



DATA NOTE

The genome sequence of the Common Spotted Hoverfly, *Eupeodes luniger* (Meigen, 1822) [version 1; peer review: 3 approved]

William L.S. Hawkes^{1,2}, Karl R. Wotton²,
University of Oxford and Wytham Woods Genome Acquisition Lab,
Darwin Tree of Life Barcoding collective,
Wellcome Sanger Institute Tree of Life programme,
Wellcome Sanger Institute Scientific Operations: DNA Pipelines collective,
Tree of Life Core Informatics collective, Darwin Tree of Life Consortium

¹Swiss Ornithological Institute, Sempach, Switzerland

²Centre for Ecology and Conservation, University of Exeter, Penryn, England, UK

v1 First published: 18 Oct 2023, 8:472
<https://doi.org/10.12688/wellcomeopenres.20144.1>
Latest published: 18 Oct 2023, 8:472
<https://doi.org/10.12688/wellcomeopenres.20144.1>

Abstract

We present a genome assembly from an individual female *Eupeodes luniger* (the Common Spotted Hoverfly; Arthropoda; Insecta; Diptera; Syrphidae). The genome sequence is 616.9 megabases in span. Most of the assembly is scaffolded into 4 chromosomal pseudomolecules. The mitochondrial genome has also been assembled and is 17.45 kilobases in length.

Keywords

Eupeodes luniger, Common Spotted Hoverfly, genome sequence, chromosomal, Diptera



This article is included in the [Tree of Life](#) gateway.

Open Peer Review

Approval Status

	1	2	3
version 1			
18 Oct 2023	view	view	view

1. **De-Long Guan**, Shaanxi Normal University, Xi'an, China
2. **Lucio Navarro-Escalante** , The University of Texas at Austin, Austin, USA
3. **Adam James Reid** , University of Cambridge, Cambridge, UK

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: Hawkes WLS: Investigation, Resources, Writing – Original Draft Preparation, Writing – Review & Editing; Wotton KR: 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).

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

Copyright: © 2023 Hawkes WLS *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: Hawkes WLS, Wotton KR, University of Oxford and Wytham Woods Genome Acquisition Lab *et al.* **The genome sequence of the Common Spotted Hoverfly, *Eupeodes luniger* (Meigen, 1822) [version 1; peer review: 3 approved]** Wellcome Open Research 2023, 8:472 <https://doi.org/10.12688/wellcomeopenres.20144.1>

First published: 18 Oct 2023, 8:472 <https://doi.org/10.12688/wellcomeopenres.20144.1>

Species taxonomy

Eukaryota; Metazoa; Eumetazoa; Bilateria; Protostomia; Ecdysozoa; Panarthropoda; Arthropoda; Mandibulata; Pancrustacea; Hexapoda; Insecta; Dicondylia; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Eremoneura; Cyclorrhapha; Aschiza; Syrphoidea; Syrphidae; Syrphinae; Syrphini; Eueodes; *Eueodes*; *Eueodes luniger* (Meigen, 1822) (NCBI:txid1124559).

Background

The Common Spotted Hoverfly, *Eueodes luniger*, is a medium-sized yellow and black hoverfly that is thought to be a Batesian mimic of social wasps to protect itself from predators (Golding *et al.*, 2005). Both males and females have yellow lunate comma shaped markings on their black abdomens (Ball & Morris, 2015). In the UK this is the most common *Eueodes*, yet males especially can be difficult to separate from other *Eueodes* species (Stubbs & Falk, 2002). The females, however, are readily identifiable and have a dark Y-shaped mark in front of the ocelli (Stubbs & Falk, 2002).

Adult *E. luniger* hoverflies utilise the protein from pollen for egg maturation, making them important pollinators (Doyle *et al.*, 2020; Rotheray, 1993). The larvae of these flies are aphidophagous on a wide range of aphid species in a variety of habitats (Jamali *et al.*, 2018; Rotheray, 1993). Therefore, they play an important ecological role in the form of pest control. This makes them a potentially useful species for integrated pest management (IPM) approaches. *E. luniger* is a migrant hoverfly, as is the similar *Eueodes corollae* (Sivell *et al.*, 2023), but may have a more resident population as peaks in abundance are seen earlier in the spring as well as in the summer (Ball & Morris, 2015; Hlaváček *et al.*, 2022). This is the first whole genome sequence for *Eueodes luniger*, and we believe that the sequence described here, generated as part of the Darwin Tree of Life project, will further aid understanding of the biology and ecology of this hoverfly.

Genome sequence report

The genome was sequenced from one female *Eueodes luniger* (Figure 1) collected from Wytham Woods, Oxfordshire, UK (51.77, -1.33). A total of 40-fold coverage in Pacific Biosciences single-molecule HiFi long reads was generated. Primary assembly contigs were scaffolded with chromosome conformation Hi-C data. Manual assembly curation corrected 177 missing joins or mis-joins and removed 15 haplotypic duplications, reducing the assembly length by 0.9% and the scaffold number by 58.04%, and increasing the scaffold N50 by 120.28%.

