DATA NOTE

The genome sequence of *Molossus alvarezi* González-Ruiz, Ramírez-Pulido and Arroyo-Cabrales, 2011 (Chiroptera, Molossidae) [version 1; peer review: 2 approved]

Nancy B. Simmons[1], Melissa R. Ingala[1,2], Myrtani Pieri[3], Brian P. O'Toole[4], Jonathan L. Gray[4], Philip Philge[4], Thomas L Volkert[4], Ning Zhang[4], Linelle Abueg[5], Nadolina Brajuka[5], Erich Jarvis[5], Giulio Formenti[5], Kirsty McCaffrey[5], Meike Mai[6], Emma C. Teeling[7,8], Sonja C Vernes[6,9], The Bat Biology Foundation, The Bat1K Consortium

[1]Department of Mammalogy, Division of Vertebrate Zoology, American Museum of Natural History, New York, NY 10024, USA
[2]Fairleigh Dickinson University Department of Biological Sciences, Madison, NJ, 07940, USA
[3]Department of Life Sciences, School of Life and Health Sciences, University of Nicosia, Nicosia, Cyprus
[4]Paratus Sciences, New York, USA
[5]Vertebrate Genome Lab, The Rockefeller University, New York, USA
[6]University of St Andrews School of Biology, St Andrews, Scotland, UK
[7]University College Dublin School of Biology and Environmental Science, Dublin, Ireland
[8]Wellcome Sanger Institute, Wellcome Genome Campus, Cambridgeshire, CB10 1SA, UK
[9]Neurogenetics of Vocal Communication Group, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

v1 **First published:** 12 Sep 2024, **9**:522
https://doi.org/10.12688/wellcomeopenres.22726.1
Latest published: 12 Sep 2024, **9**:522
https://doi.org/10.12688/wellcomeopenres.22726.1

Abstract

We present a genome assembly from an individual female *Molossus alvarezi* (Chordata; Mammalia; Chiroptera; Molossidae). The genome sequence is 2.490 Gb in span. The majority of the assembly is scaffolded into 24 chromosomal pseudomolecules, with the X sex chromosomes assembled.

Keywords

Molossus alvarezi, genome sequence, chromosomal, Bat1K

Open Peer Review

Approval Status ✓ ✓

	1	2
version 1 12 Sep 2024	✓ view	✓ view

1. **Diego A Caraballo**, Universidad de Buenos Aires and CONICET, Buenos Aires, Argentina

2. **M. Alejandra Camacho**, Pontificia Universidad Catolica del Ecuador Escuela de Ciencias Biologicas, Quito, Ecuador

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

Corresponding authors: Nancy B. Simmons (simmons@amnh.org), Sonja C Vernes (scv1@st-andrews.ac.uk)

Author roles: Simmons NB: Conceptualization, Data Curation, Funding Acquisition, Investigation, Resources, Supervision, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Ingala MR**: Investigation, Resources, Writing – Review & Editing; **Pieri M** : Software, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **O'Toole BP**: Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Gray JL**: Investigation; **Philge P**: Software, Visualization, Writing – Review & Editing; **L Volkert T**: Project Administration, Supervision; **Zhang N**: Investigation; **Abueg L**: Formal Analysis; **Brajuka N**: Validation; **Jarvis E**: Supervision, Writing – Review & Editing; **Formenti G**: Supervision, Writing – Review & Editing; **McCaffrey K**: Data Curation, Formal Analysis; **Mai M**: Project Administration, Writing – Review & Editing; **Teeling EC**: Conceptualization, Data Curation, Project Administration, Supervision, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **C Vernes S**: Conceptualization, Data Curation, Project Administration, Resources, Supervision, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing;

Competing interests: BPO, JLG and NZ TLV are employees of Paratus Sciences Corporation and are option holders. PP serves as a consultant for Paratus Sciences Corporation. NBS, MP, TLV, LA, NB, EJ, GF, KM, MM, ECT and SCV declare no competing interests.

Grant information: SCV was supported by a UKRI Future Leaders Fellowship, (MR/T021985/1), an ERC Consolidator Grant (101001702; BATSPEAK), and a Max Planck Research Group awarded by the Max Planck Society). MRI was supported by a Peter Buck Postdoctoral Fellowship from the Smithsonian National Museum of Natural History. ECT is a Wellcome collaborator and the Irish Research Council Laureate Award IRCLA/2017/58 and Science Foundation Ireland Future Frontiers 19/FFP/6790. Fieldwork by NBS was supported by the Taxonomic Mammalogy Fund of the American Museum of Natural History. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.
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: Simmons NB, Ingala MR, Pieri M *et al.* **The genome sequence of *Molossus alvarezi* González-Ruiz, Ramírez-Pulido and Arroyo-Cabrales, 2011 (Chiroptera, Molossidae) [version 1; peer review: 2 approved]** Wellcome Open Research 2024, **9**:522 https://doi.org/10.12688/wellcomeopenres.22726.1

