




DATA NOTE

The genome sequence of the Bulrush Veneer, *Calamotropha paludella* (Hübner, 1824) [version 1; peer review: 1 approved, 1 approved with reservations]

Douglas Boyes¹⁺, Peter W.H. Holland ²,
University of Oxford and Wytham Woods Genome Acquisition Lab,
Darwin Tree of Life Barcoding collective,
Wellcome Sanger Institute Tree of Life programme,
Wellcome Sanger Institute Scientific Operations: DNA Pipelines collective,
Tree of Life Core Informatics collective, Darwin Tree of Life Consortium

¹UK Centre for Ecology & Hydrology, Wallingford, England, UK

²University of Oxford, Oxford, England, UK

⁺ Deceased author

V1 First published: 13 Apr 2023, 8:168
<https://doi.org/10.12688/wellcomeopenres.19297.1>

Latest published: 13 Apr 2023, 8:168
<https://doi.org/10.12688/wellcomeopenres.19297.1>

Abstract

We present a genome assembly from an individual male *Calamotropha paludella* (the Bulrush Veneer; Arthropoda; Insecta; Lepidoptera; Crambidae). The genome sequence is 742.5 megabases in span. Most of the assembly is scaffolded into 30 chromosomal pseudomolecules, including the Z sex chromosome. The mitochondrial genome has also been assembled and is 16.3 kilobases in length. Gene annotation of this assembly on Ensembl identified 21,500 protein coding genes.

Keywords




Calamotropha paludella, Bulrush Veneer, genome sequence, chromosomal, Lepidoptera



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

Open Peer Review

Approval Status  

	1	2
version 1 13 Apr 2023	 view	 view
1. Muzafar Riyaz  , Division of Taxonomy and Biodiversity, Entomology Research Institute, Loyola Collège, Tamil Nadu, India		
2. Anne duploux , University of Helsinki, Helsinki, 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: Boyes D: Investigation, Resources; Holland PWH: 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: © 2023 Boyes D *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: Boyes D, Holland PWH, University of Oxford and Wytham Woods Genome Acquisition Lab *et al.* **The genome sequence of the Bulrush Veneer, *Calamotropha paludella* (Hübner, 1824) [version 1; peer review: 1 approved, 1 approved with reservations]** Wellcome Open Research 2023, 8:168 <https://doi.org/10.12688/wellcomeopenres.19297.1>

First published: 13 Apr 2023, 8:168 <https://doi.org/10.12688/wellcomeopenres.19297.1>

Species taxonomy

Eukaryota; Metazoa; Ecdysozoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Pyraloidea; Crambidae; Crambinae; *Calamotropha*; *Calamotropha paludella* (Hübner, 1824) (NCBI:txid1371681).

Background

The Bulrush Veneer, *Calamotropha paludella*, is a small grey or light brown moth in the family Crambidae, superfamily Pyraloidea. The species is widely distributed across Europe and Asia, from Portugal to Japan (GBIF Secretariat, 2022). The moth has also been reported from Madagascar (Viette, 1971) and Australia (GBIF Secretariat, 2022).

In the UK, *C. paludella* has been reported in small numbers from many regions in the south of England and from southern Wales, but is most common in East Anglia, Hampshire and the Isle of Wight (NBN Atlas Partnership, 2021). The moth is associated with marsh and wetland habitats as the larvae are specialist feeders on great and lesser reedmace, also called bulrush: *Typha latifolia* and *T. angustifolia*.

In the UK, the adult moth is on the wing in July and August; it can be seen flying over water bodies at dusk and is also attracted to light. Eggs are laid on the edge of a leaf of *Typha* sp. and the larvae form mines within the leaf and then later in the stem of the plant. The larvae overwinter, resuming feeding in spring (Bierne, 1952). External anatomy of larval and pupal stages has been described in detail from specimens in Gifu Prefecture, Japan, where the species also overwinters as a larva or occasionally as a pupa (Funakoshi, 1989).

A complete genome sequence of *C. paludella* will facilitate research into the extreme food plant specificity of this species and contribute to the growing set of genomic resources for understanding lepidopteran evolution.