The final assembly has a total length of 616.9 Mb in 59 sequence scaffolds with a scaffold N50 of 320.0 Mb (Table 1). A summary of the assembly statistics is shown in Figure 2, while the distribution of assembly scaffolds on GC proportion and coverage is shown in Figure 3. The cumulative assembly plot in Figure 4 shows curves for subsets of scaffolds assigned to



Figure 1. Photograph of the *Eueodes luniger* (idEupLuni2) specimen used for genome sequencing.

different phyla. Most (99.79%) of the assembly sequence was assigned to 4 chromosomal-level scaffolds. Chromosome-scale scaffolds confirmed by the Hi-C data are named in order of size (Figure 5; Table 2). We did not identify the X chromosome as sequence data from the heterogametic sex was not available and homology is unreliable for sex chromosome identification in Diptera due to frequent sex chromosome turnover (Vicoso & Bachtrog, 2015). While not fully phased, the assembly deposited is of one haplotype. Contigs corresponding to the second haplotype have also been deposited. The mitochondrial genome was also assembled and can be found as a contig within the multifasta file of the genome submission.

The estimated Quality Value (QV) of the final assembly is 63.4 with *k*-mer completeness of 100%, and the assembly has a BUSCO v5.3.2 completeness of 96.6% (single = 96.0%, duplicated = 0.6%), using the diptera_odb10 reference set (*n* = 3,285).

Metadata for specimens, spectral estimates, sequencing runs, contaminants and pre-curation assembly statistics can be found at <https://links.tol.sanger.ac.uk/species/1124559>.

Methods

Sample acquisition and nucleic acid extraction

The specimen used for genome sequencing was a female *Eueodes luniger* (specimen ID Ox000159, ToLID idEupLuni2). A second female *Eueodes luniger* (specimen ID Ox000147, ToLID idEupLuni1) was used for RNA sequencing. Both specimens were netted in Wytham Woods, Oxfordshire (biological vice-county Berkshire), UK (latitude 51.77, longitude -1.33) on 2019-08-07. The specimens were collected and identified by Will Hawkes (University of Exeter) and snap-frozen on dry ice.

Table 1. Genome data for *Eupeodes luniger*, idEupLuni2.1.

Project accession data		
Assembly identifier	idEupLuni2.1	
Species	<i>Eupeodes luniger</i>	
Specimen	idEupLuni2	
NCBI taxonomy ID	1124559	
BioProject	PRJEB61505	
BioSample ID	SAMEA7520045	
Isolate information	idEupLuni2, female: head and thorax (DNA sequencing and Hi-C scaffolding) idEupLuni1: abdomen (RNA sequencing)	
Assembly metrics*		Benchmark
Consensus quality (QV)	63.4	≥ 50
k-mer completeness	100%	≥ 95%
BUSCO**	C:96.6%[S:96.0%,D:0.6%],F:0.7%,M:2.7%,n:3,285	C ≥ 95%
Percentage of assembly mapped to chromosomes	99.79%	≥ 95%
Sex chromosomes	-	localised homologous pairs
Organelles	Mitochondrial genome assembled	complete single alleles
Raw data accessions		
PacificBiosciences SEQUEL II	ERR11263503	
Hi-C Illumina	ERR11271522	
Genome assembly		
Assembly accession	GCA_951509635.1	
Accession of alternate haplotype	GCA_951508365.1	
Span (Mb)	616.9	
Number of contigs	333	
Contig N50 length (Mb)	5.2	
Number of scaffolds	59	
Scaffold N50 length (Mb)	320.0	
Longest scaffold (Mb)	320.0	

* Assembly metric benchmarks are adapted from column VGP-2020 of "Table 1: Proposed standards and metrics for defining genome assembly quality" from (Rhie *et al.*, 2021).

** BUSCO scores based on the diptera_odb10 BUSCO set using v5.3.2. C = complete [S = single copy, D = duplicated], F = fragmented, M = missing, n = number of orthologues in comparison. A full set of BUSCO scores is available at <https://blobtoolkit.genomehubs.org/view/idEupLuni2.1/dataset/CATOCI01/busco>.

The idEupLuni2 sample was prepared for DNA extraction at the Tree of Life laboratory, Wellcome Sanger Institute (WSI). The specimen was weighed and dissected on dry ice with tissue set aside for Hi-C sequencing. Head and thorax

