

PREPRINT

Author-formatted, not peer-reviewed document posted on 16/10/2025

DOI: <https://doi.org/10.3897/arphapreprints.e175020>

**The genome sequence of the Common
Brassy Ringlet, *Erebia cassioides* (Reiner
& Hohenwarth, 1792) (Lepidoptera,
Nymphalidae)**

 Camille Cornet,  Kay Lucek

The genome sequence of the Common Brassy Ringlet, *Erebia cassioides* (Reiner & Hohenwarth, 1792) (Lepidoptera, Nymphalidae)

Camille Cornet[‡], Kay Lucek[‡]

[‡] University of Neuchâtel, Neuchâtel, Switzerland

Corresponding author: Camille Cornet (camille.cornet@unine.ch)

Reviewed

v 1

Academic editor: Robert Waterhouse

Abstract

We present a chromosome-level genome assembly from a female specimen of the Common Brassy Ringlet *Erebia cassioides* (Arthropoda, Insecta, Lepidoptera, Nymphalidae). The genome consists of a primary assembly of 546 Mb and an alternate assembly of 406 Mb. The primary assembly is scaffolded into 11 chromosomes, including the Z and the W sex chromosomes. The mitochondrial genome has also been assembled, with a length of 15.19 kb.

Keywords

Erebia cassioides, genome assembly, European Reference Genome Atlas, Biodiversity Genomics Europe, Satyrinae, Common Brassy Ringlet

Introduction

The Common Brassy Ringlet *Erebia cassioides* Reiner & Hohenwarth, 1792 (Lepidoptera, Nymphalidae) has a widespread, yet localised and geographically isolated distribution across different mountain ranges in Europe, ranging from the Cantabrian Mountains in the West to the Pyrenees, the French Massif Central, the Alps over the Balkans to the Carpathian Mountains in the East (Kudrna et al. 2011). Many of these isolated occurrences have been described as distinct subspecies (Albre et al. 2008, Schmitt et al. 2016), but their evolutionary relationship and broader taxonomic context has not been resolved. For instance, *E. cassioides* has previously been proposed as the sister species of the Swiss brassy ringlet *E. tyndarus* based on few genetic markers (Peña et al. 2015, Gratton et al. 2015), but recent genomic data suggests that it is only a closely-related sibling species (Augustijnen et al. 2024). In the Alps, *E. cassioides* occurs only in the eastern and western parts, being separated by *E. tyndarus* (Sonderegger

2005, Schmitt et al. 2016), which is ecologically very similar and likely outcompetes *E. cassioides* (Lucek et al. 2020, Augustijn et al. 2022, Klečková et al. 2023). The species occurs above the tree line on grassy slopes with stones and rocks at altitudes between 1600 and 2600 metres, where the primary host plant is *Festuca* spp. (Sonderegger 2005). In the Alps, the life cycle of *E. cassioides* has been reported to take one year, where individuals overwinter as larvae and adults fly in July and August (Sonderegger 2005).

The genome of *Erebia cassioides* provides the opportunity to study different evolutionary mechanisms as *E. cassioides* belongs to the so-called *tyndarus* clade of *Erebia*, which has a very high variation in chromosome numbers (Albre et al. 2008, Lucek 2018). The variation in chromosome number is associated with an increased rate of speciation (Augustijn et al. 2024). This genome, therefore, enables insights into the genomic architecture underlying chromosomal speciation. Moreover, *E. cassioides* forms stable and very narrow contact zones with its sibling species *E. tyndarus*, where gene flow is restricted to a few first-generation hybrids (Gratton et al. 2015, Lucek et al. 2020, Augustijn and Lucek 2024), representing an advanced stage of speciation. Different potential barriers have been suggested to underlie reproductive isolation between the two species, including wing shape and genital morphology (Sonderegger 2005, Lucek et al. 2020, Augustijn et al. 2022), endosymbionts (Lucek et al. 2020, Augustijn and Lucek 2024) and cuticular hydrocarbons (Kleckova et al. 2025). However, whether chromosomal rearrangements exist between *E. cassioides* and *E. tyndarus* that could act as barriers to gene flow is unclear (Augustijn and Lucek 2024).

