DATA NOTE

The genome sequence of the Meadow Fritillary, *Melitaea parthenoides* Keferstein, 1851 (Lepidoptera: Nymphalidae)

[version 1; peer review: 2 approved]

Yannick Chittaro[1], 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:** 11 Aug 2025, **10**:436
https://doi.org/10.12688/wellcomeopenres.24718.1
Latest published: 11 Aug 2025, **10**:436
https://doi.org/10.12688/wellcomeopenres.24718.1

Abstract

We present a genome assembly from a female specimen of *Melitaea parthenoides* (Meadow Fritillary; Arthropoda; Insecta; Lepidoptera; Nymphalidae). The assembly contains two haplotypes with total lengths of 592.15 megabases and 568.40 megabases. Most of haplotype 1 (99.72%) 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 15.17 kilobases.

Keywords

Melitaea parthenoides, Meadow Fritillary, genome sequence, chromosomal, Lepidoptera

This article is included in the Tree of Life gateway.

Open Peer Review

Approval Status ✓ ✓

	1	2
version 1 11 Aug 2025	✓ view	✓ view

1. **Evan J Kipp**, University of Minnesota Twin Cities Institute on the Environment, Saint Paul, USA

2. **Arun Arumugaperumal**, Rajalakshmi Engineering College, Thandalam 602105, 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; **Lucek K**: Resources, Supervision, Writing – Review & Editing; **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). KL was 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, Lucek K, Wright CJ *et al.* **The genome sequence of the Meadow Fritillary, *Melitaea parthenoides* Keferstein, 1851 (Lepidoptera: Nymphalidae) [version 1; peer review: 2 approved]** Wellcome Open Research 2025, **10**:436 https://doi.org/10.12688/wellcomeopenres.24718.1

First published: 11 Aug 2025, **10**:436 https://doi.org/10.12688/wellcomeopenres.24718.1

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; Papilionoidea; Nymphalidae; Nymphalinae; Melitaeini; Melitaeina; *Melitaea*; *Mellicta*; *Melitaea parthenoides* Keferstein, 1851 (NCBI:txid113340)

Background

The Meadow Fritillary *Melitaea parthenoides* is endemic to Western Europe (Kudrna *et al.*, 2015). Its range includes the mountainous regions of the Iberian Peninsula (Gómez Bustillo & Fernández-Rubio, 1974), a large part of France (Lafranchis *et al.*, 2015), some parts of the western Italian Alps (Vila *et al.*, 2009), western and northern Switzerland (Ligue Suisse pour la Protection de la Nature, 1987), and southern Germany (Reinhardt *et al.*, 2020). A former presence in Austria, in northern Tyrol, seems credible (Huemer, 2013).

Meliteae parthenoides is a medium-sized Fritillary with a relatively regular black pattern on the upper surface of its wings. It closely resembles several other species in the genus, notably *M. athalia*, *M. celadussa* and *M. varia*. The male genitalia are the most reliable distinguishing feature: The tips of the valves are slightly bent downwards, usually with small ventral spines (van Oorschot & Coutsis, 2014). In females, the background colour is often less uniform, where postdiscal areas are darker and more contrasting.

Melitaea parthenoides colonises dry limestone grasslands and extensive mesophilous meadows, as well as wet meadows, especially on the edges of swamps, from the plains to over 2 400 m in some cases (Gómez Bustillo & Fernández-Rubio, 1974; Lafranchis *et al.*, 2015).

The eggs are laid in patches under the undersides of the leaves of host plants. The young caterpillars live together in a woven nest. Second generation caterpillars overwinter there after the second moult (Ligue Suisse pour la Protection de la Nature, 1987). The host plants are plantain (*Plantago lanceolata*, *P. media*), but the caterpillars sometimes eat other plants, especially after overwintering (*Veronica* spp., *Rhinanthus* spp., *Melampyrum* spp., Reinhardt *et al.*, 2020).

