Check for updates

DATA NOTE

The genome sequence of the Notch-wing Button, *Acleris emargana* (Fabricius, 1775) [version 1; peer review: awaiting peer review]

Douglas Boyes[1+], James Hammond [iD][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

[1]UK Centre for Ecology & Hydrology, Wallingford, England, UK
[2]University of Oxford, Oxford, England, UK

+ Deceased author

Abstract
We present a genome assembly from an individual male *Acleris emargana* (the Notch-wing Button; Arthropoda; Insecta; Lepidoptera; Tortricidae). The genome sequence is 691.4 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 16.34 kilobases in length. Gene annotation of this assembly on Ensembl identified 21,886 protein coding genes.

Keywords
Acleris emargana, Notch-wing Button, genome sequence, chromosomal, Lepidoptera

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.

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; **Hammond J**: Writing – Original Draft Preparation, Writing – Review & Editing;

Species taxonomy

Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda; Insecta; Dicondylia; Pterygota; Neoptera; Endopterygota; Amphiesmenoptera; Lepidoptera; Glossata; Neolepidoptera; Heteroneura; Ditrysia; Apoditrysia; Tortricoidea; Tortricidae; Tortricinae; Tortricini; *Acleris*; *Acleris emargana* (Fabricius, 1775) (NCBI:txid758706).

Background

The Notch-wing Button *Acleris emargana* (Fabricius, 1775) is a moth in the Tortricidae family. The species' vernacular name is a reference to the remarkable shape of the forewings, which show a large inward 'notch' along their leading edge. As in other members of its genus, the species is polymorphic (Bradley *et al.*, 1973). *Acleris emargana* has a Palearctic distribution, being found across northern Eurasia east to at least northern China and Altai and Buryatia in Russia (Karsholt *et al.*, 2005). The species' distribution at the eastern boundary of its range is poorly known due to confusion with its recently split sister taxon *Acleris effractana*, which has a Holarctic distribution (Karsholt *et al.*, 2005). *Acleris emargana* is widespread across the British Isles except in the Outer Hebrides where it appears to be replaced by *effractana* (Elliott *et al.*, 2018).

The species overwinters as an egg, with the larva feeding between May and July on *Salix* or *Populus*, and less commonly, *Betula*, *Corylus* or *Alnus* (Bradley *et al.*, 1973; Elliott *et al.*, 2018). The larva feeds from within a folded leaf or between spun leaves (Bradley *et al.*, 1973; Elliott *et al.*, 2018). Pupation occurs in the larval habitation or amongst moss. Adults occur between July and September, often found in damp areas where sallow (*Salix* ssp.) is established (Bradley *et al.*, 1973; Elliott *et al.*, 2018).

The genome of *Acleris emargana* 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 complete, chromosome-level genome sequence for *Acleris emargana*, based on one male specimen from Wytham Woods, Oxfordshire, UK.

Genome sequence report

The genome was sequenced from one male *Acleris emargana* (Figure 1) collected from Wytham Woods, Oxfordshire (51.77, –1.34). A total of 25-fold coverage in Pacific Biosciences single-molecule HiFi long reads and 67-fold coverage in 10X Genomics read clouds was generated. Primary assembly contigs were scaffolded with chromosome conformation Hi-C data. Manual assembly curation corrected 27 missing joins or mis-joins and removed three haplotypic duplications, reducing the assembly length by 0.99% and the scaffold number by 10.53%, and decreasing the scaffold N50 by 1.34%.

The final assembly has a total length of 691.4 Mb in 33 sequence scaffolds with a scaffold N50 of 22.8 Mb (Table 1). Most (99.98%) of the assembly sequence was assigned to

Figure 1. Photograph of the *Acleris emargana* (ilAclEmar1) specimen used for genome sequencing.

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 58.5 with *k*-mer completeness of 99.99%, and the assembly has a BUSCO v5.3.2 completeness of 97.8% (single = 97.1%, duplicated = 0.7%), 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/758706.

Genome annotation report

The *Acleris emargana* genome assembly (GCA_927399475.2) was annotated using the Ensembl rapid annotation pipeline (Table 1; https://rapid.ensembl.org/Acleris_emargana_GCA_927399445.1/Info/Index). The resulting annotation includes 22,022 transcribed mRNAs from 21,886 protein-coding genes.