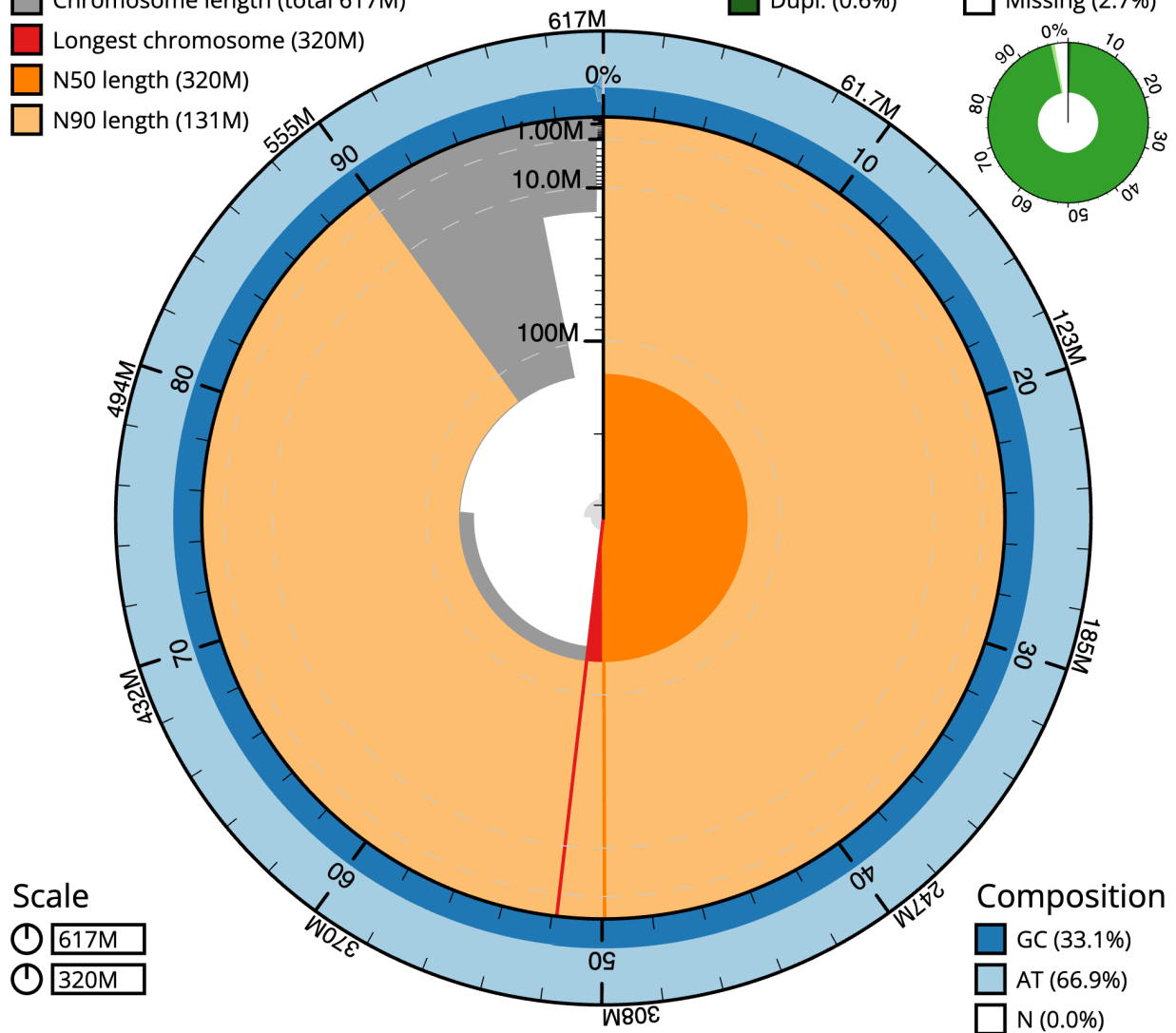
tissue was disrupted using a Nippi Powermasher fitted with a BioMasher pestle. DNA was extracted at the WSI Scientific Operations core using the Qiagen MagAttract HMW DNA kit, according to the manufacturer's instructions.

Chromosome statistics

- Log10 chromosome count (total 60)
- Chromosome length (total 617M)
- Longest chromosome (320M)
- N50 length (320M)
- N90 length (131M)

BUSCO diptera_odb10 (3285)

- Comp. (96.6%)
- Frag. (0.7%)
- Dupl. (0.6%)
- Missing (2.7%)



Dataset: CATOCI01

Figure 2. Genome assembly of *Eupeodes luniger*, idEupLuni2.1: metrics. The BlobToolKit Snailplot shows N50 metrics and BUSCO gene completeness. The main plot is divided into 1,000 size-ordered bins around the circumference with each bin representing 0.1% of the 616,908,419 bp assembly. The distribution of scaffold lengths is shown in dark grey with the plot radius scaled to the longest scaffold present in the assembly (319,955,300 bp, shown in red). Orange and pale-orange arcs show the N50 and N90 scaffold lengths (319,955,300 and 131,323,115 bp), respectively. The pale grey spiral shows the cumulative scaffold count on a log scale with white scale lines showing successive orders of magnitude. The blue and pale-blue area around the outside of the plot shows the distribution of GC, AT and N percentages in the same bins as the inner plot. A summary of complete, fragmented, duplicated and missing BUSCO genes in the diptera_odb10 set is shown in the top right. An interactive version of this figure is available at <https://blobtoolkit.genomehubs.org/view/idEupLuni2.1/dataset/CATOCI01/snail>.

RNA was extracted from abdomen tissue of idEupLuni1 in the Tree of Life Laboratory at the WSI using TRIzol, 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.

