Check for updates

DATA NOTE

The genome sequence of the silver stretch spider, *Tetragnatha montana* (Simon, 1874) (Araneae: Tetragnathidae) [version 1; peer review: awaiting peer review]

Alistair P. McGregor[1], Lauren Sumner-Rooney[2], Richard Burkmar[3], Anna Schoenauer[4],
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

[1]Department of Biosciences, Durham University, Durham, England, UK
[2]Museum für Naturkunde, Leibniz-Institut für Evolutions- und Biodiversitätsforschung, Berlin, Germany
[3]UK Centre for Ecology & Hydrology, Bailrigg, England, UK
[4]Oxford Brookes University, Oxford, England, UK

v1 **First published:** 22 May 2024, **9**:288
https://doi.org/10.12688/wellcomeopenres.21782.1
Latest published: 22 May 2024, **9**:288
https://doi.org/10.12688/wellcomeopenres.21782.1

Abstract
We present a genome assembly from an individual female *Tetragnatha montana* (the silver stretch spider; Arthropoda; Arachnida; Araneae; Tetragnathidae). The genome sequence is 784.7 megabases in span. Most of the assembly is scaffolded into 13 chromosomal pseudomolecules, including the X sex chromosome. The mitochondrial genome has also been assembled and is 15.49 kilobases in length.

Keywords
Tetragnatha montana, silver stretch spider, genome sequence, chromosomal, Araneae

This article is included in the Tree of Life gateway.

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.

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

Author roles: McGregor AP: Investigation, Resources, Writing – Original Draft Preparation, Writing – Review & Editing; **Sumner-Rooney L**: Investigation, Resources; **Burkmar R**: Investigation, Resources; **Schoenauer A**: Investigation, Resources;

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.

How to cite this article: McGregor AP, Sumner-Rooney L, Burkmar R *et al.* **The genome sequence of the silver stretch spider, *Tetragnatha montana* (Simon, 1874) (Araneae: Tetragnathidae) [version 1; peer review: awaiting peer review]** Wellcome Open Research 2024, **9**:288 https://doi.org/10.12688/wellcomeopenres.21782.1

First published: 22 May 2024, **9**:288 https://doi.org/10.12688/wellcomeopenres.21782.1

Species taxonomy
Eukaryota; Opisthokonta; Metazoa; Eumetazoa; Bilateria; Protostomia; Ecdysozoa; Panarthropoda; Arthropoda; Chelicerata; Arachnida; Araneae; Araneomorphae; Entelegynae; Orbiculariae; Araneoidea; Tetragnathidae; *Tetragnatha*; *Tetragnatha montana* (Simon, 1874) (NCBI:txid500645).

Background
The tetragnathid *Tetragnatha montana*, commonly known as the silver stretch spider, was first described by Eugène Simon in 1874 and first recorded in the UK in 1895 (British Arachnological Society, 2024; Simon, 1874; World Spider Catalog, 2024).

T. montana is one of six species of *Tetragnatha* found in the UK; alongside *T. extensa* (common long-jawed orbweb spider), *T. pinicola*, *T. obtusa*, *T. nigrita* and *T. striata* (striped long-jawed orbweb spider) (Bee *et al.*, 2020). Tetragnathids, or long-jawed orbweb spiders, including *T. montana*, are characterised by their elongated chelicerae, legs, and bodies. *T. montana* is Palaearctic and is widespread in central and western Europe (World Spider Catalogue, 2024). *T. montana*, together with *T. extensa*, are the most common and widely distributed tetragnathids in the UK, although *T. montana* has a wider range and is more numerous in the south than the north (Bee *et al.*, 2020; British Arachnological Society, 2024).

T. montana spins simple orb webs in grass and bushes, usually near open water, and can also be found on trees (Bee *et al.*, 2020; British Arachnological Society, 2024). Like other species of this genus, *T. montana* pose cryptically with their legs and bodies stretched out along grass, stems, or branches to hide when alarmed.

The *T. montana* genome was sequenced as part of the Darwin Tree of Life Project. This chromosomal-level assembly of the *T. montana* genome is from a female specimen (specimen ID SAMEA7520006) collected on 29/7/2019 from Wytham Woods near Oxford, UK. The *T. montana* genome will provide new insights into the biology of this spider, for example how sperm competition may have shaped the mating system of this and other *Tetragnatha* species (e.g. Yoward & Oxford, 2014). The genome of this tetragnathid will also provide a valuable comparison to the genomes of spiders in other families to help better understand spider and arachnid genome evolution (e.g. Aase-Remedios *et al.*, 2023).

