DATA NOTE

The genome sequence of the Twin-spotted Quaker, *Anorthoa munda* (Denis & Schiffermüller, 1775) [version 1; peer review: awaiting peer review]

Douglas Boyes[1+], Owen T. Lewis[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

v1 **First published:** 30 Jun 2023, **8**:284
https://doi.org/10.12688/wellcomeopenres.19625.1
Latest published: 30 Jun 2023, **8**:284
https://doi.org/10.12688/wellcomeopenres.19625.1

Abstract

We present a genome assembly from an individual male *Anorthoa munda* (the Twin-spotted Quaker; Arthropoda; Insecta; Lepidoptera; Noctuidae). The genome sequence is 938.6 megabases in span. Most of the assembly is scaffolded into 27 chromosomal pseudomolecules, including the Z sex chromosome. The mitochondrial genome has also been assembled and is 15.33 kilobases in length. Gene annotation of this assembly on Ensembl identified 22,894 protein coding genes.

Keywords

Anorthoa munda, Twin-spotted Quaker, genome sequence, chromosomal, Lepidoptera

This article is included in the Tree of Life gateway.

Open Peer Review

Approval Status *AWAITING PEER REVIEW*

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

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

Author roles: Boyes D: Investigation, Resources; **Lewis OT**: 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) 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.

How to cite this article: Boyes D, Lewis OT, University of Oxford and Wytham Woods Genome Acquisition Lab *et al.* **The genome sequence of the Twin-spotted Quaker, *Anorthoa munda* (Denis & Schiffermüller, 1775) [version 1; peer review: awaiting peer review]** Wellcome Open Research 2023, **8**:284 https://doi.org/10.12688/wellcomeopenres.19625.1

First published: 30 Jun 2023, **8**:284 https://doi.org/10.12688/wellcomeopenres.19625.1

Species taxonomy

Eukaryota; Metazoa; Eumetazoa; Bilateria; Protostomia; Ecdysozoa; Panarthropoda; Arthropoda; Mandibulata; Pancrustacea; Hexapoda; Insecta; Dicondylia; Pterygota; Neoptera; Endopterygota; Amphiesmenoptera; Lepidoptera; Glossata; Neolepidoptera; Heteroneura; Ditrysia; Obtectomera; Noctuoidea; Noctuidae; Hadeninae; Anorthoa (Denis & Schiffermüller, 1775) (NCBI:txid988060).

Background

The Twin-spotted Quaker (*Anorthoa munda*) is the only representative of its genus in the UK. This species was formerly placed in the genus *Orthosia* alongside several similar spring-flying noctuid moth species known in English as "Quakers" or "Drabs", and the validity of the genus *Anorthoa* has been challenged (Craik, 2018).

Anorthoa munda has a wide distribution in woodland habitats throughout much of Britain and Ireland (Waring *et al.*, 2017). Formerly absent from much of Scotland, it has expanded its distribution there in recent decades, and is also flying significantly earlier in the year than in the 1970s (Randle *et al.*, 2019). Its distribution extends across temperate areas of Europe, Russian and Asia, eastwards to Japan (GBIF Secretariat, 2023).

Anorthoa munda is a variable species, with the ground colour of the wings varying from a pale buff colour to reddish-brown. The two black median dots on each forewing are distinctive but are occasionally reduced or absent. There is a single annual generation, with adults on the wing in March and April when they are attracted to light traps and can be found feeding on the blossom of sallow trees (*Salix* spp.) (Waring *et al.*, 2017). The caterpillars feed at night on a wide variety of trees and other woody plants, including ash (*Fraxinus excelsior*), oak (*Quercus* spp.), poplar (*Populus* spp.) and willow (*Salix* spp.) (Henwood *et al.*, 2020). These moths spend the winter as pupae within cocoons formed underground (Henwood *et al.*, 2020).