Methods

Sample acquisition and nucleic acid extraction

The specimen used for DNA sequencing was a male *Acleris emargana* (specimen ID Ox000806, individual ilAclEmar1), collected from Wytham Woods, Oxfordshire (biological vice-county Berkshire), UK (latitude 51.77, longitude –1.34) on 2020-08-01 using a light trap. The specimen used for Hi-C sequencing (specimen ID Ox000951, individual ilAclEmar2) was collected from the same location on 2020-09-08. Both

Table 1. Genome data for *Acleris emargana*, ilAclEmar1.2.

Project accession data		
Assembly identifier	ilAclEmar1.2	
Species	*Acleris emargana*	
Specimen	ilAclEmar1	
NCBI taxonomy ID	758706	
BioProject	PRJEB48047	
BioSample ID	SAMEA7746613	
Isolate information	ilAclEmar1, male: whole organism (DNA sequencing) ilAclEmar2: whole organism (Hi-C scaffolding)	
Assembly metrics*		***Benchmark***
Consensus quality (QV)	58.5	*≥ 50*
k-mer completeness	99.99%	*≥ 95%*
BUSCO**	C:97.8%[S:97.1%,D:0.7%], F:0.5%,M:1.7%,n:5,286	*C ≥ 95%*
Percentage of assembly mapped to chromosomes	99.98%	*≥ 95%*
Sex chromosomes	Z chromosome	*localised homologous pairs*
Organelles	Mitochondrial genome assembled	*complete single alleles*
Raw data accessions		
PacificBiosciences SEQUEL IIe	ERR7123967	
10X Genomics Illumina	ERR7113542, ERR7113544, ERR7113543, ERR7113545	
Hi-C Illumina	ERR7167663	
Genome assembly		
Assembly accession	GCA_927399475.2	
Accession of alternate haplotype	GCA_927399445.1	
Span (Mb)	691.4	
Number of contigs	72	
Contig N50 length (Mb)	16.7	
Number of scaffolds	33	
Scaffold N50 length (Mb)	22.8	
Longest scaffold (Mb)	101.2	
Genome annotation		
Number of protein-coding genes	21,886	
Number of gene transcripts	22,022	

* 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 v5.3.2. C = complete [S = single copy, D = duplicated], F = fragmented, M = missing, n = number of orthologues in com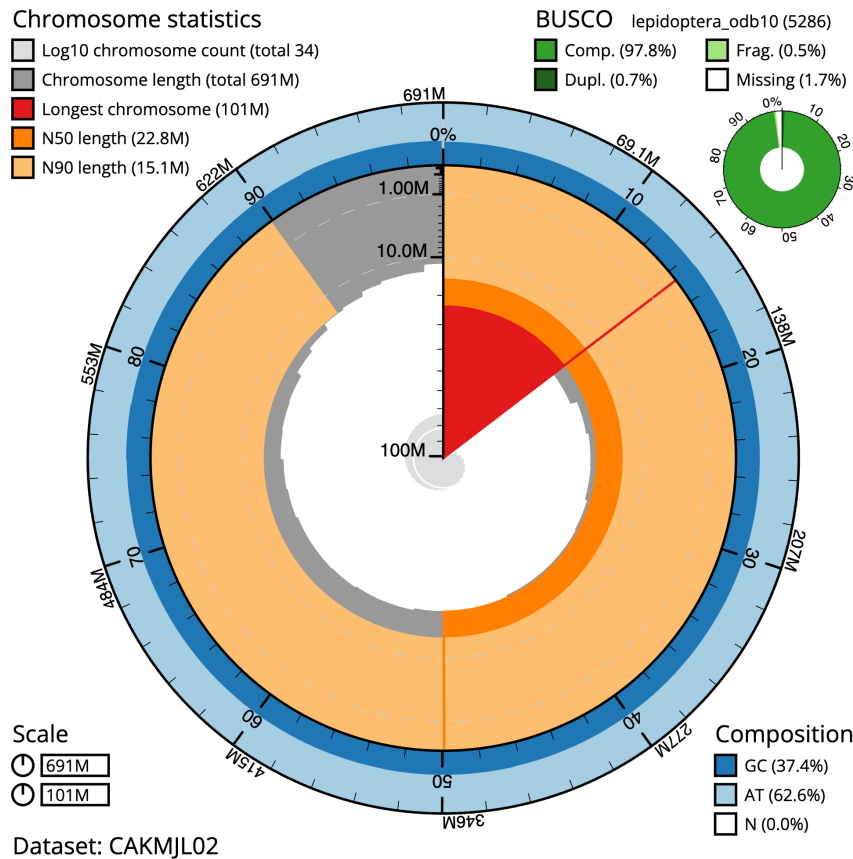parison. A full set of BUSCO scores is available at https://blobtoolkit.genomehubs.org/view/ilAclEmar1.2/dataset/CAKMJL02/busco.