First published: 12 Sep 2024, **9**:522 https://doi.org/10.12688/wellcomeopenres.22726.1

Species taxonomy

Eukaryota; Metazoa; Eumetazoa; Bilateria; Deuterostomia; Chordata; Craniata; Vertebrata; Gnathostomata; Teleostomi; Euteleostomi; Sarcopterygii; Dipnotetrapodomorpha; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Boreoeutheria; Laurasiatheria; Chiroptera; Yangochiroptera; Molossidae; Molossinae; *Molossus*; *Molossus alvarezi* (*Loureiro et al., 2019; Loureiro et al., 2020; Meredith et al., 2011; Simmons & Cirranello, 2024; Teeling et al., 2005*).

Introduction

Molossid bats are swift aerial insectivores that are distributed throughout the world. As shown in Figure 1, they comprise two subfamilies, the South American endemic Tomopeatinae and the cosmopolitan Molossinae (Eger, 2008b), the latter consisting of 21 genera and 131 species (Simmons & Cirranello, 2024). Within this group, the genus *Molossus* comprises 15 species distributed broadly across the Neotropics (Loureiro *et al.*, 2019; Simmons & Cirranello, 2024). *Molossus alvarezi* was described for the first time in 2011 by Gonzalez-Ruiz *et al.* (González-Ruiz *et al.*, 2011) based on specimens from the Yucatán Peninsula of Mexico, but this species is now known to range from the Yucatán through Belize, Guatemala, and Honduras south to Colombia, Venezuela, Trinidad, Surinam, French Guiana, and Peru (Simmons & Cirranello, 2024).

Molossus alvarezi as shown in Figure 2 is a medium-sized molossid that is most morphologically similar to *Molossus sinaloae* (*Eger, 2008a*), with which it was previously synonymized, although the two species do not appear to be particularly close relatives based on recent phylogenies (Loureiro *et al.*, 2020). Forearm length of *M. alvarezi* varies from 42-48 mm; dorsal fur is chocolate brown while the venter is slightly paler and grayish (González-Ruiz *et al.*, 2011; Miller, 2003). A key feature for distinguishing these bats in the field is that dorsal fur of *M. alvarezi* is bicolored with a white base that extends

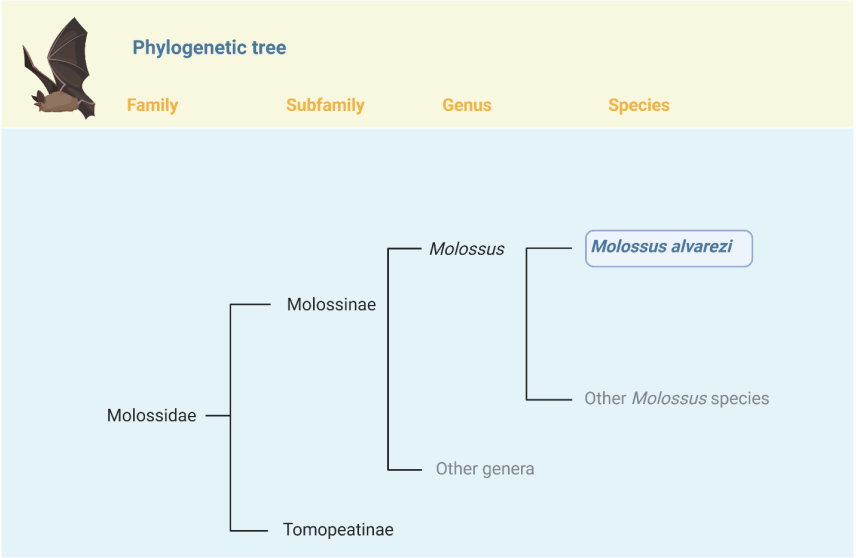

Figure 1. Position of *Molossus alvarezi* in the phylogeny of Family Molossidae. *Molossus alvarezi* is one of 15 species currently recognized in the genus *Molossus* (*Loureiro et al., 2019; Loureiro et al., 2020; Simmons & Cirranello, 2024*). *Molossus* belongs to the Subfamily Molossinae, which currently includes 20 genera and 133 species (Simmons & Cirranello, 2024). Within *Molossus*, *M. alverezi* occupies a basal branch and is sister to a large clade of other species from Central and South America (Loureiro *et al.*, 2020). Figure created with Biorender.com.

