigs using bwa-mem2 (Vasimuddin *et al.*, 2019). Contigs were further scaffolded with Hi-C data in YaHS (Zhou *et al.*, 2023), using the --break option for handling potential

misassemblies. Thescaffolded assemblies were evaluated using Gfastats (Formenti *et al.*, 2022), BUSCO (Manni *et al.*, 2021) and MERQURY.FK(Rhie *et al.*, 2020).

The mitochondrial genome was assembled using MitoHiFi (Uliano-Silva *et al.*, 2023), which runs MitoFinder (Allio *et al.*, 2020) and uses these annotations to select the final mitochondrial contig and to ensure the general quality of the sequence.

Assembly curation

The assembly was decontaminated using the Assembly Screen for Cobionts and Contaminants (ASCC) pipeline. TreeVal was used to generate the flat files and maps for use in curation. Manual curation was conducted primarily in PretextView and HiGlass (Kerpedjiev *et al.*, 2018). Scaffolds were visually inspected and corrected as described by Howe *et al.* (2021). Manual corrections included 23 breaks, 42 joins, and removal of 16 haplotypic duplications. The curation process is documented at https://gitlab.com/wtsi-grit/rapid-curation. PretextSnapshot was used to generate a Hi-C contact map of the final assembly.

Assembly quality assessment

Chromosomal painting was performed using lep_busco_painter using Merian elements, which represent the 32 ancestral linkage groups in Lepidoptera (Wright *et al.*, 2024). Painting was based on gene locations from the lepidoptera_odb10 BUSCO analysis and chromosome lengths from the genome index produced using SAMtools faidx (Danecek *et al.*, 2021). Each complete BUSCO (including both single-copy and duplicated BUSCOs) was assigned to a Merian element using a reference database, and coloured positions were plotted along chromosomes drawn to scale.

The Merqury.FK tool (Rhie *et al.*, 2020), run in a Singularity container (Kurtzer *et al.*, 2017), was used to evaluate *k*-mer completeness and assembly quality for both haplotypes using the *k*-mer databases (k = 31) computed prior to genome assembly. The analysis outputs included assembly QV scores and completeness statistics.

The genome was analysed using the BlobToolKit pipeline, a Nextflow implementation of the earlier Snakemake BlobToolKit pipeline (Challis *et al.*, 2020). The pipeline aligns PacBio reads using minimap2 (Li, 2018) and SAMtools (Danecek *et al.*, 2021) to generate coverage tracks. Simultaneously, it queries the GoaT database (Challis *et al.*, 2023) to identify relevant BUSCO lineages and runs BUSCO (Manni *et al.*, 2021). For the three domain-level BUSCO lineages, BUSCO genes are aligned to the UniProt Reference Proteomes database (Bateman *et al.*, 2023) using DIAMOND blastp (Buchfink *et al.*, 2021). The genome is divided into chunks based on the density of BUSCO genes from the closest taxonomic lineage, and each chunk is aligned to the UniProt Reference Proteomes database with DIAMOND blastx. Sequences without hits are chunked using seqtk and aligned to the NT database with blastn (Altschul *et al.*, 1990). The BlobToolKit suite consolidates all outputs into a blobdir for visualisation. The BlobToolKit pipeline was developed using nf-core tooling (Ewels *et al.*, 2020) and MultiQC (Ewels *et al.*, 2016), with package management via Conda and Bioconda (Grüning *et al.*, 2018), and containerisation through Docker (Merkel, 2014) and Singularity (Kurtzer *et al.*, 2017).

Genome sequence report

Sequence data

The genome of a specimen of *Panolis flammea* was sequenced using Pacific Biosciences single-molecule HiFi long reads, generating 19.03 Gb (gigabases) from 1.86 million reads, which were used to assemble the genome. GenomeScope2.0 analysis estimated the haploid genome size at 705.69 Mb, with a heterozygosity of 0.83% and repeat content of 31.11%. These estimates guided expectations for the assembly. Based on the estimated genome size, the sequencing data provided approximately 26× coverage. Hi-C sequencing produced 140.59 Gb from 931.08 million reads, which were used to scaffold the assembly. Table 1 summarises the specimen and sequencing details.

Assembly statistics

The genome was assembled into two haplotypes using Hi-C phasing. Haplotype 1 was curated to chromosome level, while haplotype 2 was assembled to scaffold level. The final assembly has a total length of 709.26 Mb in 77 scaffolds, with 192 gaps, and a scaffold N50 of 24.46 Mb (Table 2).

Most of the assembly sequence (99.83%) was assigned to 31 chromosomal-level scaffolds, representing 30 autosomes and the Z sex chromosome. Chromosome Z was assigned based on synteny to the genome of *Agriopis aurantiaria* (GCA_914767915.1) (Boyes *et al.*, 2023). These chromosome-level

Table 2. Genome assembly statistics.

