



DATA NOTE

The genome sequence of the centipede *Strigamia acuminata* (Leach, 1816) [version 1; peer review: 3 approved]

Gregory D. Edgecombe¹, Duncan Sivell¹,
Natural History Museum 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

¹Natural History Museum, London, England, UK

V1 First published: 21 Sep 2023, 8:420
<https://doi.org/10.12688/wellcomeopenres.19941.1>
Latest published: 21 Sep 2023, 8:420
<https://doi.org/10.12688/wellcomeopenres.19941.1>

Abstract

We present a genome assembly from an individual male *Strigamia acuminata* (centipede; Arthropoda; Chilopoda; Geophilomorpha; Geophilidae; Linotaeniinae). The genome sequence is 237.5 megabases in span. Most of the assembly is scaffolded into 11 chromosomal pseudomolecules, including the X and Y sex chromosomes. The mitochondrial genome has also been assembled and is 15.07 kilobases in length.

Keywords

Strigamia acuminata, centipede, genome sequence, chromosomal, Geophilomorpha



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

Open Peer Review

Approval Status



	1	2	3
version 1 21 Sep 2023	 view	 view	 view

1. **Jerome H.L Hui**, Chinese University of Hong Kong, Hong Kong, China
2. **Maurijn van der Zee** , Leiden University, Sylviusweg, The Netherlands
3. **Ken Kraaijeveld** , Vrije University Amsterdam, Amsterdam, The Netherlands

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: **Edgecombe GD:** Investigation, Resources, Writing – Original Draft Preparation, Writing – Review & Editing; **Sivell D:** 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) 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 Edgecombe GD *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: Edgecombe GD, Sivell D, Natural History Museum Genome Acquisition Lab *et al.* **The genome sequence of the centipede *Strigamia acuminata* (Leach, 1816) [version 1; peer review: 3 approved]** Wellcome Open Research 2023, 8:420 <https://doi.org/10.12688/wellcomeopenres.19941.1>

First published: 21 Sep 2023, 8:420 <https://doi.org/10.12688/wellcomeopenres.19941.1>

Species taxonomy

Eukaryota; Metazoa; Eumetazoa; Bilateria; Protostomia; Ecdysozoa; Panarthropoda; Arthropoda; Mandibulata; Myriapoda; Chilopoda; Pleurostigmophora; Epimorpha; Geophilomorpha; Geophilidae; Linotaeniinae; *Strigamia*; *Strigamia acuminata* (Leach, 1816) (NCBI:txid1255758).

Background

Strigamia acuminata is a western Palaearctic species distributed throughout most of continental Europe, from the Iberian Peninsula to the Caucasus and Volga basin, as well as Sicily, Crete and Great Britain (Bonato *et al.*, 2012; Bonato *et al.*, 2016; Bonato *et al.*, 2023). Integrative taxonomy suggests that European populations of *S. acuminata* form a complex of at least two species (Bonato *et al.*, 2023). In the UK, *S. acuminata* is known only from England and Wales, the largest numbers of records being from the southeast and east of England. The following occurrence data are summarised from Barber (2022). It is usually an inland species although some coastal records are known. It is collected throughout the year but has a patchy and unpredictable occurrence. Records are overwhelmingly in rural sites, with only a few urban occurrences. In Britain, as in Europe, *S. acuminata* is most common in deciduous woodland, and is less frequently found in mixed woodland and wetland. Its most typical microsites are under dead wood or in leaf litter. Most collections are from lowland sites (<200 m), but it has been found at up to 1,013 m in Snowdonia, north Wales.

Like the other two species of *Strigamia* in the UK, *S. acuminata* is distinguished in the field from most other British Geophilomorpha by its reddish colouration. Adults reach a body length of 30 mm in England and Wales but up to 40 mm in some European populations. A leg number of 37 to 41 pairs (versus 47 or more pairs) provides the simplest basis for its distinction from the other two British congeners, *S. crassipes* and *S. maritima* (Barber, 2008). Males have 37 or 39 leg pairs and females 41 in the UK; in other parts of the geographic range 39–43 pairs encompass the variation (Bonato *et al.*, 2023). As in other congeners, the ultimate legs of females are slender and those of males swollen. In Germany, *S. acuminata* has been observed to group feed on the millipede *Julus* (Weil, 1958).