Genome sequence report

The genome was sequenced from one male *C. paludella* specimen (Figure 1) collected from Wytham Woods, Oxfordshire

(biological vice-county: Berkshire), UK (latitude 51.77, longitude −1.34). A total of 35-fold coverage in Pacific Biosciences single-molecule HiFi long reads and 125-fold coverage in 10X Genomics read clouds were generated. Primary assembly contigs were scaffolded with chromosome conformation Hi-C data. Manual assembly curation corrected 11 missing joins or mis-joins and removed two haplotypic duplications, reducing the scaffold number by 18.87% and increasing the scaffold N50 by 2.33%.

The final assembly has a total length of 742.5 Mb in 43 sequence scaffolds with a scaffold N50 of 26.9 Mb (Table 1). Most (99.86%) of the assembly sequence was assigned to 30 chromosomal-level scaffolds, representing 29 autosomes, and the Z sex chromosome. Chromosome-scale scaffolds confirmed by the Hi-C data are named in order of size (Figure 2–Figure 5; Table 2). 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 59.2 with *k*-mer based completeness of 100%, and the assembly has a BUSCO v5.3.2 completeness of 98.7% (single = 98.2%, duplicated = 0.5%), using the lepidoptera_odb10 reference set (*n* = 5,286).

Metadata for specimens, spectral estimates, sequencing runs, contaminants and pre-curation assembly statistics can be found at <https://links.tol.sanger.ac.uk/species/1371681>.

Genome annotation report

The *Calamotropha paludella* genome assembly GCA_927399485.1 (ilCalPalu1.1) was annotated using the Ensembl rapid annotation pipeline (Table 1; https://rapid.ensembl.org/Calamotropha_paludella_GCA_927399485.1/Info/Index/). The resulting annotation includes 21,650 transcribed mRNAs from 21,500 protein-coding genes.

Methods

Sample acquisition and nucleic acid extraction

A male *Calamotropha paludella* specimen (specimen number Ox000818, ToLID ilCalPalu1) was used for genome sequencing and Hi-C scaffolding. This specimen was collected from Wytham Woods, Oxfordshire (biological vice-county: Berkshire), UK (latitude 51.77, longitude −1.34) on 1 August 2020 using a light trap. A second specimen (specimen number Ox001672, ToLID ilCalPalu2) was used for RNA sequencing. This specimen was collected from a fen habitat in Wytham Woods (latitude 51.77, longitude −1.31) on 17 July 2021, also using a light trap. Both specimens were collected and identified by Douglas Boyes (University of Oxford) and then snap-frozen from live on dry ice.

DNA was extracted at the Tree of Life laboratory, Wellcome Sanger Institute (WSI). The ilCalPalu1 sample was weighed and dissected on dry ice with head and thorax tissue set aside



Figure 1. Photograph of the *Calamotropha paludella* (ilCalPalu1) specimen used for genome sequencing.

Table 1. Genome data for *Calamotropha paludella*, ilCalPalu1.1.

Project accession data		
Assembly identifier	ilCalPalu1.1	
Species	<i>Calamotropha paludella</i>	
Specimen	ilCalPalu1	
NCBI taxonomy ID	1371681	
BioProject	PRJEB47466	
BioSample ID	SAMEA7746625	
Isolate information	ilCalPalu1, abdomen (genome sequencing); head and thorax (Hi-C scaffolding) ilCalPalu2, head and thorax (RNA sequencing)	
Assembly metrics*		Benchmark
Consensus quality (QV)	59.2	≥ 50
k-mer completeness	100%	≥ 95%
BUSCO**	C:98.7%[S:98.2%,D:0.5%], F:0.3%,M:0.9%,n:5,286	C ≥ 95%
Percentage of assembly mapped to chromosomes	99.86%	≥ 95%
Sex chromosomes	Z chromosome	localised homologous pairs
Organelles	Mitochondrial genome assembled	complete single alleles
Raw data accessions		
PacificBiosciences SEQUEL II	ERR6808070, ERR6909089	
10X Genomics Illumina	ERR6747932-ERR6747939	
Hi-C Illumina	ERR6747940	
PolyA RNA-Seq Illumina	ERR10123655	
Genome assembly		
Assembly accession	GCA_927399485.1	
Accession of alternate haplotype	GCA_927399455.1	
Span (Mb)	742.5	
Number of contigs	60	
Contig N50 length (Mb)	23.9	
Number of scaffolds	43	
Scaffold N50 length (Mb)	26.9	
Longest scaffold (Mb)	48.0	
Genome annotation		
Number of protein-coding genes	21,500	
Number of gene transcripts	21,650	

* 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 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/ilCalPalu1.1/dataset/CAKMJK01/busco>.