Assembly name	ilPanFlam1.hap1.1	ilPanFlam1.hap2.1
Assembly accession	GCA_964261395.1	GCA_964261445.1
Assembly level	chromosome	scaffold
Span (Mb)	709.26	713.60
Number of chromosomes	31	N/A
Number of contigs	269	301
Contig N50	5.96 Mb	5.16 Mb
Number of scaffolds	77	100
Scaffold N50	24.46 Mb	24.47 Mb
Longest scaffold length (Mb)	32.08	N/A
Sex chromosomes	Z	N/A
Organelles	Mitochondrial genome: 15.46 kb	N/A

scaffolds, confirmed by Hi-C data, are named according to size (Figure 2; Table 3). Chromosome painting with Merian elements illustrates the distribution of orthologues along chromosomes and highlights patterns of chromosomal evolution relative to Lepidopteran ancestral linkage groups (Figure 3).

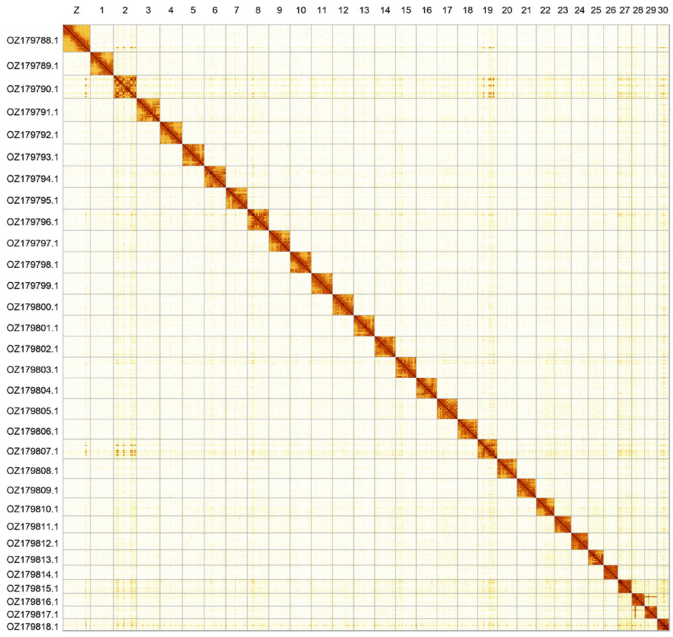

Figure 2. Hi-C contact map of the *Panolis flammea* genome assembly. Assembled chromosomes are shown in order of size and labelled along the axes. The plot was generated using PretextSnapshot.

Table 3. Chromosomal pseudomolecules in the haplotype 1 genome assembly of *Panolis flammea* ilPanFlam1.

INSDC accession	Molecule	Length (Mb)	GC%	Assigned Merian elements
OZ179789.1	1	27.19	38	M2
OZ179790.1	2	27.11	38	M23
OZ179791.1	3	27.02	38.50	M17;M20
OZ179792.1	4	26.08	38.50	M1
OZ179793.1	5	25.73	38	M9
OZ179794.1	6	25.28	38	M11
OZ179795.1	7	25.20	37.50	M18
OZ179796.1	8	25.02	38.50	M14
OZ179797.1	9	24.86	37.50	M12
OZ179798.1	10	24.82	38	M8
OZ179799.1	11	24.75	38	M7
OZ179800.1	12	24.70	38	M5
OZ179801.1	13	24.46	38	M3
OZ179802.1	14	24.33	38	M10
OZ179803.1	15	24.30	38	M21
OZ179804.1	16	24.14	38	M6
OZ179805.1	17	24.13	38	M22
OZ179806.1	18	23.42	37.50	M16
OZ179807.1	19	22.95	38.50	M19
OZ179808.1	20	22.88	38	M4
OZ179809.1	21	22.70	38	M15
OZ179810.1	22	21.17	38	M24
OZ179811.1	23	19.81	38	M26
OZ179812.1	24	19.66	38	M13
OZ179813.1	25	17.96	38	M27
OZ179814.1	26	16.77	38	M28
OZ179815.1	27	16.07	39	M31
OZ179816.1	28	15.05	39.50	M29
OZ179817.1	29	14.66	38.50	M25
OZ179818.1	30	13.75	40.50	M30
OZ179788.1	Z	32.08	38	MZ

The mitochondrial genome was also assembled. This sequence is included as a contig in the multifasta file of the genome submission and as a standalone record.