Figure 2. *Molossus alvarezi*. Adult individuals of *Molossus alvarezi* from Belize [Photo **A** taken by Brock and Sherri Fenton and photos **B** and **C** taken by Charles Francis].

more than half the hair length (González-Ruiz *et al.*, 2011; Miller, 2003). Other craniodental features also distinguish this species from congeners (González-Ruiz *et al.*, 2011).

The diet, behavior, and echolocation calls of *M. alvarezi* are not known, although we presume that this species is an aerial insectivore that uses FM echolocation calls like other molossid bats (Jung *et al.*, 2014). The species *M. alvarezi* is currently classified as Data Deficient in the IUCN Red List of Threatened Species (Barquez *et al.*, 2015).

Genome sequence report

The genome was sequenced from a single female *M. alvarezi* (field number BZ-4, catalog number AMNH:Mammalogy:280599) collected at Lamanai Outpost Lodge, Orange Walk District, Belize (17.75156 N, 88.65376 W) on 8 November 2021. A total of 51x-fold coverage in Pacific Biosciences Hi-Fi long reads (contig N50 63 Mb) was generated after removal of all reads shorter than 10kb. Primary assembly contigs were scaffolded with chromosome confirmation Hi-C data. The final assembly has a total length of 2.48 Gb in 504 sequence scaffolds with a scaffold N50 of 112.65 Mb (Table 1). The assembly has a BUSCO (Simao *et al.*, 2015) completeness of [98.2]% using the Laurasiatheria reference set. The assembly was fully phased and both haplotypes are deposited. Chromosomal pseudomolecules in the genome assembly of *M. alvarezi* are shown in Table 2.

Methods

The *M. alvarezi* specimen was a female individual captured on an American Museum of Natural History (AMNH) field expedition to the Lamanai area in the Orange Walk District of

Table 1. Genome data for *Molossus alvarezi*.

Project accession data	
Assembly identifier	GCA_037157525.1
Species	*Molossus alvarezi*
Specimen	mMolAlv2
NCBI taxonomy ID	1552295
BioProject	Accession: PRJNA944206; Bat1K: Accession: PRJNA489245; ID: 489245
BioSample ID	SAMN31836526
Isolate information	Female - Muscle
Genome assembly	
Assembly accession	GCA_037157525.1
Accession of Alternative haplotype	GCA_037157525.1 (Primary), GCA_037176705.1 (Alternative)
Span (Mb)	2,395.14
Number of contigs	1,369
Contig N50 length (Mb)	4,060.33
Number of scaffolds	187
Scaffold N50 length (Mb)	113,916.77
Longest scaffold (Mb)	270,994.32

* BUSCO scores based on the laurasiatheria_odb10 set using v5.0.0. C= complete [S= single copy, D=duplicated], F=fragmented, M=missing, n=number of orthologues in comparison.

* *Molossus alvarezi* BUSCO scores based on laurasiatheria_odb10 BUSCO set v5.3.2.

Table 2. Chromosomal pseudomolecules in the genome assembly of *Molossus alvarezi*. ENA accession Chromosome Size (Mb) GC%. The chromosome number of *Molossus alvarezi* is 2n=24.

ENA accession	Chromosome	Size (Mb)	GC%
SUPER_1	1	278.8	0.4081
SUPER_X	2	150.3	0.3943
SUPER_2	3	134.3	0.4173
SUPER_3	4	131.9	0.4318
SUPER_4	5	120.3	0.4204
SUPER_5	6	120.1	0.4252
SUPER_6	7	119.8	0.4041
SUPER_7	8	116.9	0.409
SUPER_8	9	112.7	0.4059
SUPER_9	10	112.3	0.3984
SUPER_10	11	110.2	0.4077
SUPER_11	12	108.7	0.4366
SUPER_12	13	103.9	0.4288
SUPER_13	14	99.1	0.431
SUPER_14	X	81.9	0.4308
SUPER_15	15	77.4	0.4153
SUPER_16	16	74.9	0.4378
SUPER_17	17	73.5	0.4595
SUPER_18	18	69.7	0.4397
SUPER_19	19	61.1	0.4462
SUPER_20	20	58.6	0.4645
SUPER_21	21	39.4	0.4726
SUPER_22	22	31.5	0.4538
SUPER_23	23	17.1	0.4738

Belize. The bat was caught in a ground-level mist net set in the gardens at Lamanai Outpost Lodge (17.75156 N, 88.65376 W). All efforts were made to minimize any distress or suffering by the animal. The individual sampled was subjected to minimal handling after capture, and it was held in a clean cloth bag after capture as per best practices for field containment of bats (Kunz & Parsons, 2009). After species identification, the individual was euthanized humanely by trained personnel the same night it was captured. Capture and sampling were conducted under Belize Forest Department Permit FD/WL/1/21, and samples were exported under Belize Forest Department permit FD/WL/7/22(08). The individual sampled was identified as *M. alvarezi* based on morphological traits and measurements described by Gonzalez-Ruiz *et al.* (González-Ruiz *et al.*, 2011).