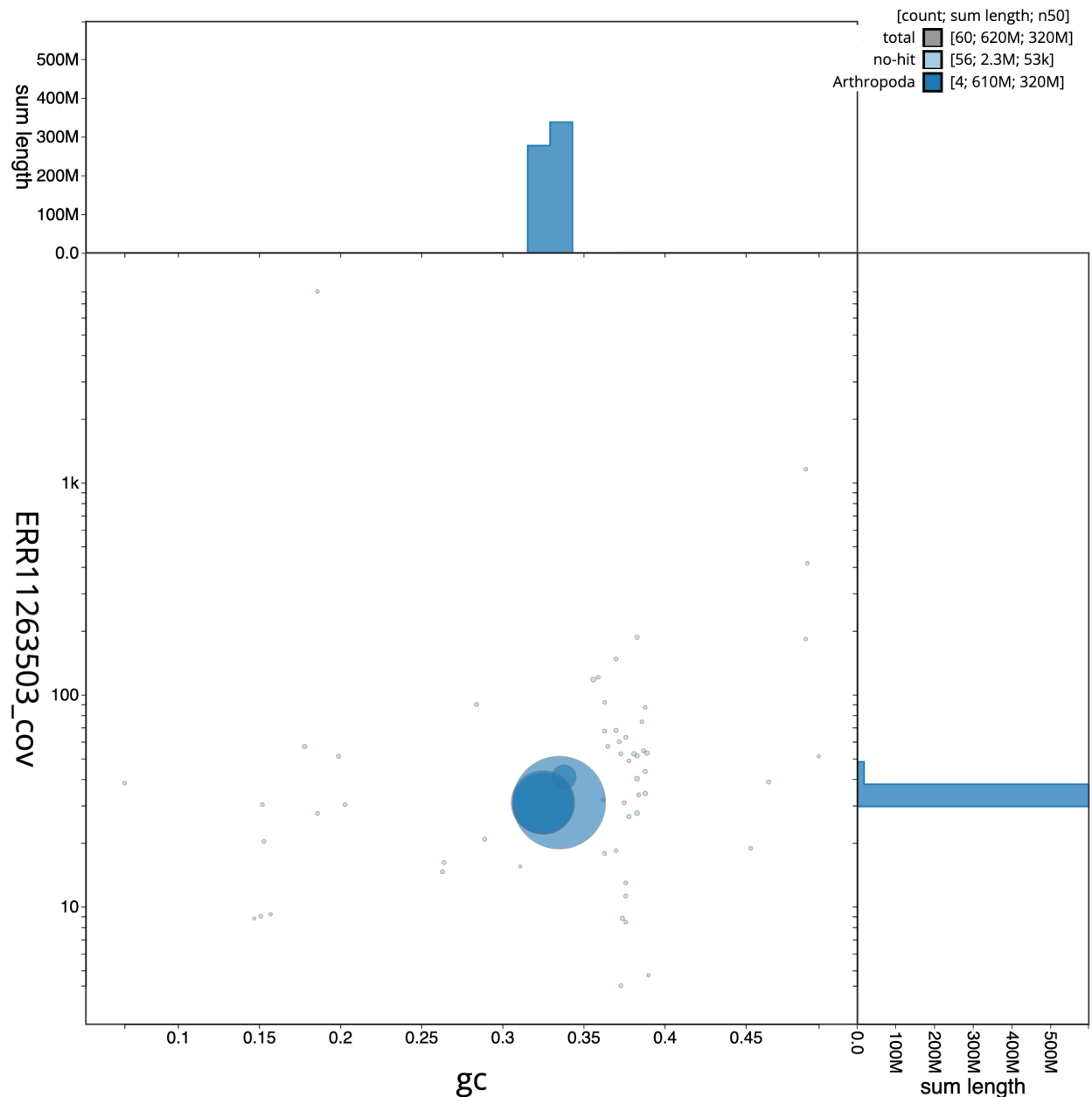


Figure 3. Genome assembly of *Eupeodes luniger*, idEupLuni2.1: BlobToolKit GC-coverage plot. Scaffolds are coloured by phylum. Circles are sized in proportion to scaffold length. Histograms show the distribution of scaffold length sum along each axis. An interactive version of this figure is available at <https://blobtoolkit.genomehubs.org/view/idEupLuni2.1/dataset/CATOCI01/blob>.

Sequencing

Pacific Biosciences HiFi circular consensus DNA sequencing libraries were constructed according to the manufacturers' instructions. Poly(A) RNA-Seq libraries were constructed using the NEB Ultra II RNA Library Prep kit. DNA and RNA sequencing was performed by the Scientific Operations core at the WSI on Pacific Biosciences SEQUEL II (HiFi) and Illumina NovaSeq 6000 (RNA-Seq) instruments. Hi-C data

were also generated from remaining tissue of idEupLuni2 using the Arima2 kit and sequenced on the Illumina NovaSeq 6000 instrument.

Genome assembly, curation and evaluation

Assembly was carried out with Hifiasm (Cheng *et al.*, 2021) and haplotypic duplication was identified and removed with purge_dups (Guan *et al.*, 2020). The assembly was then

