



## DATA NOTE

# The genome sequence of the Lesser Horseshoe Bat, *Rhinolophus hipposideros* (Bechstein, 1800) (Chiroptera: Rhinolophidae)

[version 1; peer review: 3 approved, 1 not approved]

Ine Alvarez van Tussenbroek<sup>1</sup>, Sonja C. Vernes<sup>id</sup><sup>1</sup>, Haris Nicolaou<sup>2</sup>, Emma C. Teeling<sup>id</sup><sup>1,3,4</sup>, Meike Mai<sup>id</sup><sup>1</sup>, Myrtani Pieri<sup>5</sup>, Bat1K Consortium, 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

<sup>1</sup>School of Biology, University of St Andrews, St Andrews, Scotland, UK<sup>2</sup>Department of Forests, Ministry of Agriculture, Rural Development and Environment, Nicosia, Cyprus<sup>3</sup>University College Dublin, Dublin, Leinster, Ireland<sup>4</sup>Tree of Life Programme, Wellcome Sanger Institute, Hinxton, England, UK<sup>5</sup>School of Life and Health Sciences, University of Nicosia, Nicosia, Cyprus

**V1** First published: 10 Apr 2026, 11:212  
<https://doi.org/10.12688/wellcomeopenres.26210.1>  
 Latest published: 10 Apr 2026, 11:212  
<https://doi.org/10.12688/wellcomeopenres.26210.1>

## Abstract

We present a genome assembly from an individual female *Rhinolophus hipposideros* (Lesser Horseshoe Bat; Chordata; Mammalia; Chiroptera; Rhinolophidae). The assembly contains two haplotypes with total lengths of 2 171.31 megabases and 2 240.90 megabases. Most of haplotype 1 (96.52%) is scaffolded into 29 chromosomal pseudomolecules, including the X sex chromosome. Most of haplotype 2 (93.46%) is scaffolded into 29 chromosomal pseudomolecules, including the X sex chromosome. The mitochondrial genome has also been assembled, with a length of 16.85 kilobases. Gene annotation of this assembly on Ensembl identified 18 700 protein-coding genes. This assembly was generated as part of the Darwin Tree of Life project, which produces reference genomes for eukaryotic species found in Britain and Ireland.

## Keywords

*Rhinolophus hipposideros*; Lesser Horseshoe Bat; genome sequence; chromosomal; Chiroptera

## Open Peer Review

Approval Status

|                  | 1                    | 2                    | 3                    | 4                    |
|------------------|----------------------|----------------------|----------------------|----------------------|
| <b>version 1</b> |                      |                      |                      |                      |
| 10 Apr 2026      | <a href="#">view</a> | <a href="#">view</a> | <a href="#">view</a> | <a href="#">view</a> |

1. **Graham Etherington**<sup>id</sup>, Earlham Institute, Norwich, UK
2. **Xiuguang Mao**<sup>id</sup>, East China Normal University, Shanghai, China
3. **Jiang Feng**, Northeast Normal University, Changchun, China
4. **Michael Hiller**<sup>id</sup>, Senckenberg Research Institute, Frankfurt, Germany

Any reports and responses or comments on the



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

article can be found at the end of the article.



This article is included in the [Bat1K genomes](#) collection.

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

**Author roles:** **Alvarez van Tussenbroek I:** Investigation, Resources, Writing – Review & Editing; **Vernes SC:** Investigation, Resources, Writing – Original Draft Preparation; **Nicolaou H:** Investigation, Resources; **Teeling EC:** Writing – Review & Editing; **Mai M:** Resources, Writing – Review & Editing; **Pieri M:** Writing – Original Draft Preparation;

**Competing interests:** No competing interests were disclosed.

**Grant information:** This work was supported by Wellcome through core funding to the Wellcome Sanger Institute (220540) and the Darwin Tree of Life Discretionary Award [218328, <https://doi.org/10.35802/218328>]. SCV was supported by a UKRI Future Leaders Fellowship (MR/T021985/1) and ERC Consolidator Grant (101001702; BATSPEAK). ECT is supported by Irish Research Council Laureate Award (IRCLA/2017/58) and Science Foundation Ireland Future Frontiers (19/FFP/6790). *The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

**Copyright:** © 2026 Alvarez van Tussenbroek I *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:** Alvarez van Tussenbroek I, Vernes SC, Nicolaou H *et al.* **The genome sequence of the Lesser Horseshoe Bat, *Rhinolophus hipposideros* (Bechstein, 1800) (Chiroptera: Rhinolophidae) [version 1; peer review: 3 approved, 1 not approved]** Wellcome Open Research 2026, 11:212 <https://doi.org/10.12688/wellcomeopenres.26210.1>

**First published:** 10 Apr 2026, 11:212 <https://doi.org/10.12688/wellcomeopenres.26210.1>

## Species taxonomy

Eukaryota; Opisthokonta; Metazoa; Eumetazoa; Bilateria; Deuterostomia; Chordata; Craniata; Vertebrata; Gnathostomata; Teleostomi; Euteleostomi; Sarcopterygii; Dipnotetrapodomorpha; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Boreoeutheria; Laurasiatheria; Chiroptera; Yinpterochiroptera; Rhinolophoidea; Rhinolophidae; Rhinolophinae; *Rhinolophus*; *Rhinolophus hipposideros* (Bechstein, 1800) (NCBI:txid77218).

## Background

The lesser horseshoe bat, *R. hipposideros*, is a small, western Palearctic bat species, typically weighing 4–9 g, with a forearm length of 35–42 mm (Wilson & Mittermeier, 2019). It is one of the smallest members of the genus *Rhinolophus* in Europe and is readily identified by its distinctive horseshoe-shaped noseleaf, which it uses to emit highly directional echolocation calls. Its fur is soft, grey-brown dorsally and more pale ventrally. Identification in the field is primarily based on facial morphology and cranial morphology (e.g. Kargopoulos *et al.*, 2023), noseleaf structure, and characteristic echolocation call frequency, although the last can overlap with *Rhinolophus mehelyi* in some regions (Salsamendi *et al.*, 2006). In such cases, advanced image-based morphometrics and genetic tools can provide improved accuracy (Cao *et al.*, 2023; Heller & von Helversen, 1989).

The global distribution of *R. hipposideros* is extensive across Europe, North Africa, and the Middle East, with confirmed records from Iberia, the British Isles, Central Europe, the Balkans, Anatolia, and the Maghreb, as reported in GBIF.org and phylogeographic studies (Dool *et al.*, 2013). The species recolonised Europe from multiple glacial refugia following the Last Glacial Maximum, maintaining persistent genetic lineages in the Ibero-Maghreb and Anatolia/Middle East regions (Dool *et al.*, 2013). Within this broad range, *R. hipposideros* exhibits flexibility in roost selection but remains sensitive to microclimatic stability and prey availability. It is assessed as Least Concern on the IUCN Red List. However, there are population declines across parts of its range linked to roost disturbance, landscape fragmentation, pesticide use, and decreasing insect abundance. The species is noted as regionally extinct in the Netherlands and Luxembourg, emphasising its vulnerability in intensively farmed or urbanised regions.

*R. hipposideros* is adapted to temperate environments with moderate climatic variation. It depends on warm and stable roosting conditions during the breeding season and utilizes torpor during periods of low insect activity. Its high-frequency echolocation calls, typically ranging between 105 and 115 kHz, allow for the detection of small prey in cluttered woodland and hedgerow habitats (Heller & von Helversen, 1989). Foraging behavior is characterised by short, slow, and highly manoeuvrable flights close to vegetation, reflecting adaptation to complex acoustic environments. These traits contribute to the species' ability to exploit diverse landscapes, from Mediterranean woodlands to temperate forests, though its sensitivity to roost microclimate and insect density limits resilience to environmental change.

The species shows considerable regional variation in roosting ecology. In northern and western Europe, maternity colonies are mainly found in buildings such as attics or lofts, where stable temperature conditions can be maintained throughout the reproductive period. In contrast, in southern and eastern parts of its range, natural underground roosts such as caves, abandoned mines, and cellars are more frequently used. Maternity colonies form in spring, typically consisting of several dozen females, each giving birth to a single pup annually. Non-invasive genetic monitoring has revealed dynamic changes in colony composition across the breeding season, with pregnant and nonpregnant females dominating before parturition, and mothers with dependent offspring prevailing afterward (Zarzoso-Lacoste *et al.*, 2018). Seasonal use of multiple roosts within the same area is also common, indicating behavioral flexibility in response to environmental or anthropogenic pressures.