The animal was euthanized by isoflurane inhalation (<1 ml to moisten cotton ball), a humane approved method that rapidly causes unconsciousness and eventually death upon inhalation. Bats euthanized by this method are rendered unconscious within seconds due to their high respiration rate, and death occurs within a minute or two with no significant suffering by the animal. Tissues were removed from the subject individual immediately following euthanasia and were flash-frozen in a liquid nitrogen dry shipper, with the cold chain maintained from field to museum to laboratory. All data were recorded and reported in accordance with the ARRIVE guidelines – see data availability section and Table 1. All work was conducted with approval by the AMNH Institutional Animal Care and Use Committee (AMNHIACUC-20191212).

DNA was extracted using Nanobind extraction from muscle tissue following the Circulomics Nanobind HMW DNA Extraction Protocol. Pacific Biosciences HiFi libraries were constructed according to the manufacturer's instructions. Hi-C data was generated using the Arima Hi-C+ High Coverage kit from the same muscle tissue sample. Sequencing was performed by the Genomic Operations DNA Pipelines at Paratus Sciences on Pacific Biosciences Sequel IIe (HiFi reads) and Illumina NextSeq 2000 (Hi-C) instruments.

Assembly was carried out following the Vertebrate Genome Project Galaxy pipeline v2.0 (Larivière *et al.*, 2024). A brief synopsis of the method is as follows: Genome size was estimated using GenomeScope2 (Ranallo-Benavidez *et al.*, 2020). Hifiasm with Hi-C phasing was used for genome assembly (Cheng *et al.*, 2022). The quality of the assembly was evaluated using Merqury (Nurk *et al.*, 2020) and BUSCO (Manni *et al.*, 2021). Scaffolding with Hi-C data (Rao *et al.*, 2014) was carried out with YaHS (Zhou *et al.*, 2023). PretextView was implemented to generate a Hi-C contact map (Figure 3). Figure 4–Figure 6 were generated using BlobToolKit (Challis *et al.*, 2020). All bioinformatics software utilised for the *M. alvarezi* analysis are depicted in Table 3.

Ethics and consent

All work was conducted with approval by the AMNH Institutional Animal Care and Use Committee (AMNHIACUC-20191212).

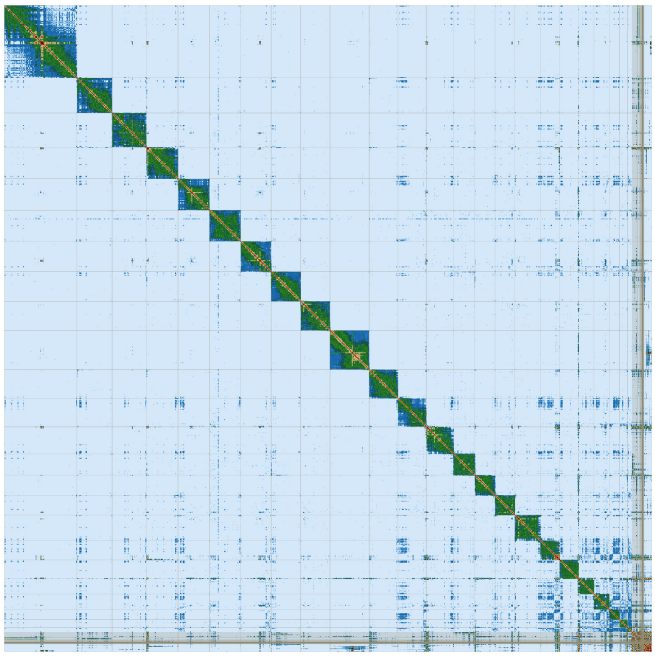

Figure 3. Hi-C Contact Map of the *Molossus alvarezi* assembly with 24 chromosomes, visualized using PretextView.

