



DATA NOTE

The genome sequence of the Blood-vein moth, *Timandra*

comae Schmidt, 1931

[version 1; peer review: 3 approved]

Denise C. Wawman ¹,

University of Oxford and Wytham Woods Genome Acquisition Lab,
Darwin Tree of Life Barcoding collective,
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, Darwin Tree of Life Consortium

¹Edward Grey Institute, Department of Biology, University of Oxford, Oxford, England, UK

V1 First published: 01 Mar 2024, 9:100
<https://doi.org/10.12688/wellcomeopenres.21088.1>

Latest published: 01 Mar 2024, 9:100
<https://doi.org/10.12688/wellcomeopenres.21088.1>

Abstract

We present a genome assembly from an individual male *Timandra comae* (the Blood-vein; Arthropoda; Insecta; Lepidoptera; Geometridae). The genome sequence is 334.4 megabases in span. Most of the assembly is scaffolded into 31 chromosomal pseudomolecules, including the Z sex chromosome. The mitochondrial genome has also been assembled and is 15.91 kilobases in length.

Keywords







Timandra comae, Blood-vein moth, genome sequence, chromosomal, Lepidoptera



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

Open Peer Review

Approval Status 

	1	2	3
version 1 01 Mar 2024	 view	 view	 view
1. Niclas Backström  ,	Uppsala University, Uppsala, Sweden		
2. Kay Lucek  ,	University of Neuchâtel, Neuchâtel, Switzerland		
3. Marko Mutanen  ,	University of Oulu, Oulu, Finland		

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: Wawman DC: Investigation, Resources, 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, <https://doi.org/10.35802/206194>] and the Darwin Tree of Life Discretionary Award [218328, <https://doi.org/10.35802/218328>]. *The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

Copyright: © 2024 Wawman DC *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: Wawman DC, University of Oxford and Wytham Woods Genome Acquisition Lab, Darwin Tree of Life Barcoding collective *et al.* **The genome sequence of the Blood-vein moth, *Timandra comae* Schmidt, 1931 [version 1; peer review: 3 approved]** Wellcome Open Research 2024, 9:100 <https://doi.org/10.12688/wellcomeopenres.21088.1>

First published: 01 Mar 2024, 9:100 <https://doi.org/10.12688/wellcomeopenres.21088.1>



DATA NOTE

The genome sequence of the Blood-vein moth, *Timandra comae* Schmidt, 1931

Denise C. Wawman ¹,

University of Oxford and Wytham Woods Genome Acquisition Lab,
Darwin Tree of Life Barcoding collective,
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, Darwin Tree of Life Consortium

¹Edward Grey Institute, Department of Biology, University of Oxford, Oxford, England, UK

v1 First published: N/A, N/A: N/A N/A
Latest published: N/A, N/A: N/A N/A

Open Peer Review

Approval Status AWAITING PEER REVIEW

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

Abstract

We present a genome assembly from an individual male *Timandra comae* (the Blood-vein; Arthropoda; Insecta; Lepidoptera; Geometridae). The genome sequence is 334.4 megabases in span. Most of the assembly is scaffolded into 31 chromosomal pseudomolecules, including the Z sex chromosome. The mitochondrial genome has also been assembled and is 15.91 kilobases in length.

Keywords

Timandra comae, Blood-vein moth, genome sequence, chromosomal, Lepidoptera



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

Corresponding author: Darwin Tree of Life Consortium (mark.blaxter@sanger.ac.uk)

Author roles: Wawman DC: Investigation, Resources, 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, <https://doi.org/10.35802/206194>] and the Darwin Tree of Life Discretionary Award [218328, <https://doi.org/10.35802/218328>]. *The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

Copyright: © 2024 Wawman DC *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: Wawman DC, University of Oxford and Wytham Woods Genome Acquisition Lab, Darwin Tree of Life Barcoding collective *et al.* **The genome sequence of the Blood-vein moth, *Timandra comae* Schmidt, 1931** Wellcome Open Research , : <https://doi.org/>

First published: N/A, **N/A:** N/A N/A

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; Geometroidea; Geometridae; Sterrhinae; *Timandra*; *Timandra comae* Schmidt, 1931 (NCBI:txid190366).