*R. hipposideros* exhibits notable chromosomal diversity, with three karyotypic variants ( $2n = 54, 56, 58$ ) distributed geographically across Europe and Asia Minor (Kacprzyk *et al.*, 2016; Puerma *et al.*, 2008; Volleth *et al.*, 2013). This variation provides an important model for studying chromosomal evolution and speciation within the genus *Rhinolophus*. The species also demonstrates resource partitioning through subtle variation in echolocation frequency bands, facilitating coexistence with other rhinolophid bats occupying the same habitats (Heller & von Helversen, 1989). Ecologically, *R. hipposideros* is a voracious predator of small dipterans and lepidopterans, many of which are agricultural pests, contributing to ecosystem services such as pest control and biodiversity maintenance in Mediterranean and temperate agroecosystems (Baroja *et al.*, 2019). The species' foraging activity is strongly influenced by habitat structure and availability of sheltered flight corridors, making it a sensitive indicator of landscape integrity and ecological connectivity.

While several high-profile sequencing efforts have focused on related species such as *Rhinolophus ferrumequinum* (Jebb *et al.*, 2020), the distinctive karyotypes and rDNA organisation patterns in *R. hipposideros* underline its evolutionary uniqueness (Kacprzyk *et al.*, 2016; Puerma *et al.*, 2008; Volleth *et al.*, 2013). A complete genome sequence for *R. hipposideros* will support a broad range of applications, including improved species identification and delimitation,

conservation genetics, and deeper understanding of adaptive traits such as immunity, echolocation, and hibernation physiology. Furthermore, genome-scale data will facilitate monitoring of population structure and gene flow across fragmented landscapes, providing valuable insights for conservation management and disease surveillance.

Here we present a chromosome-level genome sequence for *R. hipposideros*. The sequence is based on tissue obtained from a male collected in an abandoned shed in Kalopanayiotis, Cyprus. This sampling is part of the [Bat1K Project](#) ([Teeling et al., 2018](#)) and the Darwin Tree of Life Project (DTOL). The Bat1K is a collaborative effort to sequence all extant bat species, and DTOL aims to sequence all named eukaryotic species in the Atlantic Archipelago of Britain and Ireland.

## Methods

### Sample acquisition

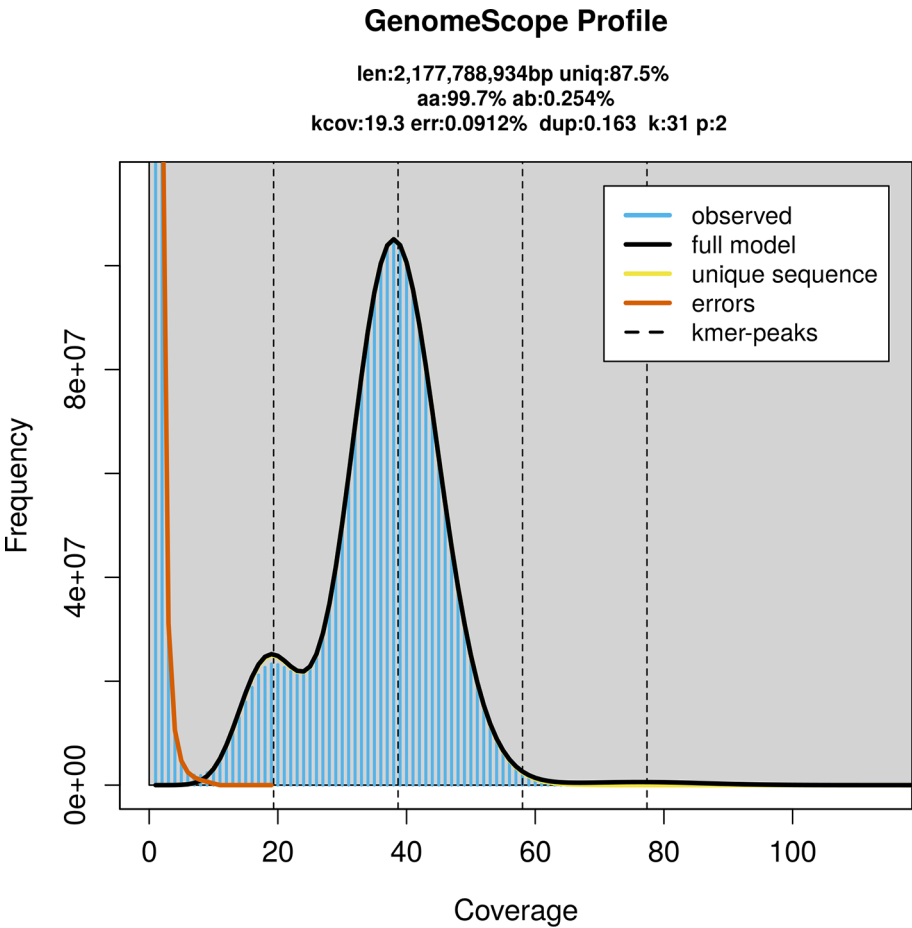
The specimen used for genome sequencing was an adult female *R. hipposideros* (specimen ID SAN00002493, ToLID mRhiHip2; [Figure 1](#)), collected from Department of Biological Sciences, University of Cyprus, Nicosia, Cyprus (latitude 35.1464, longitude 33.4098) on 2022-12-05. A group of approximately 60 *R. hipposideros* were observed roosting in an abandoned toolshed near a river, in Kalopanayiotis, Cyprus (latitude 35.1464, longitude 33.4098) on 2022-12-05. While this species tends to breed and hibernate at different locations, this particular group uses the same space for both. One female individual was collected from this group. It was transported to a laboratory in Nicosia, where it was euthanised, dissected and multiple organs were extracted. Samples were placed in Eppendorf tubes, snap frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  until shipping for sequencing. The specimen was collected by Ine Alvarez Van Tussenbroek and identified by Haris Nicolaou. The same specimen was used for RNA sequencing.

### Nucleic acid extraction

Protocols for high molecular weight (HMW) DNA extraction developed at the Wellcome Sanger Institute (WSI) Tree of Life Core Laboratory are available on [protocols.io](#) ([Howard et al., 2025](#)). The mRhiHip2 sample was weighed and [triaged](#) to determine the appropriate extraction protocol. Tissue from the heart was homogenised by [powermashing](#) using a PowerMasher II tissue disruptor. HMW DNA was extracted using the [Automated MagAttract v2](#) protocol. DNA was sheared into an average fragment size of 12–20 kb following the [Megaruptor@3 for LI PacBio](#) protocol. Sheared DNA was purified by [automated SPRI](#) (solid-phase reversible immobilisation). The concentration of the sheared and purified DNA was assessed using a Nanodrop spectrophotometer and Qubit Fluorometer using the Qubit dsDNA High Sensitivity



**Figure 1.** Lesser horseshoe bat, *Rhinolophus hipposideros*. Individuals of *R. hipposideros*. (A-B) Specimen was collected from a group roosting in an abandoned toolshed in Cyprus (Photos taken by Ine Alvarez van Tussenbroek).



**Figure 2. Frequency distribution of *k*-mers generated using GenomeScope2.** The plot shows observed and modelled *k*-mer spectra, providing estimates of genome size, heterozygosity, and repeat content based on unassembled sequencing reads.

Assay kit. Fragment size distribution was evaluated by running the sample on the FemtoPulse system. For this sample, the final post-shearing DNA had a Qubit concentration of 9.96 ng/μL and a yield of 1 294.80 ng.

RNA was extracted from brain tissue of mRhiHip2 in the Tree of Life Laboratory at the WSI using the [RNA Extraction: Automated MagMax™ mirVana protocol](#). The RNA concentration was assessed using a Nanodrop spectrophotometer

**Table 1. Specimen and sequencing data for BioProject PRJEB75441.**

| Platform                      | PacBio HiFi    | Hi-C               | RNA-seq            |
|-------------------------------|----------------|--------------------|--------------------|
| ToLID                         | mRhiHip2       | mRhiHip2           | mRhiHip2           |
| Specimen ID                   | SAN00002493    | SAN00002493        | SAN00002493        |
| BioSample (source individual) | SAMEA113980802 | SAMEA113980802     | SAMEA113980802     |
| BioSample (tissue)            | SAMEA113980896 | SAMEA113980894     | SAMEA113980901     |
| Tissue                        | heart          | muscle             | brain              |
| Instrument                    | Revio          | Illumina NovaSeq X | Illumina NovaSeq X |
| Run accessions                | ERR13033477    | ERR13063106        | ERR13493949        |
| Read count total              | 7.91 million   | 1 062.69 million   | 93.25 million      |
| Base count total              | 86.90 Gb       | 160.47 Gb          | 14.08 Gb           |



and a Qubit Fluorometer using the Qubit RNA Broad-Range Assay kit. Analysis of the integrity of the RNA was done using the Agilent RNA 6000 Pico Kit and Eukaryotic Total RNA assay.