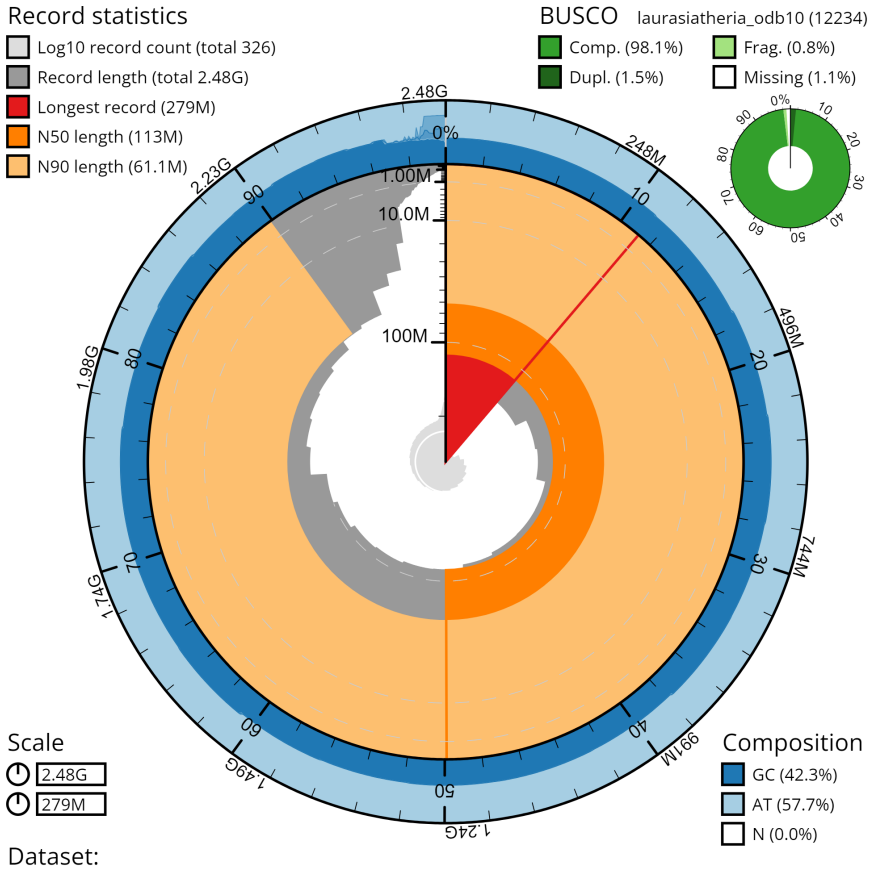

Figure 4. Genome assembly metrics generated using blobtoolkit for the *Molossus alvarezi* genome assembly. The larger snail plot depicts scaffold statistics including N50 length (bright orange) and base composition (blue). The smaller plot shows BUSCO completeness in green.

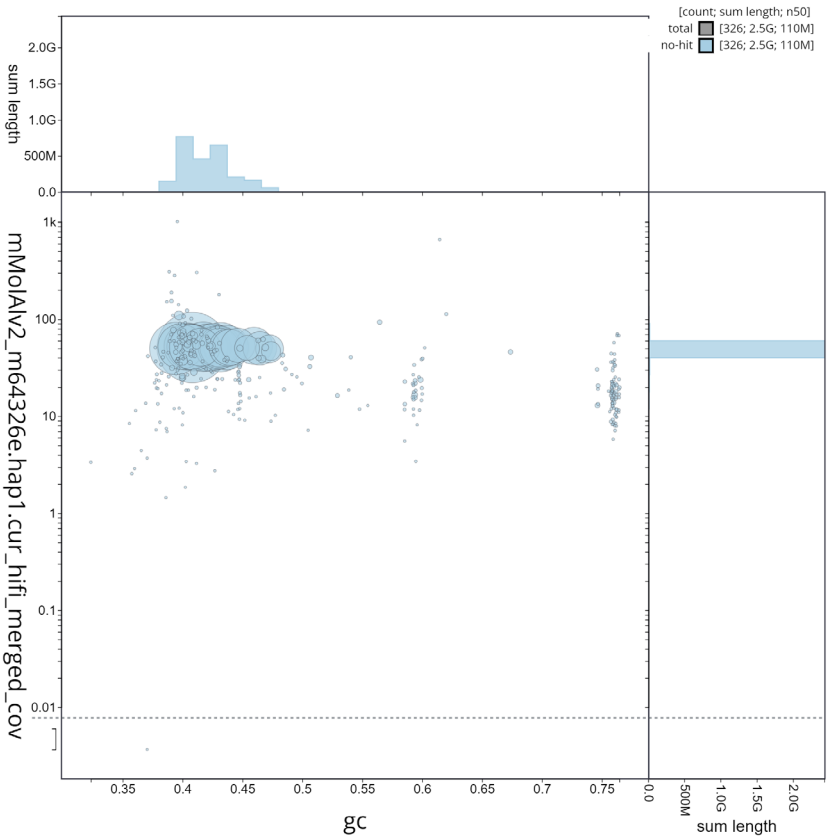

Figure 5. GC coverage plot generated for the *Molossus alvarezi* assembly using blobtoolkit. Individual chromosomes and scaffolds are represented by each circle. The circles are sized in proportion to chromosome/scaffold length. Histograms show the sum length of chromosome/scaffold size along each axis. Color of circles indicate taxonomic hits of each Phylum represented in the assembly.

