Check for updates

DATA NOTE

The genome sequence of the Water Veneer, *Acentria ephemerella* (Denis & Schiffermüller, 1775) [version 1; peer review: awaiting peer review]

Douglas Boyes[1+], Peter Mulhair [ID][2],
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]UK Centre for Ecology & Hydrology, Wallingford, England, UK
[2]University of Oxford, Oxford, England, UK

[+] Deceased author

Open Peer Review

Approval Status *AWAITING PEER REVIEW*

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

Abstract
We present a genome assembly from an individual male *Acentria ephemerella* (the Water Veneer; Arthropoda; Insecta; Lepidoptera; Crambidae). The genome sequence is 340.8 megabases in span. Most of the assembly is scaffolded into 31 chromosomal pseudomolecules, including the Z sex chromosome. The mitochondrial genome has also been assembled and is 15.35 kilobases in length. Gene annotation of this assembly on Ensembl identified 17,748 protein coding genes.

Keywords
Acentria ephemerella, Water Veneer moth, genome sequence, chromosomal, Lepidoptera

This article is included in the Tree of Life gateway.

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

Author roles: Boyes D: Investigation, Resources; **Mulhair P**: Writing – Original Draft Preparation;

How to cite this article: Boyes D, Mulhair P, University of Oxford and Wytham Woods Genome Acquisition Lab *et al.* **The genome sequence of the Water Veneer,** *Acentria ephemerella* **(Denis & Schiffermüller, 1775) [version 1; peer review: awaiting peer review]** Wellcome Open Research 2024, **9**:134 https://doi.org/10.12688/wellcomeopenres.21099.1

First published: 08 Mar 2024, **9**:134 https://doi.org/10.12688/wellcomeopenres.21099.1

Species taxonomy

Eukaryota; Opisthokonta; Metazoa; Eumetazoa; Bilateria; Protostomia; Ecdysozoa; Panarthropoda; Arthropoda; Mandibulata; Pancrustacea; Hexapoda; Insecta; Dicondylia; Pterygota; Neoptera; Endopterygota; Amphiesmenoptera; Lepidoptera; Glossata; Neolepidoptera; Heteroneura; Ditrysia; Obtectomera; Pyraloidea; Crambidae; Nymphulinae; *Acentria*; *Acentria ephemerella* (Denis & Schiffermüller, 1775) (NCBI:txid1666818).

Background

Acentria ephemerella (Water Veneer moth) is a crambid moth, and belongs to the Acentropinae (Nymphulinae) subfamily. All species within this subfamily have adapted to life in freshwater habitats, with similar life history traits to caddisflies including aquatic larvae and terrestrial adults (Léger *et al.*, 2021; Pabis, 2018; Regier *et al.*, 2012). The water veneer moth is a visually unassuming species within the subfamily, with light grey forewings, however, has a remarkable life history. Adult females are usually brachypterous, with shortened, non-functioning wings and spend their entire life under water. However, another, less common female morph exists which has a normal wing phenotype and is terrestrial, similar to the males but larger in size (Berg, 1942). The aquatic, wingless females have additional hairs on their second and thirds legs which are used for underwater swimming. Further sexual dimorphisms between males and aquatic females have been observed in the structure and function of the compound eyes (Lau *et al.*, 2007). All adults of this species have an under-developed proboscis, likely a result of adult males living only a day or two solely to breed, where they mate with aquatic females at the water's surface. On warm, humid nights adult males, which are attracted to light, can be found in large swarms, sometimes in the thousands.

This species is native to Europe where it is widespread, but it has also spread to North America. It is usually found in lentic freshwater habitats, such as ponds, lakes and, marshes, where the larvae feed on a range of aquatic plants such as pondweeds (*Potomogeton* spp.) and Canadian waterweed (*Elodea canadensis*) (Gross *et al.*, 2002). The larvae, which are case-building in late instars, generally overwinter in cocoons on macrophytes. *A. ephemerella* has been suggested as a useful biocontrol agent against non-native pest plants such as Eurasian watermilfoil, which is prefers over native species (Batra, 1977; Gross *et al.*, 2001; Johnson *et al.*, 1997).

The genome of the water veneer, *Acentria ephemerella*, 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 *Acentria ephemerella*, based on one male specimen from Wytham Woods, Oxfordshire.