### PacBio HiFi library preparation and sequencing

Library preparation and sequencing were performed at the WSI Scientific Operations core. Libraries were prepared using the SMRTbell Prep Kit 3.0 (Pacific Biosciences, California, USA), following the manufacturer's instructions. The kit includes reagents for end repair/A-tailing, adapter ligation, post-ligation SMRTbell bead clean-up, and nuclease treatment. Size selection and clean-up were performed using diluted AMPure PB beads (Pacific Biosciences). DNA concentration was quantified using a Qubit Fluorometer v4.0 (ThermoFisher Scientific) and the Qubit 1X dsDNA HS assay kit. Final library fragment size was assessed with the Agilent Femto Pulse Automated Pulsed Field CE Instrument (Agilent Technologies) using the gDNA 55 kb BAC analysis kit.

The sample was sequenced on a Revio instrument (Pacific Biosciences). The prepared library was normalised to 2 nM, and 15 µL was used for making complexes. Primers were annealed and polymerases bound to generate circularised complexes, following the manufacturer's instructions. Complexes were purified using 1.2X SMRTbell beads, then diluted to the Revio loading concentration (200–300 pM) and spiked with a Revio sequencing internal control. The sample was sequenced on a Revio 25 M SMRT cell. The SMRT Link software (Pacific Biosciences), a web-based workflow manager, was used to configure and monitor the run and to carry out primary and secondary data analysis.

### Hi-C

#### Sample preparation and crosslinking

The Hi-C sample was prepared from 20–50 mg of frozen muscle tissue of the mRhiHip2 sample using the Arima-HiC v2 kit (Arima Genomics). Following the manufacturer's instructions, tissue was fixed and DNA crosslinked using TC buffer to a final formaldehyde concentration of 2%. The tissue was homogenised using the Diagnocine Power Masher-II. Crosslinked DNA was digested with a restriction enzyme master mix, biotinylated, and ligated. Clean-up was performed with SPRISelect beads before library preparation. DNA concentration was measured with the Qubit Fluorometer (Thermo Fisher Scientific) and Qubit HS Assay Kit. The biotinylation percentage was estimated using the Arima-HiC v2 QC beads.

#### Hi-C library preparation and sequencing

Biotinylated DNA constructs were fragmented using a Covaris E220 sonicator and size selected to 400–600 bp using SPRISelect beads. DNA was enriched with Arima-HiC v2 kit Enrichment beads. End repair, A-tailing, and adapter ligation were carried out with the NEBNext Ultra II DNA Library Prep Kit (New England Biolabs), following a modified protocol where library preparation occurs while DNA remains bound to the Enrichment beads. Library amplification was performed using KAPA HiFi HotStart mix and a custom Unique Dual Index (UDI) barcode set (Integrated DNA Technologies). Depending on sample concentration and biotinylation percentage determined at the crosslinking stage, libraries were amplified with 10–16 PCR cycles. Post-PCR clean-up was performed with SPRISelect beads. Libraries

**Table 2. Genome assembly statistics.**

| Assembly name                | mRhiHip2.hap1.1         | mRhiHip2.hap2.2 |
|------------------------------|-------------------------|-----------------|
| Assembly accession           | GCA_964194185.1         | GCA_964194215.2 |
| Assembly level               | chromosome              | chromosome      |
| Span (Mb)                    | 2 171.31                | 2 240.90        |
| Number of chromosomes        | 29                      | 29              |
| Number of contigs            | 2 349                   | 1 036           |
| Contig N50                   | 2.01 Mb                 | 44.95 Mb        |
| Number of scaffolds          | 722                     | 980             |
| Scaffold N50                 | 88.36 Mb                | 89.07 Mb        |
| Longest scaffold length (Mb) | 127.72                  | 128.19          |
| Sex chromosomes              | X                       | X               |
| Organelles                   | Mitochondrion: 16.85 kb | -               |

were quantified using the AccuClear Ultra High Sensitivity dsDNA Standards Assay Kit (Biotium) and a FLUOstar Omega plate reader (BMG Labtech).

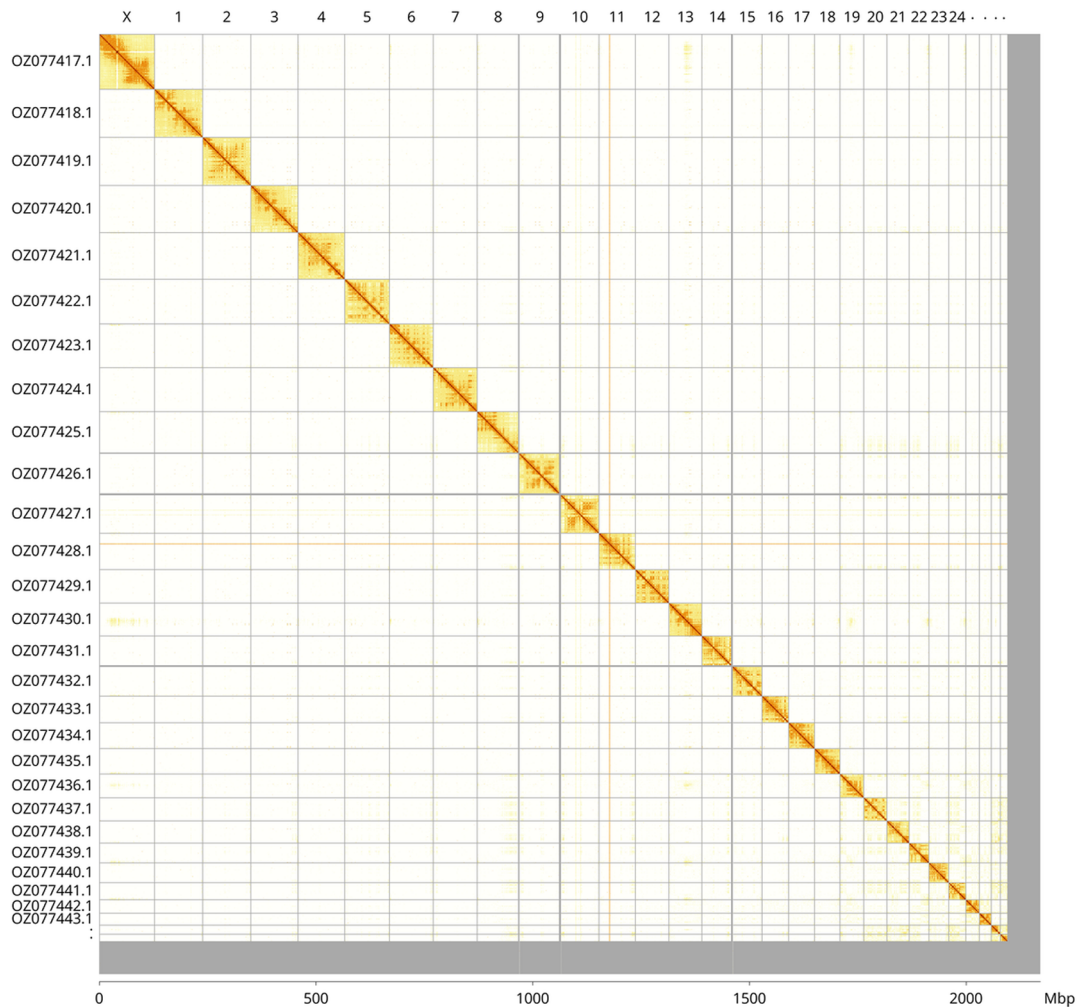
Prior to sequencing, libraries were normalised to 10 ng/μL. Normalised libraries were quantified again to create equimolar and/or weighted 2.8 nM pools. Pool concentrations were checked using the Agilent 4200 TapeStation (Agilent) with High Sensitivity D500 reagents before sequencing. Sequencing was performed using paired-end 150 bp reads on the Illumina NovaSeq X.