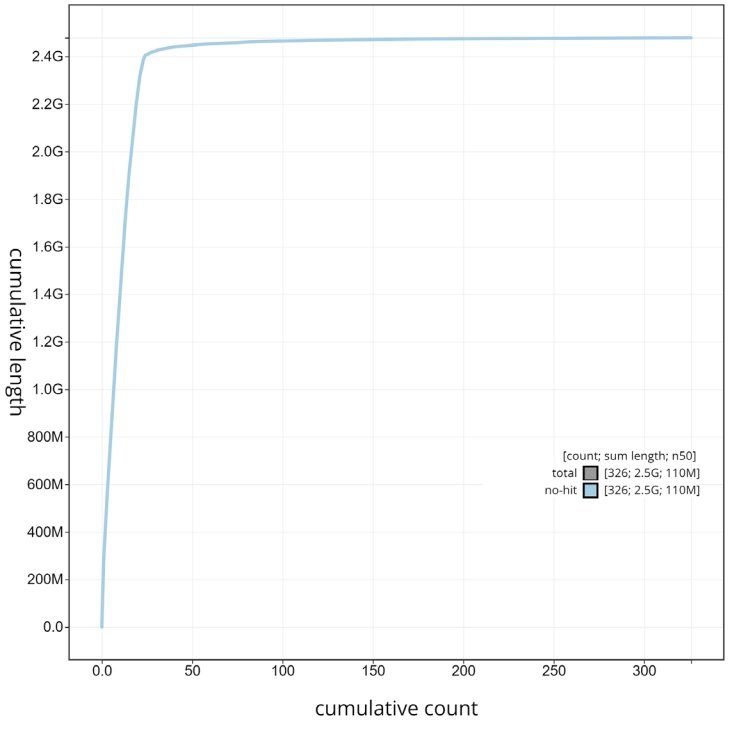

Figure 6. Cumulative sequence plot generated for the *Molossus alvarezi* assembly using blobtoolkit. The grey line shows the cumulative length for all chromosomes/scaffolds in the assembly. Colored lines represent Phylum represented in the assembly.

Table 3. Software tools used.

Software tool	Version	Source
bamUtil	1.0.15	https://genome.sph.umich.edu/wiki/BamUtil:_bam2FastQ
MultiQC	1.13	https://github.com/ewels/MultiQC
Genomescope	2.0	https://github.com/tbenavi1/genomescope2.0
hifiasm	0.19.3	https://github.com/chhylp123/hifiasm
purge_dups	1.2.6	https://github.com/dfguan/purge_dups
BUSCO	5.3.2	https://busco.ezlab.org/
Merqury	1.3	https://github.com/marbl/merqury
Assembly-stats	17.02	https://github.com/rjchallis/assembly-stats
Arima-HiC Mapping Pipeline	-	https://github.com/ArimaGenomics/mapping_pipeline
YaHS	1.1	https://github.com/c-zhou/yahs
HiGlass	1.11.7	https://github.com/higlass/higlass
samtools	1.9	https://www.htslib.org/
PretextView	-	https://github.com/sanger-tol/PretextView/tree/master
BUSCO	5.7.0	https://busco.ezlab.org/
BlobToolKit	4.3.5	https://github.com/blobtoolkit/blobtoolkit
pbmm2	1.13.1	https://github.com/PacificBiosciences/pbmm2
Blast	2.15.0+	https://blast.ncbi.nlm.nih.gov/Blast.cgi

Data availability

The *M. alvarezi* genome sequencing initiative is part of the Bat1K genome sequencing project. The genome assembly is released openly for reuse. The *M. alvarezi* genomic data is available on GenomeArk in this link: https://www.genomeark.org/bat1k-curated-assembly/Molossus_alvarezi.html Underlying data may be available for non-commercial research purposes upon request. Please email info@batbio.org for more information.

The genome assembly can be found in the European Nucleotide Archive: Molossus alvarezi (Alvarez's mastiff bat). Accession number: GCA_037157525.1, https://www.ebi.ac.uk/ena/browser/view/GCA_037157525.1 (BAT1K, 2024a)

In the NCBI database, the BioProject for Molossus alvarezi isolate: mMolAlv1 (Alvarez's mastiff bat) is listed under Accession number: PRJNA944206, https://www.ncbi.nlm.nih.gov/bioproject/PRJNA944206.

This project is part of the broader Bat1K BioProject PRJNA489245. (BAT1K, 2024b)

Data accession identifiers are reported in Table 1.

References

BAT1K: **The genome sequence of *Molossus alvarezi* (Alvarez's mastiff bat).** European Nucleotide Archive. [dataset], accession number GCA_ 037157525.1, 2024a.
https://www.ebi.ac.uk/ena/browser/view/GCA_037157525.1

BAT1K: ***Molossus alvarezi* isolate: mMolAlv1 (Alvarez's mastiff bat).** NCBI BioProject. [dataset], accession number PRJNA944206, 2024b.
https://www.ncbi.nlm.nih.gov/bioproject/PRJNA944206

Barquez R, Rodriguez B, Miller B, *et al.*: **Molossus rufus. The IUCN red list of**

threatened species 2015. 2015; e. T13644A22107969.