A genome sequence for *Anorthoa munda* will contribute to a growing data set of resources for understanding lepidopteran biology. The genome of *Anorthoa munda* 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 *Anorthoa munda*, based on one male specimen from Wytham Woods, Oxfordshire, UK.

Genome sequence report

The genome was sequenced from one male *Anorthoa munda* (Figure 1) collected from Woods, Oxfordshire, UK (51.77, –1.34). A total of 27-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 75 missing joins or mis-joins and removed 23 haplotypic duplications, reducing the assembly length by 2.15% and the scaffold number by 20.83%, and decreasing the scaffold N50 by 2.74%.

Figure 1. Photograph of the *Anorthoa munda* (ilAnoMund1) specimen used for genome sequencing.

The final assembly has a total length of 938.6 Mb in 38 sequence scaffolds with a scaffold N50 of 32.4 Mb (Table 1). Most (99.94%) of the assembly sequence was assigned to 27 chromosomal-level scaffolds, representing 26 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 63.9 with *k*-mer completeness of 100%, and the assembly has a BUSCO v5.3.2 completeness of 99.0% (single = 98.1%, duplicated = 0.9%), 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/988060.

Genome annotation report

The *Anorthoa munda* genome assembly (GCA_945859665.1) was annotated using the Ensembl rapid annotation pipeline (Table 1; https://rapid.ensembl.org/Anorthoa_munda_GCA_945859665.1/Info/Index). The resulting annotation includes 23,114 transcribed mRNAs from 22,894 protein-coding genes.

Methods

Sample acquisition and nucleic acid extraction

A male *Anorthoa munda* (specimen ID Ox001097, individual ilAnoMund1) was collected from Wytham Woods, Oxfordshire (biological vice-county Berkshire), UK (latitude

Table 1. Genome data for *Anorthoa munda*, ilAnoMund1.1.

Project accession data		
Assembly identifier	ilAnoMund1.1	
Species	*Anorthoa munda*	
Specimen	ilAnoMund1	
NCBI taxonomy ID	988060	
BioProject	PRJEB53934	
BioSample ID	SAMEA10107020	
Isolate information	ilAnoMund1, male: thorax (DNA sequencing), head (Hi-C scaffolding), abdomen (RNA sequencing)	
Assembly metrics*		***Benchmark***
Consensus quality (QV)	63.9	*≥ 50*
k-mer completeness	100%	*≥ 95%*
BUSCO**	C:99.0%[S:98.1%,D:0.9%], F:0.3%,M:0.7%,n:5,286	*C ≥ 95%*
Percentage of assembly mapped to chromosomes	99.94%	*≥ 95%*
Sex chromosomes	Z chromosome	*localised homologous pairs*
Organelles	Mitochondrial genome assembled	*complete single alleles*
Raw data accessions		
PacificBiosciences SEQUEL II	ERR9902015	
Hi-C Illumina	ERR9904199	
PolyA RNA-Seq Illumina	ERR10890689	
Genome assembly		
Assembly accession	GCA_945859665.1	
Accession of alternate haplotype	GCA_945859655.1	
Span (Mb)	938.6	
Number of contigs	124	
Contig N50 length (Mb)	14.2	
Number of scaffolds	38	
Scaffold N50 length (Mb)	32.4	
Longest scaffold (Mb)	66.4	
Genome annotation		
Number of protein-coding genes	22,894	
Number of gene transcripts	23,114	

* 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 comparison. A full set of BUSCO scores is available at https://blobtoolkit.genomehubs.org/view/ilAnoMund1.1/dataset/CAMAOL01/busco.