Assembly quality metrics

For haplotype 1, the estimated QV is 64.9, and for haplotype 2, 65.0. When the two haplotypes are combined, the assembly achieves an estimated QV of 65.0. The *k*-mer completeness is 81.28% for haplotype 1, 81.32% for haplotype 2, and 99.56% for the combined haplotypes (Figure 4). BUSCO analysis using the lepidoptera_odb10 reference set (n = 5 286) (Kriventseva *et al.*, 2019) identified 98.8% of the expected gene set (single = 98.2%, duplicated = 0.6%) for haplotype 1. The snail plot in Figure 5 summarises the scaffold length distribution and other

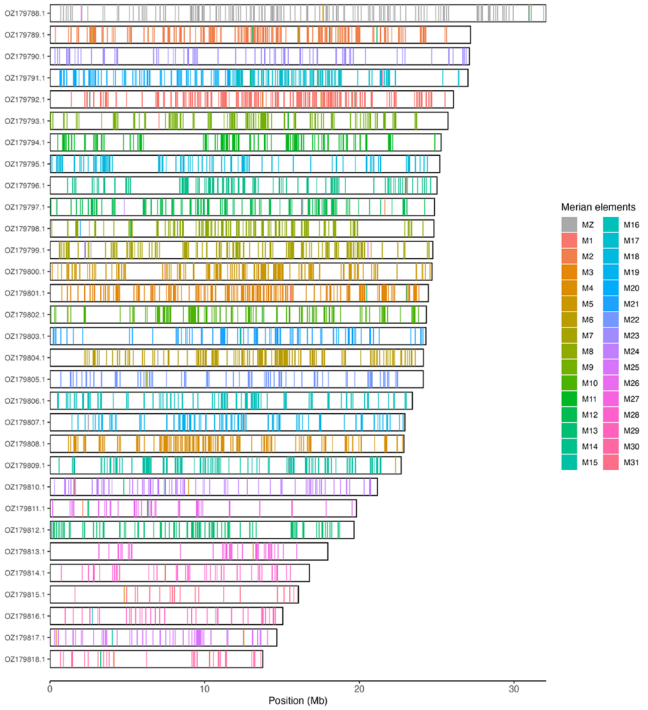

Figure 3. Merian elements painted across chromosomes in the ilPanFlam1.hap1.1 assembly of *Panolis flammea*. Chromosomes are drawn to scale, with the positions of orthologues shown as coloured bars. Each orthologue is coloured by the Merian element that it belongs to. All orthologues which could be assigned to Merian elements are shown.

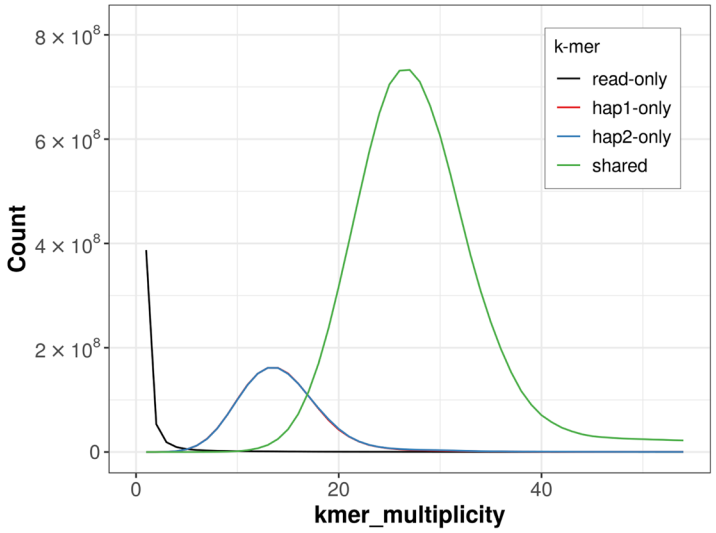

Figure 4. Evaluation of *k*-mer completeness using MerquryFK. This plot illustrates the recovery of *k*-mers from the original read data in the final assemblies. The horizontal axis represents *k*-mer multiplicity, and the vertical axis shows the number of *k*-mers. The black curve represents *k*-mers that appear in the reads but are not assembled. The green curve (the homozygous peak) corresponds to *k*-mers shared by both haplotypes and the red and blue curves (the heterozygous peaks) show *k*-mers found only in one of the haplotypes.