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, Jarvis ED, Fedrigo O, *et al.*: **Haplotype-resolved assembly of diploid genomes without parental data.** *Nat Biotechnol.* 2022; **40**(9): 1332–1335.
PubMed Abstract | Publisher Full Text | Free Full Text

Eger JL: **Family Molossidae.** Chicago, The University of Chicago Press, 2008a.

Eger JL: **Mammals of South America. Marsupials, xenarthrans, shrews, and bats, Family Molossidae.** Chicago, The University of Chicago Press, 2008b.
Publisher Full Text

González-Ruiz N, Ramírez-Pulido J, Arroyo-Cabrales J: **A new species of mastiff bat (Chiroptera: Molossidae: *Molossus*) from Mexico.** *Mamm Biol.* 2011; **76**(4): 461–469.
Publisher Full Text

Jung K, Molinari J, Kalko EKV: **Driving factors for the evolution of species-specific echolocation call design in new world free-tailed bats (Molossidae).** *PLoS One.* 2014; **9**(1): e85279.
PubMed Abstract | Publisher Full Text | Free Full Text

Kunz TH, Parsons S: **Ecological and behavioral methods for the study of bats.** Baltimore, MD, USA, Johns Hopkins University Press, 2009.
Reference Source

Larivière D, Abueg L, Brajuka N, *et al.*: **Scalable, accessible and reproducible reference genome assembly and evaluation in Galaxy.** *Nat Biotechnol.* 2024; **42**(3): 367–370.
PubMed Abstract | Publisher Full Text

Loureiro LO, Engstrom MD, Lim BK: **Single Nucleotide Polymorphisms (SNPs) provide unprecedented resolution of species boundaries, phylogenetic relationships, and genetic diversity in the mastiff bats (*Molossus*).** *Mol Phylogenet Evol.* 2020; **143**: 106690.
PubMed Abstract | Publisher Full Text

Loureiro LO, Engstrom MD, Lim BK: **Not all *Molossus* are created equal: genetic variation in the mastiff bat reveals diversity masked by conservative morphology.** *Acta Chiropt.* 2019; **21**(1): 51–64.
Publisher Full Text

Manni M, Berkeley MR, Seppey M, *et al.*: **BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes.** *Mol Biol Evol.* 2021; **38**(10): 4647–4654.
PubMed Abstract | Publisher Full Text | Free Full Text

Meredith RW, Janečka JE, Gatesy J, *et al.*: **Impacts of the Cretaceous Terrestrial Revolution and KPg extinction on mammal diversification.** *Science.* 2011; **334**(6055): 521–524.
PubMed Abstract | Publisher Full Text

Miller BM: **Community ecology of the non-phyllostomid bats of Northwestern Belize, with a landscape level assessment of the bats of Belize.** University of Kent, 2003.
Reference Source

Nurk S, Walenz BP, Rhie A, *et al.*: **HiCanu: accurate assembly of segmental duplications, satellites, and allelic variants from high-fidelity long reads.** *Genome Res.* 2020; **30**(9): 1291–1305.
PubMed Abstract | Publisher Full Text | Free 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.
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

Simao 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

Simmons NB, Cirranello AL: **Bat Species of the World: a taxonomic and geographic database.** Version 1.5. 2024.

Teeling EC, Springer MS, Madsen O, *et al.*: **A molecular phylogeny for bats illuminates biogeography and the fossil record.** *Science.* 2005; **307**(5709): 580–584.
PubMed Abstract | Publisher Full Text

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

Open Peer Review

Current Peer Review Status: ✔ ✔

Version 1

Reviewer Report 14 October 2024

https://doi.org/10.21956/wellcomeopenres.25029.r99870

✔ **M. Alejandra Camacho**
Facultad de Ciencias Exactas y Naturales, Pontificia Universidad Catolica del Ecuador Escuela de Ciencias Biologicas, Quito, Pichincha, Ecuador

Dear Editor and Authors,

The genome assembly presented in this manuscript provides bases for future studies on the species' evolutionary history, genetic diversity, and conservation status. The use of advanced sequencing technologies demonstrates a high level of technical rigor, and the fully phased assembly represents a key milestone for bat genomics. I believe that with some minor clarifications and adjustments, this work will make an important contribution to the field.

