DATA NOTE

The genome sequence of the meadow brown, *Maniola jurtina* (Linnaeus, 1758) [version 1; peer review: 2 approved, 1 approved with reservations]

Konrad Lohse, Jamie Weir, Darwin Tree of Life Barcoding collective, Wellcome Sanger Institute Tree of Life programme, Wellcome Sanger Institute Scientific Operations: DNA Pipelines collective, Tree of Life Core Informatics collective, Darwin Tree of Life Consortium

1 Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, UK

Abstract

We present a genome assembly from an individual female *Maniola jurtina* (the meadow brown; Arthropoda; Insecta; Lepidoptera; Nymphalidae). The genome sequence is 402 megabases in span. The complete assembly is scaffolded into 30 chromosomal pseudomolecules, with the W and Z sex chromosome assembled. Gene annotation of this assembly on Ensembl has identified 12,502 protein coding genes.

Keywords

Maniola jurtina, meadow brown, genome sequence, chromosomal

This article is included in the Tree of Life gateway.

Open Peer Review

Approval Status

| 1 | 2 | 3 |
|---|---|---|
| ? | ✔ | ✔ |

Version 1

05 Nov 2021

1. Julio Rozas, University of Barcelona, Barcelona, Spain

2. Violaine Llaurens, Université des Antilles Muséum National d'Histoire Naturelle, Paris, France

CNRS, Paris, France

3. Zdenek F. Fric, Biology Centre CAS, České Budějovice, Czech Republic

Any reports and responses or comments on the article can be found at the end of the article.
Corresponding author: Darwin Tree of Life Consortium (mark.blaxter@sanger.ac.uk)

Author roles: Lohse K: Investigation, Resources, Writing – Review & Editing; Weir J: Writing – Original Draft Preparation, Writing – Review & Editing;

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by Wellcome through core funding to the Wellcome Sanger Institute (206194) and the Darwin Tree of Life Discretionary Award (218328). KL is supported by a NERC fellowship (NE/L011522/1) and an ERC grant (ModelGenom Land 757648). JW is supported by an EastBio PhD studentship.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2021 Lohse K et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Lohse K, Weir J, Darwin Tree of Life Barcoding collective et al. The genome sequence of the meadow brown, Maniola jurtina (Linnaeus, 1758) [version 1; peer review: 2 approved, 1 approved with reservations] Wellcome Open Research 2021, 6:296 https://doi.org/10.12688/wellcomeopenres.17304.1

First published: 05 Nov 2021, 6:296 https://doi.org/10.12688/wellcomeopenres.17304.1
Species taxonomy
Eukaryota; Metazoa; Ecdysozoa; Arthropoda; Hexapoda; Insecta; Protostomia; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Papilionoidea; Nymphalidae; Satyrinae; Satyrini; Maniolina; Maniola; *Maniola jurtina* (Linnaeus, 1758) (NCBI: txid191418).