The species is bivoltine at low altitudes (May to June, then August to September) and in warmer locations, whereas at high altitudes or in humid locations it generally forms only one generation in June to July (Ebert & Rennwald, 1991; Gómez Bustillo & Fernández-Rubio, 1974; Lafranchis *et al.*, 2015; Ligue Suisse pour la Protection de la Nature, 1987).

Although it is still fairly widespread in the south and west of its range it is threatened by the intensification of agricultural practices, such as eutrophication due to the regular application of synthetic fertilisers or manure, and the increase in the number of annual cuttings, as well as overgrazing. This has led to its decline in several regions, and the species is considered threatened in some regions (for example in northern France, Lafranchis *et al.*, 2015) or even countries, such as Germany (Reinhardt & Bolz, 2011) and Switzerland (Wermeille *et al.*, 2014).

This reference genome presented here will help to unravel the complex genetic structure revealed by mitochondrial DNA (Dapporto *et al.*, 2022). It will also help to conduct future conservation genomic research, especially in the threatened part of the species range. The sequence data was derived from a female specimen (Figure 1) collected from Dittingen BL, Switzerland.

Methods

Sample acquisition

The specimen used for genome sequencing was an adult female *Melitaea parthenoides* (specimen ID SAN28000138, ToLID ilMelPart1; Figure 1), collected from Dittingen BL, Switzerland (latitude 47.4431, longitude 7.4918; elevation 500 m) on 03/06/2023. 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 ilMelPart1 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

Figure 1. Top: Live specimen of *Melitaea parthenoides*; Bottom: Voucher image of the *Melitaea parthenoides* (ilMelPart1) specimen used for genome sequencing.

of the sheared and purified DNA was assessed using a Nanodrop spectrophotometer and Qubit Fluorometer using the Qubit dsDNA High Sensitivity Assay kit. 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 50.77 ng/μL and a yield of 2 386.19 ng, with a fragment size of 17.0 kb. The 260/280 spectrophotometric ratio was 1.88, and the 260/230 ratio was 2.43.

PacBio HiFi library preparation and sequencing

Library preparation and sequencing were performed at the WSI Scientific Operations core. Libraries were prepared using the SMRTbell Prep Kit 3.0 (Pacific Biosciences, California, USA), following 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.

Hi-C

Sample preparation and crosslinking

The Hi-C sample was prepared from 20–50 mg of frozen head tissue of the ilMelPart1 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 AccuClear 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

Table 1. Specimen and sequencing data for BioProject.

Platform	PacBio HiFi	Hi-C
ToLID	ilMelPart1	ilMelPart1
Specimen ID	SAN28000138	SAN28000138
BioSample (source individual)	SAMEA115110004	SAMEA115110004
BioSample (tissue)	SAMEA115110026	SAMEA115110025
Tissue	thorax	head
Sequencing platform and model	Revio	Illumina NovaSeq X
Run accessions	ERR13605496	ERR13602159
Read count total	1.40 million	727.80 million
Base count total	15.66 Gb	109.90 Gb

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. The scaffolded 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 7 breaks and 20 joins. 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 *Melitaea parthenoides* was sequenced using Pacific Biosciences single-molecule HiFi long reads, generating 15.66 Gb (gigabases) from 1.40 million reads, which were used to assemble the genome. GenomeScope2.0 analysis estimated the haploid genome size at 573.87 Mb, with a heterozygosity of 1.19% and repeat content of 39.00%. These estimates guided expectations for the assembly. Based on the estimated genome size, the sequencing data provided approximately 26× coverage. Hi-C sequencing produced 109.90 Gb from 727.80 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 592.15 Mb in 69 scaffolds, with 116 gaps, and a scaffold N50 of 20.37 Mb (Table 2).

Most of the assembly sequence (99.72%) was assigned to 31 chromosomal-level scaffolds, representing 30 autosomes and the Z sex chromosome. Chromosome Z was assigned based on Hi-C signal. The species appears to be ZO. The chromosome-level 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).

Table 2. Genome assembly statistics.