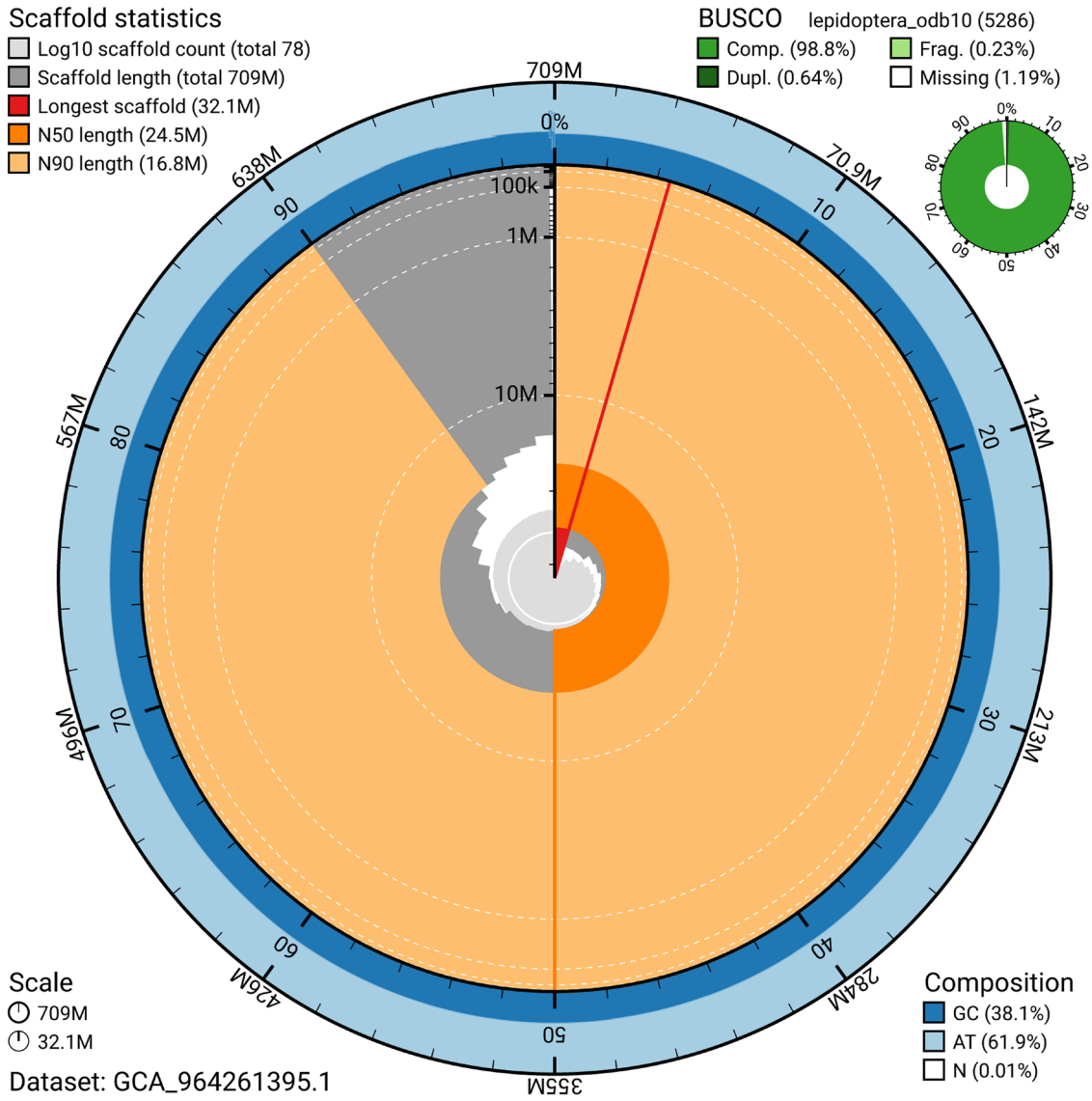

Figure 5. Assembly metrics for ilPanFlam1.hap1.1. The BlobToolKit snail plot provides an overview of assembly metrics and BUSCO gene completeness. The circumference represents the length of the whole genome sequence, and the main plot is divided into 1,000 bins around the circumference. The outermost blue tracks display the distribution of GC, AT, and N percentages across the bins. Scaffolds are arranged clockwise from longest to shortest and are depicted in dark grey. The longest scaffold is indicated by the red arc, and the deeper orange and pale orange arcs represent the N50 and N90 lengths. A light grey spiral at the centre shows the cumulative scaffold count on a logarithmic scale. A summary of complete, fragmented, duplicated, and missing BUSCO genes in the set is presented at the top right. An interactive version of this figure can be accessed on the BlobToolKit viewer.

assembly statistics for haplotype 1. The blob plot in Figure 6 shows the distribution of scaffolds by GC proportion and coverage for haplotype 1.

Table 4 lists the assembly metric benchmarks adapted from Rhie *et al.* (2021) the Earth BioGenome Project Report on Assembly Standards September 2024. The EBP metric, calculated for the haplotype 1, is **6.C.Q64**, meeting the recommended reference standard.

