Check for updates

DATA NOTE

ERGA-BGE reference genome of the European hamster (*Cricetus cricetus*), an IUCN critically endangered species of the Eurasian steppe

[version 1; peer review: awaiting peer review]

Virág Tünde Nyíri[1], Gábor Sramkó[1,2], Tamás Cserkész[3], Lajos Szatmári[3], Astrid Böhne[4], Rita Monteiro[4], Thomas Marcussen[5,6], Rebekah A. Oomen[5-9], Torsten H. Struck[6], José Melo-Ferreira[10-12], João Pimenta[10,12], Marta Gut[13,14], Laura Aguilera[13,14], Francisco Câmara Ferreira[13,14], Fernando Cruz[13,14], Jèssica Gómez-Garrido[13,14], Tyler S. Alioto[13,14], Fergal Martin[15], Swati Sinha[15], Leanne Haggerty[15], Chiara Bortoluzzi[16]

[1]University of Debrecen, Debrecen, Hungary
[2]HUN-REN–UD Conservation Biology Research Group, Debrecen, Hungary
[3]Hungarian Natural History Museum, Budapest, Hungary
[4]Leibniz Institute for the Analysis of Biodiversity Change, Museum Koenig Bonn, Bonn, 53113, Germany
[5]University of Oslo Centre for Ecological and Evolutionary Synthesis, Oslo, Oslo, Norway
[6]Natural History Museum, P.O. Box 1172, Blindern, University of Oslo, Oslo, 0318, Norway
[7]Tjärnö Marine Laboratory, University of Gothenburg, Gothenburg, Sweden
[8]Centre for Coastal Research, University of Agder, Kristiansand, Norway
[9]Department of Biological Sciences, University of New Brunswick Saint John, Saint John, Canada
[10]CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Universidade do Porto, Vairão, Portugal
[11]Departamento de Biologia, Faculdade de Ciências da Universidade do Porto, Porto, Portugal
[12]BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Vairão, Portugal
[13]Universitat de Barcelona, Barcelona, Spain
[14]Centro Nacional de Analisi Genomico (CNAG), Barcelona, Spain
[15]European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, England, UK
[16]Swiss Institute of Bioinformatics, Environmental Bioinformatics group, Amphipôle, Quartier UNIL-Sorge, Lausanne, 1015, Switzerland

v1 **First published:** 09 Jun 2026, **6**:196
https://doi.org/10.12688/openreseurope.23911.1
Latest published: 09 Jun 2026, **6**:196
https://doi.org/10.12688/openreseurope.23911.1

Abstract

The common hamster is currently one of the fastest-declining mammals in Europe, with its protection facing multiple challenges due to the complexity of the threats posed by factors such as population fluctuations, agricultural intensification, and reduced reproductive

Open Peer Review

Approval Status *AWAITING PEER REVIEW*

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

output. Overcoming these challenges requires population genomic data to inform targeted conservation measures. Therefore, the *Cricetus cricetus* reference genome will serve as an important genomic resource for conservation genomic works aiming to halt the decline of this critically endangered species of the western part of the Eurasian steppe. Consistently with the reported karyotype, a total of 12 contiguous chromosomal pseudomolecules (10 autosomes and 2 sex chromosomes) and one mitochondrial genome were assembled from the genome sequence. This chromosome-level assembly encompasses 2.5 Gb, composed of 179 contigs and 49 scaffolds, with contig and scaffold N50 values of 47.7 Mb and 305.5 Mb, respectively.

Keywords
Agricultural habitat, Black-bellied hamster, Common hamster, Cricetidae, genome assembly, genome annotation, habitat shift, European Reference Genome Atlas, Biodiversity Genomics Europe, Earth BioGenome Project, European hamster, Rodentia



This article is included in the Horizon Europe gateway.



This article is included in the Genome Reports from the Biodiversity Genomics Europe Project collection.