Genome sequence report

The genome was sequenced from one male *Acentria ephemerella* (Figure 1) collected from Wytham Woods, Oxfordshire,

Figure 1. Photograph of the *Acentria ephemerella* (ilAceEphe1) specimen used for genome sequencing.

UK (51.77, –1.34). A total of 64-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 two missing joins or mis-joins.

The final assembly has a total length of 340.8 Mb in 35 sequence scaffolds with a scaffold N50 of 12.3 Mb (Table 1). The snailplot in Figure 2 provides a summary of the assembly statistics, while the distribution of assembly scaffolds on GC proportion and coverage is shown in Figure 3. The cumulative assembly plot in Figure 4 shows curves for subsets of scaffolds assigned to different phyla. Most (99.94%) of the assembly sequence was assigned to 31 chromosomal-level scaffolds, representing 30 autosomes and the Z sex chromosome. Chromosome-scale scaffolds confirmed by the Hi-C data are named in order of size (Figure 5; Table 2). 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 68.0 with *k*-mer completeness of 100.0%, and the assembly has a BUSCO v5.3.2 completeness of 98.8% (single = 98.4%, duplicated = 0.3%), using the lepidoptera_odb10 reference set (*n* = 5,286).

Metadata for specimens, barcode results, spectra estimates, sequencing runs, contaminants and pre-curation assembly statistics are given at https://links.tol.sanger.ac.uk/species/1666818.

Genome annotation report

The *Acentria ephemerella* genome assembly (GCA_943193645.1) was annotated using the Ensembl rapid annotation pipeline at the European Bioinformatics Institute (EBI). The resulting annotation includes 17,946 transcribed mRNAs from 17,748

Table 1. Genome data for *Acentria ephemerella*, **ilAceEphe1.1.**

Project accession data		
Assembly identifier	ilAceEphe1.1	
Species	*Acentria ephemerella*	
Specimen	ilAceEphe1	
NCBI taxonomy ID	1666818	
BioProject	PRJEB52477	
BioSample ID	SAMEA10978991	
Isolate information	ilAceEphe1, male: whole organism (DNA sequencing) ilAceEphe2: whole organism (Hi-C sequencing)	
Assembly metrics*		*Benchmark*
Consensus quality (QV)	68.0	*≥ 50*
k-mer completeness	100.0%	*≥ 95%*
BUSCO**	C:98.8%[S:98.4%,D:0.3%], F:0.3%,M:1.0%,n:5,286	*C ≥ 95%*
Percentage of assembly mapped to chromosomes	99.94%	*≥ 95%*
Sex chromosomes	ZZ	*localised homologous pairs*
Organelles	Mitochondrial genome: 15.35 kb	*complete single alleles*
Raw data accessions		
PacificBiosciences SEQUEL II	ERR9745004	
Hi-C Illumina	ERR9682482	
PolyA RNA-Seq Illumina	ERR10123695	
Genome assembly		
Assembly accession	GCA_943193645.1	
Accession of alternate haplotype	GCA_943193655.1	
Span (Mb)	340.8	
Number of contigs	38	
Contig N50 length (Mb)	12.3	
Number of scaffolds	35	
Scaffold N50 length (Mb)	12.3	
Longest scaffold (Mb)	15.31	
Genome annotation		
Number of protein-coding genes	17,748	
Number of gene transcripts	17,946	