This reference genome contributes to the European Reference Genome Atlas (ERGA) goals of coordinating the production of high-quality genome sequences that represent the eukaryotic biodiversity in Europe (Mazzoni et al. 2023).

Materials & Methods

The genome assembly strategy consisted of assembly with PacBio HiFi reads and scaffolding with Hi-C data from the same female individual.

Sample and Sampling Information

One adult *Erebia cassioides* female was collected by Kay Lucek on the 10 August 2022 by hand-netting near Grindelwald, Switzerland (46.654582°N, 8.025222°E). A female, the heterogametic sex in Lepidoptera, was sampled to also assemble the *W* chromosome. Samples were kept alive until flash-freezing and storage at -80°C. Species identification was conducted visually in the field.

Vouchering Information

The wings of the specimen were deposited at the Muséum d'Histoire Naturelle de Neuchâtel, Neuchâtel, Switzerland (voucher ID MHNN-65-8969). Wing pictures of another female from the same population are presented in Fig. 1.

Data Availability

The genome sequence of *Erebia cassioides* was produced under the European Reference Genome Atlas Switzerland (ERGA-CH) framework. The umbrella BioProject for this species is PRJEB98373. The primary genome assembly is available on the European Nucleotide Atlas (ENA) under accession number GCA_976986335 and the alternate assembly under accession GCA_976984905. The metadata and associated raw sequencing reads for the samples used for genome assembly and scaffolding (assigned Tree of Life ID iEreCass2) are available under BioSample accession SAMEA120234452. The code used to produce the assembly is available on GitHub (<https://github.com/camille-cornet/ErebiaGenomeAssembly>) and archived on Zenodo (Cornet and Lucek 2025).

Genetic Information

The genome of *Erebia cassioides* is diploid with 10 chromosome pairs as estimated by karyotyping studies ($2n = 20$; De Lesse (1960), Robinson (1971)). Before sequencing, genome size estimates were approximately 500 Mb, based on two pre-existing chromosome-level *Erebia* assemblies (K. Lohse et al. 2022, O. Lohse et al. 2022) and a contig level assembly of another *E. cassioides* female of the same population (Augustijnen and Lucek 2024). The sex chromosome system was expected to be ZZ/ZW (female heterogamy) as for most butterflies (Traut et al. 2007).

DNA/RNA Processing

For genome assembly, high-molecular-weight (HMW) DNA was extracted from half a thorax using the Qiagen MagAttract HMW DNA Kit (Qiagen, Hombrechtikon, Switzerland), following the manufacturer's instructions.

For assembly scaffolding, Hi-C data was generated using the head of the same individual that was used for PacBio sequencing. After grinding in liquid nitrogen, prior to library preparation, crosslinking was performed using the Arima High Coverage HiC kit (Arima Genomics, San Diego, CA, USA), following the manufacturer's instructions.

Library Preparation & Sequencing

For genome assembly, PacBio sequencing was performed by the Genomics Technologies Facility (GTF, Lausanne, Switzerland) on the PacBio Sequel IIe, generating 1.63 million HiFi reads or 16.4 Gb of data. Average HiFi read length was 10 kb.

For assembly scaffolding, Hi-C library preparation following crosslinking was performed using the High Coverage HiC kit (Arima Genomics, San Diego, CA, USA), following the manufacturer's instructions. Sequencing on an Illumina NovaSeq 6000 (300 cycles, 150

bp long paired-end reads) was performed by the Next Generation Sequencing Platform (NGS Platform, Bern, Switzerland), resulting in 155 million read pairs.

Genome Assembly Methods

After quality control of the HiFi reads using NanoPlot v.1.32.1 (De Coster and Rademakers 2023), reads were assembled into contigs using hifiasm v.0.16.0-r369 (Cheng et al. 2021) in primary assembly mode, with the settings -D 10 -N 200 to improve assembly of repetitive regions. Haplotypic duplication was then removed from the primary assembly and an alternate assembly was produced, using purgedups v.1.2.5 (Guan et al. 2020) with default settings.