Assembly name	ilMelPart1.hap1.1	ilMelPart1.hap2.1
Assembly accession	GCA_964261285.1	GCA_964261255.1
Assembly level	chromosome	scaffold
Span (Mb)	592.15	568.40
Number of chromosomes	31	N/A
Number of contigs	185	212
Contig N50	6.74 Mb	6.24 Mb
Number of scaffolds	69	98
Scaffold N50	20.37 Mb	20.13 Mb
Longest scaffold length (Mb)	27.92	N/A
Sex chromosomes	Z	N/A
Organelles	Mitochondrial genome: 15.17 kb	N/A

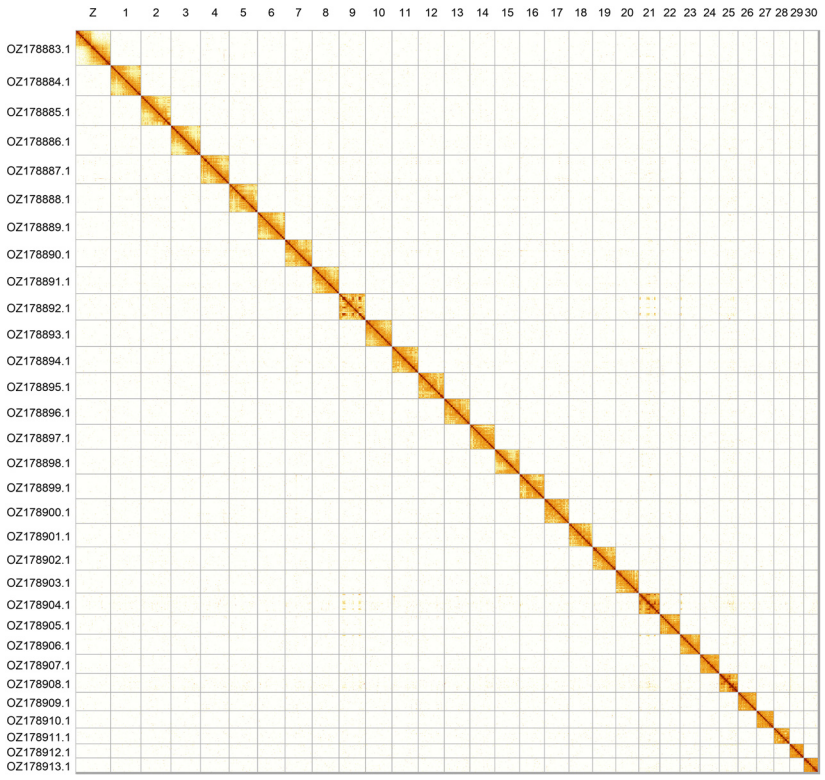

Figure 2. Hi-C contact map of the *Melitaea parthenoides* genome assembly. Assembled chromosomes are shown in order of size and labelled along the axes. The plot was generated using PretextSnapshot.

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 63.7, and for haplotype 2, 63.7. When the two haplotypes are combined, the assembly

Table 3. Chromosomal pseudomolecules in the haplotype 1 genome assembly of *Melitaea parthenoides* ilMelPart1.

INSDC accession	Molecule	Length (Mb)	GC%	Assigned Merian elements
OZ178884.1	1	24.13	34	M2
OZ178885.1	2	23.90	34.50	M17;M20
OZ178886.1	3	23.42	34.50	M1
OZ178887.1	4	22.76	34	M8
OZ178888.1	5	22.72	34.50	M3
OZ178889.1	6	21.73	34	M9
OZ178890.1	7	21.58	34	M6
OZ178891.1	8	21.43	34	M12
OZ178892.1	9	21.08	34.50	M23
OZ178893.1	10	20.99	34.50	M7
OZ178894.1	11	20.93	34	M5
OZ178895.1	12	20.66	34	M18
OZ178896.1	13	20.37	34	M16
OZ178897.1	14	19.96	34.50	M10
OZ178898.1	15	19.71	34	M4
OZ178899.1	16	19.57	34	M21
OZ178900.1	17	19.51	34	M22
OZ178901.1	18	18.84	34	M15
OZ178902.1	19	18.45	34.50	M11
OZ178903.1	20	18.23	35	M14
OZ178904.1	21	16.78	36.50	M30
OZ178905.1	22	16.06	34	M13
OZ178906.1	23	16.03	34.50	M26
OZ178907.1	24	15.18	35	M19
OZ178908.1	25	14.90	37	M31
OZ178909.1	26	14.74	34.50	M24
OZ178910.1	27	13.72	34.50	M28
OZ178911.1	28	12.61	34.50	M27
OZ178912.1	29	11.33	35	M25
OZ178913.1	30	11.26	35.50	M29
OZ178883.1	Z	27.92	34	MZ