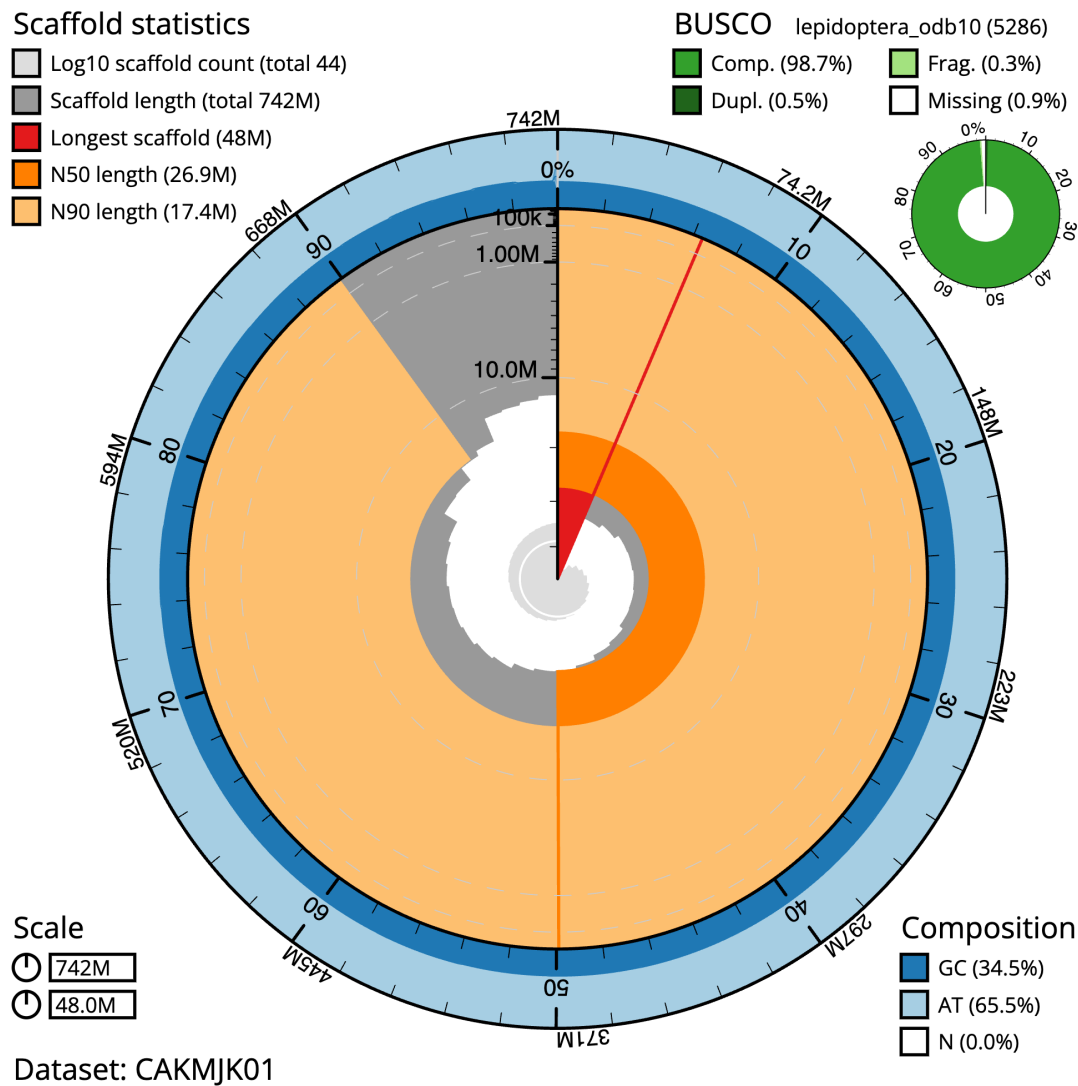


Figure 2. Genome assembly of *Calamotropha paludella*, iCalPalu1.1: metrics. The BlobToolKit Snailplot shows N50 metrics and BUSCO gene completeness. The main plot is divided into 1,000 size-ordered bins around the circumference with each bin representing 0.1% of the 742,481,776 bp assembly. The distribution of scaffold lengths is shown in dark grey with the plot radius scaled to the longest scaffold present in the assembly (47,976,261 bp, shown in red). Orange and pale-orange arcs show the N50 and N90 scaffold lengths (26,936,083 and 17,375,879 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/iCalPalu1.1/dataset/CAKMJK01/snail>.