For scaffolding of the draft primary assembly into chromosomes, the quality of the Hi-C reads was checked using FastQC v.0.12.1 (<https://github.com/s-andrews/FastQC>) and five bases were trimmed off the 5' end of each read to reach a better mapping rate. Poly-G tails were trimmed while also applying a minimal length filter of 120 bp and a PHRED quality filter of 30 using fastp v.0.23.4 (Chen 2023). The reads were mapped to the primary assembly using the Arima mapping pipeline v.03 (https://github.com/ArimaGenomics/mapping_pipeline) with a minimum mapping quality of 1. The assembly was scaffolded using YaHS v.1.2 (Zhou et al. 2023) with default settings and manual curation was performed using Juicebox v.1.11.08 (Durand et al. 2016). The scaffolded assembly was then decontaminated using BlobToolKit v.4.3 (Challis et al. 2020) and tiara v.1.0.3 (Karlicki et al. 2022) to remove any bacterial or mitochondrial scaffold. The mitogenome was assembled from the raw HiFi reads using MitoHiFi v.3.0.0 (Uliano-Silva et al. 2023). The HiFi reads were then mapped back to the scaffolded assembly to close gaps with tsgapcloser v.1.1.1 (Xu et al. 2020). Scaffolds corresponding to haplotypic duplication or unplaced repetitive regions were then manually removed. The quality of the final primary assembly was assessed to confirm that it reaches the Earth Biogenome Project (EBP) standards (Rhie et al. 2021), using BUSCO v.5.7.1 (Manni et al. 2021) for functional completeness with the Lepidoptera OrthoDB v.10 dataset (Waterhouse et al. 2013), gfastats v.1.3.6 (Formenti et al. 2022) for contiguity and Merqury v.1.3 (Rhie et al. 2020) for *k*-mer completeness, base pair quality and false duplication rates.

The Z and W sex chromosomes were identified based on the coverage of the HiFi reads (generated from a female individual) and Illumina reads from a male individual from the same population (Augustijn et al. 2024) mapped back to the primary assembly using minimap2 v.2.21 (Li 2018) and bwa mem v.0.7.17 (Li 2013), respectively, followed by samtools coverage v.1.14 (Danecek et al. 2021) with a maximum depth of 100 and a minimum read and mapping quality of 30. The correspondance between the chromosomes of *E. cassiodes* and Merian elements, i.e. ancestral linkage groups in Lepidoptera (Wright et al. 2024), was assessed using the lep_buscoPainter tool v.1.0.0 (https://github.com/charlottewright/lep_buscoPainter), based on the BUSCO output.

Results

Here, we report the assembly statistics for the chromosome-level primary assembly of *Erebia cassioides*, which meet the EBP standards (Rhie et al. 2021).

Genome Assembly

The primary assembly of *Erebia cassioides* has a total length of 546,119,514 bp, distributed in nine autosomes, one Z chromosome, one W chromosome and the mitogenome (Table 1, Figs 2, 3). The GC content is 36.8%. The draft assembly, comprising 23 contigs, had a contig N50 of 45.1 Mb and L50 of 5. The primary assembly has a scaffold N50 of 55.7 Mb, L50 of five and twelve gaps (total gap length: 1,200 bp). The gene content completeness analysis of the assembly resulted in a BUSCO completeness score of 98.8% (98.2% single-copy and 0.6% duplicated, Fig. 2). The alternate assembly has a total length of 405,746,844 bp distributed in 2046 scaffolds, with a contig N50 of 515 kb. Its BUSCO completeness score is 79.2% (78.4% single-copy and 0.8% duplicated). Combining the primary and alternate assemblies, the *k*-mer completeness score is 97.9%, false duplication rate 1.7% and the assembly has a base quality value of 60.1 (Fig. 4). The chromosome-level scaffolds, confirmed by Hi-C data (Fig. 3), are named according to size (Table 1). Chromosome painting with Merian elements illustrates the distribution of orthologues along chromosomes and highlights patterns of chromosomal evolution relative to Lepidopteran ancestral linkage groups (Table 1, Fig. 5).