achieves an estimated QV of 63.7. The *k*-mer completeness is 77.06% for haplotype 1, 73.81% for haplotype 2, and 99.31% for the combined haplotypes (Figure 4). BUSCO analysis using the lepidoptera_odb10 reference set (n = 5 286) (Kriventseva *et al.*, 2019) identified 98.5% of the expected gene set (single = 97.8%, duplicated = 0.8%) for haplotype 1. The snail plot in Figure 5 summarises the scaffold length distribution and other assembly statistics for haplotype 1. The blob

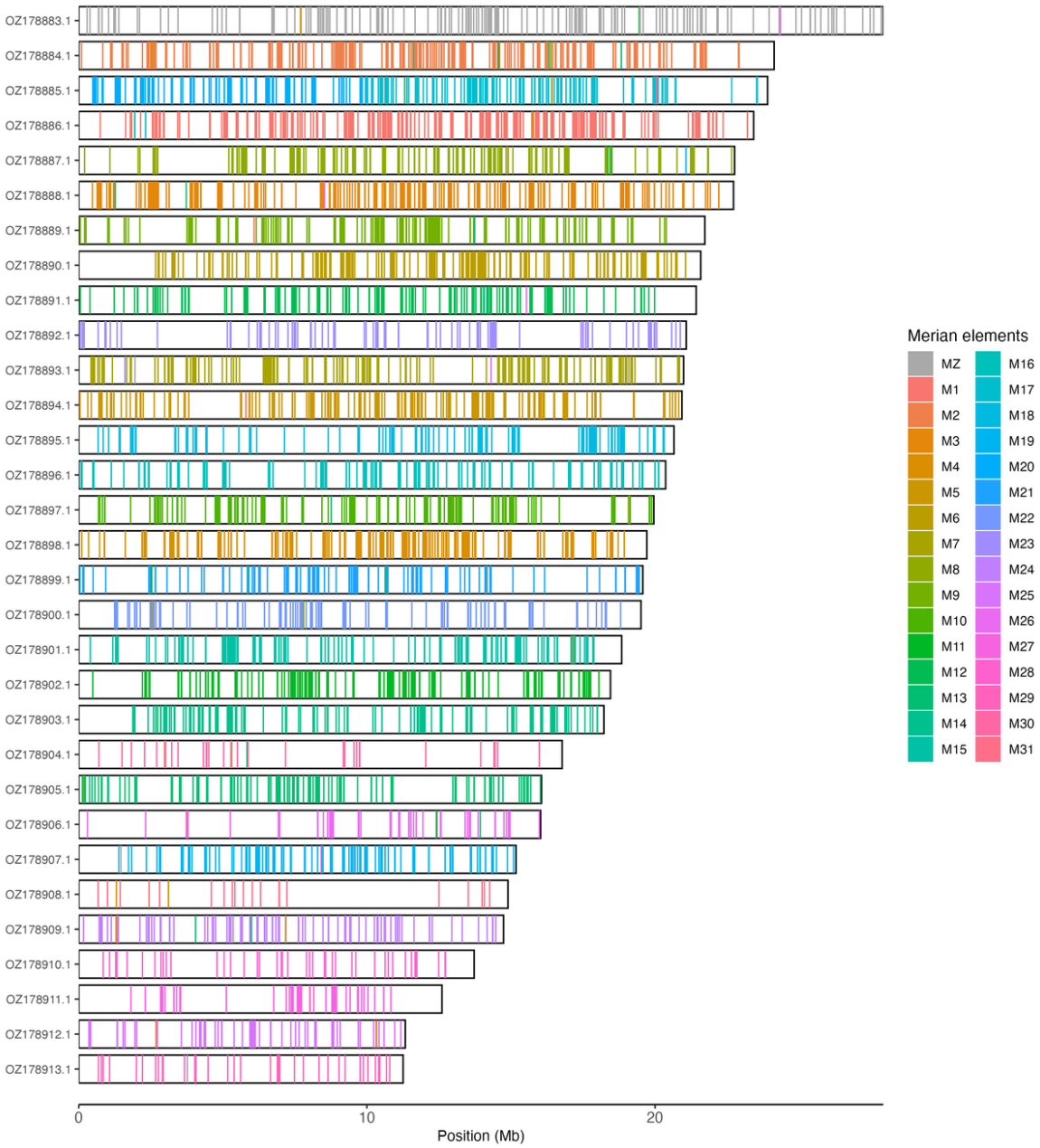

Figure 3. Merian elements painted across chromosomes in the ilMelPart1.hap1.1 assembly of *Melitaea parthenoides*. Chromosomes are drawn to scale, with the positions of