### RNA library preparation and sequencing

Libraries were prepared using the NEBNext® Ultra™ II Directional RNA Library Prep Kit for Illumina (New England Biolabs), following the manufacturer's instructions. Poly(A) mRNA in the total RNA solution was isolated using oligo (dT) beads, converted to cDNA, and uniquely indexed; 14 PCR cycles were performed. Libraries were size-selected to produce fragments between 100–300 bp. Libraries were quantified, normalised, pooled to a final concentration of 2.8 nM, and diluted to 150 pM for loading. Sequencing was carried out on the Illumina NovaSeq X, generating paired-end reads.

### Genome assembly

Prior to assembly of the PacBio HiFi reads, a database of  $k$ -mer counts ( $k = 31$ ) was generated from the filtered reads using FastK. GenomeScope2 (Ranallo-Benavidez *et al.*, 2020) was used to analyse the  $k$ -mer frequency distributions, providing estimates of genome size, heterozygosity, and repeat content.



**Figure 3. Hi-C contact map of the *Rhinolophus hipposideros* genome assembly mRhiHip2.hap1.1.** Assembled chromosomes are shown in order of size and labelled along the axes, with a megabase scale shown below. The plot was generated using PretextSnapshot.

The HiFi reads were assembled using Hifiasm in Hi-C phasing mode (Cheng *et al.*, 2021, 2022), producing two haplotypes. Hi-C reads (Rao *et al.*, 2014) were mapped to the primary contigs using bwa-mem2 (Vasimuddin *et al.*, 2019). Contigs were further scaffolded with Hi-C data in YaHS (Zhou *et al.*, 2023), using the --break option for handling potential misassemblies. The scaffolded assemblies were evaluated using Gfastats (Formenti *et al.*, 2022), BUSCO (Manni *et al.*, 2021) and MERQURY.FK (Rhie *et al.*, 2020).

The mitochondrial genome was assembled using MitoHiFi (Uliano-Silva *et al.*, 2023).

### Assembly curation

The assembly was decontaminated using the Assembly Screen for Cobionts and Contaminants (ASCC) pipeline. TreeVal was used to generate the flat files and maps for use in curation. Manual curation was conducted primarily in PretextView and HiGlass (Kerpedjiev *et al.*, 2018). Scaffolds were visually inspected and corrected as described by Howe *et al.* (2021). Manual corrections included five breaks, 13 joins, and removal of five haplotypic duplications. This reduced the scaffold count by 1.6%. The curation process is described at <https://gitlab.com/wtsi-grit/rapid-curation>. PretextViewSnapshot was used to generate a Hi-C contact map of the final assembly.

**Table 3. Chromosomal pseudomolecules in the haplotype 1 genome assembly of *Rhinolophus hipposideros* mRhiHip2 (haplotype 2 also at chromosome level).**

| INSDC accession | Molecule | Length (Mb) | GC%   |
|-----------------|----------|-------------|-------|
| OZ077418.1      | 1        | 111.16      | 38.50 |
| OZ077419.1      | 2        | 110.87      | 38    |
| OZ077420.1      | 3        | 109.18      | 38.50 |
| OZ077421.1      | 4        | 107.82      | 37.50 |
| OZ077422.1      | 5        | 102.78      | 40.50 |
| OZ077423.1      | 6        | 101.43      | 38.50 |
| OZ077424.1      | 7        | 101.14      | 38.50 |
| OZ077425.1      | 8        | 96.75       | 41.50 |
| OZ077426.1      | 9        | 95.76       | 40    |
| OZ077427.1      | 10       | 88.36       | 41    |
| OZ077428.1      | 11       | 83.88       | 41    |
| OZ077429.1      | 12       | 77.65       | 40.50 |
| OZ077430.1      | 13       | 75.91       | 39.50 |
| OZ077431.1      | 14       | 70.31       | 41.50 |
| OZ077432.1      | 15       | 68.45       | 41.50 |
| OZ077433.1      | 16       | 61.46       | 41    |
| OZ077434.1      | 17       | 60.11       | 39.50 |
| OZ077435.1      | 18       | 58.31       | 38.50 |
| OZ077436.1      | 19       | 54.60       | 42.50 |
| OZ077437.1      | 20       | 53.93       | 45    |
| OZ077438.1      | 21       | 50.94       | 45.50 |
| OZ077439.1      | 22       | 45.73       | 43.50 |
| OZ077440.1      | 23       | 45.64       | 43.50 |
| OZ077441.1      | 24       | 39.88       | 45    |
| OZ077442.1      | 25       | 30.68       | 41    |
| OZ077443.1      | 26       | 27.77       | 41.50 |
| OZ077444.1      | 27       | 20.99       | 49    |
| OZ077445.1      | 28       | 16.57       | 46    |
| OZ077417.1      | X        | 127.72      | 40    |



### Assembly quality assessment

The Merqury.FK tool (Rhie *et al.*, 2020) was run in a Singularity container (Kurtzer *et al.*, 2017) to evaluate  $k$ -mer completeness and assembly quality for both haplotypes using the  $k$ -mer database ( $k = 31$ ) computed prior to genome assembly. The analysis outputs included assembly QV scores and completeness statistics.

The genome was analysed using the BlobToolKit pipeline, a Nextflow implementation of the earlier Snakemake version (Challis *et al.*, 2020). The pipeline aligns PacBio reads using minimap2 (Li, 2018) and SAMtools (Danecek *et al.*, 2021) to generate coverage tracks. It runs BUSCO (Manni *et al.*, 2021) using lineages identified from the NCBI Taxonomy (Schoch *et al.*, 2020). For the three domain-level lineages, BUSCO genes are aligned to the UniProt Reference Proteomes database (Bateman *et al.*, 2023) using DIAMOND blastp (Buchfink *et al.*, 2021). The genome is divided into chunks based on the density of BUSCO genes from the closest taxonomic lineage, and each chunk is aligned to the UniProt Reference Proteomes database with DIAMOND blastx. Sequences without hits are chunked using seqtk and aligned to the NT database with blastn (Altschul *et al.*, 1990). The BlobToolKit suite consolidates all outputs into a blobdir for visualisation. The BlobToolKit pipeline was developed using nf-core tooling (Ewels *et al.*, 2020) and MultiQC (Ewels *et al.*, 2016), with containerisation through Docker (Merkel, 2014) and Singularity (Kurtzer *et al.*, 2017).

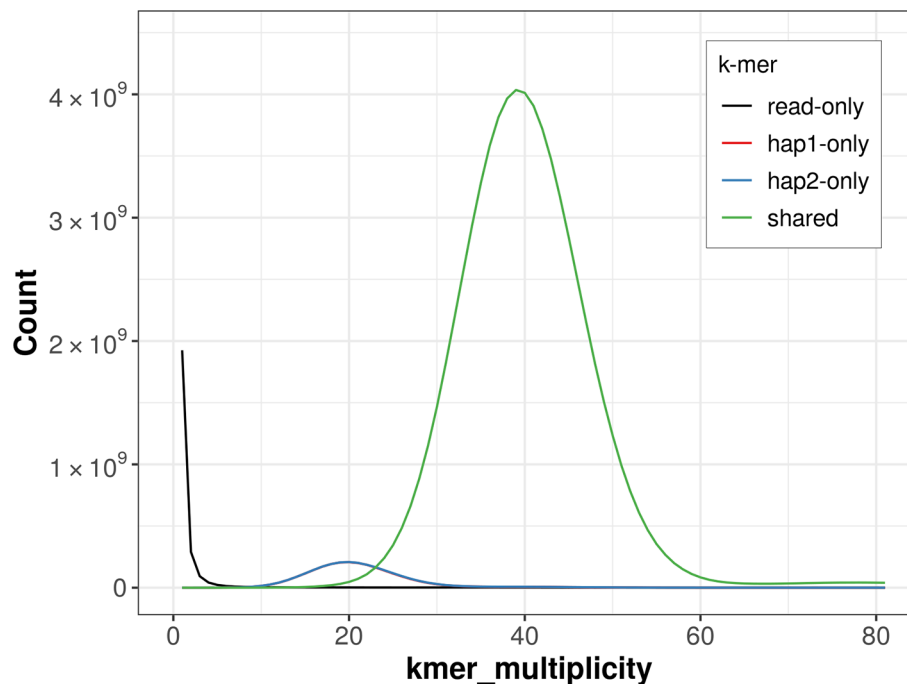
### Genome sequence report

#### Sequence data

PacBio sequencing of the *R. hipposideros* specimen generated 86.90 Gb (gigabases) from 7.91 million reads, which were used to assemble the genome. GenomeScope2.0 analysis estimated the haploid genome size at 2 177.79 Mb, with a heterozygosity of 0.25% and repeat content of 12.51% (Figure 2). These estimates guided expectations for the assembly. Based on the estimated genome size, the sequencing data provided approximately 39 $\times$  coverage. Hi-C sequencing produced 160.47 Gb from 1 062.69 million reads, which were used to scaffold the assembly. RNA sequencing data were also generated and are available in public sequence repositories. Table 1 summarises the specimen and sequencing details.