Acknowledgements

This work was supported by the Swiss National Science Foundation (SNSF) Grant ID 202869. We thank the Next Generation Sequencing Platform in Bern for their outstanding support in conducting Hi-C sequencing. We are grateful to the Wellcome Sanger Institute, in particular the Tree of Life Programme, for granting us access to their computational infrastructure to perform the genome assembly and quality control. We further thank the ERGA Community, in particular ERGA-CH, for the useful community resources and guidelines that helped us produce and publish this reference genome.

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Albre J, Gers C, Legal L (2008) Molecular phylogeny of the *Erebia tyndarus* (Lepidoptera, Rhopalocera, Nymphalidae, Satyrinae) species group combining CoxII and ND5 mitochondrial genes: A case study of a recent radiation. *Molecular Phylogenetics and Evolution* 47 (1): 196-210. <https://doi.org/10.1016/j.ympev.2008.01.009>
- Augustijn H, Patsiou T, Lucek K (2022) Secondary contact rather than coexistence - *Erebia* butterflies in the Alps. *Evolution* 76 (11): 2669-2686. <https://doi.org/10.1111/evo.14615>
- Augustijn H, Lucek K (2024) Beyond gene flow: (non)-parallelism of secondary contact in a pair of highly differentiated sibling species. *Molecular Ecology* 33 (17). <https://doi.org/10.1111/mec.17488>
- Augustijn H, Bätischer L, Cesanek M, Chkhartishvili T, Dincă V, Iankoshvili G, Ogawa K, Vila R, Klopstein S, de Vos J, Lucek K (2024) A macroevolutionary role for chromosomal fusion and fission in *Erebia* butterflies. *Science Advances* 10 (16). <https://doi.org/10.1126/sciadv.adl0989>
- Challis R, Richards E, Rajan J, Cochrane G, Blaxter M (2020) BlobToolKit – Interactive Quality Assessment of Genome Assemblies. *G3 Genes|Genomes|Genetics* 10 (4): 1361-1374. <https://doi.org/10.1534/g3.119.400908>
- Cheng H, Concepcion G, Feng X, Zhang H, Li H (2021) Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nature Methods* 18 (2): 170-175. <https://doi.org/10.1038/s41592-020-01056-5>
- Chen S (2023) Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *iMeta* 2 (2). <https://doi.org/10.1002/imt2.107>
- Cornet C, Lucek K (2025) *Erebia* genome assembly code. 1.2. Zenodo. Release date: 2025-10-10. URL: <https://doi.org/10.5281/zenodo.17314723>
- Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H (2021) Twelve years of SAMtools and BCFtools. *GigaScience* 10 (2). <https://doi.org/10.1093/gigascience/giab008>
- De Coster W, Rademakers R (2023) NanoPack2: population-scale evaluation of long-read sequencing data. *Bioinformatics* 39 (5). <https://doi.org/10.1093/bioinformatics/btad311>
- De Lesse H (1960) Spéciation et variation chromosomique chez les Lépidoptères Rhopalocères. *Ann. Soc. Nat., Zool.* 1-223. URL: <https://cir.nii.ac.jp/crid/1571698600201581056>
- Durand N, Robinson J, Shamim M, Machol I, Mesirov J, Lander E, Aiden EL (2016) Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. *Cell systems* 3 (1): 99-101. <https://doi.org/10.1016/j.cels.2015.07.012>
- Formenti G, Abueg L, Brajuka A, Brajuka N, Gallardo-Alba C, Giani A, Fedrigo O, Jarvis ED (2022) Gfastats: conversion, evaluation and manipulation of genome sequences using assembly graphs. *Bioinformatics* 38 (17): 4214-4216. <https://doi.org/10.1093/bioinformatics/btac460>
- Gratton P, Trucchi E, Trasatti A, Riccarducci G, Marta S, Allegrucci G, Cesaroni D, Sbordoni V (2015) Testing classical species properties with contemporary data: how 'bad species' in the brassy ringlets (*Erebia tyndarus* complex, Lepidoptera) turned good. *Systematic Biology* 65 (2): 292-303. <https://doi.org/10.1093/sysbio/syv087>