orthologues shown as coloured bars. Each orthologue is coloured by the Merian element that it belongs to. All orthologues which could be assigned to Merian elements are shown.

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.Q63**, 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

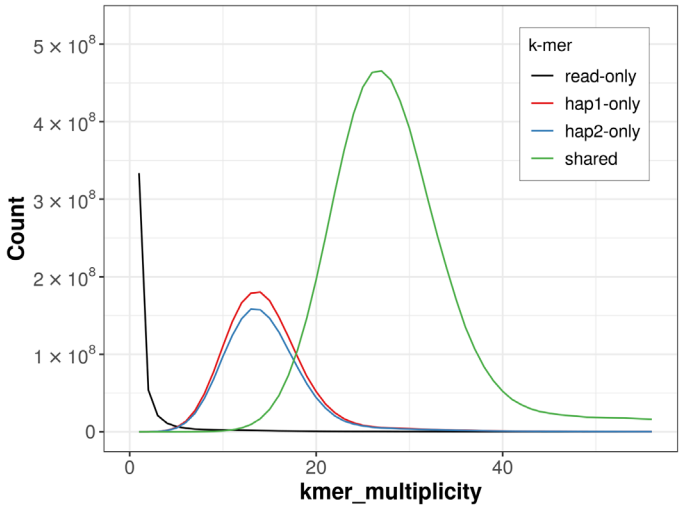

Figure 4. Evaluation of *k*-mer completeness using MerquryFK. This plot illustrates the recovery of *k*-mers from the original read data in the final assemblies. The horizontal axis represents *k*-mer multiplicity, and the vertical axis shows the number of *k*-mers. The black curve represents *k*-mers 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.