Figure 2. Genome assembly of *Acleris emargana*, ilAclEmar1.2: 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 691,448,267 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 (101,224,011 bp, shown in red). Orange and pale-orange arcs show the N50 and N90 scaffold lengths (22,775,158 and 15,148,969 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/ilAclEmar1.2/dataset/CAKMJL02/snail.

specimens were collected and identified by Douglas Boyes (University of Oxford) and preserved on dry ice.

DNA was extracted at the Tree of Life laboratory, Wellcome Sanger Institute (WSI). The ilAclEmar1 sample was weighed and dissected on dry ice with tissue set aside for Hi-C sequencing. Tissue from the whole organism 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.

Sequencing
Pacific Biosciences HiFi circular consensus and 10X Genomics read cloud DNA sequencing libraries were constructed according to the manufacturers' instructions. DNA sequencing was performed by the Scientific Operations core at the WSI on Pacific Biosciences SEQUEL II (HiFi) and Illumina NovaSeq 6000 (10X) instruments. Hi-C data were also generated from whole organism tissue of ilAclEmar2 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). One round of polishing was
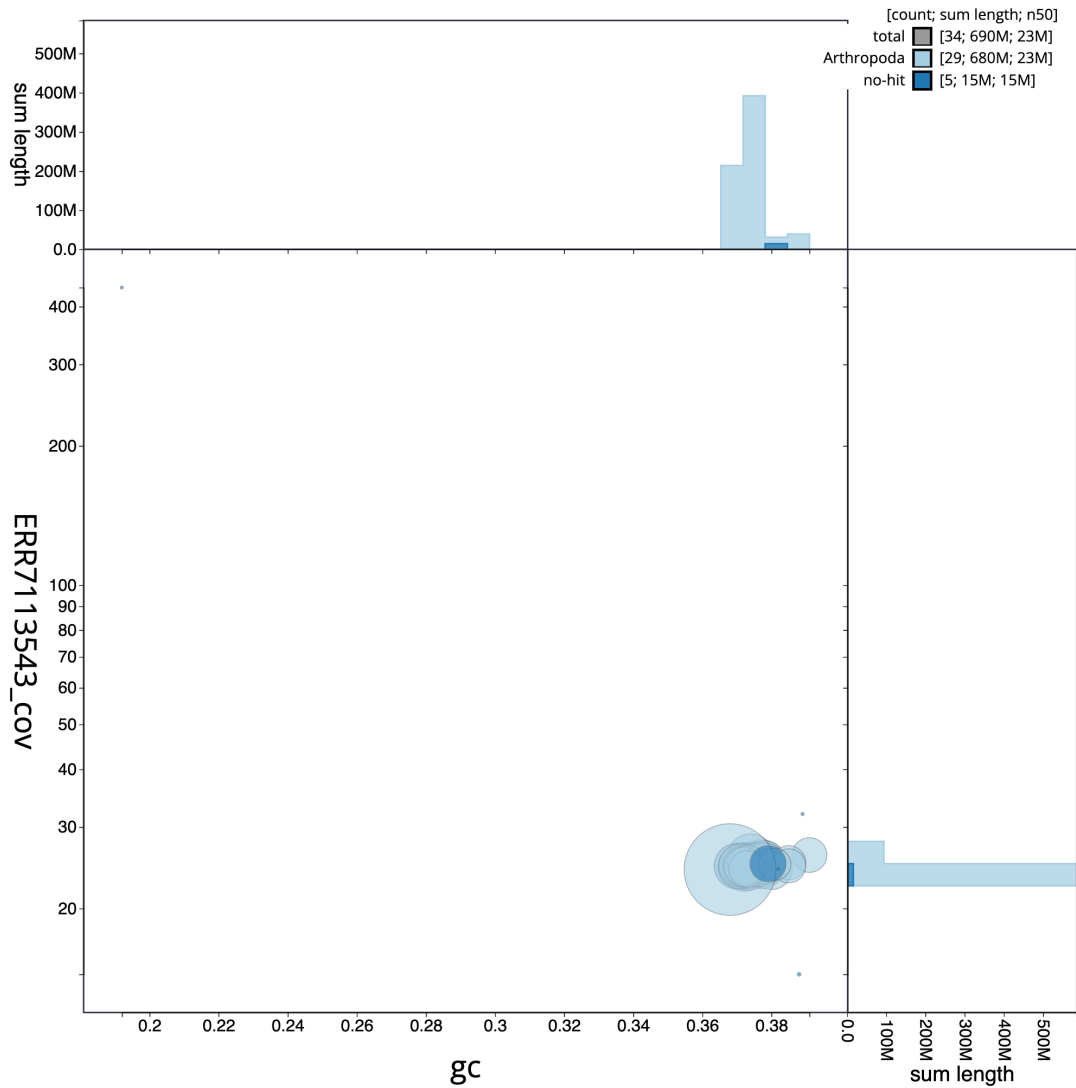
Figure 3. Genome assembly of *Acleris emargana*, ilAclEmar1.2: BlobToolKit GC-coverage plot. Scaffolds are coloured by phylum. Circles are sized in proportion to scaffold length. Histograms show the distribution of scaffold length sum along each axis. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/ilAclEmar1.2/dataset/CAKMJL02/blob.