#### Assembly statistics

The genome was assembled into two haplotypes using Hi-C phasing. Haplotype 1 was curated to chromosome level, while haplotype 2 was assembled to scaffold level. The final assembly has a total length of 2 171.31 Mb in 722 scaffolds, with 1 627 gaps, and a scaffold N50 of 88.36 Mb (Table 2).



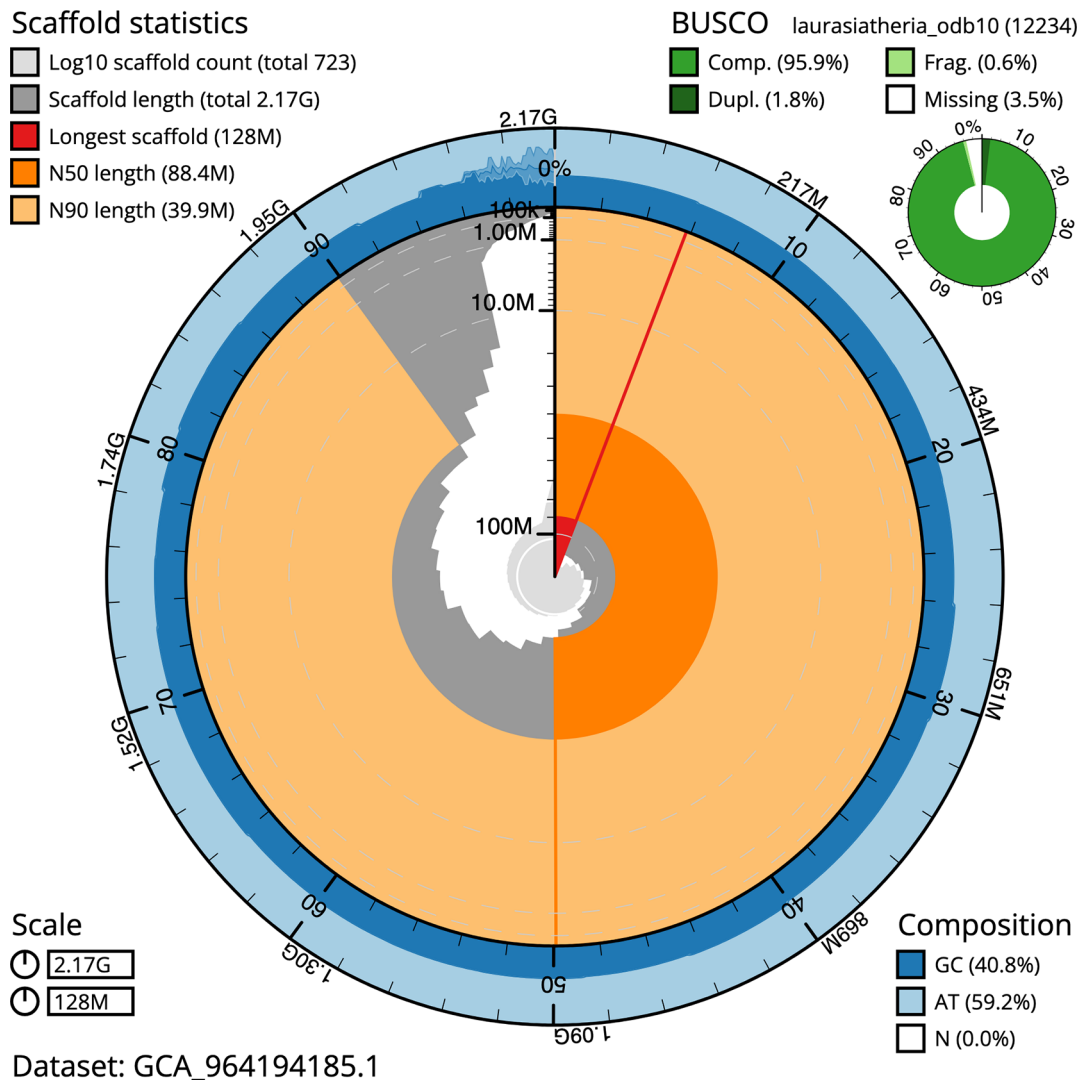
**Figure 4.** Evaluation of  $k$ -mer completeness using MerquryFK. This plot illustrates the recovery of  $k$ -mers from the original read data in the final assemblies. The horizontal axis represents  $k$ -mer multiplicity, and the vertical axis shows the number of  $k$ -mers. The black curve represents  $k$ -mers that appear in the reads but are not assembled. The green curve corresponds to  $k$ -mers shared by both haplotypes, and the red and blue curves show  $k$ -mers found only in one of the haplotypes.

Most of the haplotype 1 assembly sequence (96.52%) was assigned to 29 chromosomal-level scaffolds, representing 28 autosomes and the X sex chromosome. These chromosome-level scaffolds, confirmed by Hi-C data, are named according to size (Figure 3; Table 3). X chromosome identified based on synteny with *Eptesicus nilssonii* (GCA\_951640355.1) (van der Kooij *et al.*, 2023).

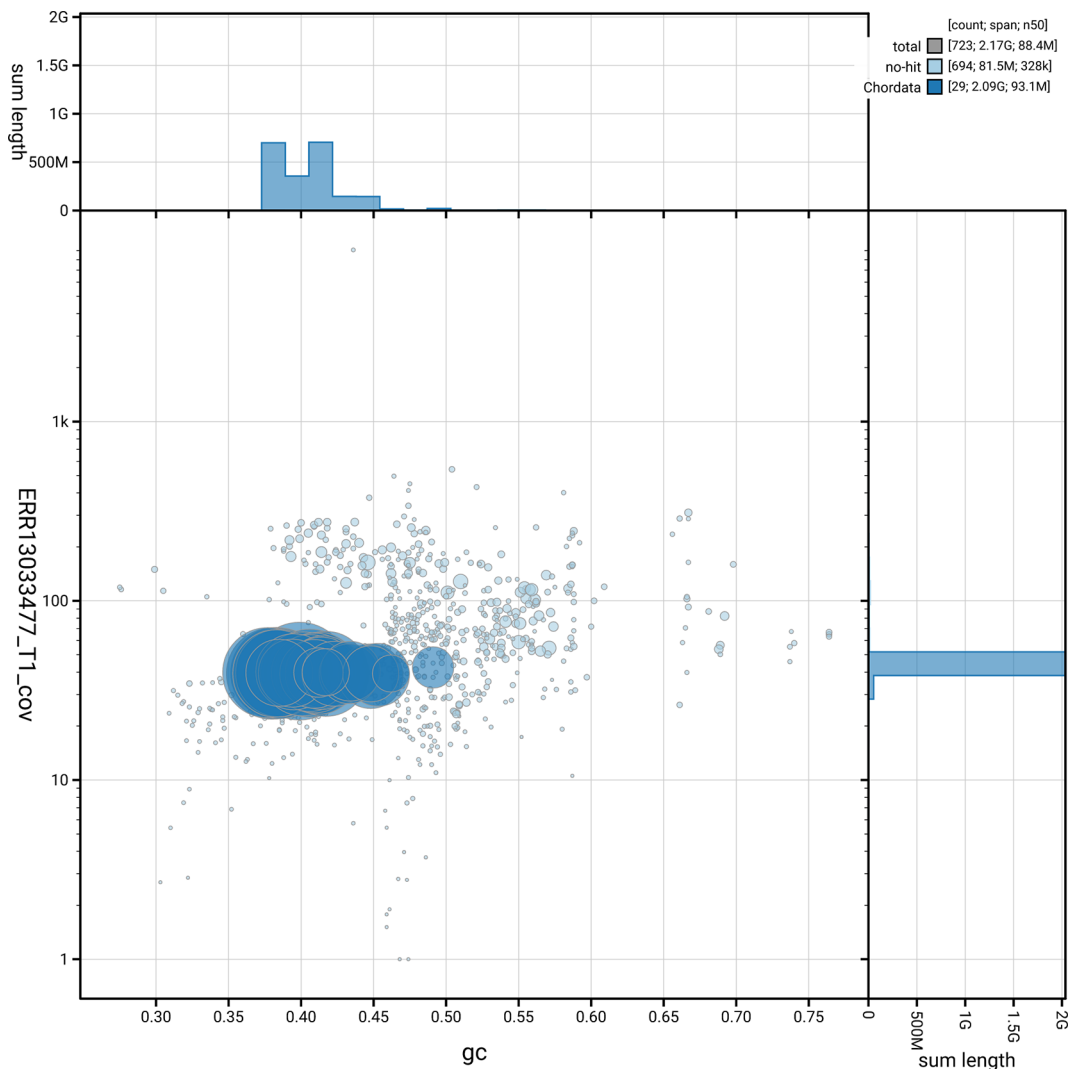
The mitochondrial genome was also assembled (length 16.85 kb, OZ077446.1). This sequence is included as a contig in the multifasta file of the genome submission and as a standalone record.