- Guan D, McCarthy SA, Wood J, Howe K, Wang Y, Durbin R (2020) Identifying and removing haplotypic duplication in primary genome assemblies. *Bioinformatics* 36 (9): 2896-2898. <https://doi.org/10.1093/bioinformatics/btaa025>
- Karlicki M, Antonowicz S, Karnkowska A (2022) Tiarā: deep learning-based classification system for eukaryotic sequences. *Bioinformatics* 38 (2): 344-350. <https://doi.org/10.1093/bioinformatics/btab672>
- Klecková I, Clancy MV, Cornet C, Alvarez N, Matos-Maravá P, Lucek K (2025) Co-Occurring Sister Taxa of Mountain Butterflies Exhibit Distinct Cuticular Hydrocarbon Profiles. *Ecology and Evolution* 15 (9): 72027. <https://doi.org/10.1002/ece3.72027>
- Klečková I, Klečka J, Fric ZF, Česánek M, Dutoit L, Pellissier L, Matos-Maravá P (2023) Climatic Niche Conservatism and Ecological Diversification in the Holarctic Cold-Dwelling Butterfly Genus *Erebia*. *Insect Systematics and Diversity* 7 (1). <https://doi.org/10.1093/isd/ixad002>
- K. Lohse K, Hayward A, Laetsch D, Vila R, Lucek K, 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 (2022) The genome sequence of the Arran brown, *Erebia ligea* (Linnaeus, 1758). Wellcome Open Research 7 <https://doi.org/10.12688/wellcomeopenres.18115.1>
- Kudrna O, Harpke A, Lux K, Pennerstorfer J, Schweiger O, Settele J, Wiemers M (2011) Distribution atlas of butterflies in Europe. Gesellschaft für Schmetterlingsschutz, Halle.
- Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv. arXiv:1303.3997. <https://doi.org/10.48550/arXiv.1303.3997>
- Li H (2018) Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34 (18): 3094-3100. <https://doi.org/10.1093/bioinformatics/bty191>
- Lucek K (2018) Evolutionary Mechanisms of Varying Chromosome Numbers in the Radiation of *Erebia* Butterflies. *Genes* 9 (3). <https://doi.org/10.3390/genes9030166>
- Lucek K, Butlin RK, Patsiou T (2020) Secondary contact zones of closely-related *Erebia* butterflies overlap with narrow phenotypic and parasitic clines. *Journal of evolutionary biology* 33 (9): 1152-1163. <https://doi.org/10.1111/jeb.13669>
- Manni M, Berkeley MR, Seppely M, Simão FA, Zdobnov EM (2021) BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. *Molecular Biology and Evolution* 38 (10): 4647-4654. <https://doi.org/10.1093/molbev/msab199>
- Mazzoni C, Ciofi C, Waterhouse R (2023) Biodiversity: an atlas of European reference genomes. *Nature* 619 (7969): 252-252. <https://doi.org/10.1038/d41586-023-02229-w>
- O. Lohse O, Lohse K, Augustijnen H, Lucek K (2022) The genome sequence of the scotch argus butterfly, *Erebia aethiops* (Esper, 1777). Wellcome Open Research 7 <https://doi.org/10.12688/wellcomeopenres.17927.1>
- Peña C, Witthauer H, Klečková I, Fric Z, Wahlberg N (2015) Adaptive radiations in butterflies: evolutionary history of the genus *Erebia* (Nymphalidae: Satyrinae). *Biological Journal of the Linnean Society* 116 (2): 449-467. <https://doi.org/10.1111/bij.12597>
- Rhie A, Walenz B, Koren S, Phillippy A (2020) Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies. *Genome Biology* 21 (1). <https://doi.org/10.1186/s13059-020-02134-9>
- Rhie A, McCarthy S, Fedrigo O, Damas J, Formenti G, Koren S, Uliano-Silva M, Chow W, Fungtammasan A, Kim J, Lee C, Ko BJ, Chaisson M, Gedman G, Cantin L, Thibaud-Nissen F, Haggerty L, Bista I, Smith M, Haase B, Mountcastle J, Winkler S, Paez S,