Genome sequence report
The genome was sequenced from a female *Tetragnatha montana* (Figure 1) collected from Wytham Woods, Oxfordshire, UK (51.77, –1.33). A total of 26-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 39 missing joins or mis-joins and removed 13 haplotypic duplications, reducing the assembly length by 0.83% and the scaffold number by 11.83%, and increasing the scaffold N50 by 0.76%.

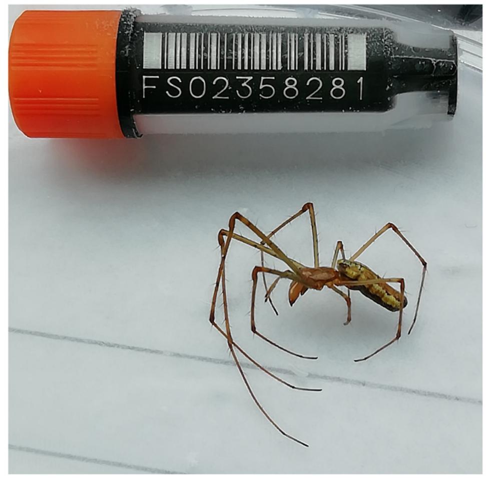

Figure 1. Photograph of the *Tetragnatha montana* (qqTetMont2) specimen used for genome sequencing.

The final assembly has a total length of 784.7 Mb in 81 sequence scaffolds with a scaffold N50 of 61.1 Mb (Table 1). The snail plot 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 (99.3%) of the assembly sequence was assigned to 13 chromosomal-level scaffolds, representing 11 autosomes and the X sex chromosome. Chromosome-scale scaffolds confirmed by the Hi-C data are named in order of size (Figure 5; Table 2). Chromosomes X_1 and X_2 were assigned by synteny to *Metellina segmentata* (GCA_947359465.1) (Henriques *et al.*, 2023). 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 62.0 with *k*-mer completeness of 100.0%, and the assembly has a BUSCO v completeness of 97.6% (single = 91.8%, duplicated = 5.8%), using the arachnida_odb10 reference set (n = 2,934).

Metadata for specimens, BOLD barcode results, spectra estimates, sequencing runs, contaminants and pre-curation assembly statistics are given at https://links.tol.sanger.ac.uk/species/500645.

Methods
Sample acquisition and nucleic acid extraction
A female *Tetragnatha montana* (specimen ID Ox000088, ToLID qqTetMont2) was collected from Wytham Great Wood, Oxfordshire (biological vice-county Berkshire), UK (latitude 51.77, longitude –1.33) on 2019-07-29 by netting. The specimen was collected and identified by Alistair McGregor, Lauren Sumner-Rooney, Richard Burkmar and Anna Schoenauer

Table 1. Genome data for *Tetragnatha montana*, qqTetMont2.1.

Project accession data		
Assembly identifier	qqTetMont2.1	
Species	*Tetragnatha montana*	
Specimen	qqTetMont2	
NCBI taxonomy ID	500645	
BioProject	PRJEB63621	
BioSample ID	SAMEA7520006	
Isolate information	qqTetMont2, female: cephalothorax (PacBio DNA and Illumina Hi-C sequencing)	
Assembly metrics*		***Benchmark***
Consensus quality (QV)	62.0	*≥ 50*
k-mer completeness	100.0%	*≥ 95%*
BUSCO**	C:97.6%[S:91.8%,D:5.8%], F:0.6%,M:1.8%,n:2,934	*C ≥ 95%*
Percentage of assembly mapped to chromosomes	99.3%	*≥ 95%*
Sex chromosomes	XX	*localised homologous pairs*
Organelles	Mitochondrial genome: 15.49 kb	*complete single alleles*
Raw data accessions		
PacificBiosciences Sequel IIe	ERR11641071	
Hi-C Illumina	ERR11641146	
Genome assembly		
Assembly accession	GCA_963680715.1	
Accession of alternate haplotype	GCA_963680735.1	
Span (Mb)	784.7	
Number of contigs	333	
Contig N50 length (Mb)	4.7	
Number of scaffolds	81	
Scaffold N50 length (Mb)	61.1	
Longest scaffold (Mb)	64.62	

* 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 arachnida_odb10 BUSCO set using version v5.4.3. 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/Tetragnatha_montana/dataset/GCA_963680715.1/busco.

(Oxford Brookes University). After identification and imaging at Oxford Brookes University, the spider was barcoded and flash-frozen in liquid nitrogen at the University of Oxford.