Background

The Blood-vein *Timandra comae* is a moth in the family Geometridae and, as such, has a thin body and wide triangular wings. It is distinguished from other moths in the United Kingdom by the pale creamy-buff wings with a straight pink or brownish-red line which runs from the tip of the forewing to its trailing edge and across the hindwing. The trailing edge of both wings is pinkish-red and there are varying amounts of fine dark speckling across both wings (Waring *et al.*, 2017). However, there is a very similar species *Timandra griselata*, found in Scandinavia, that has greyish speckling on a whitish forewing background, rather than the more yellowish colour seen in *T. comae*: there are slight differences in the female genitalia, but no differences between those of the males, and the two species are separable by DNA sequencing of the mitochondrial COI locus (Öunap *et al.*, 2005).

In the United Kingdom *T. comae*'s range is mainly restricted to the south of the country, as far north as the Scottish Borders, its having recently reached Cumberland and Northumbria: although there are records from Scotland these moths are considered immigrants, with those found in Orkney and Shetland thought to be Scandinavian in origin (Waring *et al.*, 2017).

Timandra comae has two generations a year, although it may have three in the south, where these generations may overlap, and overwinters as a larva (Waring *et al.*, 2017). The larvae feed on Docks *Rumex* spp., Oraches *Atriplex* spp., Common Sorrel *Rumex acetosa*, Sheep's Sorrel *Rumex acetosella*, and Knotgrass *Polygonum aviculare* (Kurze *et al.*, 2018; Skinner & Wilson, 2009; Waring *et al.*, 2017). The application of nitrogen fertiliser, increasing the nitrogen levels in Sheep's Sorrel, has been shown to decrease the survival rate of the larvae of *T. comae* by more than two-thirds, compared to the control group, and was thought to be a possible mechanism leading to the decline of some species (Kurze *et al.*, 2018).

We present a chromosomally complete genome sequence for *Timandra comae*, based on one specimen collected using a mercury vapour light trap in a rural garden in the hamlet of Bratton, near Minehead, in Somerset, as part of the Darwin Tree of Life Project.

Genome sequence report

The genome was sequenced from one male *Timandra comae* (Figure 1) collected from Bratton, Somerset, UK (51.20, -3.51). A total of 84-fold coverage in Pacific Biosciences single-molecule HiFi long reads was generated. Primary assembly



Figure 1. Photograph of the *Timandra comae* (ilTimComa1) specimen used for genome sequencing.

contigs were scaffolded with chromosome conformation Hi-C data. Manual assembly curation corrected 19 missing joins or mis-joins and removed 5 haplotypic duplications, reducing the assembly length by 0.45% and the scaffold number by 11.11%.

The final assembly has a total length of 334.4 Mb in 39 sequence scaffolds with a scaffold N50 of 11.5 Mb (Table 1). The snailplot in Figure 2 provides a summary of the assembly statistics, 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 different phyla. Most (99.92%) of the assembly sequence was assigned to 31 chromosomal-level scaffolds, representing 30 autosomes and the Z sex chromosome. Chromosome-scale scaffolds confirmed by the Hi-C data are named in order of size (Figure 5; Table 2). The Z chromosome was identified based on synteny with *Cyclophora punctaria* (GCA_951394245.1). 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 64.8 with *k*-mer completeness of 100.0%, and the assembly has a BUSCO v5.3.2 completeness of 98.3% (single = 98.1%, duplicated = 0.2%), using the lepidoptera_odb10 reference set (*n* = 5,286).

Metadata for specimens, barcode results, spectra estimates, sequencing runs, contaminants and pre-curation assembly statistics are given at <https://links.tol.sanger.ac.uk/species/190366>.

Table 1. Genome data for *Timandra comae*, iTimComa1.1.

Project accession data		
Assembly identifier	iTimComa1.1	
Species	<i>Timandra comae</i>	
Specimen	iTimComa1	
NCBI taxonomy ID	190366	
BioProject	PRJEB63432	
BioSample ID	SAMEA112226457	
Isolate information	iTimComa1, male: head and thorax (DNA and Hi-C sequencing)	
Assembly metrics*		Benchmark
Consensus quality (QV)	64.8	≥ 50
k-mer completeness	100.0%	≥ 95%
BUSCO**	C:98.3%[S:98.1%,D:0.2%], F:0.5%,M:1.2%,n:5,286	C ≥ 95%
Percentage of assembly mapped to chromosomes	99.92%	≥ 95%
Sex chromosomes	ZZ	localised homologous pairs
Organelles	Mitochondrial genome: 15.91 kb	complete single alleles
Raw data accessions		
PacificBiosciences SEQUEL II	ERR11593800	
Hi-C Illumina	ERR11606316	
Genome assembly		
Assembly accession	GCA_958496195.1	
Accession of alternate haplotype	GCA_958496285.1	
Span (Mb)	334.4	
Number of contigs	92	
Contig N50 length (Mb)	6.7	
Number of scaffolds	39	
Scaffold N50 length (Mb)	11.5	
Longest scaffold (Mb)	19.46	

* 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 lepidoptera_odb10 BUSCO set using version 5.3.2. 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/iTimComa1_1/dataset/iTimComa1_1/busco.