- Howard J, Vernes S, Lama T, Grutzner F, Warren W, Balakrishnan C, Burt D, George J, Biegler M, Iorns D, Digby A, Eason D, Robertson B, Edwards T, Wilkinson M, Turner G, Meyer A, Kautt A, Franchini P, Detrich HW, Svardal H, Wagner M, Naylor GP, Pippel M, Malinsky M, Mooney M, Simbirsky M, Hannigan B, Pesout T, Houck M, Misuraca A, Kingan S, Hall R, Kronenberg Z, Sović I, Dunn C, Ning Z, Hastie A, Lee J, Selvaraj S, Green R, Putnam N, Gut I, Ghurye J, Garrison E, Sims Y, Collins J, Pelan S, Torrance J, Tracey A, Wood J, Dagnew R, Guan D, London S, Clayton D, Mello C, Friedrich S, Lovell P, Osipova E, Al-Ajli F, Secomandi S, Kim H, Theofanopoulou C, Hiller M, Zhou Y, Harris R, Makova K, Medvedev P, Hoffman J, Masterson P, Clark K, Martin F, Howe K, Flicek P, Walenz B, Kwak W, Clawson H, Diekhans M, Nassar L, Paten B, Kraus RS, Crawford A, Gilbert MTP, Zhang G, Venkatesh B, Murphy R, Koepfli K, Shapiro B, Johnson W, Di Palma F, Marques-Bonet T, Teeling E, Warnow T, Graves JM, Ryder O, Haussler D, O'Brien S, Korlach J, Lewin H, Howe K, Myers E, Durbin R, Phillippy A, Jarvis E (2021) Towards complete and error-free genome assemblies of all vertebrate species. *Nature* 592 (7856): 737-746. <https://doi.org/10.1038/s41586-021-03451-0>
- Robinson R (1971) *Lepidoptera Genetics*. Pergamon Press <https://doi.org/10.1016/B978-0-08-006659-2.50005-5>
 - Schmitt T, Louy D, Zimmermann E, Habel JC (2016) Species radiation in the Alps: multiple range shifts caused diversification in Ringlet butterflies in the European high mountains. *Organisms Diversity & Evolution* 16: 791-808. <https://doi.org/10.1007/s13127-016-0282-6>
 - Sonderegger P (2005) Die Erebien der Schweiz (Lepidoptera: Satyrinae, Genus *Erebia*). Peter Sonderegger, Brügg bei Biel. <https://doi.org/10.5281/zenodo.7409160>
 - Traut W, Sahara K, Marec F (2007) Sex Chromosomes and Sex Determination in Lepidoptera. *Sexual Development* 1 (6): 332-346. <https://doi.org/10.1159/000111765>
 - Uliano-Silva M, Ferreira JN, Krashennikova K, Blaxter M, Mieszkowska N, Hall N, Holland P, Durbin R, Richards T, Kersey P, Hollingsworth P, Wilson W, Twyford A, Gaya E, Lawniczak M, Lewis O, Broad G, Martin F, Hart M, Barnes I, Formenti G, Abueg L, Torrance J, Myers E, Durbin R, Blaxter M, McCarthy S, Darwin Tree of Life Consortium (2023) MitoHiFi: a python pipeline for mitochondrial genome assembly from PacBio high fidelity reads. *BMC Bioinformatics* 24 (1). <https://doi.org/10.1186/s12859-023-05385-y>
 - Waterhouse R, Tegenfeldt F, Li J, Zdobnov E, Kriventseva E (2013) OrthoDB: a hierarchical catalog of animal, fungal and bacterial orthologs. *Nucleic Acids Research* 41 (D1). <https://doi.org/10.1093/nar/gks1116>
 - Wright C, Stevens L, Mackintosh A, Lawniczak M, Blaxter M (2024) Comparative genomics reveals the dynamics of chromosome evolution in Lepidoptera. *Nature Ecology & Evolution* 1-14. <https://doi.org/10.1038/s41559-024-02329-4>
 - Xu M, Guo L, Gu S, Wang O, Zhang R, Peters BA, Fan G, Liu X, Xu X, Deng L, Zhang Y (2020) TGS-GapCloser: A fast and accurate gap closer for large genomes with low coverage of error-prone long reads. *GigaScience* 9 (9). <https://doi.org/10.1093/gigascience/giaa094>
 - Zhou C, McCarthy SA, Durbin R (2023) YaHS: yet another Hi-C scaffolding tool. *Bioinformatics* 39 (1). <https://doi.org/10.1093/bioinformatics/btac808>