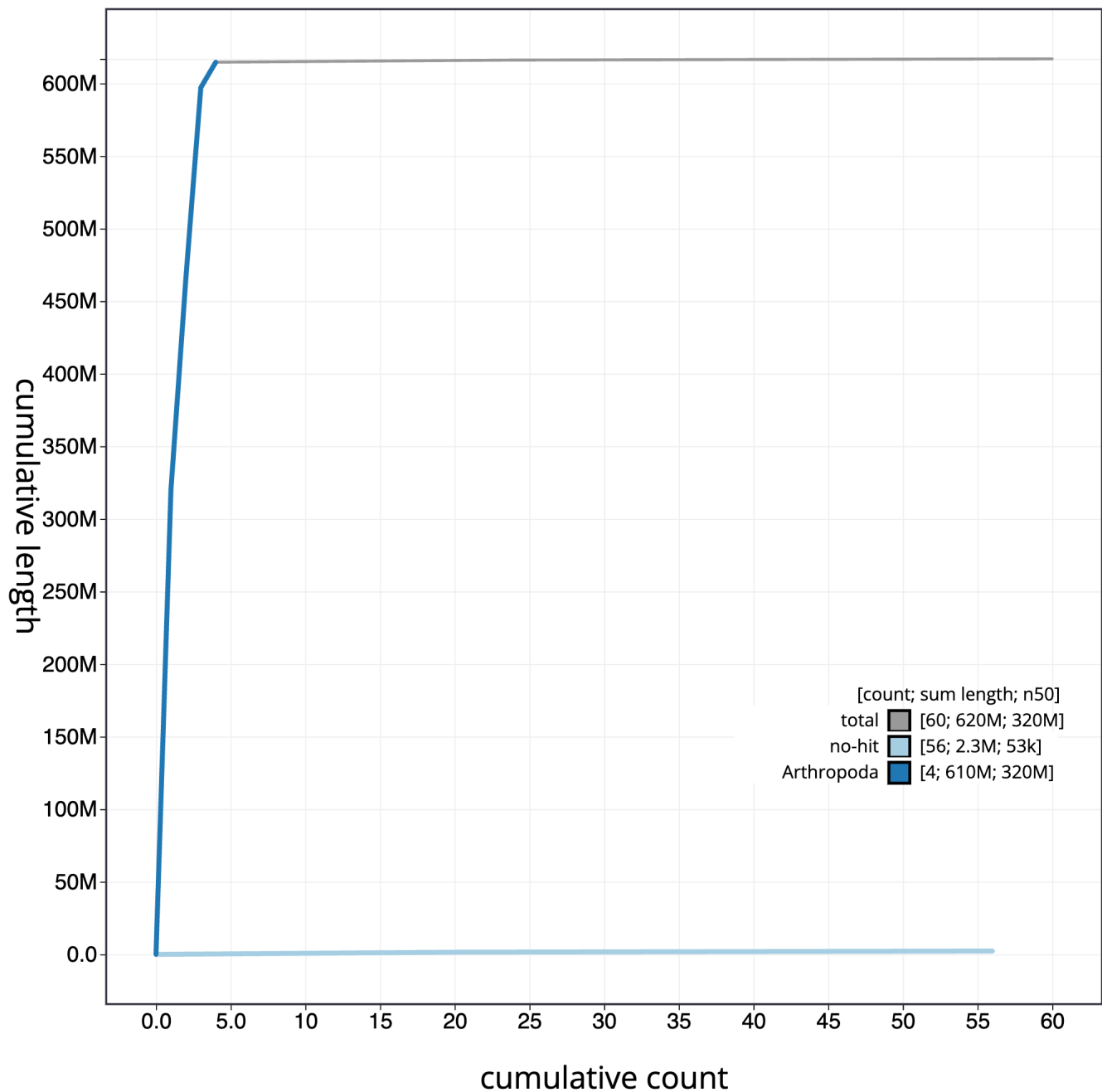


Figure 4. Genome assembly of *Eupeodes luniger*, idEupLuni2.1: BlobToolKit cumulative sequence plot. The grey line shows cumulative length for all scaffolds. Coloured lines show cumulative lengths of scaffolds assigned to each phylum using the buscogenes taxrule. An interactive version of this figure is available at <https://blobtoolkit.genomehubs.org/view/idEupLuni2.1/dataset/CATOC101/cumulative>.

scaffolded with Hi-C data (Rao *et al.*, 2014) using YaHS (Zhou *et al.*, 2023). The assembly was checked for contamination and corrected as described previously (Howe *et al.*, 2021). Manual curation was performed using HiGlass (Kerpedjiev *et al.*, 2018) and Pretext (Harry, 2022). The mitochondrial genome was assembled using MitoHiFi (Uliano-Silva *et al.*, 2023), which runs

MitoFinder (Allio *et al.*, 2020) or MITOS (Bernt *et al.*, 2013) and uses these annotations to select the final mitochondrial contig and to ensure the general quality of the sequence.