Methods

Sample acquisition and nucleic acid extraction

A male *Timandra comae* (specimen ID OX002231, ToLID iTimComa1) was collected from Bratton, Somerset, UK

(latitude 51.20, longitude -3.51) on 2022-06-20 using a light trap. The specimen was collected and identified by Denise Wawman (University of Oxford), and then preserved on dry ice.

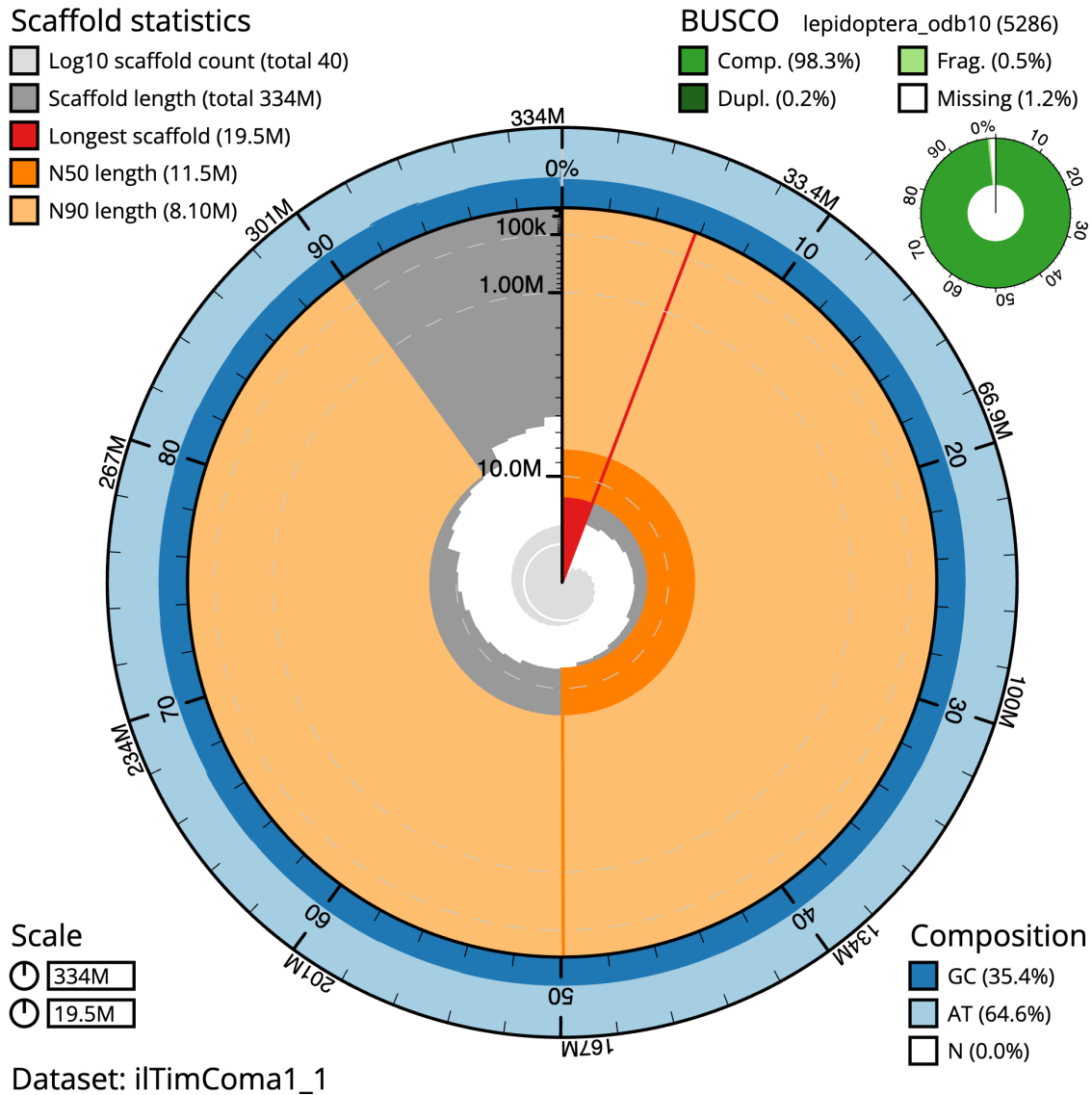


Figure 2. Genome assembly of *Timandra comae*, ilTimComa1.1: metrics. The BlobToolKit Snailplot shows N50 metrics and BUSCO gene completeness. The main plot is divided into 1,000 size-ordered bins around the circumference with each bin representing 0.1% of the 334,373,444 bp assembly. The distribution of scaffold lengths is shown in dark grey with the plot radius scaled to the longest scaffold present in the assembly (19,455,470 bp, shown in red). Orange and pale-orange arcs show the N50 and N90 scaffold lengths (11,490,666 and 8,097,011 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 lepidoptera_odb10 set is shown in the top right. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/ilTimComa1_1/dataset/ilTimComa1_1/snail.

The workflow for high molecular weight (HMW) DNA extraction at the Wellcome Sanger Institute (WSI) includes a sequence of core procedures: sample preparation; sample homogenisation, DNA extraction, fragmentation, and clean-up. In sample preparation, the ilTimComa1 sample was weighed and dissected on dry ice (Jay *et al.*, 2023). Tissue from the head and

thorax was homogenised using a PowerMasher II tissue disruptor (Denton *et al.*, 2023a). HMW DNA was extracted in the WSI Scientific Operations core using the Automated MagAttract v2 protocol (Oatley *et al.*, 2023). The DNA was sheared into an average fragment size of 12–20 kb in a Megaruptor 3 system with speed setting 31 (Bates *et al.*, 2023). Sheared