for Hi-C sequencing. Abdomen 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. Low molecular weight DNA was removed from a 20 ng aliquot of extracted DNA using the 0.8X AMPure XP purification kit prior to 10X Chromium sequencing; a minimum of 50 ng DNA was submitted for 10X sequencing. 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.

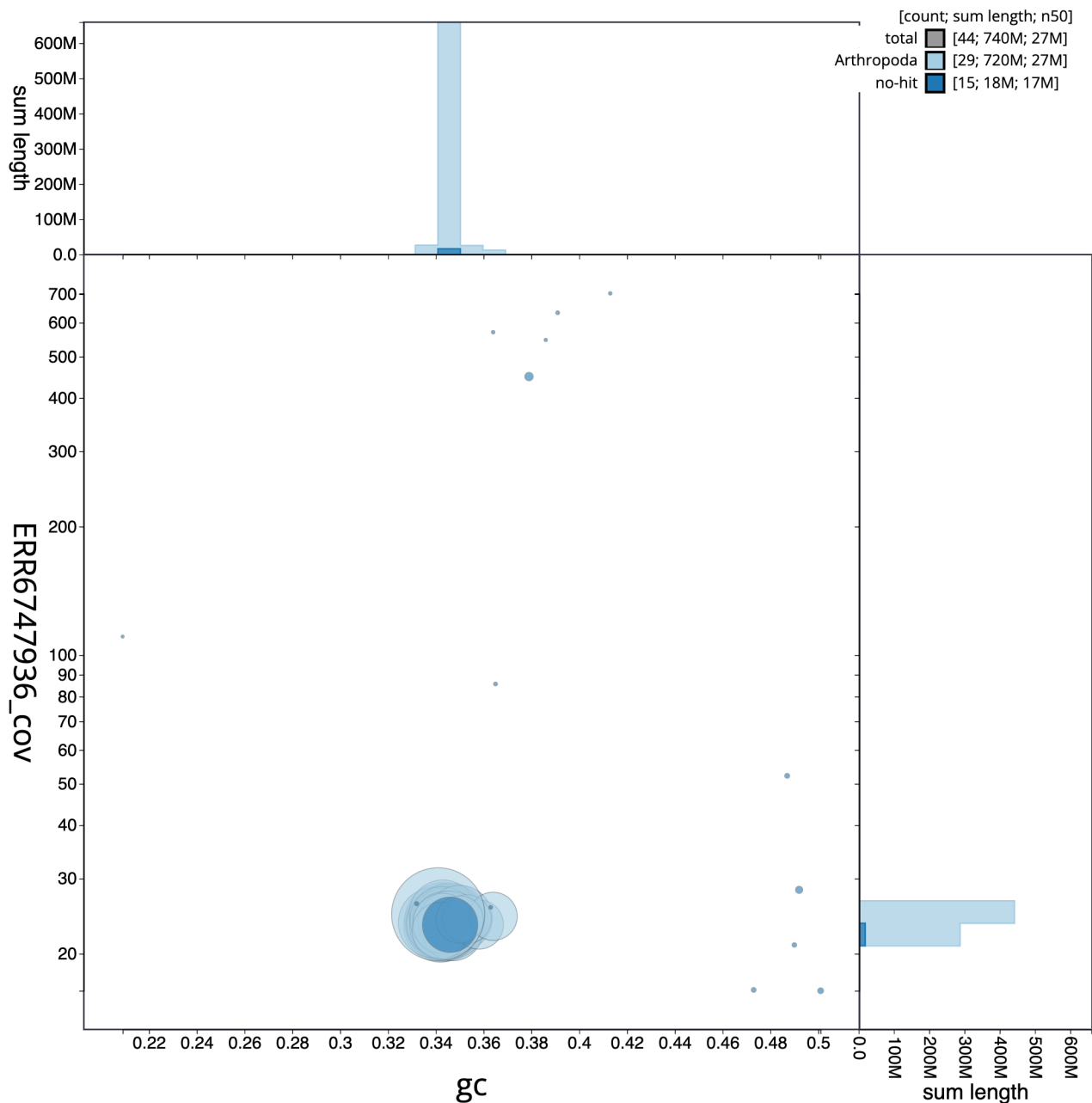


Figure 3. Genome assembly of *Calamotropha paludella*, ilCalPalu1.1: BlobToolKit GC-coverage plot. Scaffolds are coloured by phylum. Circles are sized in proportion to scaffold length. Histograms show the distribution of scaffold length sum along each axis. An interactive version of this figure is available at <https://blobtoolkit.genomehubs.org/view/ilCalPalu1.1/dataset/CAKMJK01/blob>.

RNA was extracted from head and thorax tissue of ilCalPalu2 in the Tree of Life Laboratory at the WSI using TRIzol, according to the manufacturer's instructions. RNA was then eluted in 50 µl RNase-free water and its concentration assessed using a Nanodrop spectrophotometer and Qubit Fluorometer using the Qubit RNA Broad-Range (BR) Assay kit. Analysis of the integrity of the RNA was done using Agilent RNA 6000 Pico Kit and Eukaryotic Total RNA assay.

Sequencing

Pacific Biosciences HiFi circular consensus and 10X Genomics read cloud DNA sequencing libraries were constructed according to the manufacturers' instructions. Poly(A) RNA-Seq libraries were constructed using the NEB Ultra II RNA Library Prep kit. DNA and RNA sequencing were performed by the Scientific Operations core at the WSI on Pacific Biosciences SEQUEL II (HiFi) and Illumina NovaSeq 6000 (RNA-Seq).