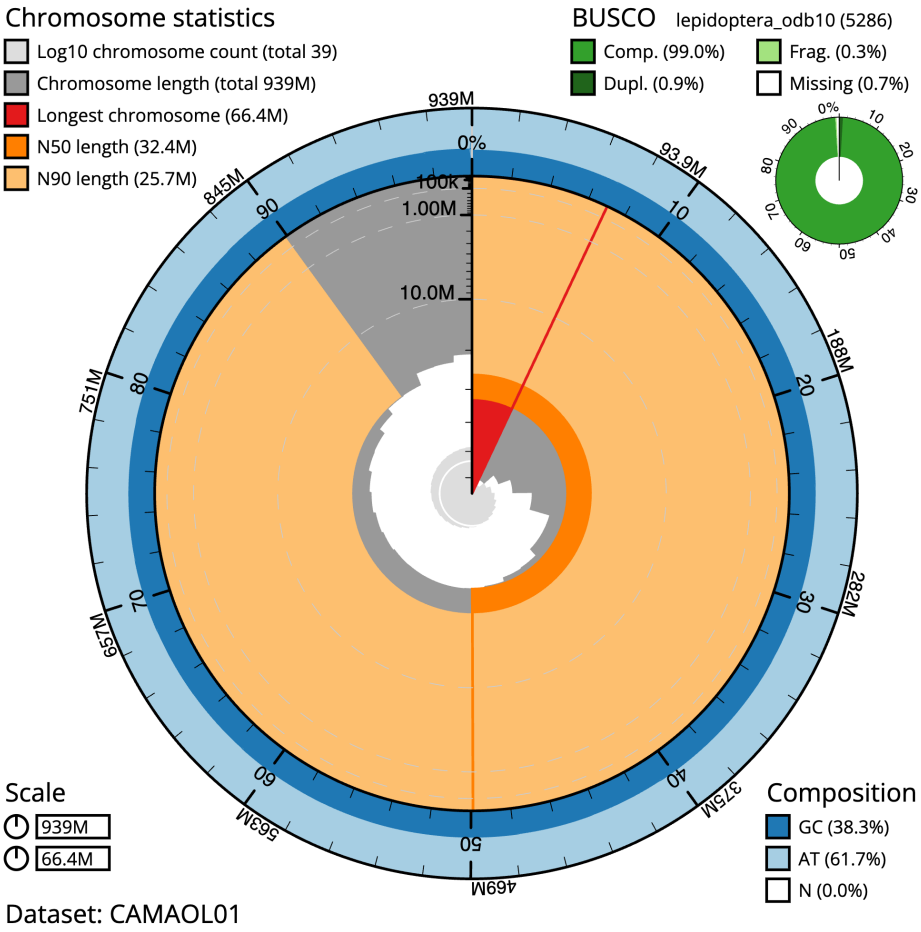

Figure 2. Genome assembly of *Anorthoa munda*, ilAnoMund1.1: metrics. The BlobToolKit Snailplot shows N50 metrics and BUSCO gene completeness. The main plot is divided into 1,000 size-ordered bins around the circumference with each bin representing 0.1% of the 938,660,760 bp assembly. The distribution of scaffold lengths is shown in dark grey with the plot radius scaled to the longest scaffold present in the assembly (66,421,374 bp, shown in red). Orange and pale-orange arcs show the N50 and N90 scaffold lengths (32,402,751 and 25,736,201 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/ilAnoMund1.1/dataset/CAMAOL01/snail.

51.77, longitude –1.34) on 2021-03-31, using a light trap. The specimen was 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 ilAnoMund1 sample was weighed and dissected on dry ice with tissue set aside for Hi-C sequencing. Thorax tissue was cryogenically disrupted to a fine powder using a Covaris cryoPREP Automated Dry Pulveriser, receiving multiple impacts. 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.

RNA was extracted from abdomen tissue of ilAnoMund1 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.