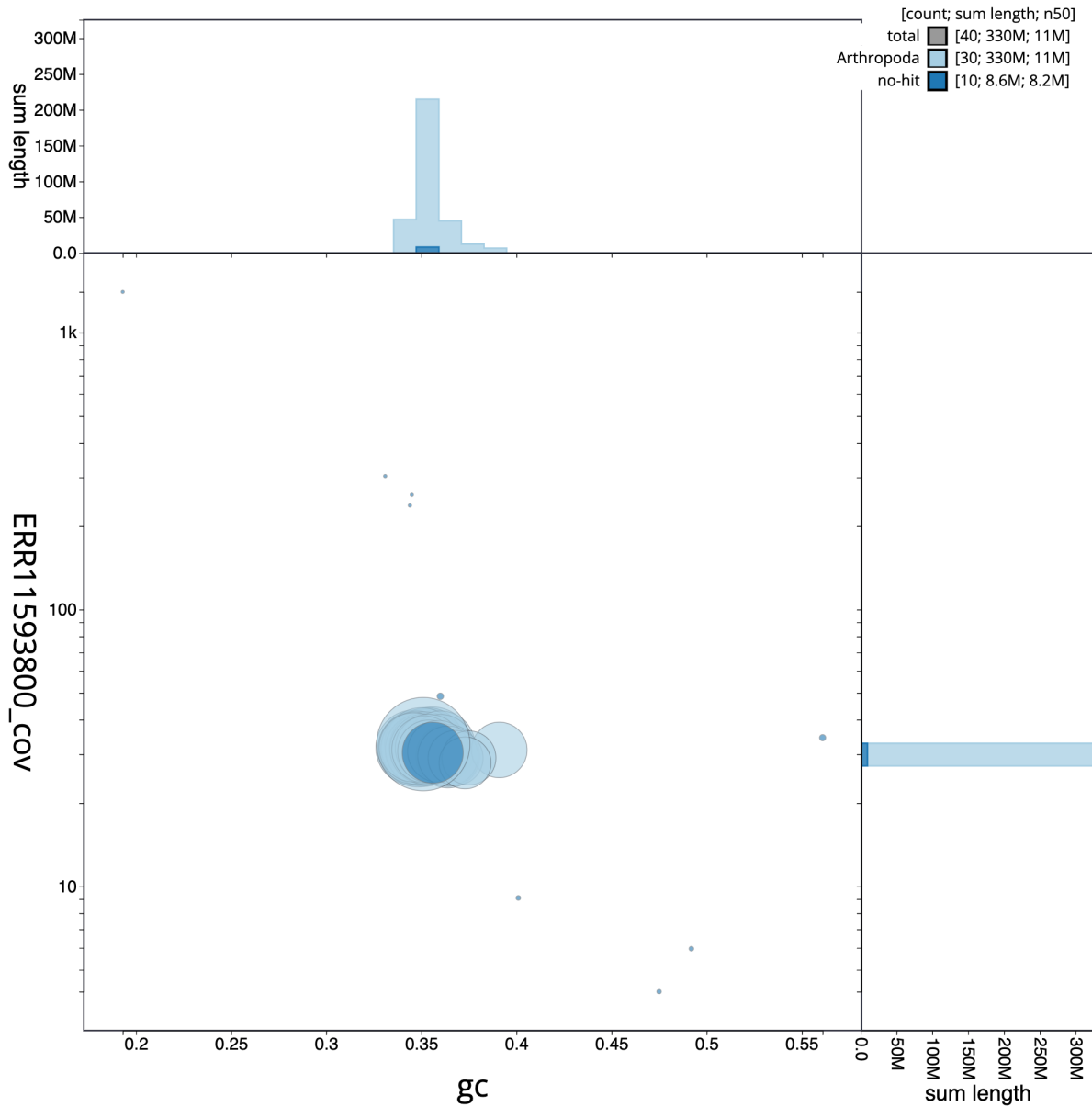


Figure 3. Genome assembly of *Timandra comae*, iTimComa1.1: BlobToolKit GC-coverage plot. Sequences are coloured by phylum. Circles are sized in proportion to sequence length. Histograms show the distribution of sequence length sum along each axis. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/iTimComa1_1/dataset/iTimComa1_1/blob.

DNA was purified by solid-phase reversible immobilisation (Strickland *et al.*, 2023): in brief, the method employs a 1.8X ratio of AMPure PB beads to sample to eliminate shorter fragments and concentrate the DNA. The concentration of the sheared and purified DNA was assessed using a Nanodrop spectrophotometer and Qubit Fluorometer and Qubit dsDNA High Sensitivity Assay kit. Fragment size distribution was evaluated by running the sample on the FemtoPulse system.

Protocols developed by the WSI Tree of Life laboratory are publicly available on protocols.io (Denton *et al.*, 2023b).

Sequencing

Pacific Biosciences HiFi circular consensus DNA sequencing libraries were constructed according to the manufacturers' instructions. DNA sequencing was performed by the Scientific Operations core at the WSI on a Pacific Biosciences SEQUEL