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

** BUSCO scores based on the lepidoptera_odb10 BUSCO set using version 5.3.2. C = complete [S = single copy, D = duplicated], F = fragmented, M = missing, n = number of orthologues in comparison. A full set of BUSCO scores is available at https://blobtoolkit.genomehubs.org/view/CALPDK01.1/dataset/CALPDK01.1/busco.
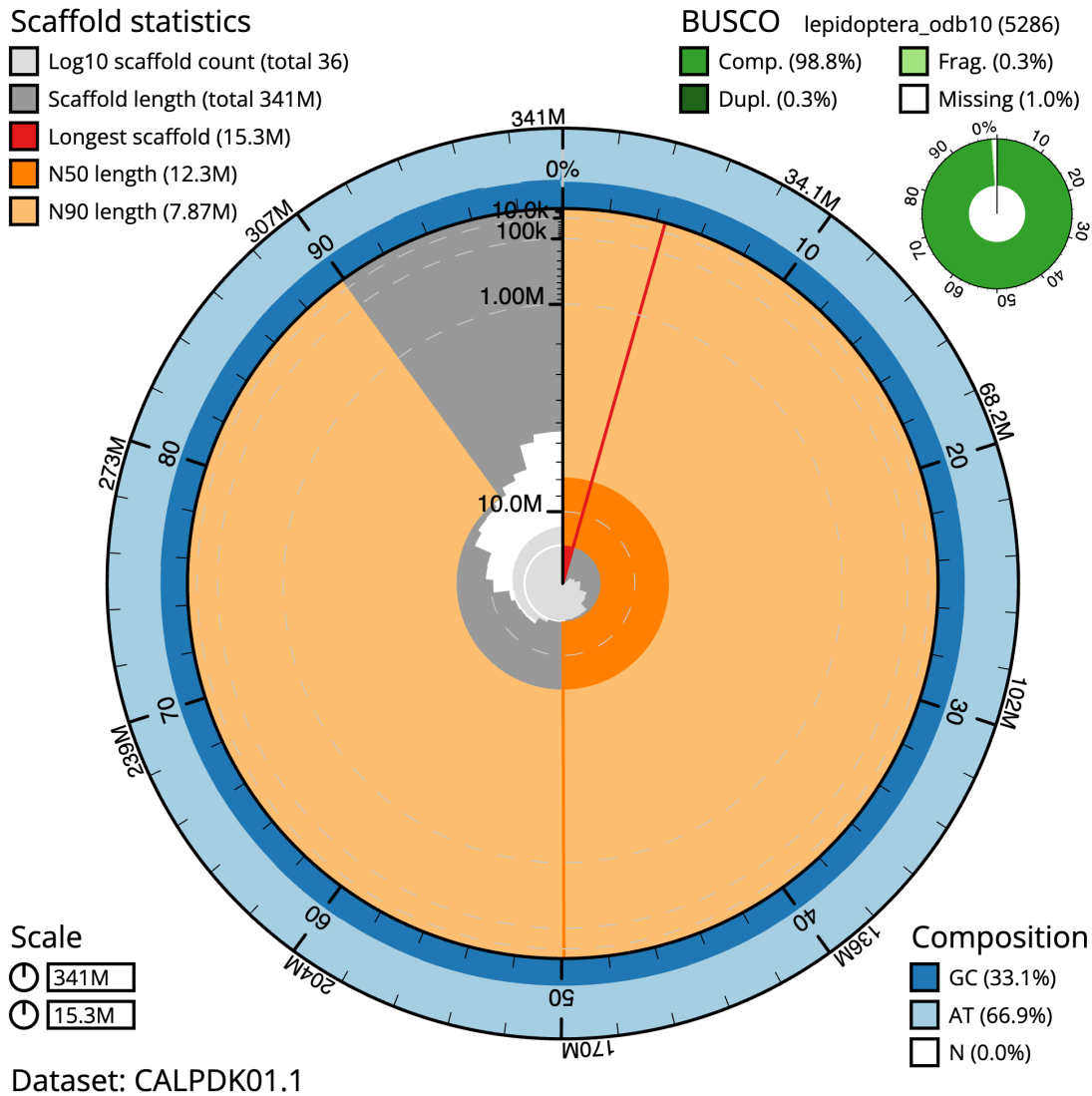
Figure 2. Genome assembly of *Acentria ephemerella*, ilAceEphe1.1: metrics. The BlobToolKit Snailplot shows N50 metrics and BUSCO gene completeness. The main plot is divided into 1,000 size-ordered bins around the circumference with each bin representing 0.1% of the 340,792,766 bp assembly. The distribution of scaffold lengths is shown in dark grey with the plot radius scaled to the longest scaffold present in the assembly (15,308,483 bp, shown in red). Orange and pale-orange arcs show the N50 and N90 scaffold lengths (12,251,206 and 7,870,291 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/CALPDK01.1/dataset/CALPDK01.1/snail.

protein-coding genes (Table 1; https://rapid.ensembl.org/Acentria_ephemerella_GCA_943193645.1/Info/Index).

Methods

Sample acquisition and nucleic acid extraction

The *Acentria ephemerella* specimens used for genome sequencing (specimen ID Ox001727, ToLID ilAceEphe1) and Hi-C sequencing (specimen ID Ox001728, ToLID ilAceEphe2) were collected from Wytham Woods, Oxfordshire (biological vice-county Berkshire), UK (latitude 51.77, longitude –1.34)

on 2021-07-17 using a light trap. Both specimens were collected and identified by Douglas Boyes (University of Oxford) and preserved on dry ice.