Introduction
The meadow brown *Maniola jurtina* is a common, Palearctic butterfly occurring throughout Europe, the Middle East, and North Africa ([gbif.org](https://gbif.org), 2021). Both widespread and often abundant, the species is associated with almost any grassy habitats ([South, 1906](https://academic.oup.com/jesz/asia/article/4/1/34/2035230)), reaching highest densities in areas where grazing or other pressures keep the sward at an intermediate height ([Maitland Emmet & Heath, 1989](https://doi.org/10.1080/00253158909381843)). Although *M. jurtina* is consistently univoltine, emergence occurs over a prolonged period in summer, which varies in length geographically and with habitat types ([Brakefield, 1987](https://www.tandfonline.com/doi/abs/10.1080/00918109.1987.11781637)). In many Mediterranean populations, females aestivate during the hottest months of the year ([Scali, 1971](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1437165/)). Eggs are laid singly or in small clusters, both on individual blades of grass or loose into a suitable tuft ([Maitland Emmet & Heath, 1989](https://doi.org/10.1080/00253158909381843)), preferentially on *Poa, Agrostis and Lolium*. Larvae overwinter, but do not undergo true diapause, and feed intermittently in warm spells. Pupae show considerable variation in colouration which is affected by light and temperature ([Brakefield, 1979](https://www.tandfonline.com/doi/abs/10.1080/00918109.1979.10671106)). The species exhibits a great deal of phenotypic variation both within and between populations ([Thomson, 1969](https://www.tandfonline.com/doi/abs/10.1080/00918109.1969.10671193)). Four sub-species are known from the British Isles – ssp. *splendida, insularis, iernes*, and *cassiteridum* – although the validity of these taxa is questionable ([Weir & Others, 2016](https://academic.oup.com/zoolinnean/article/8/1/149/5029485)), since they seem to be phenotypic extremes at opposing ends of clines ([Maitland Emmet & Heath, 1989](https://doi.org/10.1080/00253158909381843)). In their pioneering work in ecological genetics, Ford and Dowdeswell considered the evolutionary factors shaping variation in the spot patterning of the underside of the hindwings in *M. jurtina*, initially on the Isles of Scilly, then the British mainland (reviewed in ([Ford, 1964](https://www.tandfonline.com/doi/abs/10.1080/00166360.1964.10497199)) and ([Dowdeswell, 1981](https://www.tandfonline.com/doi/abs/10.1080/00918109.1981.10671240)). Several early studies ([Bigger, 1960](https://www.tandfonline.com/doi/abs/10.1080/00253158009525189); [Federley, 1938](https://www.tandfonline.com/doi/abs/10.1080/00918109.1938.10671140); [Lorković, 1941](https://www.tandfonline.com/doi/abs/10.1080/00918109.1941.10671187)), summarised in ([Robinson, 1971](https://www.tandfonline.com/doi/abs/10.1080/00918109.1971.10671280)), report a karyotype of 29 chromosomes. The genome size has been estimated as 367.3 Mb ([Mackintosh et al., 2019](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6785217/)). We note the publication of a *de novo* genome assembly of *M. jurtina* by ([Singh et al., 2020](https://academic.oup.com/zoolinnean/article/8/1/149/5029485)) and believe that the sequence described here, generated as part of the Darwin Tree of Life project, will further aid understanding of the biology of this butterfly.

Genome sequence report
The genome was sequenced from a female *M. jurtina* ([ilManJurt1; Figure 1A, B](https://biorxiv.org/content/10.1101/2020.03.26.993069v2.full.pdf)) collected from Aberlady Bay, East Lothian, Scotland, UK (latitude 56.019964, longitude -2.85808). Hi-C data were generated from another individual ([ilManJurt3; Figure 1E, F](https://biorxiv.org/content/10.1101/2020.03.26.993069v2.full.pdf)) collected from East Linton, East Lothian, Scotland, UK (latitude 55.977161, longitude -2.667545). A total of 76-fold coverage in Pacific Biosciences single-molecule long reads (N50 14 kb) and 88-fold coverage in 10X Genomics read clouds were generated. Primary assembly contigs were scaffolded with chromosome conformation Hi-C data. Manual assembly curation corrected 24 missing/misjoins and removed two haplotypic duplications, reducing the assembly size by 1.67% and scaffold number by 24.39%.

The final assembly has a total length of 402 Mb in 30 sequence scaffolds with a scaffold N50 of 15 Mb (Table 1). Of the assembly sequence, 100% was assigned to 30 chromosomal-level scaffolds, representing 28 autosomes (numbered by sequence length), and the W and Z sex chromosome (Figure 2–Figure 5; Table 2). The assembly has a BUSCO ([Simão et al., 2015](https://academic.oup.com/g�l/advance-article/doi/10.1093/g�l/rgy054/4836277)) completeness of 98.3% using the lepidoptera_odb9 reference set. While not fully phased, the assembly deposited is of one haplotype. Contigs corresponding to the second haplotype have also been deposited.

Gene annotation
The Ensembl gene annotation system ([Aken et al., 2016](https://academic.oup.com/g�l/article/7/1/26/4836272)) was used to generate annotation for the *Maniola jurtina* assembly ([GCA_905333055.1](https://ww2.ensembl.org/Maniola_jurtina/Info/Show?db=maniolina&taxid=191418&sp=Maniola%20jurtina)], see [https://rapid.ensembl.org/Maniola_jurtina_GCA_905333055.1](https://rapid.ensembl.org/Maniola_jurtina_GCA_905333055.1)]; Table 1). The annotation was created primarily through alignment of transcriptomic data to the genome, with gap filling via protein to genome alignments of a select set of proteins from UniProt ([UniProt Consortium, 2019](https://www.uniprot.org)) and OrthoDB ([Kriventseva et al., 2008](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5260585/)). Prediction tools, CPC2 ([Kang et al., 2017](https://academic.oup.com/g�l/article/7/1/26/4836272)) and RNAsamba ([Camargo et al., 2020](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7592873/)), were used to aid determination of protein coding genes.