Figure 1 (Phylogeny): The figure, as presented, is not sufficiently informative. It does not accurately convey the phylogenetic position of *Molossus alvarezi* within Molossidae, as it lacks representation of other Molossus species. While the legend mentions that "within *Molossus*, *M. alvarezi* occupies a basal branch and is sister to a large clade of other species from Central and South America," this key phylogenetic relationship is neither shown nor easily inferred from the figure. Including additional *Molossus* species in the phylogeny would provide the necessary context to support the legend's claim and improve the overall clarity and scientific value of the figure.

Genome Assembly (Table 1): The genome assembly report provides a detailed overview of the sequencing and assembly process; however, there are a few discrepancies or missing details that should be addressed for clarity and consistency. For instance, the text mentions a total assembly length of 2.48 Gb in 504 scaffolds, while Table 1 indicates 187 scaffolds, potentially causing confusion for the reader. Additionally, the reported scaffold N50 in the text is 112.65 Mb, which differs from the value listed in Table 1 (113.92 Mb). Such differences, albeit small, should be reconciled.

BUSCO Scores (Table 1): The description of the BUSCO scores in the text mentions specific categories (C = complete, S = single copy, D = duplicated, F = fragmented, M = missing), but these are not reflected or explained in **Table 1.** Including a column or at least a footnote in the table that

breaks down these BUSCO scores by category would significantly improve the clarity and informativeness of the table. As it stands, the reader is left without the specific breakdown that is crucial for understanding the quality of the assembly in terms of completeness and potential duplications or gaps. Additionally, it would be helpful to see the percentage values for these categories, as only a general percentage of completeness (98.2%) is mentioned in the text, but not in the table.

Figure 3: The Hi-C Contact Map is a valuable visualization, but it would benefit from additional context or explanation to assist readers in interpreting the data. Currently, the figure is missing guidance on how to read the contact frequencies across the chromosomes, and the **axes** are not labeled, which makes it more difficult to interpret. Maybe genomic coordinates or chromosome numbers or positions (in base pairs). Including this information would help readers understand which regions of the genome are involved in the interactions shown. It would be important to highlight key chromosomal regions or explaining what dense contact regions signify.

Figure 4: This figure could be improved by incorporating the detailed BUSCO score breakdown that was listed in Table 1. Including the categories (C = complete, S = single copy, D = duplicated, F = fragmented, M = missing) within the figure or its legend would provide more immediate insight into the quality and completeness of the genome assembly. This would make the figure more comprehensive by allowing readers to assess the assembly's quality directly from the visual data, without having to cross-reference Table 1.

Figure 5: The GC coverage plot is informative, but it might benefit from clearer labeling of the axes and a brief explanation of the significance of different GC content levels for the chromosomes or scaffolds.

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: Diversity, systematics, and taxonomy of neotropical bats

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 07 October 2024

https://doi.org/10.21956/wellcomeopenres.25029.r99868

Diego A Caraballo
Universidad de Buenos Aires and CONICET, Buenos Aires, Argentina

The manuscript submitted by Simmons et al. presents the genome assembly of a female *Molossus alvarezi*. Using a methodology that generates long reads, along with a technique to capture the spatial proximity between different regions of the genome, the authors produced a high-quality assembly totaling 2.48 Gb in length, with a mean depth of 51x, organized into 24 chromosomal pseudomolecules. The assembly is fully phased and exhibits a BUSCO completeness of 98.2% relative to the Laurasiatheria reference set. Together, these metrics confirm that it is a high-quality genome assembly. The manuscript is well-written, provides all necessary information regarding the methods and software employed, and makes the generated assembly readily available.

Minor comment:
In the introduction and Figure 1, the authors state that there are 15 extant species of Molossus. However, according to Loureiro et al. (2020), there were initially 14 species. Two additional species described later, M. melini (Montani et al. 2021) and M. paranaensis (Chambi et al. 2024), bring the actual total to 16 species.

References
1. Loureiro LO, Engstrom MD, Lim BK: Single nucleotide polymorphisms (SNPs) provide unprecedented resolution of species boundaries, phylogenetic relationships, and genetic diversity in the mastiff bats (Molossus).*Mol Phylogenet Evol*. 2020; **143**: 106690 PubMed Abstract | Publisher Full Text
2. Montani M, Tomasco I, Barberis I, Romano M, et al.: A new species ofMolossus (Chiroptera: Molossidae) from Argentina. *Journal of Mammalogy*. 2021; **102** (5): 1426-1442 Publisher Full Text
3. Chambi Velasquez M, Pavé R, Argoitia M, Schierloh P, et al.: Revisiting Molossus (Mammalia: Chiroptera: Molossidae) diversity: Exploring southern limits and revealing a novel species in Argentina. *Vertebrate Zoology*. 2024; **74**: 397-416 Publisher Full Text

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: Phylogenetics, Phylogenomics, Molecular Evolution, Viral Zoonoses, Bats, Rodents

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.