The workflow for high molecular weight (HMW) DNA extraction at the Wellcome Sanger Institute (WSI) includes a sequence of core procedures: sample preparation; sample homogenisation, DNA extraction, fragmentation, and clean-up. In sample preparation, the ilAceEphe1 sample was weighed and dissected on dry ice (Jay *et al.*, 2023). Whole organism
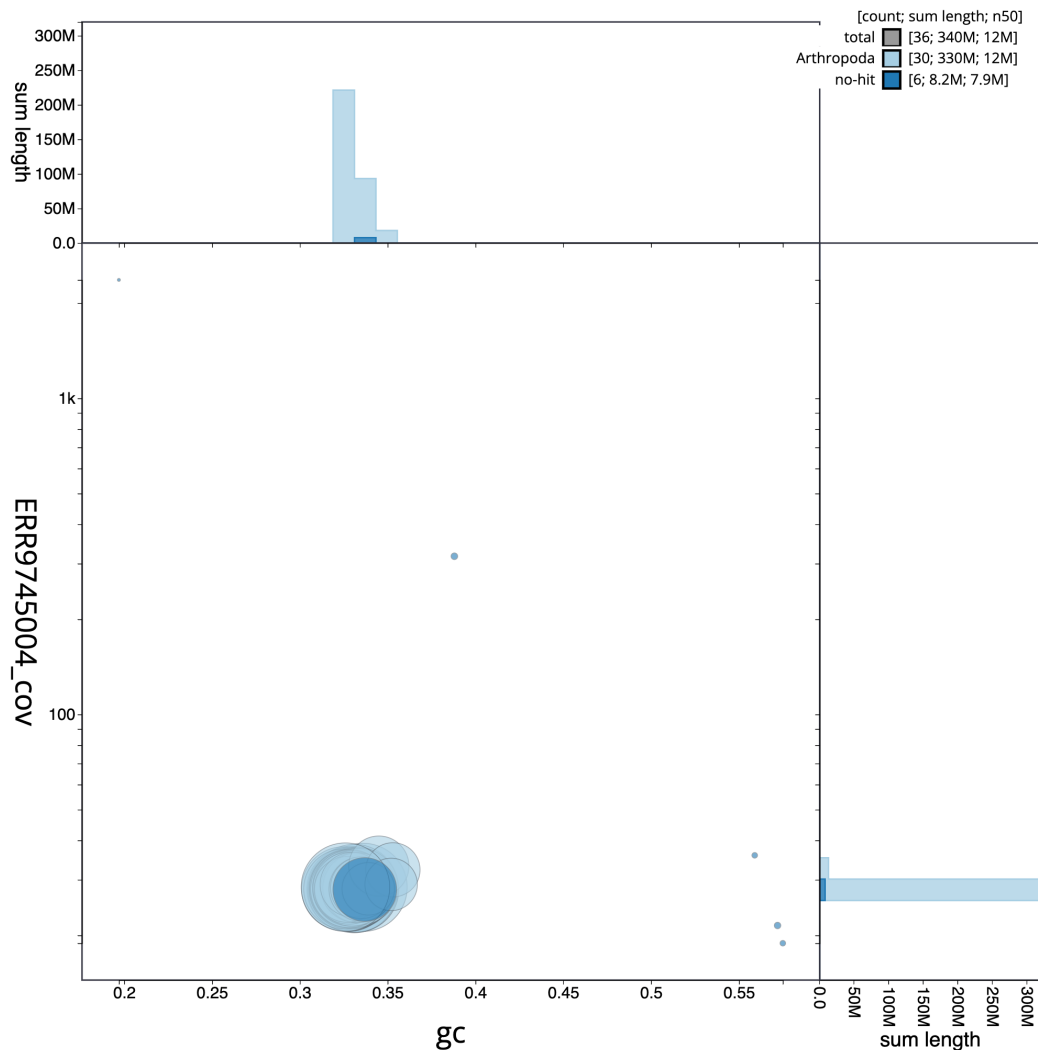
Figure 3. Genome assembly of *Acentria ephemerella*, ilAceEphe1.1: BlobToolKit GC-coverage plot. Scaffolds are coloured by phylum. Circles are sized in proportion to scaffold length. Histograms show the distribution of scaffold length sum along each axis. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/CALPDK01.1/dataset/CALPDK01.1/blob.