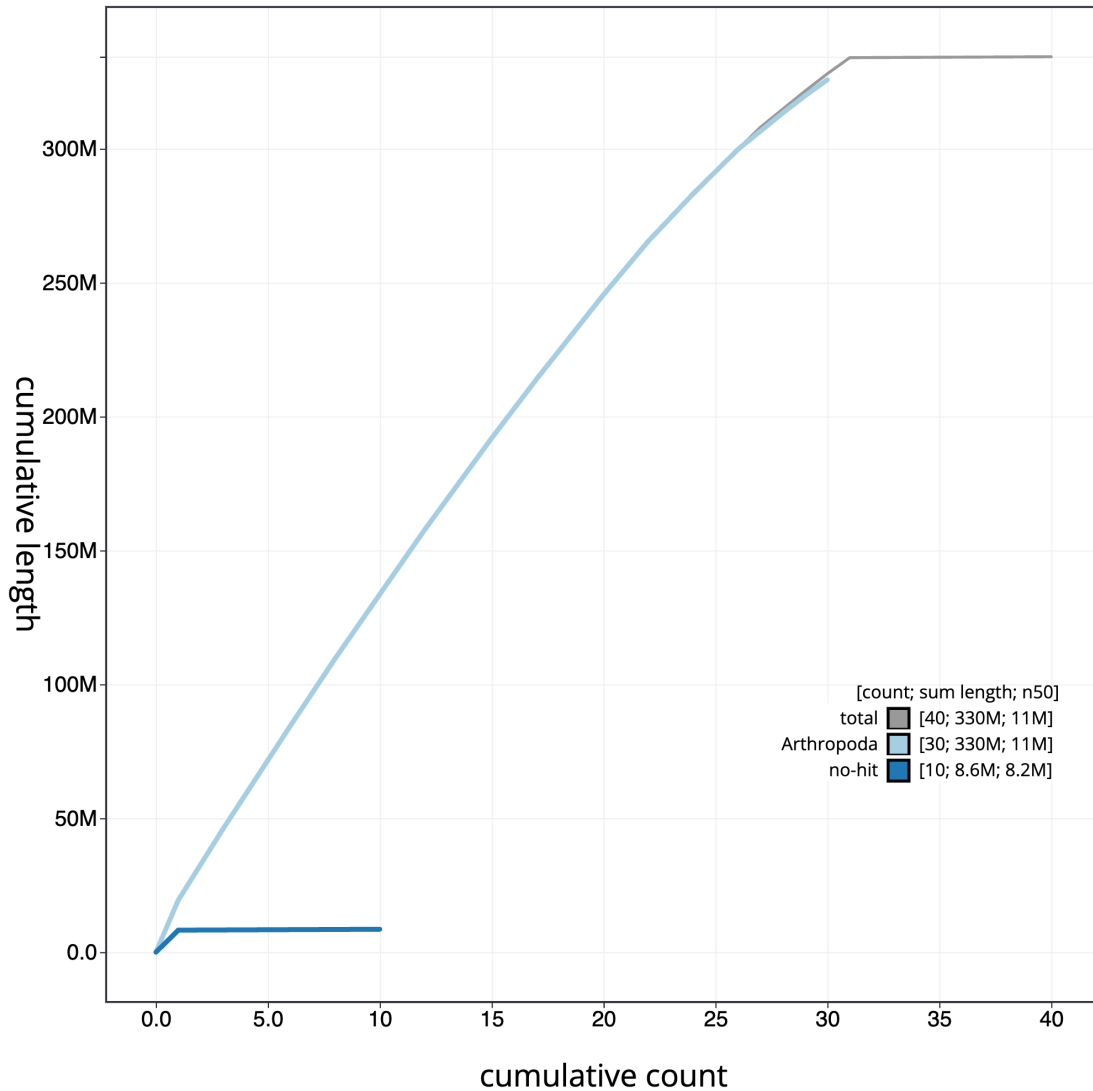


Figure 4. Genome assembly of *Timandra comae*, iTimComa1.1: BlobToolKit cumulative sequence plot. The grey line shows cumulative length for all sequences. Coloured lines show cumulative lengths of sequences assigned to each phylum using the buscogenes taxrule. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/iTimComa1_1/dataset/iTimComa1_1/cumulative.

II instrument. Hi-C data were also generated from remaining head and thorax tissue of iTimComa1 