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 YaHS (Zhou *et al.*, 2023). The assembly was checked for contamination and corrected as described previously (Howe *et al.*, 2021). Manual curation was performed using HiGlass (Kerpedjiev *et al.*, 2018) and Pretext (Harry, 2022). The mitochondrial genome was assembled using MitoHiFi (Uliano-Silva *et al.*, 2023), which runs MitoFinder (Allio *et al.*, 2020) or MITOS (Bernt *et al.*, 2013) and uses these

annotations to select the final mitochondrial contig and to ensure the general quality of the sequence.

A Hi-C map for the final assembly was produced using bwa-mem2 (Vasimuddin *et al.*, 2019) in the Cooler file format (Abdennur & Mirny, 2020). To assess the assembly 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).
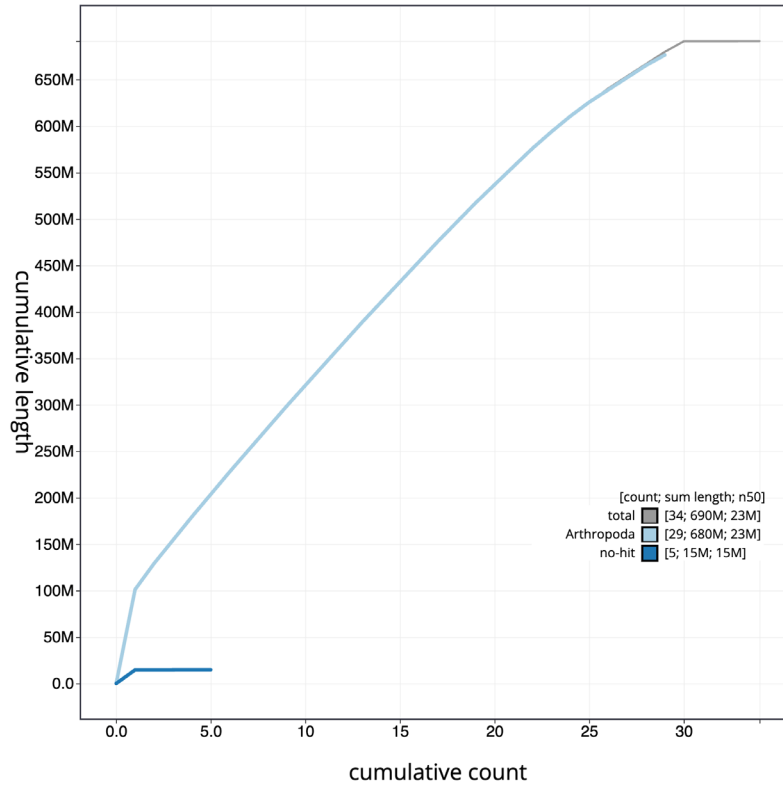
Figure 4. Genome assembly of *Acleris emargana*, **ilAclEmar1.2: BlobToolKit cumulative sequence plot.** The grey line shows cumulative length for all scaffolds. Coloured lines show cumulative lengths of scaffolds assigned to each phylum using the buscogenes taxrule. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/ilAclEmar1.2/dataset/CAKMJL02/cumulative.
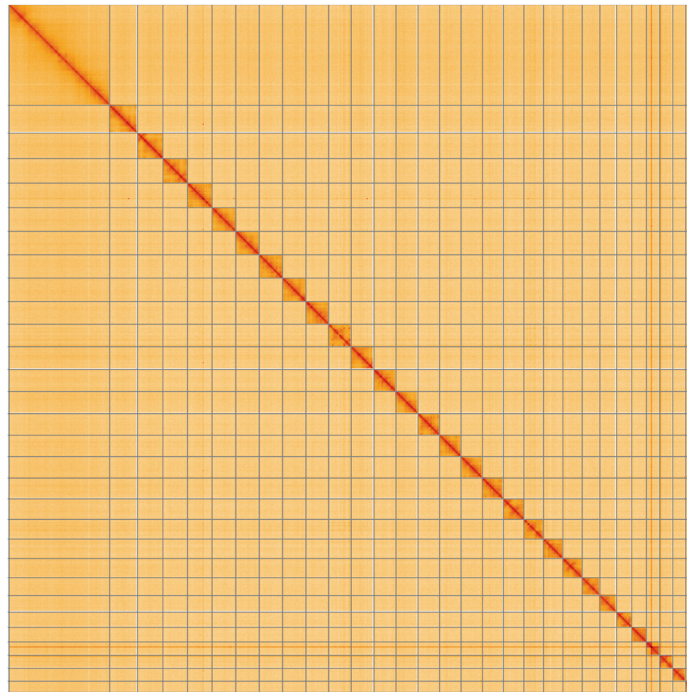