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

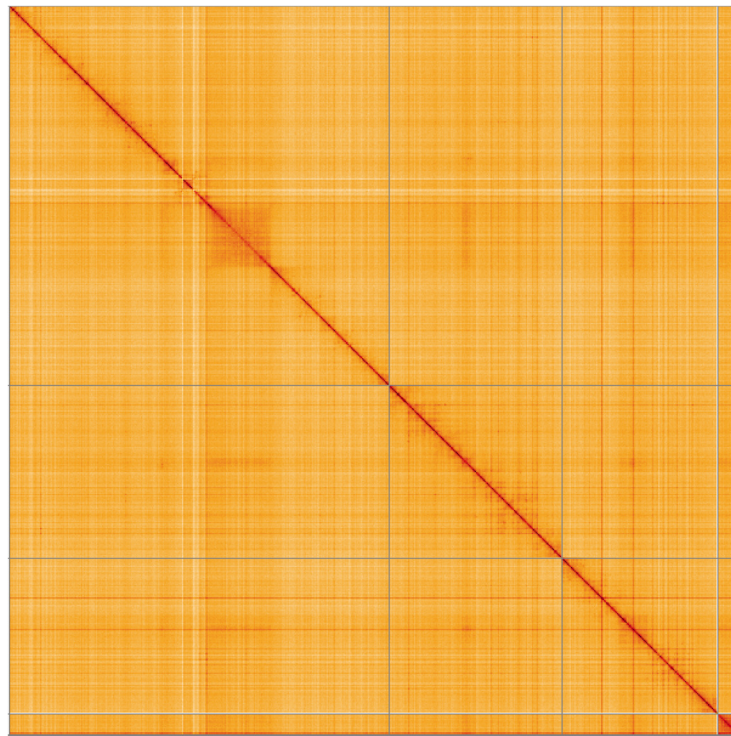


Figure 5. Genome assembly of *Eupeodes luniger*, idEupLuni2.1: Hi-C contact map of the idEupLuni2.1 assembly, visualised using HiGlass. Chromosomes are shown in order of size from left to right and top to bottom. An interactive version of this figure may be viewed at https://genome-note-higlass.tol.sanger.ac.uk/l/?d=OEGHJP_7S0uMmC-mnthHlrA.

Table 2. Chromosomal pseudomolecules in the genome assembly of *Eupeodes luniger*, idEupLuni2.

INSDC accession	Chromosome	Length (Mb)	GC%
OX608073.1	1	319.96	33.5
OX608074.1	2	145.83	32.5
OX608075.1	3	131.32	32.5
OX608076.1	4	17.52	34.0
OX608077.1	MT	0.02	18.5

(Abdennur & Mirny, 2020). To assess the assembly metrics, the k -mer completeness and QV consensus quality values were calculated in Merqury (Rhie *et al.*, 2020). This work was done using Nextflow (Di Tommaso *et al.*, 2017) DSL2 pipelines “sanger-tol/readmapping” (Surana *et al.*, 2023a) and “sanger-tol/genomenote” (Surana *et al.*, 2023b). The genome was analysed within the BlobToolKit environment (Challis *et al.*, 2020) and BUSCO scores (Manni *et al.*, 2021; Simão *et al.*, 2015) were calculated.

Table 3 contains a list of relevant software tool versions and sources.

Wellcome Sanger Institute – Legal and Governance

The materials that have contributed to this genome note have been supplied by a Darwin Tree of Life Partner. The submission of materials by a Darwin Tree of Life Partner is subject to the ‘**Darwin Tree of Life Project Sampling Code of Practice**’, which can be found in full on the Darwin Tree of Life website [here](#). By agreeing with and signing up to the Sampling Code of Practice, the Darwin Tree of Life Partner agrees they will meet the legal and ethical requirements and standards set out within this document in respect of all samples acquired for, and supplied to, the Darwin Tree of Life Project.

Further, the Wellcome Sanger Institute employs a process whereby due diligence is carried out proportionate to the nature of the materials themselves, and the circumstances under which they have been/are to be collected and provided for use. The purpose of this is to address and mitigate any potential legal and/or ethical implications of receipt and use of the materials as part of the research project, and to ensure that in doing so we align with best practice wherever possible. The overarching areas of consideration are:

Table 3. Software tools: versions and sources.

Software tool	Version	Source
BlobToolKit	4.1.7	https://github.com/blobtoolkit/blobtoolkit
BUSCO	5.3.2	https://gitlab.com/ezlab/busco
Hifiasm	0.16.1-r375	https://github.com/chhylp123/hifiasm
HiGlass	1.11.6	https://github.com/higlass/higlass
Merquy	MerquyFK	https://github.com/thegenemyers/MERQUERY.FK
MitoHiFi	3	https://github.com/marcelauliano/MitoHiFi
PretextView	0.2	https://github.com/wtsi-hpag/PretextView
purge_dups	1.2.5	https://github.com/dfguan/purge_dups
sanger-tol/genomenote	v1.0	https://github.com/sanger-tol/genomenote
sanger-tol/readmapping	1.1.0	https://github.com/sanger-tol/readmapping/tree/1.1.0
YaHS	1.2a.2	https://github.com/c-zhou/yahs

- Ethical review of provenance and sourcing of the material
- Legality of collection, transfer and use (national and international)

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

Data availability

European Nucleotide Archive: *Eupeodes luniger* (common spotted Eupeodes). Accession number PRJEB61505; <https://identifiers.org/ena.embl/PRJEB61505>. (Wellcome Sanger Institute, 2023) The genome sequence is released openly for reuse. The *Eupeodes luniger* 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. The genome will be annotated using available RNA-Seq data and presented through the [Ensembl](#) pipeline at the European

Bioinformatics Institute. Raw data and assembly accession identifiers are reported in [Table 1](#).

Author information

Members of the University of Oxford and Wytham Woods Genome Acquisition Lab are listed here: <https://doi.org/10.5281/zenodo.4789928>.

Members of the Darwin Tree of Life Barcoding collective are listed here: <https://doi.org/10.5281/zenodo.4893703>.

Members of the Wellcome Sanger Institute Tree of Life programme are listed here: <https://doi.org/10.5281/zenodo.4783585>.

Members of Wellcome Sanger Institute Scientific Operations: DNA Pipelines collective are listed here: <https://doi.org/10.5281/zenodo.4790455>.