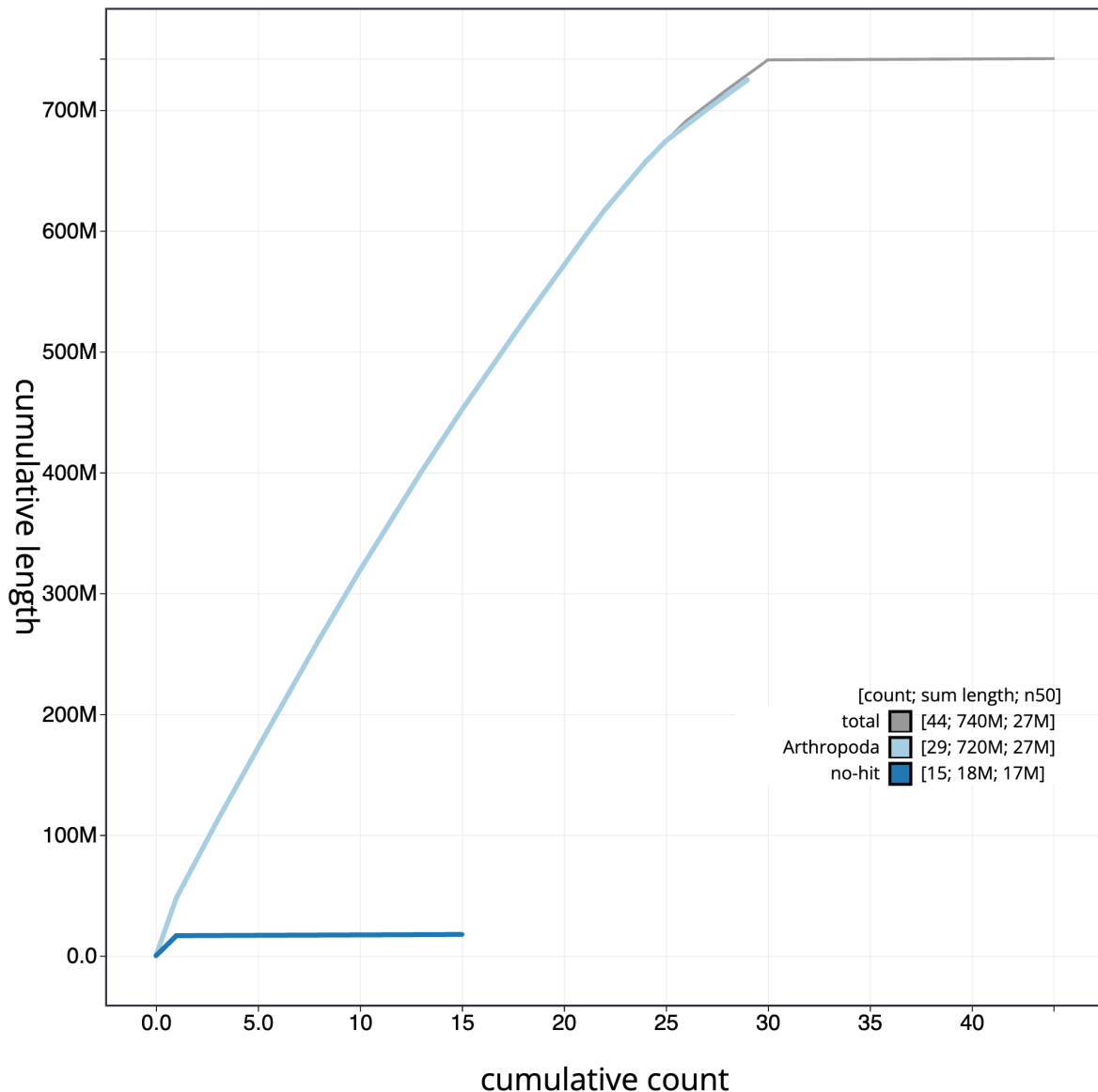


Figure 4. Genome assembly of *Calamotropha paludella*, ilCalPalu1.1: BlobToolKit cumulative sequence plot. The grey line shows cumulative length for all scaffolds. Coloured lines show cumulative lengths of scaffolds assigned to each phylum using the buscogenes taxrule. An interactive version of this figure is available at <https://blobtoolkit.genomehubs.org/view/ilCalPalu1.1/dataset/CAKMJK01/cumulative>.

and 10X) instruments. Hi-C data were also generated from head and thorax tissue of ilCalPalu1 using the Arima v2 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). One round of polishing was performed by aligning 10X Genomics read data to the assembly with Long Ranger ALIGN, calling variants with FreeBayes (Garrison & Marth, 2012). The assembly was then scaffolded with Hi-C data (Rao *et al.*, 2014) using SALSA2

(Ghurye *et al.*, 2019). The assembly was checked for contamination 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.*, 2022), which performed annotation using MitoFinder (Allio *et al.*, 2020). To evaluate the assembly, MerquyFK was used to estimate consensus quality (QV) scores and *k*-mer completeness (Rhie *et al.*, 2020). The genome was analysed and BUSCO scores (Manni *et al.*, 2021; Simão *et al.*, 2015) were calculated within the BlobToolKit environment (Challis *et al.*, 2020). Table 3 contains a list of software tool versions and sources.