Few whole genome sequences for centipedes have been generated, including those of *Strigamia maritima* (GCA_000239455.1) (Chipman *et al.*, 2014) and of *Lithobius niger* (GCA_023313725.1), *Rhysida immarginata* (GCA_023313115.1) and *Thereuonema tuberculata* (GCA_023159025.1) (So *et al.*, 2022). The MetaInvert database at Senckenberg Görlitz and the LOEWE Centre for Translational Biodiversity Genomics offers draft genome sequences for a broad spectrum of soil invertebrates, encompassing 19 Chilopoda, including *S. acuminata*, aiming to extend the taxonomic scope for soil metagenomic studies (Collins *et al.*, 2023).

The genome of *Strigamia acuminata* was sequenced as part of the Darwin Tree of Life Project, a collaborative effort to sequence all named eukaryotic species in the Atlantic

Archipelago of Britain and Ireland. Here we present a chromosomally complete genome sequence for *Strigamia acuminata*, based on one adult male specimen from Bookham Common, Surrey. The genome is of interest for comparison with the most comprehensively annotated centipede genome, and the first to be sequenced, the closely related *S. maritima* (Chipman *et al.*, 2014), which inhabits the littoral zone rather than woodland.

Genome sequence report

The genome was sequenced from one adult male *Strigamia acuminata* (Figure 1) collected from Bookham Common, Leatherhead, England (51.29, -0.38). A total of 98-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 99 missing joins or mis-joins and removed 23 haplotypic duplications, reducing the assembly length by 1.68% and the scaffold number by 27.06%, and increasing the scaffold N50 by 48.44%.

The final assembly has a total length of 237.5 Mb in 158 sequence scaffolds with a scaffold N50 of 19.3 Mb (Table 1). Most (94.83%) of the assembly sequence was assigned to 11 chromosomal-level scaffolds, representing 9 autosomes and the X and Y sex chromosomes. Chromosome-scale scaffolds confirmed by the Hi-C data are named in order of size

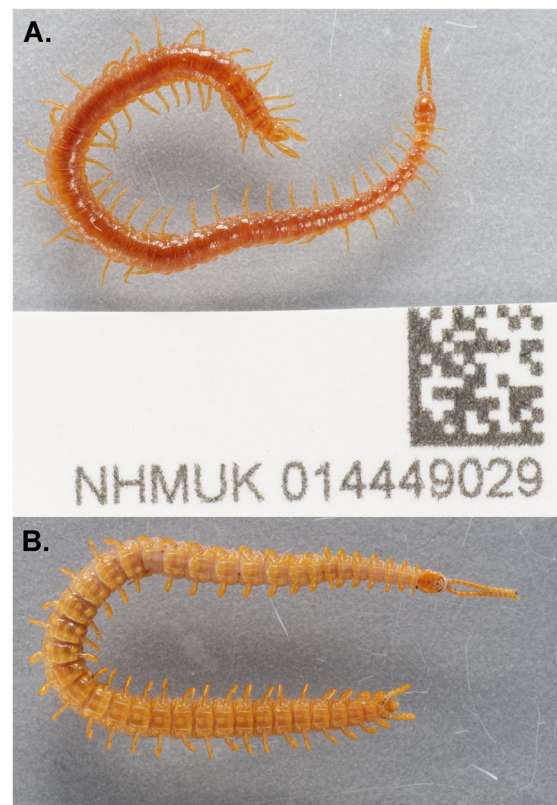


Figure 1. Photograph of the *Strigamia acuminata* specimen (NHMUK014449029) used for genome sequencing. **A**, dorsal view; **B**, ventral view.

Table 1. Genome data for *Strigamia acuminata*, qcStrAcum1.1.