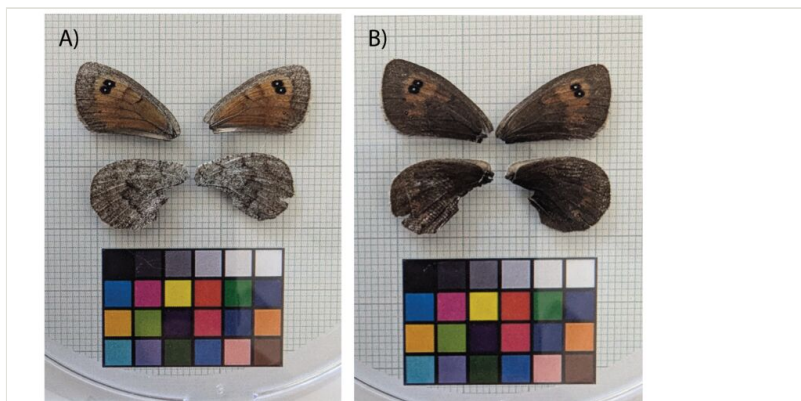


Figure 1.

Fore- and hindwings of a female *Erebia cassioides* individual from the same population as the one used for genome assembly and scaffolding (sampled in Grindelwald, Switzerland), with ventral (A) and dorsal (B) view. The wings are deposited at the Muséum d'Histoire Naturelle de Neuchâtel, Neuchâtel, Switzerland (voucher ID MHNN-65-8783).

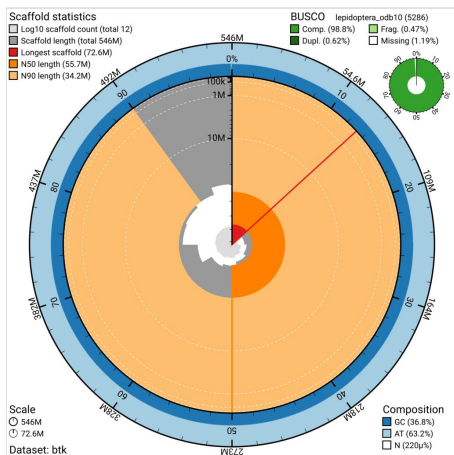


Figure 2.

Genome assembly metrics of *Erebia cassioides*. The main plot is divided into 1,000 size-ordered bins around the circumference with each bin representing 0.1% of the 546,119,514 bp assembly. The distribution of chromosome lengths is shown in dark grey with the plot radius scaled to the longest chromosome present in the assembly (72.6 Mb, shown in red). Orange and pale-orange arcs show the N50 and N90 chromosome lengths (55.7 and 34.2 Mb), respectively. The pale grey spiral shows the cumulative chromosome 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. The plot was generated using BlobToolKit, along with the decontamination step.

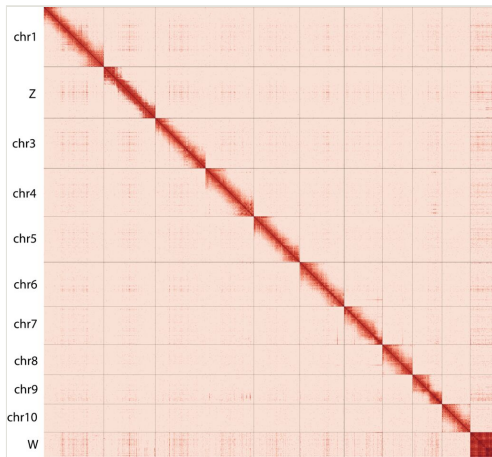


Figure 3.