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 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 (Abdennur & Mirny, 2020). To assess the assembly

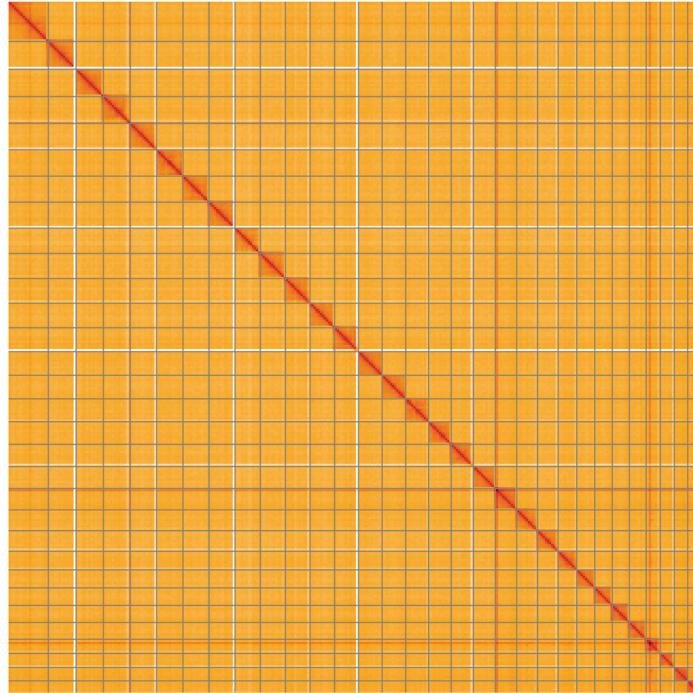


Figure 5. Genome assembly of *Timandra comae*, iTimComa1.1: Hi-C contact map of the iTimComa1.1 assembly, visualised using HiGlass. Chromosomes are shown in order of size from left to right and top to bottom. An interactive version of this figure may be viewed at https://genome-note-higlass.tol.sanger.ac.uk/l/?d=M7k_mdjfQq-QeuG4VyopUg.