Project accession data		
Assembly identifier	qcStrAcum1.1	
Species	<i>Strigamia acuminata</i>	
Specimen	qcStrAcum1	
NCBI taxonomy ID	1255758	
BioProject	PRJEB59944	
BioSample ID	SAMEA10800140	
Isolate information	qcStrAcum1, male (DNA sequencing and Hi-C data)	
Assembly metrics*		Benchmark
Consensus quality (QV)	59.6	≥ 50
<i>k</i> -mer completeness	99.99%	≥ 95%
BUSCO**	C:97.9%[S:93.6%,D:4.3%],F:1.0%,M:1.1%,n:1,013	C ≥ 95%
Percentage of assembly mapped to chromosomes	94.83%	≥ 95%
Sex chromosomes	X and Y chromosomes	<i>localised homologous pairs</i>
Organelles	Mitochondrial genome assembled	<i>complete single alleles</i>
Raw data accessions		
PacificBiosciences SEQUEL II	ERR10906091	
Hi-C Illumina	ERR10908620	
Genome assembly		
Assembly accession	GCA_949358305.1	
Accession of alternate haplotype	GCA_949357715.1	
Span (Mb)	237.5	
Number of contigs	343	
Contig N50 length (Mb)	1.8	
Number of scaffolds	158	
Scaffold N50 length (Mb)	19.3	
Longest scaffold (Mb)	67.0	

* 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 arthropoda_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/qcStrAcum1.1/dataset/CASHSZ01/busco>.

(Figure 2–Figure 5; Table 2). X and Y chromosomes in this genus are reported to be homomorphic (Green *et al.*, 2016). In this assembly, SUPER_X and SUPER_Y both have half coverage. Y was determined based on the absence of BUSCO genes (arthropoda-odb10 gene set). 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 59.6 with *k*-mer completeness of 99.99%, and the assembly

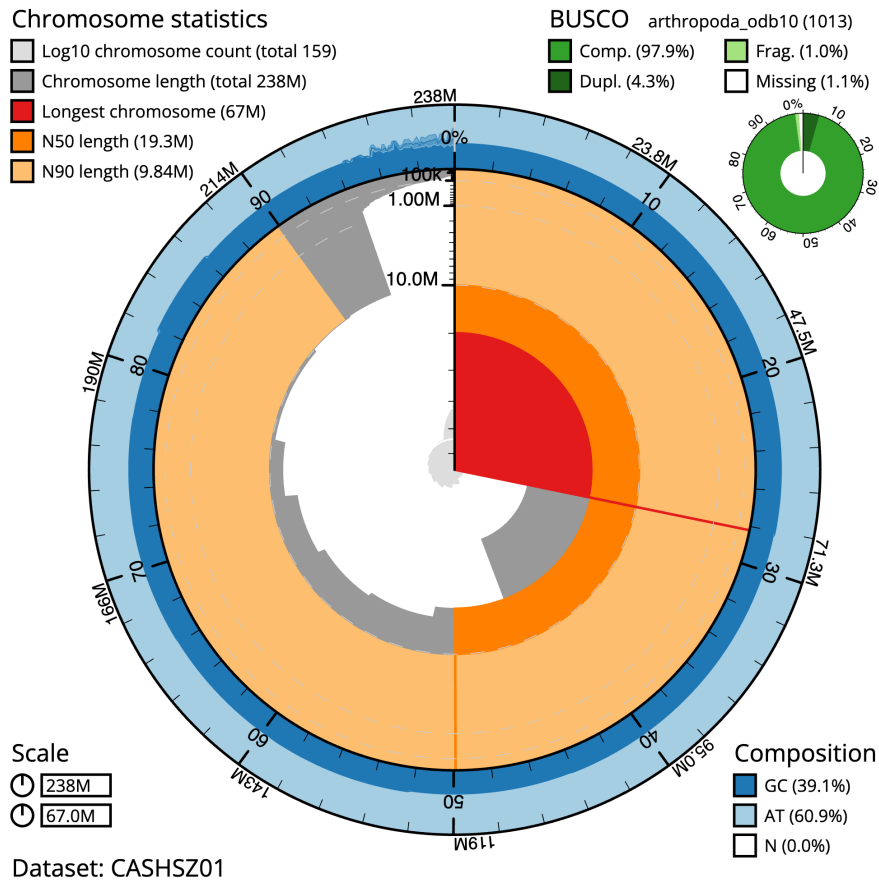


Figure 2. Genome assembly of *Strigamia acuminata*, qcStrAcum1.1: metrics. The BlobToolKit Snailplot shows N50 metrics and BUSCO gene completeness. The main plot is divided into 1,000 size-ordered bins around the circumference with each bin representing 0.1% of the 237,534,282 bp assembly. The distribution of scaffold lengths is shown in dark grey with the plot radius scaled to the longest scaffold present in the assembly (66,985,994 bp, shown in red). Orange and pale-orange arcs show the N50 and N90 scaffold lengths (19,347,077 and 9,836,499 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 arthropoda_odb10 set is shown in the top right. An interactive version of this figure is available at <https://blobtoolkit.genomehubs.org/view/qcStrAcum1.1/dataset/CASHSZ01/snail>.

has a BUSCO v5.3.2 completeness of 97.9% (single = 93.6%, duplicated = 4.3%), using the arthropoda_odb10 reference set ($n = 1,013$).