Wellcome Sanger Institute – Legal and Governance

The materials that have contributed to this genome note have been supplied by a Tree of Life collaborator. 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

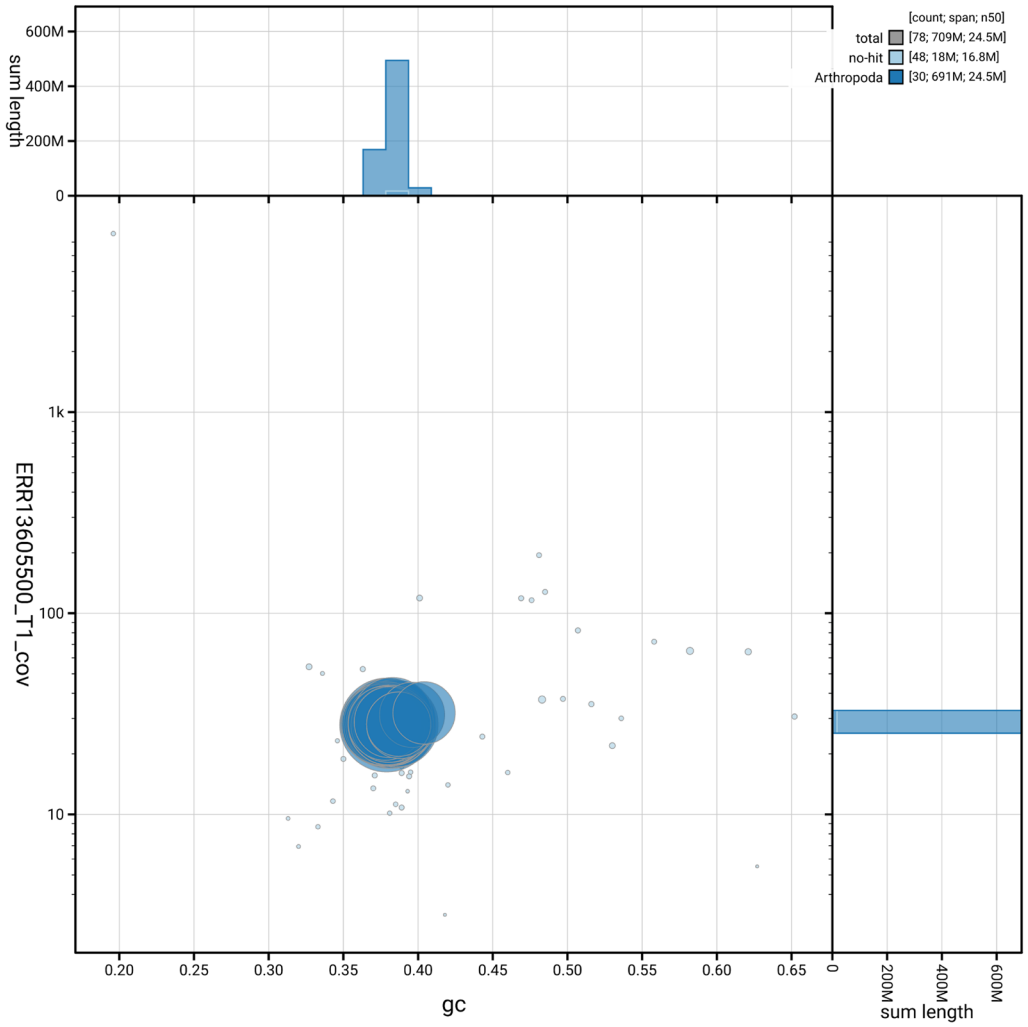

Figure 6. BlobToolKit GC-coverage plot for ilPanFlam1.hap1.1. Blob plot showing sequence coverage (vertical axis) and GC content (horizontal axis). The circles represent scaffolds, with the size proportional to scaffold length and the colour representing phylum membership. The histograms along the axes display the total length of sequences distributed across different levels of coverage and GC content. An interactive version of this figure is available on the BlobToolKit viewer.

Table 4. Earth Biogenome Project summary metrics for the *Panolis flammea* assembly.

Measure	Value	Benchmark
EBP summary (haplotype 1)	6.C.Q64	6.C.Q40
Contig N50 length	5.96 Mb	≥ 1 Mb
Scaffold N50 length	24.46 Mb	= chromosome N50
Consensus quality (QV)	Haplotype 1: 64.9; haplotype 2: 65.0; combined: 65.0	≥ 40
k-mer completeness	Haplotype 1: 81.28%; Haplotype 2: 81.32%; combined: 99.56%	≥ 95%
BUSCO	C:98.8%[S:98.2%,D:0.6%], F:0.2%,M:1.0%,n:5 286	S > 90%; D < 5%
Percentage of assembly assigned to chromosomes	99.83%	≥ 90%

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 undertaken according to a Research Collaboration Agreement or Material Transfer Agreement entered into by the Tree of Life collaborator, Genome Research Limited (operating as the Wellcome Sanger Institute), and in some circumstances, other Tree of Life collaborators.