### Assembly quality metrics

For haplotype 1, the estimated QV is 61.1, and for haplotype 2, 60.7. When the two haplotypes are combined, the assembly achieves an estimated QV of 60.9. The *k*-mer completeness is 93.50% for haplotype 1, 93.59% for haplotype 2, and 99.77% for the combined haplotypes (Figure 4).



**Figure 5. Assembly metrics for mRhiHip2.hap1.1.** The BlobToolKit snail plot provides an overview of assembly metrics and BUSCO gene completeness. The circumference represents the length of the whole genome sequence, and the main plot is divided into 1 000 bins around the circumference. The outermost blue tracks display the distribution of GC, AT, and N percentages across the bins. Scaffolds are arranged clockwise from longest to shortest and are depicted in dark grey. The longest scaffold is indicated by the red arc, and the deeper orange and pale orange arcs represent the N50 and N90 lengths. A light grey spiral at the centre shows the cumulative scaffold count on a logarithmic scale. A summary of complete, fragmented, duplicated, and missing BUSCO genes in the set is presented at the top right. An interactive version of this figure can be accessed on the [BlobToolKit viewer](#).



**Figure 6. BlobToolKit blob plot for mRhiHip2.hap1.1.** The plot shows base coverage (vertical axis) and GC content (horizontal axis). The circles represent scaffolds, with the size proportional to scaffold length and the colour representing phylum membership. The histograms along the axes display the total length of sequences distributed across different levels of coverage and GC content. An interactive version of this figure is available on the [BlobToolKit viewer](#).

**Table 4. Earth Biogenome Project summary metrics for the *Rhinolophus hipposideros* assembly.**

| Measure                                        | Value                                                      | Benchmark        |
|------------------------------------------------|------------------------------------------------------------|------------------|
| EBP summary (haplotype 1)                      | 6.C.Q61                                                    | 6.C.Q40          |
| Contig N50 length                              | 2.01 Mb                                                    | ≥ 1 Mb           |
| Scaffold N50 length                            | 88.36 Mb                                                   | = chromosome N50 |
| Consensus quality (QV)                         | Haplotype 1: 61.1; haplotype 2: 60.7; combined: 60.9       | ≥ 40             |
| <i>k</i> -mer completeness                     | Haplotype 1: 93.50%; Haplotype 2: 93.59%; combined: 99.77% | ≥ 95%            |
| BUSCO                                          | C:99.3% [S:89.8%, D:9.4%], F:0.2%, M:0.5%, n:954           | S > 90%; D < 5%  |
| Percentage of assembly assigned to chromosomes | 96.52%                                                     | ≥ 90%            |

**Notes:** The EBP summary uses log10(Contig N50); chromosome-level (C) or log10(Scaffold N50); Q (Merquary QV). BUSCO: C = complete; S = single-copy; D = duplicated; F = fragmented; M = missing; n = orthologues

**Table 5. Software versions and sources.**

| Software                   | Version             | Source                                                                                                                  |
|----------------------------|---------------------|-------------------------------------------------------------------------------------------------------------------------|
| BEDTools                   | 2.30.0              | <a href="https://github.com/arq5x/bedtools2">https://github.com/arq5x/bedtools2</a>                                     |
| BLAST                      | 2.14.0              | <a href="ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/">ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/</a> |
| BlobToolKit                | 4.3.9               | <a href="https://github.com/blobtoolkit/blobtoolkit">https://github.com/blobtoolkit/blobtoolkit</a>                     |
| BUSCO                      | 5.5.0               | <a href="https://gitlab.com/ezlab/busco">https://gitlab.com/ezlab/busco</a>                                             |
| bwa-mem2                   | 2.2.1               | <a href="https://github.com/bwa-mem2/bwa-mem2">https://github.com/bwa-mem2/bwa-mem2</a>                                 |
| Cooler                     | 0.8.11              | <a href="https://github.com/open2c/cooler">https://github.com/open2c/cooler</a>                                         |
| DIAMOND                    | 2.1.8               | <a href="https://github.com/bbuchfink/diamond">https://github.com/bbuchfink/diamond</a>                                 |
| fasta_windows              | 0.2.4               | <a href="https://github.com/tolkit/fasta_windows">https://github.com/tolkit/fasta_windows</a>                           |
| FastK                      | 1.1                 | <a href="https://github.com/thegenemyers/FASTK">https://github.com/thegenemyers/FASTK</a>                               |
| GenomeScope2.0             | 2.0.1               | <a href="https://github.com/tbenavi1/genomescope2.0">https://github.com/tbenavi1/genomescope2.0</a>                     |
| Gfastats                   | 1.3.6               | <a href="https://github.com/vgl-hub/gfastats">https://github.com/vgl-hub/gfastats</a>                                   |
| Hifiasm                    | 0.19.8-r603         | <a href="https://github.com/chhy123/hifiasm">https://github.com/chhy123/hifiasm</a>                                     |
| HiGlass                    | 1.13.4              | <a href="https://github.com/higlass/higlass">https://github.com/higlass/higlass</a>                                     |
| MerquryFK                  | 1.1.2               | <a href="https://github.com/thegenemyers/MERQURY.FK">https://github.com/thegenemyers/MERQURY.FK</a>                     |
| Minimap2                   | 2.24-r1122          | <a href="https://github.com/lh3/minimap2">https://github.com/lh3/minimap2</a>                                           |
| MitoHiFi                   | 3                   | <a href="https://github.com/marcelauliano/MitoHiFi">https://github.com/marcelauliano/MitoHiFi</a>                       |
| MultiQC                    | 1.14; 1.17 and 1.18 | <a href="https://github.com/MultiQC/MultiQC">https://github.com/MultiQC/MultiQC</a>                                     |
| Nextflow                   | 23.10.0             | <a href="https://github.com/nextflow-io/nextflow">https://github.com/nextflow-io/nextflow</a>                           |
| PretextSnapshot            | 0.0.5               | <a href="https://github.com/sanger-tol/PretextSnapshot">https://github.com/sanger-tol/PretextSnapshot</a>               |
| PretextView                | 1.0.3               | <a href="https://github.com/sanger-tol/PretextView">https://github.com/sanger-tol/PretextView</a>                       |
| samtools                   | 1.19.2              | <a href="https://github.com/samtools/samtools">https://github.com/samtools/samtools</a>                                 |
| sanger-tol/ascc            | 0.1.0               | <a href="https://github.com/sanger-tol/ascc">https://github.com/sanger-tol/ascc</a>                                     |
| sanger-tol/blobtoolkit     | 0.6.0               | <a href="https://github.com/sanger-tol/blobtoolkit">https://github.com/sanger-tol/blobtoolkit</a>                       |
| sanger-tol/curationpretext | 1.4.2               | <a href="https://github.com/sanger-tol/curationpretext">https://github.com/sanger-tol/curationpretext</a>               |
| Seqtk                      | 1.3                 | <a href="https://github.com/lh3/seqtk">https://github.com/lh3/seqtk</a>                                                 |
| Singularity                | 3.9.0               | <a href="https://github.com/sylabs/singularity">https://github.com/sylabs/singularity</a>                               |
| TreeVal                    | 1.4.0               | <a href="https://github.com/sanger-tol/treeval">https://github.com/sanger-tol/treeval</a>                               |
| YaHS                       | 1.2a.2              | <a href="https://github.com/c-zhou/yahs">https://github.com/c-zhou/yahs</a>                                             |

BUSCO analysis using the metazoa\_odb10 reference set ( $n = 954$ ) identified 99.3% of the expected gene set (single = 89.8%, duplicated = 9.4%) in haplotype 1. The snail plot in [Figure 5](#) summarises the scaffold length distribution and other assembly statistics for haplotype 1. The blob plot in [Figure 6](#) shows the distribution of scaffolds by GC proportion and coverage for haplotype 1.