Metadata for specimens, spectral estimates, sequencing runs, contaminants and pre-curation assembly statistics can be found at <https://links.tol.sanger.ac.uk/species/1255758>.

Methods

Sample acquisition and nucleic acid extraction

An adult male *Strigamia acuminata* (specimen ID NHMUK014449029, ToLID qcStrAcum1) was collected from Bookham Common, Leatherhead, England, UK (latitude 51.29, longitude -0.38) on 2021-04-20. The specimen was collected by Gregory Edgecombe and Duncan Sivell (both Natural History Museum) and identified by Gregory Edgecombe, and was then preserved in 70% ethanol.

DNA was extracted at the Tree of Life laboratory, Wellcome Sanger Institute (WSI). The qcStrAcum1 sample was weighed and dissected on dry ice with tissue set aside for Hi-C sequencing. Somatic tissue was disrupted using a Nippi Powermasher fitted with a BioMasher pestle. High molecular weight (HMW) DNA was extracted using the Qiagen MagAttract HMW DNA extraction kit. HMW DNA was sheared into an average fragment size of 12–20 kb in a Megaruptor 3 system with speed setting 30. Sheared DNA was purified by solid-phase reversible immobilisation using AMPure PB beads with a 1.8X ratio of beads to sample to remove the shorter fragments and concentrate the DNA sample. 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.