Data availability

European Nucleotide Archive: Panolis flammea (pine beauty). Accession number PRJEB79174. The genome sequence is released openly for reuse. The *Panolis flammea* genome sequencing initiative is part of the Sanger Institute Tree of Life Programme (PRJEB43745) and Project Psyche (PRJEB71705). 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 Ensembl at the European Bioinformatics Institute. Raw data and assembly accession identifiers are reported in Table 1 and Table 2.

Pipelines used for genome assembly at the WSI Tree of Life are available at https://pipelines.tol.sanger.ac.uk/pipelines. Table 5 lists software versions used in this study.

Author information

Contributors are listed at the following links:

- 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
- Project Psyche Community.

Table 5. Software versions and sources.

Software	Version	Source
BEDTools	2.30.0	https://github.com/arq5x/bedtools2
BLAST	2.14.0	ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/
BlobToolKit	4.3.9	https://github.com/blobtoolkit/blobtoolkit
BUSCO	5.5.0	https://gitlab.com/ezlab/busco
bwa-mem2	2.2.1	https://github.com/bwa-mem2/bwa-mem2
Cooler	0.8.11	https://github.com/open2c/cooler
DIAMOND	2.1.8	https://github.com/bbuchfink/diamond
fasta_windows	0.2.4	https://github.com/tolkit/fasta_windows
FastK	1.1	https://github.com/thegenemyers/FASTK
GenomeScope2.0	2.0.1	https://github.com/tbenavi1/genomescope2.0
Gfastats	1.3.6	https://github.com/vgl-hub/gfastats
GoaT CLI	0.2.5	https://github.com/genomehubs/goat-cli
Hifiasm	0.19.8-r603	https://github.com/chhylp123/hifiasm
HiGlass	1.13.4	https://github.com/higlass/higlass
MerquryFK	1.1.2	https://github.com/thegenemyers/MERQURY.FK
Minimap2	2.24-r1122	https://github.com/lh3/minimap2
MitoHiFi	3	https://github.com/marcelauliano/MitoHiFi
MultiQC	1.14; 1.17 and 1.18	https://github.com/MultiQC/MultiQC
Nextflow	23.10.0	https://github.com/nextflow-io/nextflow

Software	Version	Source
PretextSnapshot	N/A	https://github.com/sanger-tol/PretextSnapshot
PretextView	0.2.5	https://github.com/sanger-tol/PretextView
samtools	1.19.2	https://github.com/samtools/samtools
sanger-tol/ascc	0.1.0	https://github.com/sanger-tol/ascc
sanger-tol/blobtoolkit	0.6.0	https://github.com/sanger-tol/blobtoolkit
Seqtk	1.3	https://github.com/lh3/seqtk
Singularity	3.9.0	https://github.com/sylabs/singularity
TreeVal	1.2.0	https://github.com/sanger-tol/treeval
YaHS	1.2a.2	https://github.com/c-zhou/yahs

References

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

Altschul SF, Gish W, Miller W, *et al.*: **Basic Local Alignment Search Tool.** *J Mol Biol.* 1990; **215**(3): 403–410.
PubMed Abstract | Publisher Full Text

Bateman A, Martin MJ, Orchard S, *et al.*: **UniProt: the universal protein knowledgebase in 2023.** *Nucleic Acids Res.* 2023; **51**(D1): D523–D531.
PubMed Abstract | Publisher Full Text | Free Full Text

Boyes D, Mulhair PO, University of Oxford and Wytham Woods Genome Acquisition Lab, *et al.*: **The genome sequence of the scarce umber, *Agriopis aurantiaria* (Hübner, 1799).** *Wellcome Open Res.* 2023; **8**: 463.
PubMed Abstract | Publisher Full Text | Free Full Text

Buchfink B, Reuter K, Drost HG: **Sensitive protein alignments at Tree-of-Life scale using DIAMOND.** *Nat Methods.* 2021; **18**(4): 366–368.
PubMed Abstract | Publisher Full Text | Free Full Text

Carter DJ: **Pest Lepidoptera of Europe, with special reference to the British Isles**. Dordrecht: Springer Netherlands, 1984; **30**.
Reference Source