The workflow for high molecular weight (HMW) DNA extraction at the Wellcome Sanger Institute (WSI) Tree of Life Core Laboratory includes a sequence of core procedures: sample preparation; sample homogenisation, DNA extraction, fragmentation, and clean-up. The sample was prepared for DNA extraction at the WSI Tree of life Core Laboratory: the qqTetMont2 sample was weighed and dissected on dry ice (Jay *et al.*, 2023). Tissue from the cephalothorax was

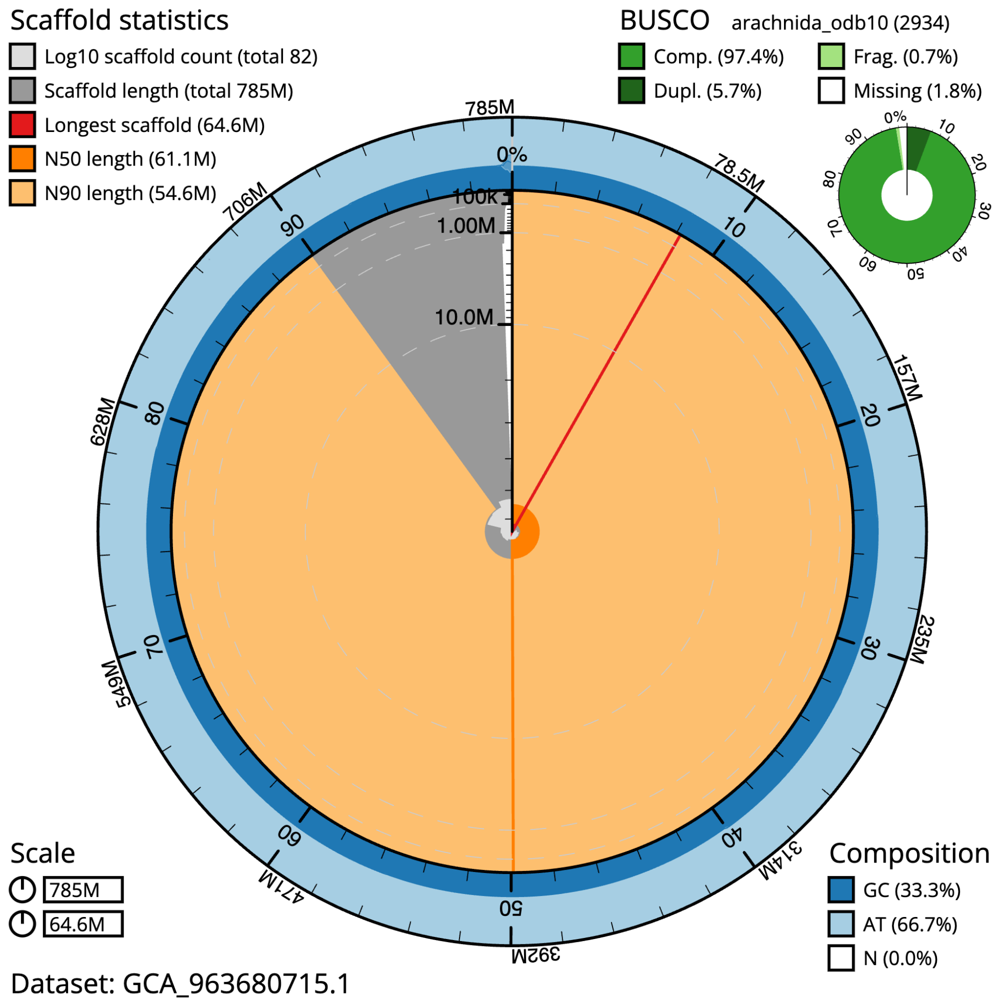

Figure 2. Genome assembly of *Tetragnatha montana*, qqTetMont2.1: metrics. The BlobToolKit snail plot shows N50 metrics and BUSCO gene completeness. The main plot is divided into 1,000 size-ordered bins around the circumference with each bin representing 0.1% of the 784,702,176 bp assembly. The distribution of sequence lengths is shown in dark grey with the plot radius scaled to the longest sequence present in the assembly (64,617,543 bp, shown in red). . Orange and pale-orange arcs show the N50 and N90 sequence lengths (61,119,082 and 54,634,059 bp), respectively. The pale grey spiral shows the cumulative sequence 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 arachnida_odb10 set is shown in the top right. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/Tetragnatha_montana/dataset/GCA_963680715.1/snail.