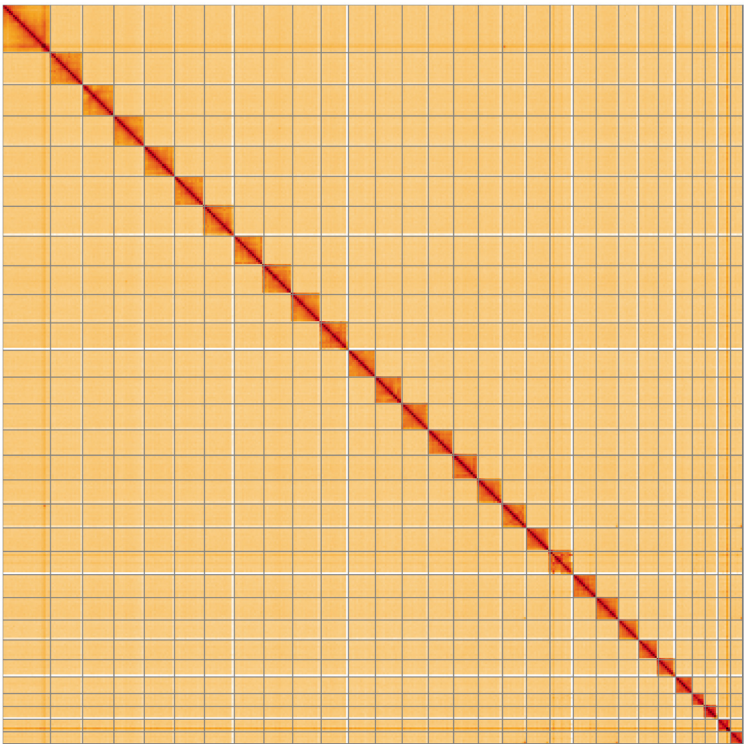


Figure 5. Genome assembly of *Calamotropha paludella*, ilCalPalu1.1: Hi-C contact map of the ilCalPalu1.1 assembly, visualised using HiGlass. Chromosomes are shown in order of size from left to right and top to bottom. An interactive version of this figure may be viewed at <https://genome-note-higlass.tol.sanger.ac.uk/l/?d=TO8E34MXRuK-y9xijSgFkQ>.

Table 2. Chromosomal pseudomolecules in the genome assembly of *Calamotropha paludella*, ilCalPalu1.

INSDC accession	Chromosome	Size (Mb)	GC%
OV656804.1	1	32.21	34.3
OV656805.1	2	31.37	34.6
OV656806.1	3	30.4	34.5
OV656807.1	4	30.27	34.5
OV656808.1	5	30.04	34.3
OV656809.1	6	29.85	34.3
OV656810.1	7	29.72	34.2
OV656811.1	8	28.85	34.4
OV656812.1	9	28.39	34.4
OV656813.1	10	27.51	34.4
OV656814.1	11	26.94	34.2
OV656815.1	12	26.8	33.9
OV656816.1	13	26.32	34.4
OV656817.1	14	25.27	34.2
OV656818.1	15	24.8	34.3

INSDC accession	Chromosome	Size (Mb)	GC%
OV656819.1	16	24.22	34.3
OV656820.1	17	23.86	34.4
OV656821.1	18	23.7	34.3
OV656822.1	19	23.32	35
OV656823.1	20	23	34.4
OV656824.1	21	22.52	34.6
OV656825.1	22	19.93	34.7
OV656826.1	23	19.75	34.3
OV656827.1	24	17.38	34.2
OV656828.1	25	16.59	34.6
OV656829.1	26	13.08	35.8
OV656830.1	27	12.86	35.3
OV656831.1	28	12.49	36.4
OV656832.1	29	12.05	35
OV656803.1	Z	47.98	34.1
OV656833.1	MT	0.02	21.2
-	unplaced	1.02	43.2

Table 3. Software tools: versions and sources.

Software tool	Version	Source
BlobToolKit	4.0.7	https://github.com/blobtoolkit/blobtoolkit
BUSCO	5.3.2	https://gitlab.com/ezlab/busco
FreeBayes	1.3.1-17-gaa2ace8	https://github.com/freebayes/freebayes
gEVAL	N/A	https://geval.org.uk/
Hifiasm	0.15.3	https://github.com/chhylp123/hifiasm
HiGlass	1.11.6	https://github.com/higlass/higlass
Long Ranger ALIGN	2.2.2	https://support.10xgenomics.com/genome-exome/software/pipelines/latest/advanced/other-pipelines
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
SALSA	2.2	https://github.com/salsa-rs/salsa

Genome annotation

The BRAKER2 pipeline (Brůna *et al.*, 2021) was used in the default protein mode to generate annotation for the *Calamotropha paludella* assembly (GCA_927399485.1) in Ensembl Rapid Release.

Ethics and compliance issues

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](#). 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. All efforts are undertaken to minimise the suffering of animals used for sequencing. 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: *Calamotropha paludella* (bulrush veneer). Accession number PRJEB47466; <https://identifiers.org/ena.embl/PRJEB47466>. (Wellcome Sanger Institute, 2022)