Challis R, Kumar S, Sotero-Caio C, *et al.*: **Genomes on a Tree (GoaT): a versatile, scalable search engine for genomic and sequencing project metadata across the eukaryotic Tree of Life [version 1; peer review: 2 approved].** *Wellcome Open Res.* 2023; **8**: 24.
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

Cheng H, Jarvis ED, Fedrigo O, *et al.*: **Haplotype-resolved assembly of diploid genomes without parental data.** *Nat Biotechnol.* 2022; **40**(9): 1332–1335.
PubMed Abstract | Publisher Full Text | Free Full Text

Danecek P, Bonfield JK, Liddle J, *et al.*: **Twelve years of SAMtools and BCFtools.** *GigaScience.* 2021; **10**(2): giab008.
PubMed Abstract | Publisher Full Text | Free Full Text

Ewels P, Magnusson M, Lundin S, *et al.*: **MultiQC: summarize analysis results for multiple tools and samples in a single report.** *Bioinformatics.* 2016; **32**(19): 3047–3048.
PubMed Abstract | Publisher Full Text | Free Full Text

Ewels PA, Peltzer A, Fillinger S, *et al.*: **The nf-core framework for community-curated bioinformatics pipelines.** *Nat Biotechnol.* 2020; **38**(3): 276–278.
PubMed Abstract | Publisher Full Text

Formenti G, Abueg L, Brajuka A, *et al.*: **Gfastats: conversion, evaluation and manipulation of genome sequences using assembly graphs.** *Bioinformatics.* 2022; **38**(17): 4214–4216.
PubMed Abstract | Publisher Full Text | Free Full Text

GBIF Secretariat: ***Panolis flammea* (Denis & Schiffermüller) 1775**. 2023.
Publisher Full Text

Grüning B, Dale R, Sjödin A, *et al.*: **Bioconda: sustainable and comprehensive software distribution for the life sciences.** *Nat Methods.* 2018; **15**(7): 475–476.
PubMed Abstract | Publisher Full Text | Free Full Text

Haynes KJ, Allstadt AJ, Klimetzek D: **Forest defoliator outbreaks under climate change: effects on the frequency and severity of outbreaks of five pine insect pests.** *Glob Chang Biol.* 2014; **20**(6): 2004–18.
PubMed Abstract | Publisher Full Text

Hicks BJ, Barbour DA, Evans HF, *et al.*: **The history and control of the pine beauty moth, *Panolis flammea* (D. & S.) Lepidoptera: Noctuidae), in Scotland from 1976 to 2000.** *Agr Forest Entomol.* 2001; **3**(3): 161–68.
Publisher Full Text

Hicks BJ, Leather SR, Watt AD: **Changing dynamics of the pine beauty moth (*Panolis flammea*) in Britain: the loss of enemy free space?** *Agric For Entomol.* 2008; **10**(3): 263–71.
Publisher Full Text

Howard C, Denton A, Jackson B, *et al.*: **On the path to reference genomes for all biodiversity: lessons learned and laboratory protocols created in the Sanger Tree of Life core laboratory over the first 2000 species.** *bioRxiv.* 2025.
Publisher Full Text

Howe K, Chow W, Collins J, *et al.*: **Significantly improving the quality of genome assemblies through curation.** *GigaScience.* 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

Klimetzek D: **Insekten-Großschädlinge an Kiefer in Nordbayern und der Pfalz: analyse und Vergleich 1810–1970**. 1979; **2**.
Reference Source

Kononenko VS, Mikkola K: ***Panolis flammea* auct. – two species (Lepidoptera: Noctuidae, Hadeninae).** *Nota Lepidopterol.* 1989; **12**: 29–35.

Kriventseva EV, Kuznetsov D, Tegenfeldt F, *et al.*: **OrthoDB v10: Sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs.** *Nucleic Acids Res.* 2019; **47**(D1): D807–D811.
PubMed Abstract | Publisher Full Text | Free Full Text

Kurtzer GM, Sochat V, Bauer MW: **Singularity: scientific containers for mobility of compute.** *PLoS One.* 2017; **12**(5): e0177459.
PubMed Abstract | Publisher Full Text | Free Full Text

Li H: **Minimap2: pairwise alignment for nucleotide sequences.** *Bioinformatics.* 2018; **34**(18): 3094–3100.
PubMed Abstract | Publisher Full Text | Free Full Text

Lowe AJ, Hicks BJ, Worley K, *et al.*: **Genetic differentiation in Scottish populations of the pine beauty moth, *Panolis flammea* (Lepidoptera: Noctuidae).** *Bull Entomol Res.* 2005; **95**(6): 517–26.
PubMed Abstract | Publisher Full Text

Majerus MEN: **Genetic control of two melanic forms of *Panolis flammea* (Lepidoptera: Noctuidae).** *Heredity.* 1982; **49**(2): 171–77.
Publisher 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