Members of the Tree of Life Core Informatics collective are listed here: <https://doi.org/10.5281/zenodo.5013541>.

Members of the Darwin Tree of Life Consortium are listed here: <https://doi.org/10.5281/zenodo.4783558>.

References

Abdennur N, Mirny LA: **Cooler: Scalable storage for Hi-C data and other genomically labeled arrays.** *Bioinformatics*. 2020; **36**(1): 311–316.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 Allio R, Schomaker-Bastos A, Romiguier J, et al.: **MitoFinder: Efficient**

automated large-scale extraction of mitogenomic data in target enrichment phylogenomics. *Mol Ecol Resour*. 2020; **20**(4): 892–905.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Ball S, Morris R: **Britain's Hoverflies: A Field Guide -Revised and Updated**

Second Edition (Vol. 17). Princeton: Princeton University Press, 2015; 17.

Bernt M, Donath A, Jühling F, *et al.*: **MITOS: Improved *de novo* metazoan mitochondrial genome annotation.** *Mol Phylogenet Evol.* 2013; **69**(2): 313–319.
[PubMed Abstract](#) | [Publisher Full Text](#)

Challis R, Richards E, Rajan J, *et al.*: **BlobToolKit - interactive quality assessment of genome assemblies.** *G3 (Bethesda).* 2020; **10**(4): 1361–1374.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Cheng H, Concepcion GT, Feng X, *et al.*: **Haplotype-resolved *de novo* assembly using phased assembly graphs with hifiasm.** *Nat Methods.* 2021; **18**(2): 170–175.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Di Tommaso P, Chatzou M, Floden EW, *et al.*: **Nextflow enables reproducible computational workflows.** *Nat Biotechnol.* 2017; **35**(4): 316–319.
[PubMed Abstract](#) | [Publisher Full Text](#)

Doyle T, Hawkes WLS, Massy R, *et al.*: **Pollination by hoverflies in the Anthropocene.** *Proc Biol Sci.* 2020; **287**(1927): 20200508.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Golding YC, Edmunds M, Ennos AR: **Flight behaviour during foraging of the social wasp *Vespula vulgaris* (Hymenoptera: Vespidae) and four mimetic hoverflies (Diptera: Syrphidae) *Sericomyia silentis*, *Myathropa florea*, *Helophilus* sp. and *Syrphus* sp.** *J Exp Biol.* 2005; **208**(Pt 23): 4523–7.
[PubMed Abstract](#) | [Publisher Full Text](#)

Guan D, McCarthy SA, Wood J, *et al.*: **Identifying and removing haplotypic duplication in primary genome assemblies.** *Bioinformatics.* 2020; **36**(9): 2896–2898.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Harry E: **PretextView (Paired REad TEXTure Viewer): A desktop application for viewing pretext contact maps.** 2022; [Accessed 19 October 2022].
[Reference Source](#)

Hlaváček A, Lučan RK, Hadrava J: **Autumnal migration patterns of hoverflies (Diptera: Syrphidae): interannual variability in timing and sex ratio.** *PeerJ.* 2022; **10**: e14393.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Howe K, Chow W, Collins J, *et al.*: **Significantly improving the quality of genome assemblies through curation.** *GigaScience.* Oxford University Press, 2021; **10**(1): g1aa153.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Jamali RA, Memom N, Shah MA, *et al.*: **Prevalence of aphidophagous hoverflies (Syrphidae: Syrphinae) in relation to their prey, green aphids (*Myzus persicae*) on brassica (*Brassica rapa oleifera*) in Dadu.** *JAPS: Journal of Animal & Plant Sciences.* 2018; **28**(5): 1447–1456.
[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](#)

Manni M, Berkeley MR, Seppely M, *et al.*: **BUSCO update: Novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes.** *Mol Biol Evol.* 2021; **38**(10): 4647–4654.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Rao SSP, Huntley MH, Durand NC, *et al.*: **A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping.** *Cell.* 2014; **159**(7): 1665–1680.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Rhie A, McCarthy SA, Fedrigo O, *et al.*: **Towards complete and error-free genome assemblies of all vertebrate species.** *Nature.* 2021; **592**(7856): 737–746.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Rhie A, Walenz BP, Koren S, *et al.*: **Merqure: Reference-free quality, completeness, and phasing assessment for genome assemblies.** *Genome Biol.* 2020; **21**(1): 245.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Rotheray G: **Colour Guide to Hoverfly Larvae (Diptera, Syrphidae).** *Dipterists Digest.* 1993; **9**: 1–155.
[Reference Source](#)

Simão FA, Waterhouse RM, Ioannidis P, *et al.*: **BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs.** *Bioinformatics.* 2015; **31**(19): 3210–3212.
[PubMed Abstract](#) | [Publisher Full Text](#)

Sivell D, Sivell O, Natural History Museum Genome Acquisition Lab, *et al.*: **The genome sequence of the Vagrant Hoverfly, *Eupeodes corollae* (Fabricius, 1794).** *Wellcome Open Res.* 2023; **8**: 112.
[Publisher Full Text](#)

Stubbs AE, Falk SJ: **British hoverflies: an illustrated identification guide.** British Entomological and Natural History Society, 2002.
[Reference Source](#)