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 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 IIe instrument. Hi-C data were also generated from remaining

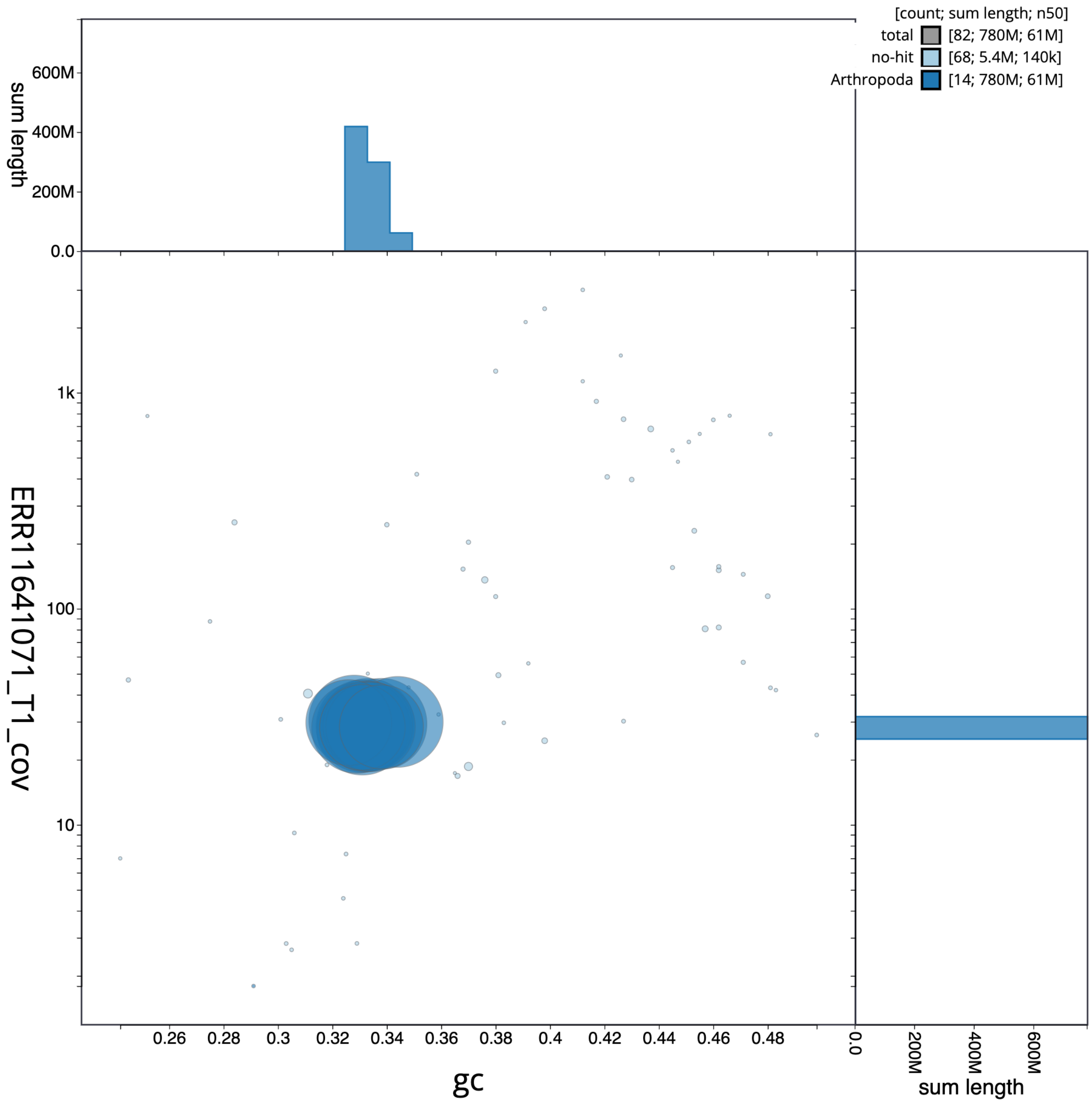

Figure 3. Genome assembly of *Tetragnatha montana*, qqTetMont2.1: BlobToolKit GC-coverage plot. Sequences are coloured by phylum. Circles are sized in proportion to sequence length. Histograms show the distribution of sequence length sum along each axis. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/Tetragnatha_montana/dataset/GCA_963680715.1/blob.

cephalothorax tissue of qqTetMont2 using the Arima2 kit and sequenced on the Illumina NovaSeq 6000 instrument.

Genome assembly and curation

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 using the TreeVal pipeline (Pointon *et al.*, 2023). Manual curation was performed using JBrowse2 (Diesh *et al.*, 2023), HiGlass (Kerpedjiev *et al.*, 2018) and PretextView (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.