Figure 5. Genome assembly of *Acleris emargana*, **ilAclEmar1.2: Hi-C contact map of the ilAclEmar1.2 assembly, visualised using HiGlass.** Chromosomes are shown in order of size from left to right and top to bottom. An interactive version of this figure may be viewed at https://genome-note-higlass.tol.sanger.ac.uk/l/?d=CxX2ql3pQpegkujExGkbMw.

Table 2. Chromosomal pseudomolecules in the genome assembly of *Acleris emargana*, ilAclEmar1.

INSDC accession	Chromosome	Length (Mb)	GC%
OV656836.1	1	27.92	37.5
OV656837.1	2	25.53	37.5
OV656838.1	3	24.76	37.0
OV656839.1	4	24.64	37.5
OV656840.1	5	23.84	37.5
OV656841.1	6	23.5	37.0
OV656842.1	7	23.48	37.0
OV656843.1	8	23.47	37.0
OV656844.1	9	22.83	37.0
OV656845.1	10	22.78	37.5
OV656846.1	11	22.78	37.5
OV656847.1	12	22.19	37.0
OV656848.1	13	22.12	37.0
OV656849.1	14	21.72	37.5
OV656850.1	15	21.54	37.5
OV656851.1	16	21.51	37.5
OV656852.1	17	21.0	38.0
OV656853.1	18	20.65	37.0
OV656854.1	19	19.74	38.0
OV656855.1	20	19.52	37.0
OV656856.1	21	19.33	38.0
OV656857.1	22	17.83	37.5
OV656858.1	23	16.79	37.0
OV656859.1	24	15.15	37.5
OV656860.1	25	14.71	38.0
OV656861.1	26	13.54	39.0
OV656862.1	27	13.03	38.5
OV656863.1	28	12.96	38.5
OV656864.1	29	11.29	38.0
OV656835.1	Z	101.22	37.0
OV656865.2	MT	0.02	19.5

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 *Acleris emargana* assembly (GCA_927399475.2) in Ensembl Rapid Release.

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.

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
FreeBayes	1.3.1-17-gaa2ace8	https://github.com/freebayes/freebayes
gEVAL	N/A	https://geval.org.uk/
Hifiasm	0.12	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
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	https://github.com/c-zhou/yahs

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: *Acleris emargana* (notch-wing button). Accession number PRJEB48047; https://identifiers. org/ena.embl/PRJEB48047. (Wellcome Sanger Institute, 2022) The genome sequence is released openly for reuse. The *Acleris emargana* 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

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**

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**

Bradley JD, Tremewan WD, Smith A: **British Tortricoid Moths: Cochylidae and Tortricidae: Tortricinae.** The Ray Society, 1973.

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**

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**

Elliott B, Langmaid JR, Palmer RM, *et al.*: **Tortricidae.** In: Langmaid, J. R., Palmer, S. M., and Young, M. R. (eds.) *A Field Guide to the Smaller Moths of Great Britain and Ireland.* The British Entomological and Natural History Society, 2018; 279.

Garrison E, Marth G: **Haplotype-based variant detection from short-read sequencing.** 2012.
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**

Karsholt O, Aarvik L, Agassiz D, *et al.*: **Acleris effractana (Hübner, 1799) - a Holarctic Tortricid.** *Nota Lepidopterol.* 2005; **28**(2): 93–102.
Reference Source

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**

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**

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.
Publisher Full Text

Uliano-Silva M, Ferreira GJRN, 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 Notch-wing Button, *Acleris emargana* (Fabricius, 1775).** European Nucleotide Archive. [dataset], accession number PRJEB48047, 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**