Methods
Sample acquisition and nucleic acid extraction
Four female *M. jurtina* samples (genome assembly, ilManJurt1; RNAseq, ilManJurt2, ilManJurt5; Hi-C, ilManJurt3) were collected and used for sequencing. Sample ilManJurt1 was caught in Aberlady, East Lothian, UK (latitude 56.019964, longitude -2.85808). Samples ilManJurt2 ([Figure 1C, D](https://biorxiv.org/content/10.1101/2020.03.26.993069v2.full.pdf)), ilManJurt3 and ilManJurt5 were caught in East Linton, East Lothian, UK (latitude 56.019964, longitude -2.85808). All samples were collected using a handnet by Konrad Lohse, University of Edinburgh, and snap-frozen in liquid nitrogen.

DNA was extracted at the Wellcome Sanger Institute (WSI) Scientific Operations core from the whole organism using the Qiagen MagAttract HMW DNA kit, according to the manufacturer’s instructions. RNA was extracted in the Tree of Life Laboratory at the WSI using TRIzol ([Invitrogen](https://www.thermofisher.com/si/home.html)), according to the manufacturer’s instructions. RNA was then eluted in 50 μl RNAse-free water and its concentration assessed using a Nanodrop spectrophotometer and Qubit Fluorometer using the Qubit RNA Broad-Range (BR) Assay kit. Analysis of the integrity of the RNA was done using Agilent RNA 6000 Pico Kit and Eukaryotic Total RNA assay.

Sequencing
Pacific Biosciences HiFi circular consensus and 10X Genomics Chromium read cloud sequencing libraries were constructed according to the manufacturers’ instructions. Poly(A)
Figure 1. Fore and hind wings of *Maniola jurtina* specimens from which the genome was sequenced. (A) Dorsal surface view of wings from specimen SC_MJ_1353 (ilManJurt1) from Aberlady, Scotland, UK, used to generate Pacific Biosciences and 10X data. (B) Ventral surface view of wings from specimen SC_MJ_1353 (ilManJurt1) from Aberlady, Scotland, UK, used to generate Pacific Biosciences and 10X data. (C) Dorsal surface view of wings from specimen SC_MJ_1360 (ilManJurt2) from East Linton, Scotland, UK, used to generate RNA-Seq data. (D) Ventral surface view of wings from specimen SC_MJ_1360 (ilManJurt2) from East Linton, Scotland, UK, used to generate RNA-Seq data. (E) Dorsal surface view of wings from specimen SC_MJ_1361 (ilManJurt3) from East Linton, Scotland, UK, used to generate Hi-C data. (F) Ventral surface view of wings from specimen SC_MJ_1361 (ilManJurt3) from East Linton, Scotland, UK, used to generate Hi-C data.
### Table 1. Genome data for *Maniola jurtina*, ilManJurt1.1.

| Project accession data                                      |
|-------------------------------------------------------------|
| Assembly identifier                                       | ilManJurt1.1                  |
| Species                                                   | *Maniola jurtina*             |
| Specimen                                                  | ilManJurt1 (genome assembly); ilManJurt2, ilManJurt5 (RNA-Seq); ilManJurt3 (Hi-C) |
| NCBI taxonomy ID                                          | NCBI:txid191418               |
| BioProject                                                | PRJEB43535                    |
| BioSample ID                                              | SAMEA7523158                  |
| Isolate information                                       | Female, whole organisms       |

### Raw data accessions

| PacBio Biosciences SEQUEL II | ERR6576323                     |
|-------------------------------|---------------------------------|
| 10X Genomics Illumina         | ERR6054518–ERR6054521           |
| Hi-C Illumina                 | ERR6054522                      |
| Illumina PolyA RNAseq        | ERR6054523, ERR6787422          |