[Table 4](#) lists the assembly metric benchmarks adapted from [Rhie \*et al.\* \(2021\)](#) and the Earth BioGenome Project Report on Assembly Standards [September 2024](#). The EBP metric, calculated for the haplotype 1, is **6.C.Q61**, meeting the recommended reference standard.

### Genome annotation report

The *R. hipposideros* genome assembly (GCA\_964194185.1) was annotated by Ensembl at the European Bioinformatics Institute (EBI). This annotation includes 39 434 transcribed mRNAs from 18 700 protein-coding and 8 002 non-coding genes. The average transcript length is 46 502.81 bp, with an average of 1.46 coding transcripts per gene and 9.14 exons per transcript. Further details of this annotation are available from the [Ensembl annotation page](#).

### Author information

Contributors are listed at the following links:

- Members of the [Wellcome Sanger Institute Tree of Life Management, Samples and Laboratory team](#)
- Members of [Wellcome Sanger Institute Scientific Operations – Sequencing Operations](#)
- Members of the [Wellcome Sanger Institute Tree of Life Core Informatics team](#)
- Members of the [Tree of Life Core Informatics collective](#)
- Members of the [Darwin Tree of Life Consortium](#)

### 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](#). 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: *R. hipposideros* (lesser horseshoe bat). Accession number [PRJEB75441](#). The genome sequence is released openly for reuse. The *R. hipposideros* genome sequencing initiative is part of the Darwin Tree of Life Project (PRJEB40665), the Sanger Institute Tree of Life Programme (PRJEB43745), the Vertebrate Genomes Project (PRJNA489243) and the Bat1K Project (PRJNA489245). All raw sequence data and the assembly have been deposited in INSDC databases. Raw data and assembly accession identifiers are reported in [Tables 1](#) and [2](#).

Production code used in genome assembly at the WSI Tree of Life is available at <https://github.com/sanger-tol>. [Table 5](#) lists software versions used in this study.

### Acknowledgements

With thanks to Elena Erotokritou for help collecting samples and the Department of Forests and the Department of Environment of the Ministry of Agriculture, Rural Development and Environment of Cyprus.

### References

- |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>Altschul SF, Gish W, Miller W, <i>et al.</i>: <b>Basic Local Alignment Search Tool</b>. <i>J. Mol. Biol.</i> 1990; <b>215</b>(3): 403–410.<br/> <a href="#">PubMed Abstract</a>   <a href="#">Publisher Full Text</a></p> <p>Baroja U, Garin I, Aihartza J, <i>et al.</i>: <b>Pest consumption in a vineyard system by the lesser horseshoe bat (<i>Rhinolophus hipposideros</i>)</b>. <i>PLOS One</i>. 2019; <b>14</b>(7): e0219265.<br/> <a href="#">PubMed Abstract</a>   <a href="#">Publisher Full Text</a>   <a href="#">Free Full Text</a></p> <p>Bateman A, Martin M-J, Orchard S, <i>et al.</i>: <b>UniProt: The Universal Protein Knowledgebase in 2023</b>. <i>Nucleic Acids Res.</i> 2023; <b>51</b>(D1): D523–D531.<br/> <a href="#">PubMed Abstract</a>   <a href="#">Publisher Full Text</a>   <a href="#">Free Full Text</a></p> | <p>Buchfink B, Reuter K, Drost H-G: <b>Sensitive protein alignments at tree-of-life scale using DIAMOND</b>. <i>Nat. Methods</i>. 2021; <b>18</b>(4): 366–368.<br/> <a href="#">PubMed Abstract</a>   <a href="#">Publisher Full Text</a>   <a href="#">Free Full Text</a></p> <p>Cao Z, Li C, Wang K, <i>et al.</i>: <b>A fast and accurate identification model for <i>Rhinolophus</i> bats based on fine-grained information</b>. <i>Sci. Rep.</i> 2023; <b>13</b>(1): 16375.<br/> <a href="#">PubMed Abstract</a>   <a href="#">Publisher Full Text</a>   <a href="#">Free Full Text</a></p> <p>Challis R, Richards E, Rajan J, <i>et al.</i>: <b>BlobToolKit – interactive quality assessment of genome assemblies</b>. <i>G3 Genes   Genomes   Genetics</i>. 2020;</p> |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|



10(4): 1361–1374.

[Publisher 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.

[Publisher Full Text](#)

Cheng H, Jarvis ED, Fedrigo O, *et al.*: **Haplotype-resolved assembly of diploid genomes without parental data.** *Nat. Biotechnol.* 2022; **40**(9): 1332–1335.

[Publisher Full Text](#)

Danecek P, Bonfield JK, Liddle J, *et al.*: **Twelve years of SAMtools and BCFtools.** *GigaScience*. 2021; **10**(2).

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

Dool SE, Puechmaile SJ, Dietz C, *et al.*: **Phylogeography and postglacial recolonization of Europe by *Rhinolophus hipposideros*: Evidence from multiple genetic markers.** *Mol. Ecol.* 2013; **22**(15): 4055–4070.

[Publisher Full Text](#)

Ewels P, Magnusson M, Lundin S, *et al.*: **MultiQC: Summarize analysis results for multiple tools and samples in a single report.** *Bioinformatics*. 2016; **32**(19): 3047–3048.

[Publisher Full Text](#)

Ewels PA, Peltzer A, Fillinger S, *et al.*: **The nf-core framework for community-curated bioinformatics pipelines.** *Nat. Biotechnol.* 2020; **38**(3): 276–278.

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

Formenti G, Abueg L, Brajuka A, *et al.*: **Gfastats: Conversion, evaluation and manipulation of genome sequences using assembly graphs.** *Bioinformatics*. 2022; **38**(17): 4214–4216.

[Publisher Full Text](#)

Heller KG, von Helversen O: **Resource partitioning of sonar frequency bands in rhinolophid bats.** *Oecologia*. 1989; **80**(2): 178–186.

[Publisher Full Text](#)

Howard C, Denton A, Jackson B, *et al.*: **On the path to reference genomes for all biodiversity: Lessons learned and laboratory protocols created in the Sanger Tree of Life core laboratory over the first 2000 species.** *bioRxiv*. 2025.

[Publisher Full Text](#)

Howe K, Chow W, Collins J, *et al.*: **Significantly improving the quality of genome assemblies through curation.** *GigaScience*. 2021; **10**(1).

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

Jebb D, Huang Z, Pippel M, *et al.*: **Six reference-quality genomes reveal evolution of bat adaptations.** *Nature*. 2020; **583**(7817): 578–584.

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

Kacprzyk J, Teeling EC, Kelleher C, *et al.*: **Wing membrane biopsies for bat cytogenetics: Finding of 2n = 54 in Irish *Rhinolophus hipposideros* (Rhinolophidae, Chiroptera, Mammalia) supports two geographically separated chromosomal variants in Europe.** *Cytogenet. Genome Res.* 2016; **148**(4): 279–283.

[Publisher Full Text](#)

Kargopoulos N, Nicolaou H, Mitsainas G, *et al.*: **A key to the skulls of the mammals of Cyprus.** *Zool. Middle East*. 2023; **69**(4): 324–347.

[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.

[Publisher Full Text](#)

Kurtzer GM, Sochat V, Bauer MW: **Singularity: Scientific containers for mobility of compute.** *PLoS One*. 2017; **12**(5): e0177459.

[Publisher Full Text](#)

Li H: **Minimap2: Pairwise alignment for nucleotide sequences.** *Bioinformatics*. 2018; **34**(18): 3094–3100.

[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.

[Publisher Full Text](#)

Merkel D: **Docker: Lightweight Linux containers for consistent development and deployment.** *Linux J*. 2014; **2014**(239).