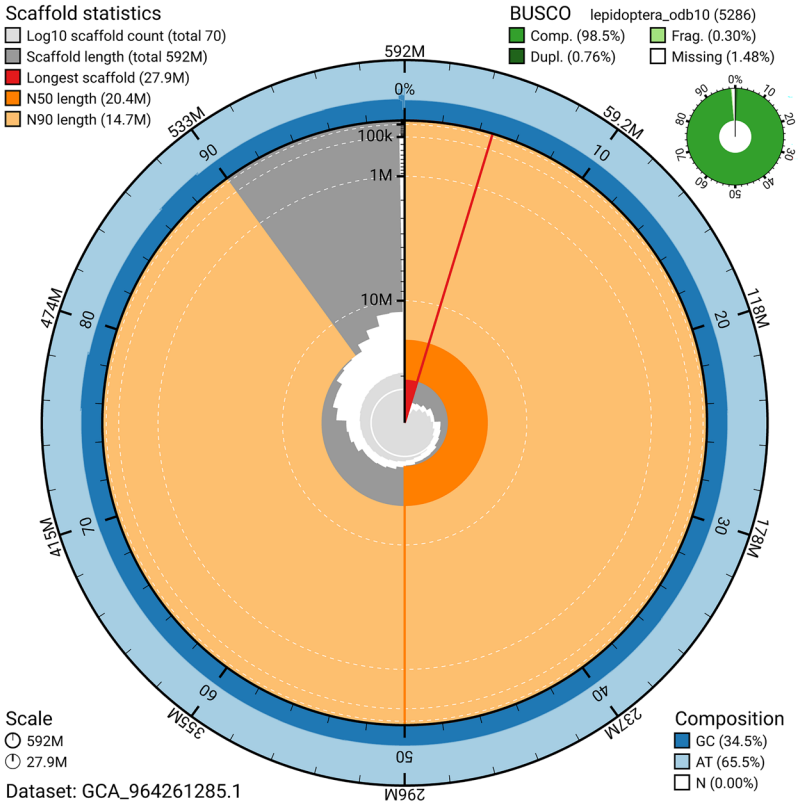

Figure 5. Assembly metrics for ilMelPart1.hap1.1. The BlobToolKit snail plot provides an overview of assembly metrics and BUSCO gene completeness. The circumference represents the length of the whole genome sequence, and the main plot is divided into 1,000 bins around the circumference. The outermost 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.

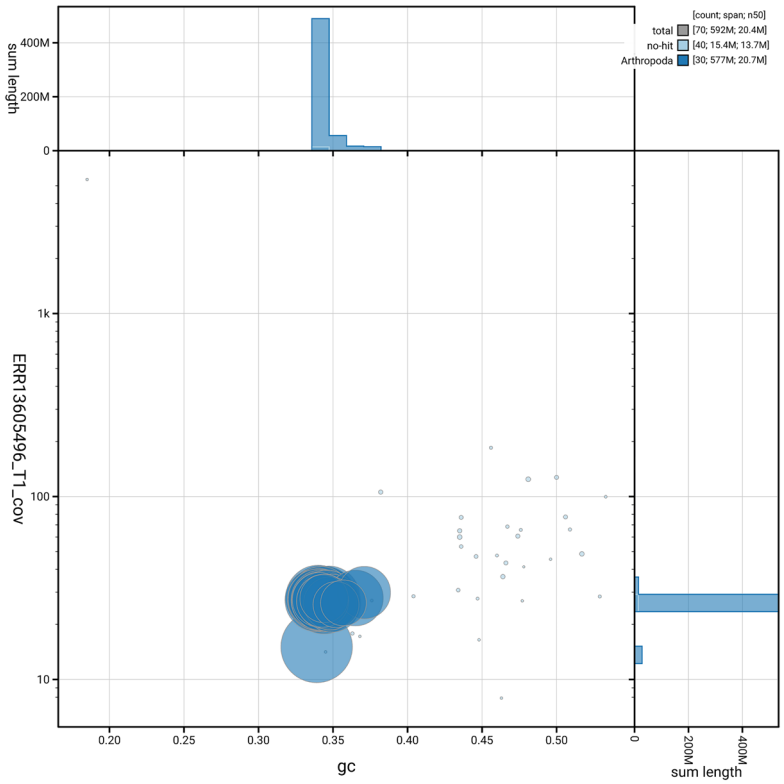

Figure 6. BlobToolKit GC-coverage plot for ilMelPart1.hap1.1. Blob plot showing sequence coverage (vertical axis) and GC content (horizontal axis). The circles represent scaffolds, with the size proportional to scaffold length and the colour representing phylum membership. The histograms along the axes 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 *Melitaea parthenoides* assembly.