Final assembly evaluation

The final assembly was post-processed and evaluated with the three Nextflow (Di Tommaso *et al.*, 2017) DSL2 pipelines “sanger-tol/readmapping” (Surana *et al.*, 2023a), “sanger-tol/genomenote” (Surana *et al.*, 2023b), and “sanger-tol/blobtoolkit” (Muffato *et al.*, 2024). The pipeline sanger-tol/readmapping

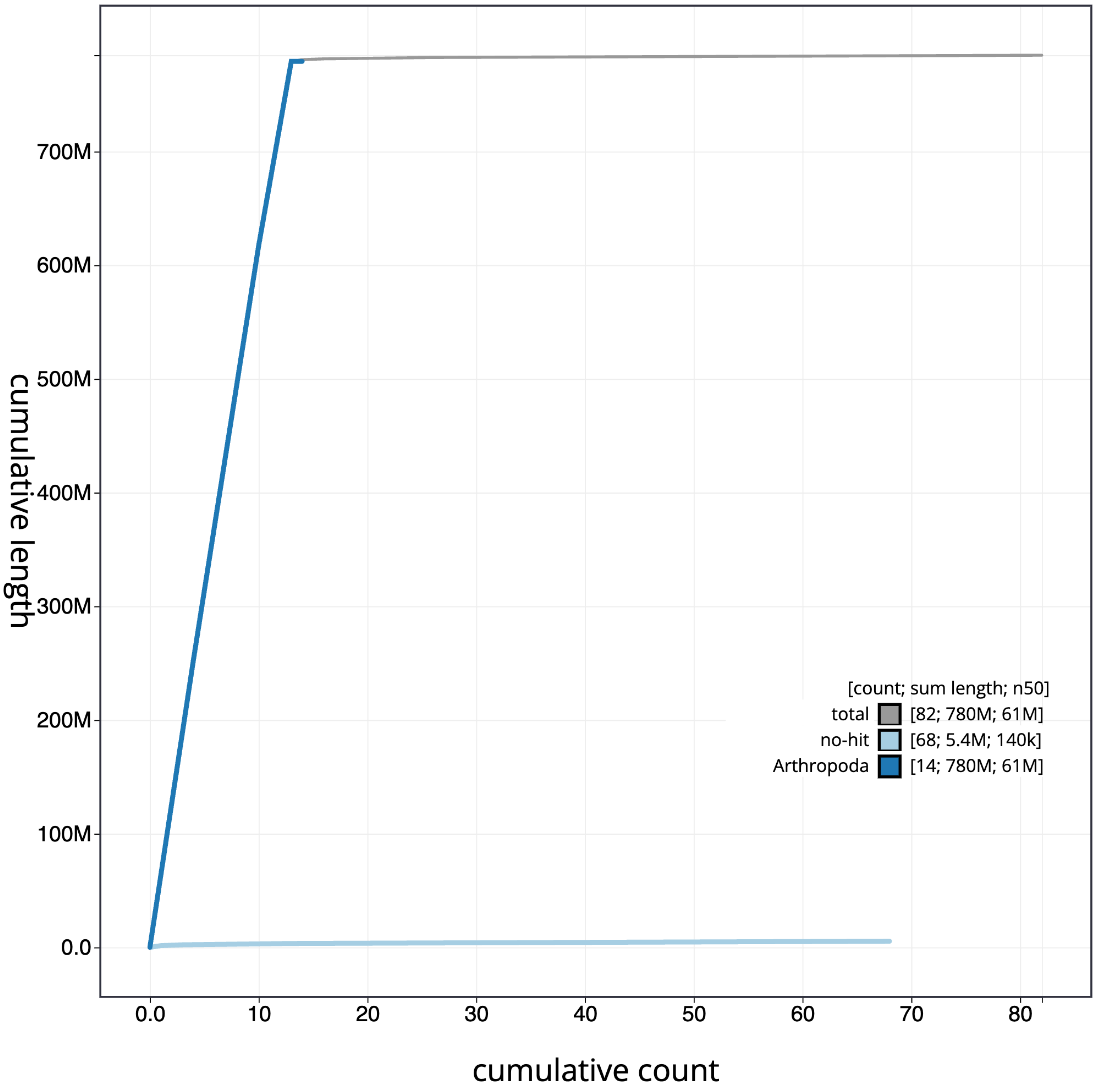

Figure 4. Genome assembly of *Tetragnatha montana*, qqTetMont2.1: BlobToolKit cumulative sequence plot. The grey line shows cumulative length for all sequences. Coloured lines show cumulative lengths of sequences assigned to each phylum using the buscogenes taxrule. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/Tetragnatha_montana/dataset/GCA_963680715.1/cumulative.

aligns the Hi-C reads with bwa-mem2 (Vasimuddin *et al.*, 2019) and combines the alignment files with SAMtools (Danecek *et al.*, 2021). The sanger-tol/genomenote pipeline transforms the Hi-C alignments into a contact map with BEDTools (Quinlan & Hall, 2010) and the Cooler tool suite (Abdennur & Mirny, 2020), which is then visualised with HiGlass (Kerpedjiev *et al.*, 2018). It also provides statistics about the assembly with the NCBI datasets (Sayers *et al.*, 2024) report, computes *k*-mer completeness and QV consensus quality values with FastK and MerquryFK, and a completeness assessment with BUSCO (Manni *et al.*, 2021).