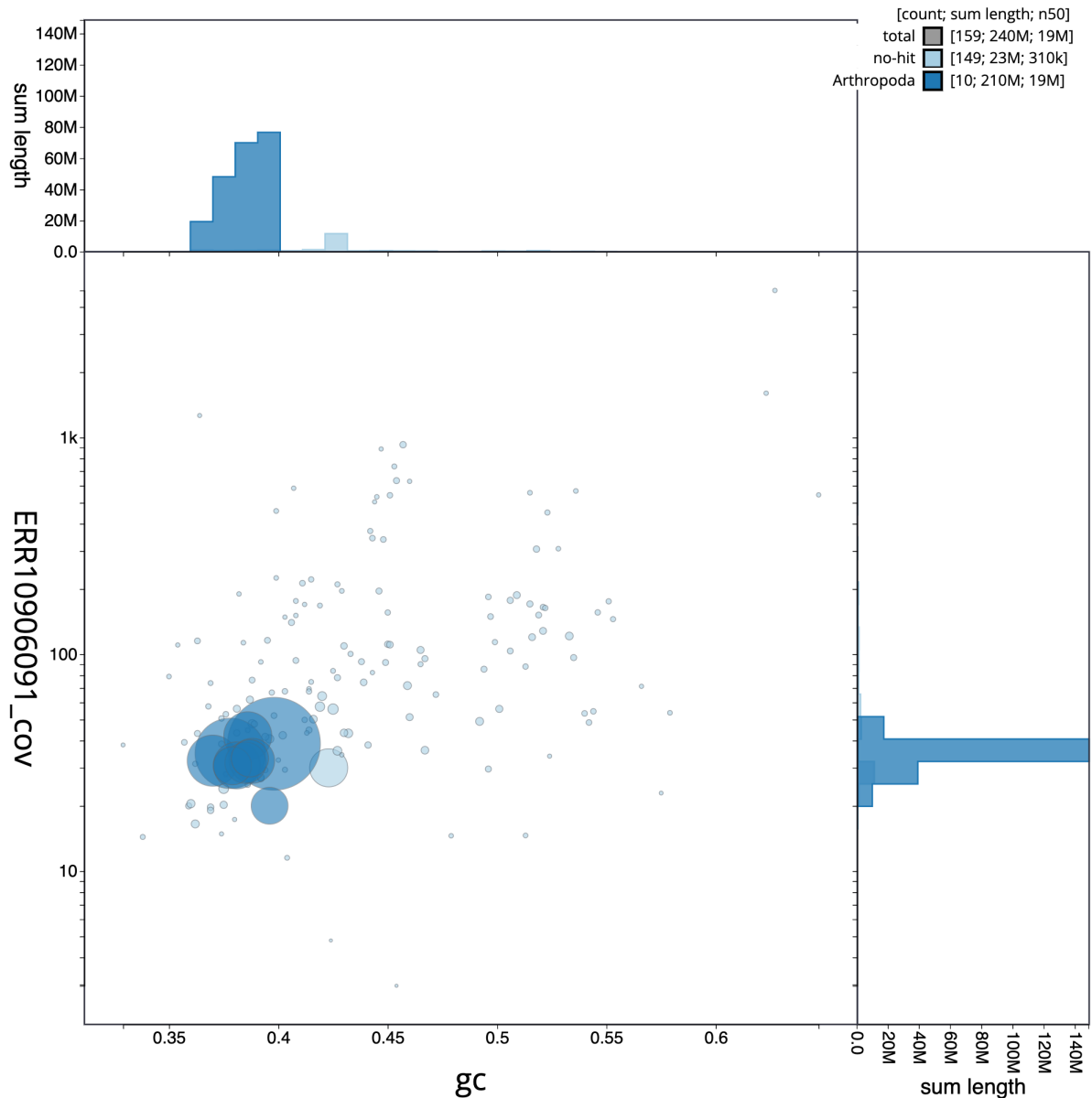


Figure 3. Genome assembly of *Strigamia acuminata*, qcStrAcum1.1: BlobToolKit GC-coverage plot. Scaffolds are coloured by phylum. Circles are sized in proportion to scaffold length. Histograms show the distribution of scaffold length sum along each axis. An interactive version of this figure is available at <https://blobtoolkit.genomehubs.org/view/qcStrAcum1.1/dataset/CASHSZ01/blob>.

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 II (HiFi) instrument. Hi-C data were also generated from tissue of qcStrAcum1 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

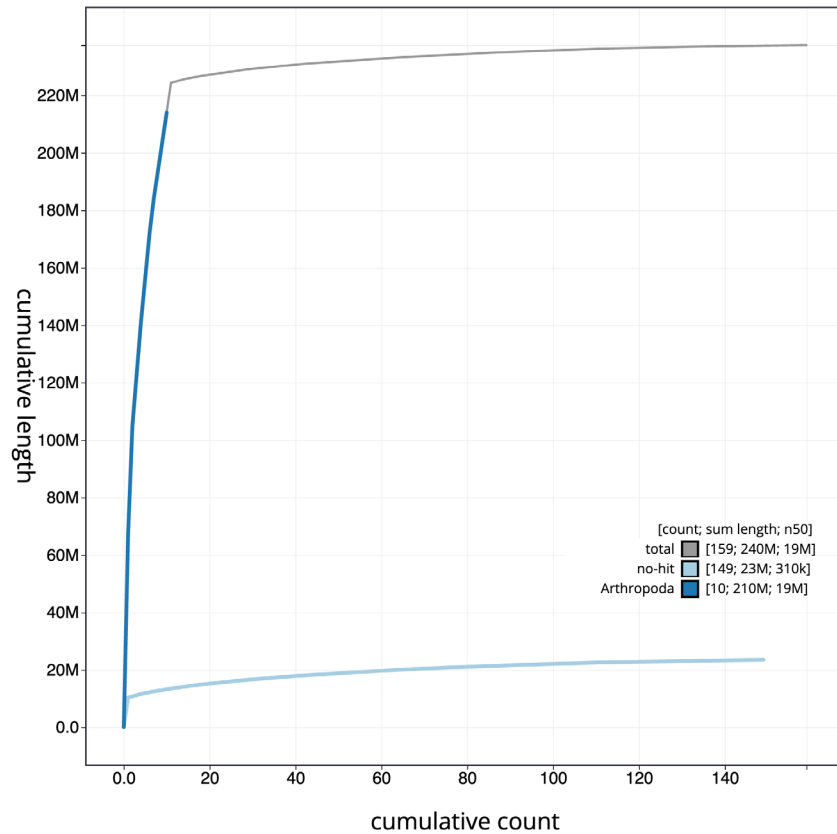


Figure 4. Genome assembly of *Strigamia acuminata*, qcStrAcum1.1: BlobToolKit cumulative sequence plot. The grey line shows cumulative length for all scaffolds. Coloured lines show cumulative lengths of scaffolds assigned to each phylum using the buscogenes taxrule. An interactive version of this figure is available at <https://blobtoolkit.genomehubs.org/view/qcStrAcum1.1/dataset/CASHSZ01/cumulative>.