Measure	Value	Benchmark
EBP summary (haplotype 1)	6.7.Q63	6.C.Q40
Contig N50 length	6.74 Mb	≥ 1 Mb
Scaffold N50 length	20.37 Mb	= chromosome N50
Consensus quality (QV)	Haplotype 1: 63.7; haplotype 2: 63.7; combined: 63.7	≥ 40
k-mer completeness	Haplotype 1: 77.06%; Haplotype 2: 73.81%; combined: 99.31%	≥ 95%
BUSCO	C:98.5%[S:97.8%,D:0.8%], F:0.3%,M:1.2%,n:5 286	S > 90%; D < 5%
Percentage of assembly assigned to chromosomes	99.72%	≥ 90%

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 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: Mellicta parthenoides (meadow fritillary). Accession number PRJEB79166. The genome sequence is released openly for reuse. The *Melitaea parthenoides* 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
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

Software	Version	Source
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

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

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

Dapporto L, Menchetti M, Vodă R, *et al.*: **The atlas of mitochondrial genetic diversity for Western Palearctic Butterflies.** *Glob Ecol Biogeogr.* 2022; **31**(11): 2184–2190.
Publisher Full Text

Ebert G, Rennwald E: **Die schmetterlinge baden-württembergs.** Band 1: Tagfalter I. Verlag Eugen Ulmer, 1991.
Reference Source

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

Gómez Bustillo MR, Fernández-Rubio F: **Mariposas de la península Ibérica. Rhopaloceros. 2 tomos.** Servicio de Publicaciones del Ministerio de Agricultura, Madrid, 1974.
Reference Source

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

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

Huemer P: **Die schmetterlinge Österreichs (Lepidoptera). Systematische und faunistische checkliste.** Tiroler Landesmuseen-Betriebsgesellschaft m. b. H., Innsbruck, Austria, 2013.
Reference Source

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

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

Kudrna O, Pennerstorfer J, Lux K: **Distribution atlas of European butterflies and skippers.** Wissenschaftlicher Verlag PEKS, 2015.
Reference Source

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

Lafranchis T, Jutzeler D, Guillosson JY, *et al.*: **La vie des papillons. Écologie, biologie et comportement des Rhopalocères de France.** Diatheo, 2015.
Reference Source

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

Ligue Suisse pour la Protection de la Nature: **Les papillons de jour et leurs biotopes. Espèces, Dangers qui les menacent, protection. Volume 1.** Fotorotar AG, 1987.
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

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

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

Reinhardt R, Bolz R: **Rote liste und gesamtartenliste der tagfalter (Rhopalocera) (Lepidoptera: Papilionoidea et Hesperioidea) Deutschlands.** In: M. Binot-Hafke, S. Balzer, N. Becker, *et al.* (eds), *Rote Liste gefährdeter Tiere, Pflanzen und Pilze Deutschlands, Band 3: Wirbellose Tiere (Teil 1), Vol. 70.* Landwirtschaftsverlag, Münster, 2011; 167–94.

Reinhardt R, Harpke A, Caspari S, *et al.*: **Verbreitungsatlas der Tagfalter und Widderchen Deutschlands.** Eugen Ulmer KG, Stuttgart (Hohenheim), 2020.
Reference Source

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**

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**

van Oorschot H, Coutsis JG: **The genus Melitaea Fabricius, 1807 (Lepidoptera: Nymphalidae, Nymphalinae). Taxonomy and systematics with special reference to the male genitalia.** Tshikolovets Publications, 2014.
Reference Source

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

Vila R, Pellecchia M, Pesce GB: **Farfalle d'Italia.** IBC Emilia-Romagna. Editrice Compositori, Bologna, 2009.
Reference Source

Wermeille E, Chittaro Y, Gonseth Y: **Liste rouge papillons diurnes et Zygènes. Espèces menacées en suisse, état 2012.** Office fédéral de l'environnement, Berne, et Centre Suisse de Cartographie de la Faune, Neuchâtel, 2014.
Reference Source

Wright CJ, Stevens L, Mackintosh A, *et al.*: **Comparative genomics reveals the dynamics of chromosome evolution in Lepidoptera.** *Nat Ecol Evol.* 2024; **8**(4): 777–790.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

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 31 October 2025

https://doi.org/10.21956/wellcomeopenres.27234.r134033

✔ **Arun Arumugaperumal**
Department of Biotechnology, Rajalakshmi Engineering College, Thandalam 602105, Chennai, Tamilnadu, India