### Genome assembly

| Assembly accession                | GCA_905333055.1                   |
|-----------------------------------|----------------------------------|
| Accession of alternate haplotype  | GCA_905333105.1                   |
| Span (Mb)                         | 402                              |
| Number of contigs                 | 53                               |
| Contig N50 length (Mb)            | 13                               |
| Number of scaffolds               | 32                               |
| Scaffold N50 length (Mb)          | 15                               |
| Longest scaffold (Mb)             | 17                               |
| BUSCO* genome score               | C:98.3%,S:97.7%,D:0.6%,F:0.3%,M:1.4%,n:5286 |

*BUSCO scores based on the lepidoptera_odb10 BUSCO set using v5.1.2. C= complete [S= single copy, D=duplicated], F=fragmented, M=missing, n=number of orthologues in comparison. A full set of BUSCO scores is available at https://blobtoolkit.genomehubs.org/view/ilManJurt1.1/dataset/CAJOSP01/busc.

RNA-Seq libraries were constructed using the NEB Ultra II RNA Library Prep kit. Sequencing was performed by the Scientific Operations core at the Wellcome Sanger Institute on Pacific Biosciences SEQUEL II (HiFi), Illumina HiSeq X (10X) and Illumina HiSeq 4000 (RNA-Seq) instruments. Hi-C data were generated using the Qiagen EpiTect Hi-C kit and sequenced on HiSeq X.

**Genome assembly**

Assembly was carried out with Hifiasm (Cheng et al., 2021). Haplotypic duplication was identified and removed with purge_dups (Guan et al., 2020). One round of polishing was performed by aligning 10X Genomics read data to the assembly with longranger align, calling variants with freebayes (Garrison & Marth, 2012). The assembly was then scaffolded with Hi-C data (Rao et al., 2014) using SALSA2 (Ghurye et al., 2019). The assembly was checked for contamination and corrected using the gEVAL system (Chow et al., 2016) as described previously (Howe et al., 2021). Manual curation (Howe et al., 2021) was performed using gEVAL, HiGlass (Kerpedjiev et al., 2018) and Pretex. The mitochondrial genome was assembled using MitoHiFi (Uliano-Silva et al., 2021).
genome was analysed and BUSCO scores generated within the BlobToolKit environment (Challis et al., 2020). Table 3 contains a list of all software tool versions used, where appropriate.

**Ethical/compliance issues**
The materials that have contributed to this genome note were 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 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.
**Figure 5.** Genome assembly of *Maniola jurtina*, ilManJurt1.1: Hi-C contact map. Hi-C contact map of the ilManJurt1.1 assembly, visualised in HiGlass. Chromosomes are shown in size order from left to right and top to bottom.

**Table 2.** Chromosomal pseudomolecules in the genome assembly of *Maniola jurtina*, ilManJurt1.1.

| INSDC accession | Chromosome | Size (Mb) | GC% |
|-----------------|------------|-----------|-----|
| HG995207.1      | 1          | 17.32     | 36.8|
| HG995209.1      | 2          | 17.19     | 36.6|
| HG995210.1      | 3          | 16.76     | 36.8|
| HG995211.1      | 4          | 16.67     | 36.8|
| HG995212.1      | 5          | 16.23     | 36.7|
| HG995213.1      | 6          | 16.20     | 36.8|
| HG995214.1      | 7          | 15.71     | 36.3|
| HG995215.1      | 8          | 15.42     | 36.9|
| HG995216.1      | 9          | 15.23     | 37  |
| HG995217.1      | 10         | 15.17     | 36.6|
| HG995218.1      | 11         | 15.11     | 36.8|
| HG995219.1      | 12         | 15.09     | 36.7|
| HG995220.1      | 13         | 14.83     | 36.9|
| HG995221.1      | 14         | 14.70     | 36.9|
| HG995222.1      | 15         | 14.15     | 36.9|

**Table 2 continued:**

| INSDC accession | Chromosome | Size (Mb) | GC% |
|-----------------|------------|-----------|-----|
| HG995223.1      | 16         | 14.03     | 36.8|
| HG995224.1      | 17         | 13.72     | 36.9|
| HG995225.1      | 18         | 13.54     | 36.9|
| HG995226.1      | 19         | 13.29     | 37  |
| HG995227.1      | 20         | 13.28     | 37  |
| HG995228.1      | 21         | 12.26     | 37  |
| HG995229.1      | 22         | 12.13     | 36.9|
| HG995230.1      | 23         | 10.05     | 37  |
| HG995231.1      | 24         | 9.43      | 37  |
| HG995232.1      | 25         | 8.57      | 37  |
| HG995233.1      | 26         | 7.75      | 38  |
| HG995235.1      | 27         | 6.95      | 37  |
| HG995236.1      | 28         | 6.66      | 38  |
| HG995234.1      | W          | 7.35      | 38  |
| HG995208.1      | Z          | 17.21     | 36.9|
| HG995237.1      | MT         | 0.02      | 20.3|
| -               | Unplaced   | 0.04      | 37.4|
### Table 3. Software tools used.