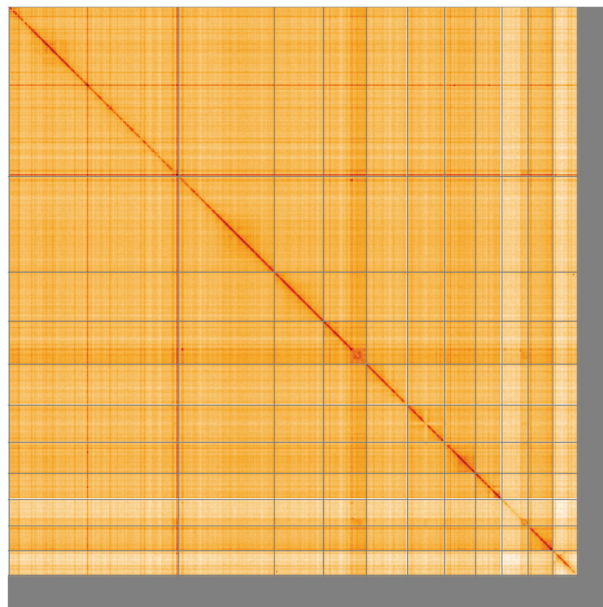


Figure 5. Genome assembly of *Strigamia acuminata*, qcStrAcum1.1: Hi-C contact map of the qcStrAcum1.1 assembly, visualised using HiGlass. Chromosomes are shown in order of size from left to right and top to bottom. An interactive version of this figure may be viewed at <https://genome-note-higlass.tol.sanger.ac.uk/I/?d=ePZD50jWQ1uKpCQEifKIEA>.

(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.

Table 2. Chromosomal pseudomolecules in the genome assembly of *Strigamia acuminata*, qcStrAcum1.

INSDC accession	Chromosome	Length (Mb)	GC%
OX442314.1	1	66.99	40.0
OX442315.1	2	37.79	38.0
OX442316.1	3	19.35	37.0
OX442317.1	4	16.99	38.5
OX442318.1	5	16.15	38.0
OX442319.1	6	14.72	39.0
OX442320.1	7	12.25	38.5
OX442321.1	8	10.36	38.0
OX442323.1	9	9.84	38.5
OX442324.1	X	9.63	39.5
OX442322.1	Y	10.34	42.5
OX442325.1	MT	0.02	37.0

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 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

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
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	1.2a	https://github.com/c-zhou/yahs

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: *Strigamia acuminata*. Accession number PRJEB59944; <https://identifiers.org/ena.embl/PRJEB59944>. (Wellcome Sanger Institute, 2023)

The genome sequence is released openly for reuse. The *Strigamia acuminata* 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 Natural History Museum Genome Acquisition Lab are listed here: <https://doi.org/10.5281/zenodo.4790042>.

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](#)
- Barber AD: **Key to the identification of British centipedes**. Field Studies Council, 2008.
- Barber AD: **Atlas of the centipedes of Britain and Ireland**. Field Studies Council, 2022.
- 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](#)
- Bonato L, Chagar A Jr, Dioguardi R, et al.: **A web resource for Chilopoda taxonomy**. *CHLOBASE*. 2016; Accessed 9 August 2023.
- Bonato L, Bortolin F, De Zen G, et al.: **Towards elucidating species diversity of European inland *Strigamia* (Chilopoda: Geophilomorpha): a first reassessment integrating multiple lines of evidence**. *Zool J Linn Soc*. 2023; zlad070. [Publisher Full Text](#)
- Bonato L, Dányi L, Socci AA, et al.: **Species Diversity of *Strigamia* Gray, 1843 (Chilopoda: Linotaeniidae): a preliminary synthesis**. *Zootaxa*. 2012; **3593**: 1–39. [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](#)
- Chipman AD, Ferrier DEK, Brena C, et al.: **The First Myriapod Genome Sequence Reveals Conservative Arthropod Gene Content and Genome Organisation in the Centipede *Strigamia maritima***. *PLoS Biol*. 2014; **12**(11): e1002005. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Collins G, Schneider C, Boštjančić LL, et al.: **MetaInvert: A new soil invertebrate genome resource provides insights into below-ground biodiversity and evolution**. *Research Square*. 2023; Accessed 17 August 2023. [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](#)
- Green JE, Dalíková M, Sahara K, et al.: **XX/XY System of Sex Determination in the Geophilomorph Centipede *Strigamia maritima***. *PLoS One*. 2016; **11**(2): e0150292. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free 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](#)
- 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](#)
- Leach WE: **XXXI. A tabular View of the external Characters of Four Classes of Animals, which Linné arranged under Insecta; with the Distribution of the Genera composing Three of these Classes into Orders, &c. and Descriptions of several New Genera and Species**. *Transactions of the Linnean Society of London*. 1816; **11**(2): 306–400. [Publisher 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.*: **Merqury: 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](#)