The authors have presented the genome sequence of *Melitaea parthenoides*, also known as the Meadow Fritillary. The assembly reported is of size 592.15 Mb and the mitogenome is of size 15.17 kb. The authors have sequenced a female specimen, and they have captured the Z chromosome. They have commented that this butterfly follows the ZO sex system. It is interesting to note that some members of the same genus have ZW sex system [(Chittaro Y, et al., 2025) - ref 1] and others have ZO system [(Wymann H, et al., 2025) - ref 2] [(Chittaro Y, et al., 2025) - ref 3]. The photograph provided in Figure 1 will be very helpful to the readers. The sequencing procedure has been performed according to the standard tree of life project protocols. A high-quality genome sequence has been reported as evident from N50 values. BUSCO completeness value of 98.5% indicates that the genome is a near-complete one with respect to conserved gene sequences present in the database. The GC blob plot also shows that there is no contaminating sequence from any other phylum.

The EBP metric has been mentioned as 6.C.Q63 in the text. But in Table 4, it is mentioned as 6.7.Q63. Kindly correct this discrepancy. After the correction, the article can be indexed.

References

1. Chittaro Y, Lucek K, Wright C, Meier J, et al.: The genome sequence of the Grisons Fritillary, Melitaea varia Herrich-Schaeffer, 1851 (Lepidoptera: Nymphalidae). *Wellcome Open Research*. 2025; **10**. Publisher Full Text
2. Wymann H, Kollross J, Lucek K, Wright C, et al.: The genome sequence of the False Heath Fritillary, Melitaea diamina (Lang, 1789) (Lepidoptera: Nymphalidae). *Wellcome Open Research*. 2025; **10**. Publisher Full Text
3. Chittaro Y, Sanchez A, Lucek K, Wright C, et al.: The genome sequence of the Nickerl's Fritillary, Melitaea aurelia Nickerl, 1850 (Lepidoptera: Nymphalidae). *Wellcome Open Research*. 2025; **10**. Publisher Full Text

Is the rationale for creating the dataset(s) clearly described?

Yes

Are the protocols appropriate and is the work technically sound?
Yes

Are sufficient details of methods and materials provided to allow replication by others?
Yes

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

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Bioinformatics; Genomics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 10 September 2025

https://doi.org/10.21956/wellcomeopenres.27234.r130207

Evan J Kipp
University of Minnesota Twin Cities Institute on the Environment, Saint Paul, Minnesota, USA

In this data note, Chittaro and colleagues describe the sequencing and assembly of the genome of *Melitaea parthenoides* (Lepidoptera: Nymphalidae), the European Meadow Fritillary. This effort represents the first genome assembly for this species and is one of a growing number of published Fritillary genomes. This high-quality genome resource will contribute to future comparative genomic and evolutionary studies across Lepidoptera and other insects.

Overall, this report is clearly written, and the methods are well described. The authors provide valuable background on the natural history of this species and highlight some threats that have led to recent population declines over certain portions of its range. Using a combination of HiFi reads and Hi-C short-reads, two haplotype assemblies were produced, both with relatively high contiguity. The primary haplotype was used to generate 31 chromosome-level scaffolds, which included the Z sex chromosome and 30 autosomes. All detailed assessments of quality metrics, including BUSCO analysis, indicate that this genome assembly is of high quality and meets the standards outlined by the Earth Biogenome Project.

Minor points:

The authors report k-mer completeness findings for both haplotypes 1 and 2, however, BUSCO

analysis was only reported for the primary haplotype. Was haplotype 2 also assessed using BUSCO against the lepidoptera_odb10 database and, if so, what were the resulting values?

I do think that some additional details are warranted surrounding the description of Merian elements in this genome. In the Methods, the authors note that each complete BUSCO was assigned to an ancestral Merian element using a 'reference database' but do not detail how this was performed or what database was used. If this was performed using publicly available tools and databases, these should be included. Additionally, for myself and other readers who may not be deeply familiar with lepidopteran genomics, it would be helpful to include an additional sentence or two describing these findings. Are the Merian element patterns presented in Figure 3 and Table 3 consistent with other Fritillary genomes or are there any noteworthy differences that may be worthy of future investigation?

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: phylogenetics, genomics

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.