The sanger-tol/blobtoolkit pipeline is a Nextflow port of the previous Snakemake Blobtoolkit pipeline (Challis *et al.*, 2020). It aligns the PacBio reads with SAMtools and minimap2 (Li, 2018) and generates coverage tracks for regions of fixed size. In parallel, it queries the GoaT database (Challis *et al.*, 2023) to identify all matching BUSCO lineages to run BUSCO (Manni *et al.*, 2021). For the three domain-level BUSCO lineage, the pipeline aligns the BUSCO genes to the Uniprot Reference Proteomes database (Bateman *et al.*, 2023) with DIAMOND (Buchfink *et al.*, 2021) blastp. The genome is also split into chunks according to the density of the BUSCO genes from the closest taxonomically lineage, and each chunk is aligned to the Uniprot Reference Proteomes database with DIAMOND blastx. Genome sequences that have no hit are then chunked with seqtk and aligned to the NT database with blastn (Altschul *et al.*, 1990). All those outputs are combined with the blobtools suite into a blobdir for visualisation.

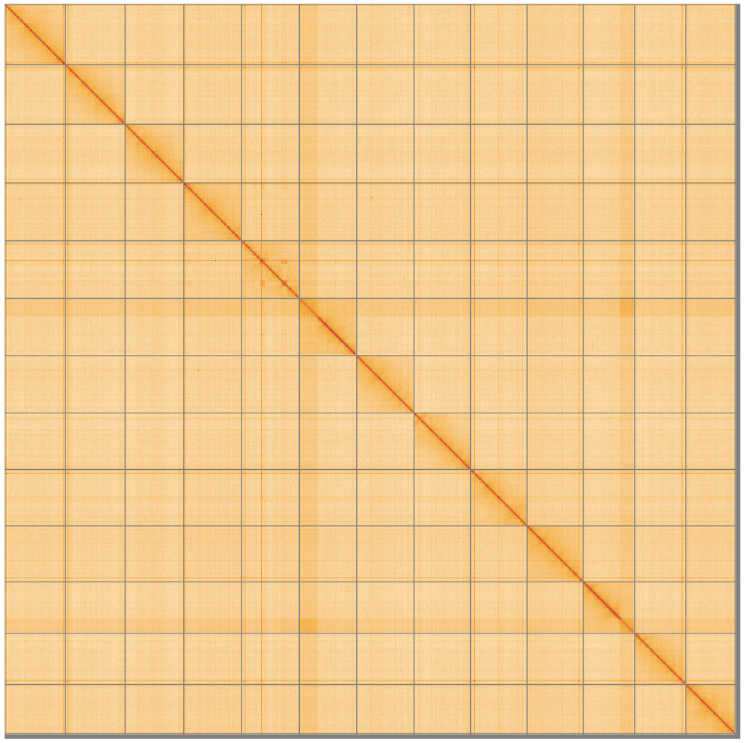

Figure 5. Genome assembly of *Tetragnatha montana*, qqTetMont2.1: Hi-C contact map of the qqTetMont2.1 assembly, visualised using HiGlass. Chromosomes are shown in order of size from left to right and top to bottom. An interactive version of this figure may be viewed at https://genome-note-higlass.tol.sanger.ac.uk/l/?d=FrlNoy89SoiJhaaO37ENNw.

Table 2. Chromosomal pseudomolecules in the genome assembly of *Tetragnatha montana*, qqTetMont2.

INSDC accession	Chromosome	Length (Mb)	GC%
OY796674.1	1	64.62	33.5
OY796675.1	2	63.81	33.0
OY796676.1	3	62.49	33.0
OY796677.1	4	61.76	34.0
OY796678.1	5	61.61	34.5
OY796680.1	6	61.12	33.0
OY796681.1	7	60.37	33.0
OY796682.1	8	60.21	33.5
OY796683.1	9	59.94	33.5
OY796685.1	10	54.63	33.0
OY796686.1	11	52.55	34.0
OY796687.1	MT	0.02	25.5
OY796684.1	X1	54.9	32.5
OY796679.1	X2	61.27	33.0

All three pipelines were developed using the nf-core tooling (Ewels *et al.*, 2020), use MultiQC (Ewels *et al.*, 2016), and make extensive use of the Conda package manager, the Bioconda initiative (Grüning *et al.*, 2018), the Biocontainers infrastructure (da Veiga Leprevost *et al.*, 2017), and the Docker (Merkel, 2014) and Singularity (Kurtzer *et al.*, 2017) containerisation solutions.