So WL, Nong W, Xie Y, *et al.*: **Myriapod genomes reveal ancestral horizontal gene transfer and hormonal gene loss in millipedes.** *Nat Commun*. 2022; **13**(1): 3010.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

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/genomenote (v1.0.dev).** *Zenodo*. 2023b; [Accessed 21 July 2023].

[Reference Source](#)

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](#)

Weil E: **Biologie der einheimatischen Geophiliden.** *Zeitschrift Für Angewandte Entomologie*. 1958; **42**: 173–209.

Wellcome Sanger Institute: **The genome sequence of the centipede *Strigamia acuminata* (Leach, 1816).** European Nucleotide Archive, [dataset], accession number PRJEB59944, 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 17 October 2023

<https://doi.org/10.21956/wellcomeopenres.22082.r67911>

© 2023 Kraaijeveld 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.

 **Ken Kraaijeveld** 

Faculty of Earth and Life Sciences, Institute of Ecological Science, Animal Ecology, Vrije University Amsterdam, Amsterdam, The Netherlands

Deep coverage Hifi reads combined with HiC yield a high-quality genome assembly that will be a valuable resource. The work is well-described and nicely illustrated.

My only request would be to provide more details on the steps taken during the manual assembly curation, as this would be useful to know for future genome assembly projects.

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: Genomics, evolution

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 05 October 2023

<https://doi.org/10.21956/wellcomeopenres.22082.r67909>

© 2023 van der Zee 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.



Maurijn van der Zee 

Institute of Biology, Leiden University, Sylviusweg, Leiden, The Netherlands

A pretty complete genome of the shorter red centipede

Data are well described, are accessible, and could be downloaded. It is not entirely clear how the 158 scaffolds were assigned to 11 chromosomes, but I guess this was mainly done using the HiC data. It is of course great that these HiC data are available. This high quality genome is a valuable supplement to the *Strigamia maritima* genome, and a useful contribution to the scientific community.

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 evo-devo, comparative genomics, evolve-and-resequence

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 05 October 2023

<https://doi.org/10.21956/wellcomeopenres.22082.r67713>

© 2023 Hui J. 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.

**Jerome H.L Hui**

Chinese University of Hong Kong, Hong Kong, China

This is an excellent genomic resource! Here, Edgecombe and colleagues report the genome sequence of centipede *Strigamia acuminata* (Leach, 1816). This species is widely recorded in England and Wales. To date, genomic resources for myriapods including centipedes are generally scarce when comparing to other arthropods such as insects and crustaceans. Molecular data of this species are also scarce prior to this report, and are mainly confined to COI sequences deposited to the NCBI database. One great achievement that I personally highly appreciate is that this is the first report of the X and Y chromosomal sequences in centipede! This new genome resource will be very useful for further studies, such as understanding their evolutionary relationships with other arthropods.

In terms of quality, this genome resource is in also excellent judging from the summary statistics, with high BUSCO numbers, high sequence continuity (scaffold N50), and majority of sequences contained on the 11 pseudochromosomes (i.e. 9 autosomes + X + Y) (plus mitochondrion). To sum up, this is an outstanding and valuable contribution.

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: I have published with Peter Holland more than three years ago, and confirm that this potential conflict of interest did not affect my ability to write an objective and unbiased review of the article.

Reviewer Expertise: Genomics, evolution, invertebrates

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.