Table 2. Chromosomal pseudomolecules in the genome assembly of *Timandra comae*, iTimComa1.

INSDC accession	Chromosome	Length (Mb)	GC%
OY292354.1	1	13.43	35.0
OY292355.1	2	13.18	35.5
OY292356.1	3	12.97	35.5
OY292357.1	4	12.74	34.5
OY292358.1	5	12.68	35.0
OY292359.1	6	12.57	35.0
OY292360.1	7	12.56	35.5
OY292361.1	8	12.19	34.5
OY292362.1	9	12.13	35.0
OY292363.1	10	11.92	35.0
OY292364.1	11	11.84	35.0
OY292365.1	12	11.49	35.0
OY292366.1	13	11.43	35.0
OY292367.1	14	11.39	34.5
OY292368.1	15	11.06	35.0

INSDC accession	Chromosome	Length (Mb)	GC%
OY292369.1	16	10.9	35.5
OY292370.1	17	10.68	36.0
OY292371.1	18	10.62	34.5
OY292372.1	19	10.32	36.0
OY292373.1	20	10.06	35.5
OY292374.1	21	9.91	35.0
OY292375.1	22	8.96	36.5
OY292376.1	23	8.71	35.5
OY292377.1	24	8.47	36.0
OY292378.1	25	8.24	35.5
OY292379.1	26	8.1	36.5
OY292380.1	27	6.84	39.0
OY292381.1	28	6.72	37.0
OY292382.1	29	6.49	37.5
OY292383.1	30	5.98	37.5
OY292353.1	Z	19.46	35.0
OY292384.1	MT	0.02	19.5

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:

- 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: *Timandra comae* (blood-vein). Accession number PRJEB63432; <https://identifiers.org/ena.embl/PRJEB63432> (Wellcome Sanger Institute, 2023). The genome sequence is released openly for reuse. The *Timandra comae* 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.7125292>.

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 Management, Samples and Laboratory team are listed here: <https://doi.org/10.5281/zenodo.10066175>.

Members of Wellcome Sanger Institute Scientific Operations: Sequencing Operations are listed here: <https://doi.org/10.5281/zenodo.10043364>.

Members of the Wellcome Sanger Institute Tree of Life Core Informatics team are listed here: <https://doi.org/10.5281/zenodo.10066637>.

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>.

Table 3. Software tools: versions and sources.

Software tool	Version	Source
BlobToolKit	4.2.1	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
Merqury	MerquryFK	https://github.com/thegenemyers/MERQURY.FK
MitoHiFi	2	https://github.com/marcelauliano/MitoHiFi
PretextView	0.2	https://github.com/wtsi-hpag/PretextView
purge_dups	1.2.3	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	yahs-1.1.91eebc2	https://github.com/c-zhou/yahs

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](#)
- Bates A, Clayton-Lucey I, Howard C: **Sanger Tree of Life HMW DNA Fragmentation: Diagenode Megaruptor® 3 for LI PacBio.** *Protocols.io.* 2023.
[Publisher Full Text](#)
- 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](#)
- Denton A, Oatley G, Cornwell C, et al.: **Sanger Tree of Life Sample Homogenisation: PowerMash.** *protocols.io.* 2023a.
[Publisher Full Text](#)
- Denton A, Yatsenko H, Jay J, et al.: **Sanger Tree of Life Wet Laboratory Protocol Collection V.1.** *protocols.io.* 2023b.
[Publisher 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](#)
- 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](#)
- 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](#)
- Jay J, Yatsenko H, Narváez-Gómez JP, et al.: **Sanger Tree of Life Sample Preparation: Triage and Dissection.** *protocols.io.* 2023.
[Publisher Full Text](#)
- Kerpedjiev P, Abdennur N, Lekschas F, et al.: **HiGlass: web-based visual exploration and analysis of genome interaction maps.** *Genome Biol.* 2018; **19**(1): 125.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Kurze S, Heinken T, Fartmann T: **Nitrogen enrichment in host plants increases the mortality of common Lepidoptera species.** *Oecologia.* 2018; **188**(4): 1227–1237.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Manni M, Berkeley MR, Seppy 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](#)
- Oatley G, Denton A, Howard C: **Sanger Tree of Life HMW DNA Extraction: Automated MagAttract v.2.** *Protocols.io.* 2023.
[Publisher Full Text](#)
- Öunap E, Viidalepp J, Saarma U: **Phylogenetic evaluation of the taxonomic status of *Timandra griseata* and *T. comae* (Lepidoptera: Geometridae: Sterrhinae).** *Eur J Entomol.* 2005; **102**(4): 607–615.
[Publisher 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.: **Merquy: Reference-free quality, completeness, and phasing assessment for genome assemblies.** *Genome Biol.* 2020; **21**(1): 245.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- 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](#)
- Skinner B, Wilson D: **Colour Identification Guide to the Moths of the British Isles.** Leiden, Netherlands: Brill, 2009.
- Strickland M, Cornwell C, Howard C: **Sanger Tree of Life Fragmented DNA clean up: Manual SPRI.** *protocols.io.* 2023.
[Publisher Full Text](#)
- Surana P, Muffato M, Qi G: **sanger-tol/readmapping: sanger-tol/readmapping v1.1.0 - Hebridean Black (1.1.0).** *Zenodo.* 2023a.
[Publisher Full Text](#)
- Surana P, Muffato M, Sadasivan Baby C: **sanger-tol/genomenote (v1.0.dev).** *Zenodo.* 2023b.
[Publisher Full Text](#)
- Uliano-Silva M, Ferreira JGRN, Krashennikova 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](#)
- Waring P, Townsend M, Lewington R: **Field Guide to the Moths of Great Britain and Ireland: Third Edition.** Bloomsbury Wildlife Guides, 2017.
[Reference Source](#)
- Wellcome Sanger Institute: **The genome sequence of the Blood-vein moth, *Timandra comae* Schmidt, 1931.** European Nucleotide Archive. [dataset], accession number PRJEB63432, 2013.
- 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 01 October 2024