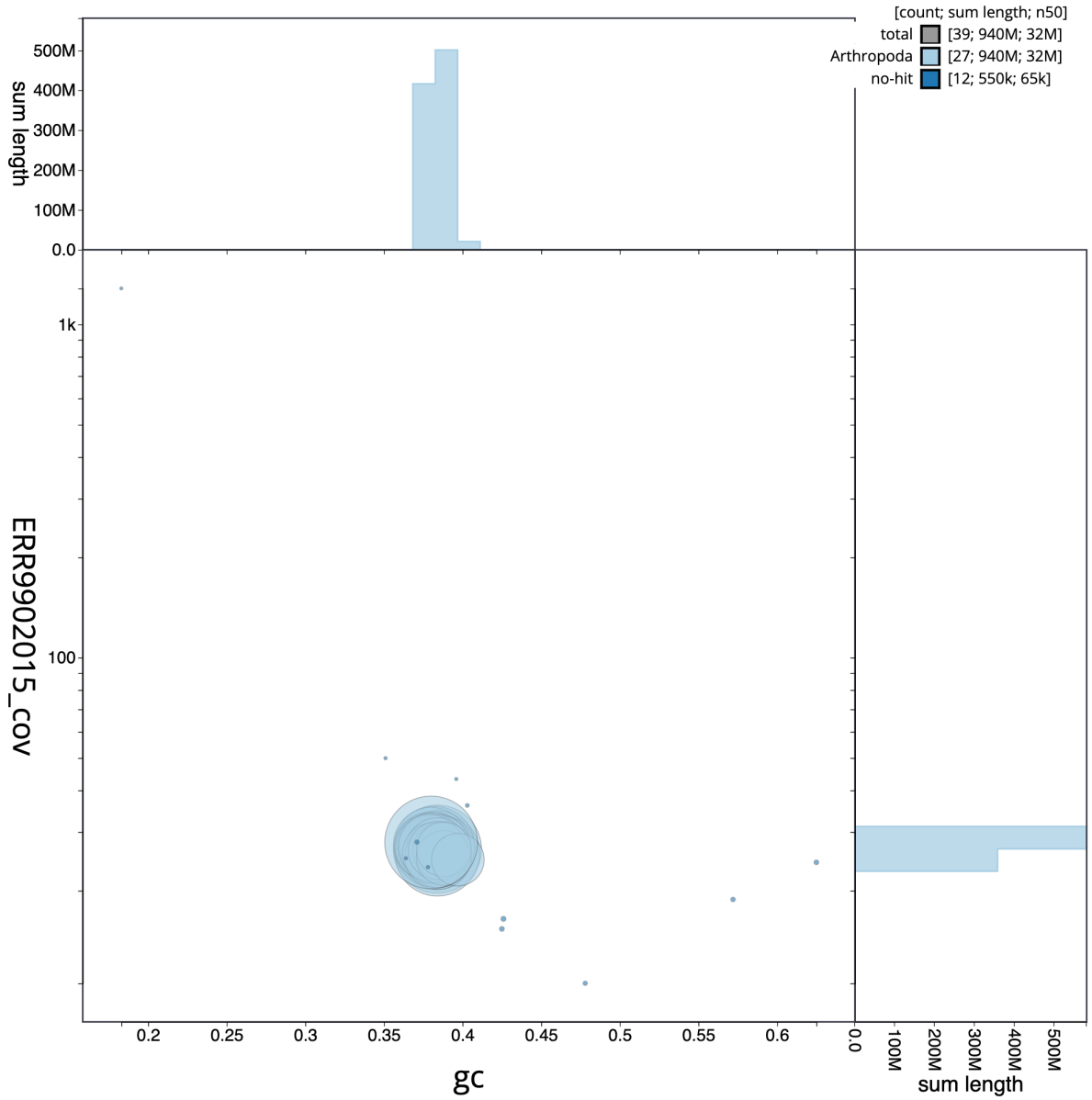

Figure 3. Genome assembly of *Anorthoa munda*, ilAnoMund1.1: BlobToolKit GC-coverage plot. Scaffolds are coloured by phylum. Circles are sized in proportion to scaffold length. Histograms show the distribution of scaffold length sum along each axis. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/ilAnoMund1.1/dataset/CAMAOL01/blob.