[Publisher Full Text](#)

Puerma E, Acosta MJ, Barragán MJ, *et al.*: **The karyotype and 5S rRNA genes from Spanish individuals of the bat species *Rhinolophus hipposideros* (Rhinolophidae; Chiroptera).** *Genetica*. 2008; **134**(3): 287–295.

[Publisher Full Text](#)

Ranallo-Benavidez TR, Jaron KS, Schatz MC: **GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes.** *Nat. Commun.* 2020; **11**(1): 1432.

[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](#)

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.

[Publisher 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).

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

Salsamendi E, Aihartza J, Goiti U, *et al.*: **Echolocation calls and morphology in Mehely's (*Rhinolophus mehelyi*) and Mediterranean (*R. euryale*) horseshoe bats: Implications for resource partitioning.** *Hystrix Ital J Mammal*. 2006; **16**.

[Publisher Full Text](#)

Schoch CL, Ciufo S, Domrachev M, *et al.*: **NCBI taxonomy: A comprehensive update on curation, resources and tools.** *Database*. 2020; **2020**: baaa062.

[Publisher Full Text](#)

Teeling EC, Vernes SC, Dávalos LM, *et al.*: **Bat biology, genomes, and the Bat1K project: To generate chromosome-level genomes for all living bat species.** *Annu. Rev. Anim. Biosci.* 2018; **6**: 23–46.

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

Uliano-Silva M, Ferreira JGRN, Krashennikova K, *et al.*: **MitoHiFi: A Python pipeline for mitochondrial genome assembly from PacBio high fidelity reads.** *BMC Bioinformatics*. 2023; **24**(1): 288.

[Publisher Full Text](#)

van der Kooij J, Vernes SC, Teeling EC, *et al.*: **The genome sequence of the northern bat, *Eptesicus nilssonii* (Keyserling & Blasius, 1839).** *Wellcome Open Res.* 2023; **8**: 362.

[Publisher Full Text](#)

Vasimuddin M, Misra S, Li H, *et al.*: **Efficient architecture-aware acceleration of BWA-MEM for multicore systems.** *2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*. IEEE; 2019; 314–24.

[Publisher Full Text](#)

Volleth M, Biedermann M, Schorcht W, *et al.*: **Evidence for two karyotypic variants of the lesser horseshoe bat (*Rhinolophus hipposideros*, Chiroptera, Mammalia) in Central Europe.** *Cytogenet. Genome Res.* 2013; **140**(1): 55–61.

[Publisher Full Text](#)

Wilson DE, Mittermeier RA: *Handbook of the Mammals of the World, Volume 9: Bats*. Lynx Edicions; 2019.

Zarzoso-Lacoste D, Jan P-L, Lehnen L, *et al.*: **Combining noninvasive genetics and a new mammalian sex-linked marker provides new tools to investigate population size, structure and individual behaviour: An application to bats.** *Mol. Ecol. Resour.* 2018; **18**(2): 217–228.

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

Zhou C, McCarthy SA, Durbin R: **YaHS: Yet another Hi-C scaffolding tool.** *Bioinformatics*. 2023; **39**(1).

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

# Open Peer Review

Current Peer Review Status:    

Version 1

Reviewer Report 23 April 2026

<https://doi.org/10.21956/wellcomeopenres.28868.r153216>

© 2026 Hiller 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.



**Michael Hiller** 

Senckenberg Research Institute, Frankfurt, Germany

This data notes represents a high-quality haplotype resolved genome assembly of *Rhinolophus hipposideros*.

The data note comprises a detailed introduction with a lot of useful information, and an accurate description of all steps to generate the genomic and transcriptomic data and to assemble the genome.

I only have one major question.

Is it not clear to me why GCA\_964194185.1 is presented as the main haplotype, which has a low contig N50 of 2,014,157 bp, potentially because of an adapter trimming issue in the pipeline. The original hap2.1 (GCA\_964194215.1) also had a contig N50 of 1,922,115 bp, but this was later improved in the hap2.2 version (GCA\_964194215.2) to 44 Mb.

Since hap2.2 was likely generated by a new haplotype-resolved assembly and dual curation, I wonder if the corresponding hap1.2 will also be released. Otherwise, the assembly methods need to be adjusted, as these 2 haplotypes most likely come from different assembly attempts.

Also, because the TOGA2-inferred gene completeness of hap2.2 (GCA\_964194215.2) is noticeably higher than hap1.1 (GCA\_964194185.1), it would make sense to present hap2.2 as more complete primary hap1 for this species, even if hap1.1 received an Ensembl gene annotation.

**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:** comparative genomics, genome assembly

**I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.**

Reviewer Report 20 April 2026

<https://doi.org/10.21956/wellcomeopenres.28868.r153220>

© 2026 Feng J. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Jiang Feng**

Northeast Normal University, Changchun, China

This manuscript presents a high-quality, chromosome-level genome assembly of *Rhinolophus hipposideros*, a species characterized by chromosomal polymorphism and ecological sensitivity but lacking a well-established reference genome. The work uses a combined strategy of PacBio HiFi and Hi-C and RNA-seq, with complete procedures for haplotype-resolved assembly, mitochondrial genome assembly, and quality evaluation, which meets current standards for vertebrate reference genomes. Overall, the manuscript is scientifically sound and requires only minor revisions to improve clarity, consistency, and completeness.

1. The abstract does not clearly state that this genome represents the first chromosome-level, haplotype-resolved genome for this species, so it is recommended to add this statement to emphasize its uniqueness.
2. There is a conflict between "the sample is male" in the last of the Background and "adult female" in the Methods and Abstract. This inconsistency must be corrected throughout the manuscript.
3. Please add the local collection permit number and ethical approval number in Cyprus, listing only the responsible government departments is insufficient.
4. Please specify the homologous species set and pipeline version used for genome annotation, and the prediction criteria for non-coding genes.
5. The sampling location is described inconsistently as "Cyprus, University of Cyprus" (Methods) and "Kalopanayiotis, Cyprus" (Background). These should be unified.
6. The binomial name *Rhinolophus hipposideros* is not italicized in the Keywords. After its first mention, the abbreviation *R. hipposideros* should be used consistently in italics.
7. The Methods mention that multiple tissues were sampled. Please explicitly list all tissues collected and used in this study.

**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:** Animal Behavioral Ecology, Animal Molecular Ecology

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

Reviewer Report 17 April 2026

<https://doi.org/10.21956/wellcomeopenres.28868.r153218>

© 2026 Mao X. 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.



**Xiuguang Mao** 

East China Normal University, Shanghai, China

This manuscript presents a high-quality chromosome-scale genome assembly for *R. hipposideros*, which provides valuable insights into chromosomal evolution within this species complex, characterized by diploid numbers of  $2n = 54, 56$ , and  $58$ . The assembly and the associated comparative analyses represent a meaningful contribution to our understanding of karyotypic diversification in this group.

For the current version of the manuscript, I have one minor comment:

In the main text, the authors state that the BUSCO assessment was performed using the metazoa\_odb10 reference set ( $n = 954$ ). However, in Figure 5, the BUSCO database is reported as laurasiatheria\_odb10 ( $n = 12,234$ ). Please verify which dataset was actually used in the analysis and update both the text and the figure accordingly.

**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:** Genome evolution, bats, phylogeny

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

Reviewer Report 16 April 2026

<https://doi.org/10.21956/wellcomeopenres.28868.r153214>

© 2026 Etherington G. 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.



**Graham Etherington** 

Earlham Institute, Norwich, UK

The authors present the first annotated genome assembly for the lesser horseshoe bat. The manuscript clearly describes both the methods employed and the results obtained from the assembly, and provides a thorough and engaging overview of the species' evolutionary and ecological background. Overall, the study is well presented. I have, however, a few minor suggestions that may further improve the clarity and completeness of the manuscript:

1. The term "*western Palaearctic*" should be capitalised as "*Western Palaearctic*".
2. It would be helpful to include the common name (*Mehely's horseshoe bat*) when first mentioning *Rhinolophus mehelyi*.
3. The authors may wish to clarify whether the 29 chromosomal pseudomolecules correspond to the expected karyotype of the sequenced individual.
4. The manuscript refers to the "distinctive rDNA organisation patterns in *Rhinolophus hipposideros*", but this point is not expanded upon. A brief explanation of what makes this organisation distinctive would improve reader understanding.

**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:** Conservation genomics, genome assembly, vertebrate genomics.

**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.**

---