Corresponding authors: Gábor Sramkó (sramko.gabor@science.unideb.hu), Chiara Bortoluzzi (chiara.bortoluzzi@sib.swiss)

Author roles: Tünde Nyíri V: Investigation, Resources, Writing – Original Draft Preparation, Writing – Review & Editing; **Sramkó G**: Investigation, Project Administration, Resources, Writing – Review & Editing; **Cserkész T**: Investigation, Resources, Writing – Review & Editing; **Szatmári L**: Investigation, Resources, Writing – Review & Editing; **Böhne A**: Methodology, Project Administration, Supervision, Writing – Review & Editing; **Monteiro R**: Methodology, Project Administration, Supervision, Writing – Review & Editing; **Marcussen T**: Methodology, Project Administration, Supervision, Writing – Review & Editing; **Oomen RA**: Methodology, Project Administration, Supervision, Writing – Review & Editing; **Struck TH**: Methodology, Project Administration, Supervision, Writing – Review & Editing; **Melo-Ferreira J**: Methodology, Project Administration, Supervision, Writing – Review & Editing; **Pimenta J**: Methodology, Project Administration, Supervision, Writing – Review & Editing; **Gut M**: Investigation, Writing – Review & Editing; **Aguilera L**: Investigation, Writing – Original Draft Preparation, Writing – Review & Editing; **Câmara Ferreira F**: Data Curation, Formal Analysis, Writing – Review & Editing; **Cruz F**: Data Curation, Formal Analysis, Writing – Review & Editing; **Gómez-Garrido J**: Data Curation, Formal Analysis, Writing – Review & Editing; **Alioto TS**: Data Curation, Formal Analysis, Supervision, Writing – Review & Editing; **Martin F**: Data Curation, Formal Analysis, Writing – Review & Editing; **Sinha S**: Data Curation, Formal Analysis, Writing – Review & Editing; **Haggerty L**: Data Curation, Formal Analysis, Writing – Review & Editing; **Bortoluzzi C**: Visualization, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: Biodiversity Genomics Europe (Grant no. 101059492) is funded by Horizon Europe under the Biodiversity, Circular Economy and Environment call (REA.B.3); co-funded by the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract numbers 22.00173 and 24.00054; and by the UK Research and Innovation (UKRI) under the Department for Business, Energy and Industrial Strategy's Horizon Europe Guarantee Scheme. This research was also made possible by the funding from the Hungarian Government and the European Union's LIFE Program under the GRASSLAND-HU Integrated project (LIFE17 IPE/HU/000018). The work was also funded by the National Research, Innovation and Development Office of Hungary via an NKFIH-OTKA grant (FK137962).
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2026 Tünde Nyíri V *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: Tünde Nyíri V, Sramkó G, Cserkész T *et al.* **ERGA-BGE reference genome of the European hamster (*Cricetus cricetus*), an IUCN critically endangered species of the Eurasian steppe [version 1; peer review: awaiting peer review]** Open Research Europe 2026, **6**:196 https://doi.org/10.12688/openreseurope.23911.1

First published: 09 Jun 2026, **6**:196 https://doi.org/10.12688/openreseurope.23911.1