| Software tool | Version | Source |
|---------------|---------|--------|
| Hifiasm       | 0.12    | Cheng et al., 2021 |
| purge_dups    | 1.2.3   | Guan et al., 2020 |
| SALSA2        | 2.2     | Ghurye et al., 2019 |
| longranger    | 2.2.2   | https://support.10xgenomics.com/genre-exome/software/pipelines/latest/advanced/other-pipelines |
| freebayes     | v1.3.1-17-gaa2ace8 | Garrison & Marth, 2012 |
| MitoHiFi      | 1.0     | https://github.com/marcelauliano/MitoHiFi |
| gEVAL         | N/A     | Chow et al., 2016 |
| HiGlass       | 1.11.6  | Kerpedjiev et al., 2018 |
| PretextView   | 0.1.x   | https://github.com/wtsi-hpag/PretextView |
| BlobToolKit   | 2.6.2   | Challis et al., 2020 |

### Data availability

European Nucleotide Archive: Maniola jurtina (meadow brown). Accession number PRJEB43535; https://identifiers.org/ena.embl/PRJEB43535.

The genome sequence is released openly for reuse. The *M. jurtina* genome sequencing initiative is part of the Darwin Tree of Life (DToL) project. All raw sequence data and the assembly have been deposited in INSDC databases. Raw data and assembly accession identifiers are reported in Table 1.

### Author information

Members of the Darwin Tree of Life Barcoding collective are listed here: https://doi.org/10.5281/zenodo.4893704.

Members of the Wellcome Sanger Institute Tree of Life programme collective are listed here: https://doi.org/10.5281/zenodo.5377053.

Members of the Wellcome Sanger Institute Scientific Operations: DNA Pipelines collective are listed here: https://doi.org/10.5281/zenodo.4790456.

Members of the Tree of Life Core Informatics collective are listed here: https://doi.org/10.5281/zenodo.5013542.

Members of the Darwin Tree of Life Consortium are listed here: https://doi.org/10.5281/zenodo.4783559.

### References

Aken BL, Ayling S, Barrett D, et al.: The Ensembl Gene Annotation System. Database (Oxford). 2016; baw093. PubMed Abstract | Publisher Full Text | Free Full Text

Bigger TRL: Chromosome Numbers of Lepidoptera. Part I. Entomologist’s Gazette. 1960; 11: 149–52.

Brakefield PM: An Experimental Study of the Maintenance of Variation in Spot Pattern in Maniola Jurtina. Thesis Ph.D., Liverpool University. 1979.

Brakefield PM: Geographical Variability In, and Temperature Effects On, the Phenology of Maniola Jurtina and Pyronia Tithonus (Lepidoptera, Satyriinae) in England and Wales. Ecol Entomol. 1987; 12(2): 139–48.

Challis R, Richards E, Rajan J, et al.: BlobToolKit - Interactive Quality Assessment of Genome Assemblies. G3 (Bethesda). 2020; 10(4): 1361-74. 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): 170-75. PubMed Abstract | Publisher Full Text | Free Full Text

Chow W, Brugger K, Caccamo M, et al.: gEVAL - a web-based browser for evaluating genome assemblies. Bioinformatics. 2016; 32(16): 2508–10. PubMed Abstract | Publisher Full Text | Free Full Text

Dowdeswell WH: The Life of the Meadow Brown. Heinemann Educational Books. 1981.

Federley H: Chromosomenzahlen Finnlän-Discher Lepidopteren. Hereditas. 1938; 24(4): 397-464. Publisher Full Text

Ford EB: Ecological Genetics. Methuen. 1964.