tissue was homogenised using a PowerMasher II tissue disruptor (Denton *et al.*, 2023a). HMW DNA was extracted using the Automated MagAttract v1 protocol (Sheerin *et al.*, 2023). DNA was sheared into an average fragment size of 12–20 kb in a Megaruptor 3 system with speed setting 30 (Todorovic *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. Poly(A) RNA-Seq libraries were constructed using the NEB Ultra II RNA Library Prep kit. DNA and RNA sequencing was performed by the Scientific Operations core at the WSI on Pacific Biosciences SEQUEL II (HiFi) and Illumina NovaSeq 6000 (RNA-Seq) instruments. Hi-C data were also generated from whole organism tissue of ilAceEphe2 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
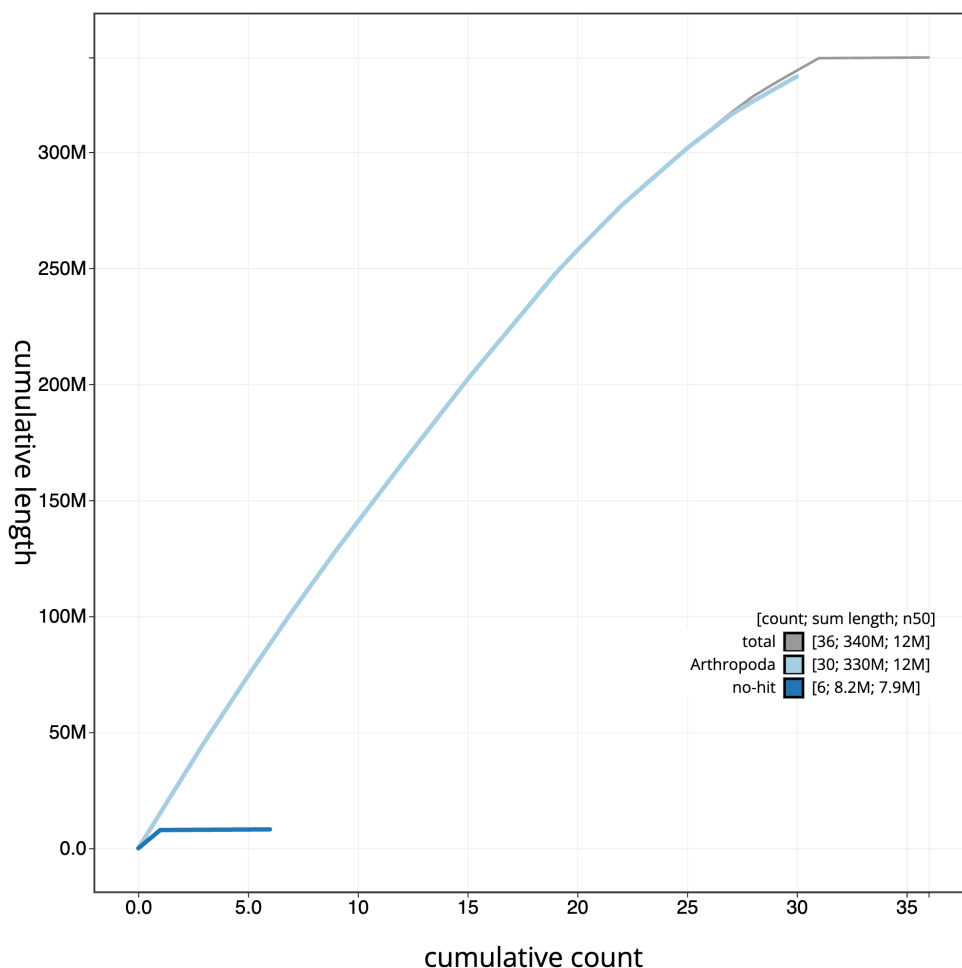
Figure 4. Genome assembly of *Acentria ephemerella*, ilAceEphe1.1: BlobToolKit cumulative sequence plot. The grey line shows cumulative length for all scaffolds. Coloured lines show cumulative lengths of scaffolds assigned to each phylum using the buscogenes taxrule. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/CALPDK01.1/dataset/CALPDK01.1/cumulative.