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.

Table 3. Software tools: versions and sources.

Software tool	Version	Source
BEDTools	2.30.0	https://github.com/arq5x/bedtools2
Blast	2.14.0	ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/
BlobToolKit	4.3.7	https://github.com/blobtoolkit/blobtoolkit
BUSCO	5.4.3	https://gitlab.com/ezlab/busco
BUSCO	5.4.3 and 5.5.0	https://gitlab.com/ezlab/busco
bwa-mem2	2.2.1	https://github.com/bwa-mem2/bwa-mem2
Cooler	0.8.11	https://github.com/open2c/cooler
DIAMOND	2.1.8	https://github.com/bbuchfink/diamond
fasta_windows	0.2.4	https://github.com/tolkit/fasta_windows
FastK	427104ea91c78c3b8b8b49f1a7d6bbeaa869ba1c	https://github.com/thegenemyers/FASTK
GoaT CLI	0.2.5	https://github.com/genomehubs/goat-cli
Hifiasm	0.19.5-r587	https://github.com/chhylp123/hifiasm
HiGlass	1.11.6	https://github.com/higlass/higlass
HiGlass	44086069ee7d4d3f6f3f0012569789ec138f42b84aa44357826c0b6753eb28de	https://github.com/higlass/higlass
MerquryFK	d00d98157618f4e8d1a9190026b19b471055b22e	https://github.com/thegenemyers/MERQURY.FK
MitoHiFi	3	https://github.com/marcelauliano/MitoHiFi
MultiQC	1.14, 1.17, and 1.18	https://github.com/MultiQC/MultiQC
NCBI Datasets	15.12.0	https://github.com/ncbi/datasets
Nextflow	23.04.0-5857	https://github.com/nextflow-io/nextflow
PretextView	0.2	https://github.com/wtsi-hpag/PretextView
purge_dups	1.2.5	https://github.com/dfguan/purge_dups
samtools	1.16.1, 1.17, and 1.18	https://github.com/samtools/samtools
sanger-tol/genomenote	1.1.1	https://github.com/sanger-tol/genomenote
sanger-tol/readmapping	1.2.1	https://github.com/sanger-tol/readmapping
Seqtk	1.3	https://github.com/lh3/seqtk
Singularity	3.9.0	https://github.com/sylabs/singularity
TreeVal	1.0.0	https://github.com/sanger-tol/treeval
YaHS	1.2a.2	https://github.com/c-zhou/yahs

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: *Tetragnatha montana* (silver stretch spider). Accession number PRJEB63621; https://identifiers.org/ena.embl/PRJEB63621 (Wellcome Sanger Institute, 2023). The genome sequence is released openly for reuse. The *Tetragnatha montana* 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.

Acknowledgements

We thank the University of Oxford for permission to collect in Wytham Woods and Peter Holland and Liam Crowley for assistance in processing this specimen for sequencing. We also thank Grace Blakeley and Luis Baudouin-Gonzalez for assisting with collecting.

References

Aase-Remedios ME, Janssen R, Leite DJ, *et al.*: **Evolution of the spider homeobox gene repertoire by tandem and Whole Genome Duplication.** *Mol Biol Evol.* 2023; **40**(12): msad239.
PubMed Abstract | Publisher Full Text | Free Full Text

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

Altschul SF, Gish W, Miller W, *et al.*: **Basic local alignment search tool.** *J Mol Biol.* 1990; **215**(3): 403–410.
PubMed Abstract | Publisher Full Text

Bateman A, Martin MJ, Orchard S, *et al.*: **UniProt: the universal protein knowledgebase in 2023.** *Nucleic Acids Res.* 2023; **51**(D1): D523–D531.
PubMed Abstract | Publisher Full Text | Free Full Text

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

Bee L, Oxford G, Smith H: **Britain's spiders - a field guide.** 2nd ed. Princeton University Press, 2020.
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

British Arachnological Society: **Summary for *Tetragnatha montana* (Araneae).** *Spider and Harvestman recording scheme website*. 2024; [Accessed 5 April 2024].
Reference Source