Garrison E, Marth G: Haplotype-Based Variant Detection from Short-Read Sequencing. arXiv: 1207.3997. 2012. Reference Source

Page 10 of 15
Open Peer Review

Current Peer Review Status:  

Version 1

Reviewer Report 15 February 2023

https://doi.org/10.21956/wellcomeopenres.19131.r54294

© 2023 Fric Z. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Zdenek F. Fric
Institute of Entomology, Biology Centre CAS, České Budějovice, Czech Republic

This work represents another published genome from the Darwin Tree of Life, now for a satyrinae butterfly *Maniola jurtina*. Thus, now not many other satyrinae butterflies are left and I hope that soon all the remaining species will follow. The rationale is clear, and the methodology is as well (despite not publishing the exact scripts for the pipeline).

The result is very welcome, *Maniola jurtina* was thoroughly studied throughout Europe regarding its biogeography and a high-quality genome will be a very important step for later genomic studies on a European scale. However, I would welcome here a short discussion comparing these results with already published results by Sing et al. (2020).¹ Here, the authors state that they are aware of the study, but looking inside, the results strongly differ. For instance, the genome size reported by Lohse & Weir states that the genome size is 402 Mb long, whereas Singh et al (2020) reported a genome size exceeding 600 Mb. So please, where is the solution of the puzzle?

References
1. Singh KS, Hosken DJ, Wedell N, Ffrench-Constant R, et al.: De Novo Genome Assembly of the Meadow Brown Butterfly, *Maniola jurtina*. G3 (Bethesda). 2020; ¹0 (5): 1477-1484 PubMed Abstract | 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?
Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Phylogeny, evolution and ecology of Lepidoptera

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 27 January 2023

https://doi.org/10.21956/wellcomeopenres.19131.r54283

© 2023 Llaurens V. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Violaine Llaurens

1 Institut de Systématique, Evolution et Biodiversité, Université des Antilles Muséum National d'Histoire Naturelle, Paris, France
2 CNRS, Paris, France

Please note that I reviewed this article together with my colleague Dr Manuela Lopez-Villavicencio who has experience in butterfly genome assembly.

This article reports the genome assembly of a Maniola jurtina female, allowing access to both the autosomes and the W and Z sexual chromosomes. This butterfly observed in Europe in particular, has raised interest for its ecology. The article clearly written and particularly pleasant to read, It shows convincing evidences for a high-quality assembly. The methods for genome assembly, quality test and Hi-C scaffolding are relevant and up-to-date. Maybe the authors could also add the default parameters used during the hifiasm run and particularly if any purging option was used. They may also provide the version of the softwares used, as well as the percentage of BUSCO proteome complete/fragmented.

The article also describes mitogenome assembly with mito-hifi, therefore providing a complete view about genomic content for this species.

Given our own interest about butterfly genomes, we would be keen on knowing the genome-wide level of heterozygosity, as well as the TE content per chromosomes and throughout the whole genome.

Overall, we think the release of this well-assembled and annotated genome is a useful contribution and we recommend the publication of this article.

Violaine Llaurens & Manuela Lopez-Villavicencio.

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:** Evolutionary ecology, population 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.

---

Lohse et al. reported the chromosome-level assembly of the meadow brown butterfly *Maniola jurtina*. This assembly was obtained using PacBio (76X coverage), 10X genomics, Hi-C data. For the annotation they also used RNAseq information. We think that this new data are valuable for the scientific community. Here there are some points, nevertheless, that the authors should consider to improve the manuscript.

**Critical Point:**
1. The Sing et al. (2020) paper also describes a genome assembly of a *Maniola jurtina* species. Despite that the assembly quality reported in this paper is lower, there are some important aspects that the authors should compare and comment. For instance, the great differences between the flow cytometry and the genome assembly for both assemblies (the assembly length in Sing et al. is nearly 50% higher than that reported here). How is the length of some close related butterfly species? There is also a great difference in the number of annotated genes between papers.

**Minor Points:**
1. Introduction: The authors comment that there are some four sub-species (perhaps questionable), but they should explain to which supposed subspecies does the sequenced
species belongs.

2. Introduction: Moreover, it would be very interesting and informative, especially for non-taxonomists or butterfly systematics, to show a phylogeny that includes (at least) the various Lepidopteran species with the genome sequenced/assembled, and for which the assembly is at the chromosomal-level quality.

3. Table 1: Despite the number of annotated genes are included in the Ensembl database, it should also be interesting to be reported in the ms.

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: Comparative 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, however I have significant reservations, as outlined above.