<https://doi.org/10.21956/wellcomeopenres.23328.r98736>

© 2024 Mutanen M. 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.



Marko Mutanen 

University of Oulu, Oulu, Finland

This article publishes a genome sequence for a looper (Geometridae) moth *Timandra comae*. *Timandra comae* belong to the subfamily Sterrhinae and within it a group that appears somewhat distinct from other Sterrhinae in Europe, except for a very close *T. griseata* which is so close to *T. comae* that their status as distinct species have sometimes been questioned. Publishing the genome therefore helps in studying this question in the future. It should be noted that the name of *T. griseata* is misspelled in the article's Background section (as griselata).

The analyzed specimen is a male.

The article makes statements of 1) the two above-mentioned species being separable in their DNA barcodes and 2) that *T. griseata* is Scandinavian. Both cannot be true, because the cluster of *griseata* is found in UK by many specimens in BOLD.

The article undoubtedly lists reported food plants of *T. comae* as mentioned in literature, but there tend to be lots of errors, and in this case I would strongly question *Atriplex* spp. being among the food plants.

As shown in the snailplot graph of Figure 2 and other graphs, the published genome is clearly of high standard. This is indicated e.g. by high BUSCO coverage (only 1.2% missing) and scaffold statistics that indicate high quality.

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: Lepidoptera diversity, taxonomy, phylogenetics, DNA barcoding, phylogenomics

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 11 September 2024

<https://doi.org/10.21956/wellcomeopenres.23328.r98735>

© 2024 Lucek K. 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.



Kay Lucek 

University of Neuchâtel, Neuchâtel, Switzerland

The authors present the chromosome level genome assembly of a male specimen of the blood vein moth *Timandra comae*. The assembly consists of 31 chromosomes including the Z chromosome. The assembly is highly complete as revealed by the high BUSCO score but not fully phased. Sequencing and genome assembly follow the current state of the art and use established methods.

Overall, the presented assembly will be of great value to study genome architecture as well as genome size evolution in Lepidoptera. I particularly appreciate the biological background given for this species, where the reference genome provides the opportunity to study the range expansion of this species in the UK and elsewhere.

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: Speciation 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 09 September 2024

<https://doi.org/10.21956/wellcomeopenres.23328.r96299>

© 2024 Backström N. 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.



Niclas Backström 

Uppsala University, Uppsala, Sweden

Minor comments:

Background, 2nd paragraph. "its having" sounds a bit odd - consider editing

Background, 4th paragraph. Unclear what "chromosomally complete" really means. Consider editing.

GSReport, 2nd paragraph. Here it is stated that the assembly contains 39 scaffolds. The number is the same as in table 1, but in Figure 2, 3 and 4 it is stated that the assembly consists of 40 scaffolds. Maybe I misunderstood something, but check if it is correct.

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 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.