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 head tissue of ilAnoMund1 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 (Kerpedjiev *et al.*, 2018) and Pretext (Harry, 2022). The mitochondrial genome was assembled using MitoHiFi (Uliano-Silva *et al.*, 2022), which runs MitoFinder (Allio *et al.*, 2020) or MITOS

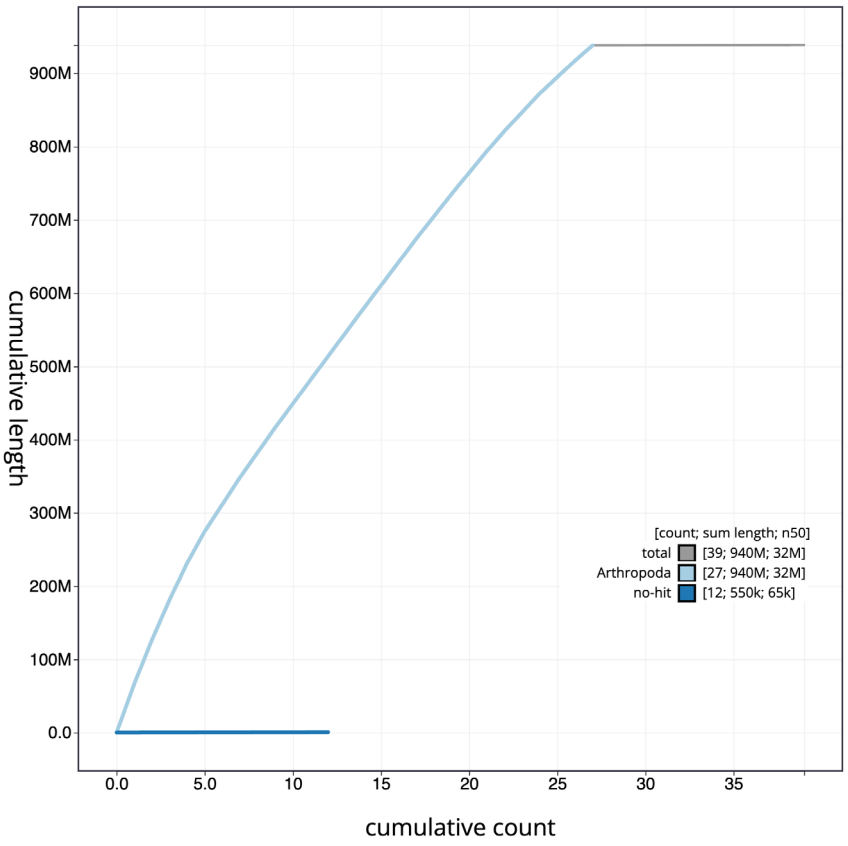

Figure 4. Genome assembly of *Anorthoa munda*, ilAnoMund1.1: BlobToolKit cumulative sequence plot. The grey line shows cumulative length for all scaffolds. Coloured lines show cumulative lengths of scaffolds assigned to each phylum using the buscogenes taxrule. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/ilAnoMund1.1/dataset/CAMAOL01/cumulative.