Merkel D: **Docker: lightweight Linux containers for consistent development and deployment.** *Linux J.* 2014; **2014**(239): 2.
Reference Source

Monasterio León Y, Aedo Elguezábal O, Escobés Jiménez R, *et al.*: **Primer registro de *Panolis flammea* ([Denis & Schiffermüller], 1775) (Lepidoptera: Noctuidae) para el continente Africano.** *Boletín de la SEA.* 2020; **66**: 273–75.
Reference Source

Pulgarin Díaz JA, Melin M, Tikkanen OP: **Thermal sum drives abundance and distribution range shift of *Panolis flammea* in Finland.** *Scand J For Res.* 2022; **37**(2): 93–98.
Publisher Full Text

Ranallo-Benavidez TR, Jaron KS, Schatz MC: **GenomeScope 2.0 and smudgeplot for reference-free profiling of polyploid genomes.** *Nat Commun.* 2020; **11**(1): 1432.
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

Robinson GS, Ackery PR, Kitching I, *et al.*: **HOSTS (from HOSTS - a database of the world's Lepidopteran hostplants)**. 2023.
Publisher Full Text

Ronkay L, Yela JL, Hreblay M: **Hadeninae II. Noctuidae Europaeae. Volume 5**. Sorø: Entomological Press, 2001.
Reference Source

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 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

Wang H, Fan X, Owada M, *et al.*: **Phylogeny, systematics and biogeography of the genus *Panolis* (Lepidoptera: Noctuidae) based on morphological and molecular evidence.** *PLoS One.* 2014; **9**(6): e90598.
PubMed Abstract | Publisher Full Text | Free Full Text

Wright CJ, Stevens L, Mackintosh A, *et al.*: **Comparative genomics reveals the dynamics of chromosome evolution in Lepidoptera.** *Nat Ecol Evol.* 2024; **8**(4): 777–90.
PubMed Abstract | Publisher Full Text | Free Full Text

Wymann HP, Rezbanyai-Reser L, Hächler M: **Die Eulenfalter der Schweiz Lepidoptera: Noctuidae, Pantheidae, Nolidae**. CSCF & SEG Neuchâtel, 2015; **28**.

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 16 September 2025

https://doi.org/10.21956/wellcomeopenres.27189.r130117

✔ **Sivasankaran Kuppusamy**
Loyola College, Chennai, India

A comprehensive chromosome-level genome of *Panolis flammea* (Denis & Schiffermuller, 1775) was sequenced, assembled and annotated using appropriate methods and softwares. The assembly contains two haplotypes with lengths of 709.26 Mb and 713.60 Mb, respectively. The sex chromosome Z was assigned based on the synteny to the genome of *Agriopis aurantiaria*.

Comments on the manuscript
In the title of the manuscript authors have given the species described year outside the parentheses like (Denis & Schiffermuller),1775. The species' descriptors' name and year can be given within the parentheses i.e., (Denis & Schiffermuller, 1775).

The assembly contains two haplotypes with more or less equal megabses. One haplotype scaffolded up to chromosome level another one is annotated up to scaffolding level only. Is there any specific reason?

The table 1 caption “Specimen details, sequencing platforms, and data yields are summarized in Table 1” was given two times in the methods. One may be deleted

Sequence annotation details like number of protein-coding genes Gene transcripts, non-coding genes have not been given in the text.

Page no 5 first column first paragraph “Thescaffolded” words were fused. The words can be separated

The research article was well prepared and the manuscript meets the necessary scientific standard and is suitable for indexing

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: Mitochondrial genome analysis of Noctuoidea moths

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 16 September 2025

https://doi.org/10.21956/wellcomeopenres.27189.r130123

Arun Arumugaperumal
Department of Biotechnology, Rajalakshmi Engineering College, Chennai, Tamil Nadu, India

The data note describes the genome sequence of the moth *Panolis flammea*, also known as the Pine Beauty. The photograph of the remains of specimen used for sequencing is shown in figure 1. But the remains does not provide a clear image of the moth. A photograph of an alive moth might provide more information about its identity. The authors have used advanced sequencing technology and the final genome assembly size is 709.26 Mb spread across 31 chromosomes. Since this is a male specimen they have identified a Z chromosome. The moth has ZW sex system. It would have been better if the authors had chosen a female specimen for sequencing. The mitochondrial genome was also assembled and is of size 15.46 kb. BUSCO analysis shows that the completeness percentage is 98.8%. The annotations were supposed to be released in Ensembl but as of now it has not been released. The accession numbers of the INSDC databases were working correctly. The article can be indexed.

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: Genomics; Bioinformatics

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.