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

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

Uliano-Silva M, Ferreira JGRN, Krashenninnikova K, *et al.*: **MitoHiFi: a python pipeline for mitochondrial genome assembly from PacBio high fidelity reads.** *BMC Bioinformatics.* 2023; **24**(1): 288.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Vasimuddin M, Misra S, Li H, *et al.*: **Efficient Architecture-Aware Acceleration of BWA-MEM for Multicore Systems.** In: *2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS).* IEEE, 2019; 314–324.
[Publisher Full Text](#)

Vicoso B, Bachtrog D: **Numerous Transitions of Sex Chromosomes in Diptera.** *PLoS Biol.* 2015; **13**(4): e1002078.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Wellcome Sanger Institute: **The genome sequence of the Common Spotted Hoverfly, *Eupeodes luniger* (Meigen, 1822).** European Nucleotide Archive. [dataset], accession number PRJEB61505, 2023.

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 28 May 2024

<https://doi.org/10.21956/wellcomeopenres.22306.r81595>

© 2024 Reid A. 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.



Adam James Reid 

University of Cambridge, Cambridge, UK

Hawkes *et al.* present a very high-quality genome assembly for the Common Spotted Hoverfly. The data and manuscript are well presented. The methods used are state-of-the-art.

Major comments: none

Minor comments: none

While the particular organism and its genome sequence are interesting in their own right, the continued generation of high quality genomes by the Darwin Tree of Life project team will underpin a great deal of future work on evolutionary biology.

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: Computational biology, genomics, transcriptomics, parasitology, developmental biology

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 25 May 2024

<https://doi.org/10.21956/wellcomeopenres.22306.r76497>

© 2024 Navarro-Escalante L. 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.



Lucio Navarro-Escalante 

Department of Molecular Biosciences, The University of Texas at Austin, Austin, Texas, USA

Overall

This paper reports the chromosomal-level genome sequences of the hoverfly *Eupeodes luniger*, a wasp-mimicking fly in the Syrphidae family. The genome was sequenced and assembled using PacBio long-reads and Hi-C Illumina reads. The mitochondrial genome is also reported here. This assembly represents a high quality full genome. In general the methods offer sufficient details and all raw data is publicly available. I support the indexing of this report and its data.

Minor comment:

When describing the use of MitoHiFi in the methods, authors should clearly indicate whether MitoFinder or MITOS was used for mitochondrial genome assembly.

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?

Partly

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

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Insect genomics, insect-microbe symbiosis, symbiont engineering

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 08 April 2024

<https://doi.org/10.21956/wellcomeopenres.22306.r76505>

© 2024 Guan D. 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.



De-Long Guan

Shaanxi Normal University, Xi'an, Shaanxi, China

Dear Editor

I have completed my review of the manuscript and I am pleased to report that I found no significant issues requiring further attention. The article is well-crafted, with a clear and concise presentation of the research topic, robust methodology, and sound analysis. The results are well-supported and the discussion is insightful, providing valuable contributions to the field.

****Major Points:****

1. ****Scientific Significance:**** The publication is a notable achievement as it presents the inaugural genome sequence for *Eupeodes luniger*, enriching our understanding of this species' biology and ecology. The genome sequence opens doors to investigate evolutionary patterns, genetic diversity, and the functional aspects of the hoverfly's genetics.
2. ****Genome Assembly Quality:**** The quality of the genome assembly is commendable, with a substantial contiguity. The meticulous use of Hi-C data for scaffolding, along with manual curation, demonstrates a high level of precision and attention to detail in the assembly process.
3. ****Ecological and Agricultural Relevance:**** The hoverfly's ecological role as a pollinator and its larvae's pest-control function by consuming aphids are well-documented. The genome sequence can serve as a foundation for research aimed at developing integrated pest management strategies, potentially revolutionizing sustainable agricultural practices.
4. ****Methodological Advancements:**** The article exemplifies the application of state-of-the-art sequencing technologies, such as Pacific Biosciences single-molecule HiFi long reads and Hi-C for chromosomal conformation capture. These techniques represent the forefront of genomic research and offer a blueprint for future genomic studies on a variety of species.

****Minor Points:****

1. ****Peer Review Status:**** The article, at the time of its initial publication, was awaiting peer review. This indicates that while the research is presented by respected institutions, it has not yet been vetted by external experts in the field. This is a critical step in the scientific process that will ultimately strengthen the credibility of the findings.
2. ****Potential for Further Analysis:**** The genome assembly lays the groundwork for in-depth functional annotation, which could reveal key genes and their roles in the hoverfly's biology. Future research could focus on annotating the genome to enhance our understanding of the species' unique characteristics and evolutionary adaptations.

3. **Comparative Genomics:** The genome sequence offers a valuable resource for comparative genomics studies with related species. Such comparisons could shed light on the genetic underpinnings of the hoverfly's distinctive traits and its adaptive evolutionary history. This represents an area ripe for further exploration.

I believe the article on the Common Spotted Hoverfly's genome sequence provides an invaluable resource for future research endeavors and has the potential to inform agricultural practices and conservation strategies.

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: Insect genomics, comparative genomics, bioinformatics, Evolutionary Biology, Acrididae, Tetrigidae, and Papilionidae.

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.