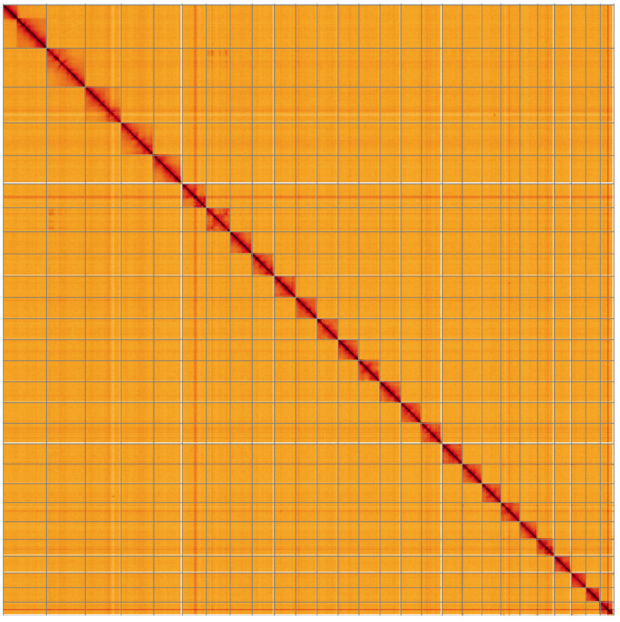

Figure 5. Genome assembly of *Anorthoa munda*, ilAnoMund1.1: Hi-C contact map of the ilAnoMund1.1 assembly, visualised using HiGlass. Chromosomes are shown in order of size from left to right and top to bottom. An interactive version of this figure may be viewed at https://genome-note-higlass.tol.sanger.ac.uk/l/?d=EUPer_6MQtixsJxzXUc3aA.

Table 2. Chromosomal pseudomolecules in the genome assembly of *Anorthoa munda*, ilAnoMund1.

INSDC accession	Chromosome	Length (Mb)	GC%
OX243894.1	1	59.81	38.5
OX243895.1	2	54.9	38.5
OX243896.1	3	50.27	38.0
OX243897.1	4	43.44	38.0
OX243898.1	5	36.76	38.0
OX243899.1	6	36.75	38.5
OX243900.1	7	34.24	38.5
OX243901.1	8	33.99	38.0
OX243902.1	9	32.83	38.0
OX243903.1	10	32.4	38.5
OX243904.1	11	32.32	38.5
OX243905.1	12	32.31	38.5
OX243906.1	13	32.29	38.5
OX243907.1	14	32.23	38.0
OX243908.1	15	31.84	38.0
OX243909.1	16	31.45	38.5
OX243910.1	17	30.78	38.0
OX243911.1	18	30.22	38.0
OX243912.1	19	29.3	38.5
OX243913.1	20	29.17	38.5
OX243914.1	21	27.52	38.0
OX243915.1	22	25.74	38.5
OX243916.1	23	25.57	38.5
OX243917.1	24	22.46	39.0
OX243918.1	25	22.35	39.0
OX243919.1	26	20.76	39.5
OX243893.1	Z	66.42	38.0
OX243920.1	MT	0.02	18.5

(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 *Anorthoa munda* assembly (GCA_945859665.1) 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	4.0.7	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

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: *Anorthoa munda* (twin-spotted quaker). Accession number PRJEB53934; https://identifiers.org/ena.embl/PRJEB53934. (Wellcome Sanger Institute, 2022)

The genome sequence is released openly for reuse. The *Anorthoa munda* 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

Brůna T, Hoff KJ, Lomsadze A: **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

Craik C: **Changes in Name Only.** *Atropos.* 2018; **60**: 24–33.

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

GBIF Secretariat: ***Anorthoa munda* (Denis & Schiffermüller), 1775**. *GBIF Backbone Taxonomy*. 2023; [Accessed 3 June 2023].
Reference Source

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

Henwood B, Sterling P, Lewington R: **Field Guide to the Caterpillars of Great Britain and Ireland.** London: Bloomsbury, 2020.
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

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

Randle Z, Evans-Hill LJ, Parsons MS, *et al.*: **Atlas of Britain & Ireland's Larger Moths.** Newbury: NatureBureau, 2019.
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, 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.
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.** *BioRxiv.* 2022.
Publisher 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

Waring P, Townsend M, Lewington R: **Field Guide to the Moths of Great Britain and Ireland: Third Edition.** Bloomsbury Wildlife Guides, 2017.

Wellcome Sanger Institute: **The genome sequence of the Twin-spotted Quaker, *Anorthoa munda* (Denis & Schiffermüller, 1775).** European Nucleotide Archive. [dataset], accession number PRJEB53934, 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