(Zhou *et al.*, 2023). The assembly was checked for contamination and corrected as described previously (Howe *et al.*, 2021). Manual curation was performed using HiGlass (Kerpedjiev *et al.*, 2018) and 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.

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.

Genome annotation
The BRAKER2 pipeline (Brůna *et al.*, 2021) was used in the default protein mode to generate annotation for the *Acentria ephemerella* assembly (GCA_943193645.1) in Ensembl Rapid Release at the EBI.

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'**,
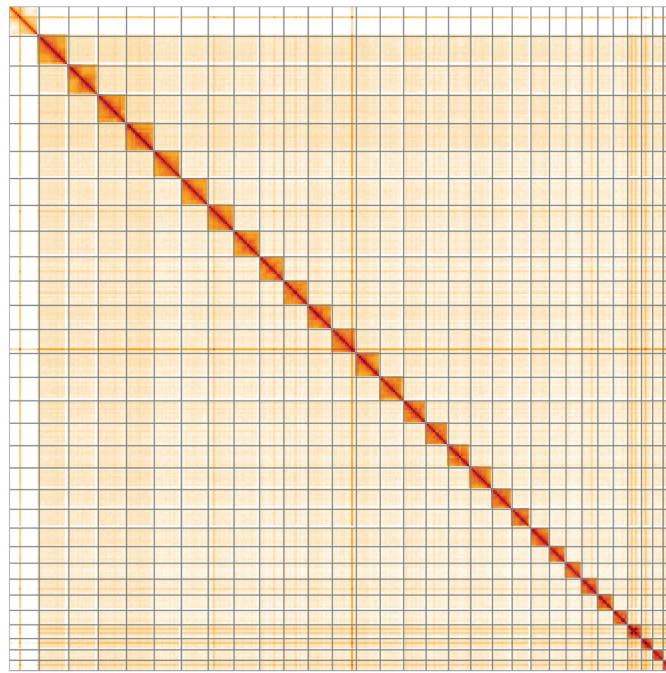
Figure 5. Genome assembly of *Acentria ephemerella*, ilAceEphe1.1: Hi-C contact map of the ilAceEphe1.1 assembly, visualised using HiGlass. Chromosomes are shown in order of size from left to right and top to bottom. An interactive version of this figure may be viewed at https://genome-note-higlass.tol.sanger.ac.uk/l/?d=EdX6IuLQSM6E-KiMJMUDVA.

Table 2. Chromosomal pseudomolecules in the genome assembly of *Acentria ephemerella*, ilAceEphe1.

INSDC accession	Chromosome	Length (Mb)	GC%
OW971875.1	1	15.2	33.5
OW971876.1	2	15.09	33.0
OW971877.1	3	14.45	33.0
OW971878.1	4	14.26	33.0
OW971879.1	5	13.84	32.5
OW971880.1	6	13.82	33.0
OW971881.1	7	13.21	33.0
OW971882.1	8	13.1	33.0
OW971883.1	9	12.48	32.5
OW971884.1	10	12.45	33.0
OW971885.1	11	12.41	33.0
OW971886.1	12	12.25	33.0
OW971887.1	13	12.24	33.0
OW971888.1	14	12.02	33.0
OW971889.1	15	11.51	33.5

INSDC accession	Chromosome	Length (Mb)	GC%
OW971890.1	16	11.49	33.0
OW971891.1	17	11.4	33.0
OW971892.1	18	11.14	33.0
OW971893.1	19	10.09	33.0
OW971894.1	20	9.6	33.0
OW971895.1	21	9.57	33.5
OW971896.1	22	8.44	32.5
OW971897.1	23	8.19	33.0
OW971898.1	24	8.14	33.5
OW971899.1	25	7.87	33.5
OW971900.1	26	7.41	33.5
OW971901.1	27	7.05	34.5
OW971902.1	28	5.74	35.5
OW971903.1	29	5.38	34.0
OW971904.1	30	5.36	35.0
OW971874.1	Z	15.31	32.5
OW971905.1	MT	0.02	20.0

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	yahs-1.1.91eebc2	https://github.com/c-zhou/yahs

which can be found in full on the Darwin Tree of Life website here. By agreeing with and signing up to the Sampling Code of Practice, the Darwin Tree of Life Partner agrees they will meet the legal and ethical requirements and standards set out within this document in respect of all samples acquired for, and supplied to, the Darwin Tree of Life Project.