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

Author information

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

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

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

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

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

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

References

- 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](#)
- Bierne BP: **British Pyralid and Plume Moths.** London: Warne and Co. 1952.
[Reference Source](#)
- Brůna T, Hoff KJ, Lomsadze A, *et al.*: **BRAKER2: Automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database.** *NAR Genom Bioinform.* 2021; **3**(1): lqaa108.
[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](#)
- Funakoshi S: **Immature Stages of *Calamotropha paludella* (Hübner)(Lepidoptera, Pyralidae) of Japan.** *Lepidoptera Science.* 1989; **40**: 235–240.
[Reference Source](#)
- Garrison E, Marth G: **Haplotype-based variant detection from short-read sequencing.** 2012.
[Reference Source](#)
- GBIF Secretariat: **Calamotropha paludella (Hübner, 1824).** *GBIF Backbone Taxonomy.* 2022; (Accessed: 22 January 2023).
[Reference Source](#)
- Ghurye J, Rhie A, Walenz BP, *et al.*: **Integrating Hi-C links with assembly graphs for chromosome-scale assembly.** *PLoS Comput Biol.* 2019; **15**(8): e1007273.
[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, Pelan S, *et al.*: **Significantly improving the quality of genome assemblies through curation.** *GigaScience.* Oxford University Press. 2021; **10**(1): gaa153.
[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](#)
- 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](#)
- NBN Atlas Partnership: **Calamotropha paludella (Hübner, [1824]) Bulrush Veneer.** *NBN Atlas.* 2021; (Accessed: 22 January 2022).
[Reference Source](#)
- 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, 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](#)
- 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](#)
- 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](#)
- Uliano-Silva M, Ferreira JGRN, Krashenninnikova K, *et al.*: **MitoHiFi: a python pipeline for mitochondrial genome assembly from PacBio High Fidelity reads.** *bioRxiv.* [Preprint]. 2022.
[Publisher Full Text](#)
- Viette P: **Les Pyralidae Crambinae des îles de l'Océan Indien occidental [Lep.].** *Bulletin de la Société entomologique de France.* 1971; **76**: 72–81.
[Reference Source](#)
- Wellcome Sanger Institute: **The genome sequence of the Bulrush Veneer, *Calamotropha paludella* (Hübner, 1824).** European Nucleotide Archive. [dataset], accession number PRJEB47466. 2022.

Open Peer Review

Current Peer Review Status:  

Version 1

Reviewer Report 10 May 2024

<https://doi.org/10.21956/wellcomeopenres.21380.r80604>

© 2024 duplouny A. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Anne duplouny

University of Helsinki, Helsinki, Finland

This report follows the format of similarly published research. It is easy to follow and provide useful information for future research on the study system and beyond.

I have only minor comments:

I would edit the introduction such that 2nd and 3rd paragraph that both start with 'In the UK' are merged. Keeping element related to the species first, and then moving on to specifics of the larval stage clearly separated. This way, the text would feel less like it is missing citation, as it is currently the case.

Figure 1: would need a scale! I realize the picture can probably not be done again with a scale (even better would be a color scale too), but maybe the legend could mention the size of the specimen, or at least of the tubes.

I found the mention of the 2 haplotypic duplications and the second haplotype being a bit surprising. More details on how those were obtained or identified would be needed to avoid misunderstanding here. There is no mention of this in the method for example, how were they identified and isolated? Was it due to the use of the second individual? please edit such that it is also accessible to a less specialized audience (young careers for example).

Thank you.

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: Ecology and 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, however I have significant reservations, as outlined above.

Reviewer Report 31 January 2024

<https://doi.org/10.21956/wellcomeopenres.21380.r71759>

© 2024 Riyaz 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.



Muzafar Riyaz 

Division of Taxonomy and Biodiversity, Entomology Research Institute, Loyola Collège, Tamil Nadu, India

The present study of the genome sequence of *Calamotropha paludella*, commonly known as the Bulrush Veneer, is undoubtedly a significant contribution to the field of genomics and entomology. The comprehensive approach employed in sequencing, assembling, and annotating the genome reflects a high level of scientific rigour. The integration of various sequencing technologies, including Pacific Biosciences HiFi and 10X Genomics read cloud DNA sequencing, along with Hi-C data, contributes to a robust genome assembly. The inclusion of various visualization tools, such as Blob Toolkit Snail plot, GC-coverage plot, cumulative sequence plot, and Hi-C contact map, enhances the accessibility of the data for both experts and non-experts. These tools provide a comprehensive overview of the genome assembly metrics, aiding in the interpretation and validation of the results. Furthermore, the extensive list of references and citations demonstrates a thorough literature review and contextualization of the study within the broader scientific community. This not only enriches the background information but also positions the research within the existing body of knowledge. In conclusion, the study presents a valuable contribution to genomic research,

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: Genomics, Phylogenetics, Lepidoptera, and Evolutionary Biology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.