Buchfink B, Reuter K, Drost HG: **Sensitive protein alignments at tree-of-life scale using DIAMOND.** *Nat Methods.* 2021; **18**(4): 366–368.
PubMed Abstract | Publisher Full Text | Free Full Text

Challis R, Kumar S, Sotero-Caio C, *et al.*: **Genomes on a Tree (GoaT): a versatile, scalable search engine for genomic and sequencing project metadata across the eukaryotic tree of life [version 1; peer review: 2 approved].** *Wellcome Open Res.* 2023; **8**: 24.
PubMed Abstract | Publisher Full Text | Free Full Text

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

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

da Veiga Leprevost F, Grüning BA, Alves Aflitos S, *et al.*: **BioContainers: an open-source and community-driven framework for software standardization.** *Bioinformatics.* 2017; **33**(16): 2580–2582.
PubMed Abstract | Publisher Full Text | Free Full Text

Danecek P, Bonfield JK, Liddle J, *et al.*: **Twelve years of SAMtools and BCFtools.** *GigaScience.* 2021; **10**(2): giab008.
PubMed Abstract | Publisher Full Text | Free Full Text

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.** *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

Diesh C, Stevens GJ, Xie P, *et al.*: **JBrowse 2: a modular genome browser with views of synteny and structural variation.** *Genome Biol.* 2023; **24**(1): 74.
PubMed Abstract | Publisher Full Text | Free Full Text

Ewels P, Magnusson M, Lundin S, *et al.*: **MultiQC: summarize analysis results for multiple tools and samples in a single report.** *Bioinformatics.* 2016; **32**(19): 3047–3048.
PubMed Abstract | Publisher Full Text | Free Full Text

Ewels PA, Peltzer A, Fillinger S, *et al.*: **The nf-core framework for community-curated bioinformatics pipelines.** *Nat Biotechnol.* 2020; **38**(3): 276–278.
PubMed Abstract | Publisher Full Text

Grüning B, Dale R, Sjödin A, *et al.*: **Bioconda: sustainable and comprehensive software distribution for the life sciences.** *Nat Methods.* 2018; **15**(7): 475–476.
PubMed Abstract | Publisher Full Text

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

Henriques S, Sivell O, Natural History Museum Genome Acquisition Lab, *et al.*: **The genome sequence of the Autumn Spider, *Metellina segmentata* (Clerck, 1757) [version 1; peer review: 1 approved].** *Wellcome Open Res.* 2023; **8**: 221.
Publisher 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

Kurtzer GM, Sochat V, Bauer MW: **Singularity: scientific containers for mobility of compute.** *PLoS One.* 2017; **12**(5): e0177459.
PubMed Abstract | Publisher Full Text | Free Full Text

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

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

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

Muffato M, Butt Z, Challis R, *et al.*: **sanger-tol/blobtoolkit: v0.3.0 – Poliwag.** 2024.
Publisher 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

Pointon DL, Eagles W, Sims Y, *et al.*: **sanger-tol/treeval v1.0.0 – Ancient Atlantis**. 2023.
Publisher Full Text

Quinlan AR, Hall IM: **BEDTools: a flexible suite of utilities for comparing genomic features.** *Bioinformatics.* 2010; **26**(6): 841–842.
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

Sayers EW, Cavanaugh M, Clark K, *et al.*: **GenBank 2024 update.** *Nucleic Acids Res.* 2024; **52**(D1): D134–D137.
PubMed Abstract | Publisher Full Text | Free Full Text

Simon E: **Les arachnides de France. Paris 1.** 1874.
Reference Source

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, Krasheninnikova K, *et al.*: **MitoHiFi: a python pipeline for mitochondrial genome assembly from PacBio high fidelity reads.** *BMC Bioinformatics.* 2023; **24**(1): 288.
PubMed Abstract | Publisher Full Text | Free Full Text

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

Wellcome Sanger Institute: **The genome sequence of the silver stretch spider, *Tetragnatha montana* (Simon, 1874) (Araneae: Tetragnathidae).** European Nucleotide Archive, [dataset], accession number PRJEB63621, 2023.

World Spider Catalog: **World Spider Catalog, *World Spider Catalog. Version 25.0. natural history museum bern***. 2024.
Publisher Full Text

Yoward PJ, Oxford GS: **Evolutionary implications of sperm competition in the silver stretch spider, *Tetragnatha montana* simon (araneae: tetragnathidae).** *Arachnology.* 2014; **16**(5): 175–182.
Publisher Full Text

Zhou C, McCarthy SA, Durbin R: **YaHS: yet another Hi-C scaffolding tool.** *Bioinformatics.* 2023; **39**(1): btac808.
PubMed Abstract | Publisher Full Text | Free Full Text