Hi-C contact map of the *Erebia cassioides* genome assembly. Assembled chromosomes are shown in order of size. The plot was generated using PretextSnapshot v.0.0.5 (<https://github.com/sanger-to/PretextSnapshot>, with options -r 4000 -c 30).

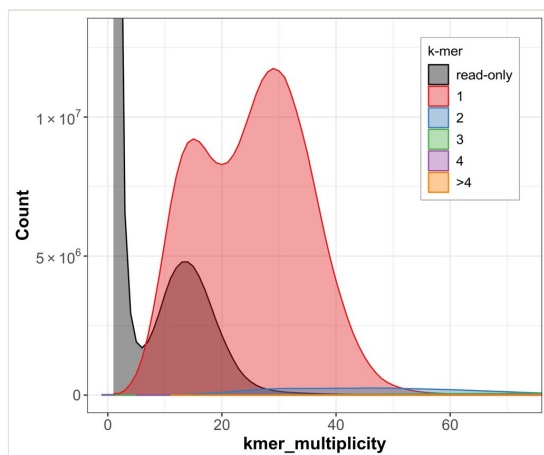


Figure 4.

Evaluation of k -mer completeness of the *Erebia cassioides* primary assembly using Merqury. This plot illustrates the recovery of k -mers from the original read data in the final primary assembly. 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 present in the primary assembly. The red curve corresponds to k -mers present in a single copy in the primary assembly, the left peak representing heterozygous k -mers and the right peak homozygous k -mers. The other curves represent k -mers present in multiple copies in the primary assembly.

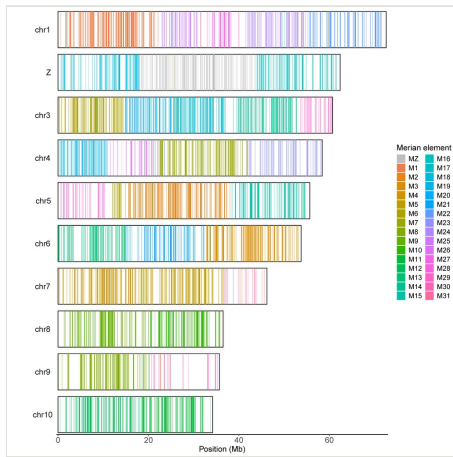


Figure 5.

Merian elements painted across chromosomes in the primary assembly of *Erebia cassioides*, using the `lep_buscoPainter` tool v.1.0.0 (https://github.com/charlottewright/lep_buscoPainter). Chromosomes are drawn to scale, with the positions of orthologues shown as coloured bars. Each orthologue is coloured by the corresponding Merian element. All orthologues which could be assigned to Merian elements are shown. The *W* chromosome is not shown as it does not correspond to any Merian element.

Table 1.
Chromosome-level scaffolds in the primary assembly of *Erebia cassioides*.

Chromosome	Length	GC%	Assigned Merian element
chr1	72.58 Mb	36.78	M1;M22;M24;M25;M27
Z	62.39 Mb	36.51	M16;M18;MZ
chr3	60.71 Mb	36.78	M6;M15;M17;M20;M29
chr4	58.42 Mb	36.75	M7;M19;M23;M26
chr5	55.68 Mb	36.76	M2;M6;M15;M28
chr6	53.76 Mb	36.75	M3;M13;M21
chr7	46.19 Mb	36.74	M4;M5;M31
chr8	36.50 Mb	36.77	M9;M10
chr9	35.71 Mb	36.99	M8;M30
chr10	34.19 Mb	36.94	M11;M12
W	29.97 Mb	37.71	
mitochondrion	15.19 kb	19.76	