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

- Ethical review of provenance and sourcing of the material

- Legality of collection, transfer and use (national and international)

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

Data availability
European Nucleotide Archive: *Acentria ephemerella* (water veneer). Accession number PRJEB52477; https://identifiers.org/ena.embl/PRJEB52477 (Wellcome Sanger Institute, 2022).

The genome sequence is released openly for reuse. The *Acentria ephemerella* 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.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.

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**

Batra SWT: **Bionomics of the aquatic moth, *Acentria niveus* (Oliver), a potential biological control agent for Eurasian watermilfoil and Hydrilla.** *J New York Entomol S.* 1977; **85**: 59–139.
Reference Source

Berg K: **Contribution to the biology of the aquatic moth *Acentropus niveus* (Oliv.).** *Vidensk Medd Dansk Naturh Foren.* 1942; **105**: 59–139.

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**

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**

Denton A, Oatley G, Cornwell C, *et al.*: **Sanger Tree of Life Sample Homogenisation: PowerMash.** *protocols.io.* 2023a.
Publisher Full Text

Denton A, Yatsenko H, Jay J, *et al.*: **Sanger Tree of Life Wet Laboratory Protocol Collection V.1.** *protocols.io.* 2023b.
Publisher Full Text

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

Gross EM, Johnson RL, Hairston NG: **Experimental evidence for changes in submersed macrophyte species composition caused by the herbivore *Acentria ephemerella* (Lepidoptera).** *Oecologia.* 2001; **127**(1): 105–114.
PubMed Abstract | **Publisher Full Text**

Gross EM, Feldbaum C, Choi C: **High abundance of herbivorous Lepidoptera larvae (*Acentria ephemerella* Denis & Schiffermüller) on submersed macrophytes in Lake Constance (Germany).** *Arch Hydrobiol.* 2002; **155**: 1–21.
Publisher Full Text

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

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

Howe K, Chow W, Collins J, *et al.*: **Significantly improving the quality of genome assemblies through curation.** *GigaScience.* Oxford University Press, 2021; **10**(1): giaa153.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Jay J, Yatsenko H, Narváez-Gómez JP, *et al.*: **Sanger Tree of Life Sample Preparation: Triage and Dissection.** *protocols.io.* 2023.
Publisher Full Text

Johnson RL, Gross EM, Hairston JNG: **Decline of the invasive submersed macrophyte *Myriophyllum spicatum* (Haloragaceae) associated with herbivory by larvae of *Acentria ephemerella* (Lepidoptera).** *Aquat Ecol.* 1997; **31**(3): 273–282.
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**

Lau TFS, Gross EM, Meyer-Rochow VB: **Sexual dimorphism and light/dark adaptation in the compound eyes of male and female *Acentria ephemerella* (Lepidoptera: Pyraloidea: Crambidae).** *Eur J Entomol.* 2007; **104**(3): 459–470.
Publisher Full Text

Léger T, Mally R, Neinhuis C, *et al.*: **Refining the phylogeny of Crambidae with complete sampling of subfamilies (Lepidoptera, Pyraloidea).** *Zool Scr.* 2021; **50**(1): 84–99.
Publisher 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**

Pabis K: **What is a moth doing under water? Ecology of aquatic and semi-aquatic Lepidoptera.** *Knowl Manag Aquat Ec.* 2018; (419): 42.
Publisher Full Text

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

Regier JC, Mitter C, Solis MA, *et al.*: **A molecular phylogeny for the Pyraloid moths (Lepidoptera: Pyraloidea) and its implications for higher-level classification.** *Syst Entomol.* 2012; **37**(4): 635–656.
Publisher 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**

Sheerin E, Sampaio F, Oatley G, *et al.*: **Sanger Tree of Life HMW DNA Extraction: Automated MagAttract v.1.** *protocols.io.* 2023.
Publisher 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**

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

Todorovic M, Sampaio F, Howard C: **Sanger Tree of Life HMW DNA Fragmentation: Diagenode Megaruptor®3 for PacBio HiFi.** *protocols.io.* 2023.
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 Water Veneer, *Acentria ephemerella* (Denis & Schiffermüller, 1775).** European Nucleotide Archive. [dataset], accession number PRJEB52477, 2022.

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**