Introduction

Cricetus cricetus was once a common member of the grasslands of the western part of the Eurasian Steppes (i.e., the naturally treeless grasslands west of the Altai Mts.) with an extension to more forested Western European regions in the Czech Republic, Germany, the Benelux states and NE France (Kryštufek et al., 2020) and in Russia on the eastern territories. This expansion might be related to the adaptation of the species to grassland-like agricultural landscapes in historical times. The phylogenetic position of the species is set in the Muroidea, where it is included in the monophyletic lineage of *Cricetinae* within *Cricetidae* that - together with the *Muridae* - forms the crown clade of muroid rodents (Steppan & Schenk, 2017). Phylogeography of the species is based on mitochondrial control region (CR) and the cytochrome b gene (*CYTB*) that identified five geographically restricted lineages putatively connected to five refugial areas: Western Europe (lineage N), Pannonian Basin (P), Caucasian (C), Pontic Steppe (E), and Middle Asia (A), of which the Pannonian lineage is basal, sister to the rest (Feoktistova et al., 2023). After the successful habitat shift to agricultural lands, we are witnessing another shift in habitat preference of the species: the species is increasingly frequently observed in urban environments (Feoktistova et al., 2013). While the global trend shows a decline, hamster populations appeared to be stable in Central-Eastern Europe. Hungary, at the center of this region, hosts the largest population in the EU, but has suffered decade-long fluctuations. During the last population increase in 2020, hamsters started to use village gardens, while disappearing from agricultural habitats. These observations suggest a potentially novel way for the species to survive, synurbanisation, a novel habitat shift triggered by an increasingly intensified agrarian environment. Still, this new potential adaptation raises new conservation challenges and underscores the need for innovative conservation actions (Szatmári et al., 2024).

Cricetus cricetus is classified as 'Critically Endangered' on the IUCN Red List (Banaszek et al. 2020), reflecting its high risk of extinction in the wild. The current classification has been changed from 'Least Concern', marking a rare case of a species changing four Red List categories, mostly due to the rapid decline of hamster populations across the whole distribution range, which threatens the species with extinction by the beginning of the 2030's. Additionally, the species is listed under Appendix II of the Bern Convention, indicating its protection and conservation as a species of European interest, requiring concerted efforts from all range states to monitor and manage its declining population. The species is also included in the 4th Annex of the Habitats Directive of the European Union, which lists species in need of strict protection and habitat management to promote their long-term survival. While it is prohibited to kill hamsters in general, Hungary is an exception, where, with the largest hamster population in the EU, the species is listed in Annex V, allowing member countries to decide how to manage populations.

The hamster is a keystone burrow-dwelling species. It is well-established that native small mammals play essential and irreplaceable roles in the ecosystem (Delibes-Mateos et al., 2011) by, for example, creating microhabitats and improving soil structure, which in turn improves precipitation absorption and ventilation, thus facilitating nutrient circulation (H□drzak et al., 2021). Additionally, the European hamster, as a primary consumer in the food chain, serves as important prey for several grassland-inhabiting top predators (e.g., the saker falcon and the imperial eagle). The ecological and conservation importance of this species calls for urgent action to ensure its long-term preservation in Europe. Based on flow cytometry, its genome ranges from 3.36 Gb to 3.49 Gb (Gregory 2002), with a karyotype of 2n = 22 (Gamperl et al., 1976). Here we provide a high-quality reference genome of a male specimen, which can serve as an important genomic resource for future conservation genomic works (Theissinger et al., 2023).

The generation of this reference resource was coordinated by the European Reference Genome Atlas' (ERGA) initiative Biodiversity Genomics Europe (BGE) project, supporting ERGA's aims of promoting transnational cooperation to promote advances in the application of genomics technologies to protect and restore biodiversity (Mazzoni et al., 2023).

Materials & Methods

ERGA's sequencing strategy includes Oxford Nanopore Technology (ONT) and/or Pacific Biosciences (PacBio) for long-read sequencing, along with Hi-C sequencing for chromosomal architecture, Illumina Paired-End (PE) for polishing (i.e. recommended for ONT-only assemblies), and RNA sequencing for transcriptomic profiling, to facilitate genome assembly and annotation.

Sample and sampling information

On 7th of March 2024, one adult male individual of the species *Cricetus cricetus* was sampled in an alfalfa field near the settlement of Balmazújváros, Hungary, using a box trap by Lajos Szatmári, Virág Nyíri, and Gábor Sramkó. Sampling was performed under permit PE-KTFO/1568–18/2020 issued by Pest County Government Office. In Hungary, the species is only partially protected, as population control in the eastern part of the country may be carried out without a permit from March 1 to May 31. Skin material was sampled and kept at −80 °C until DNA extraction.

Vouchering information

Physical reference materials for the here sequenced specimen have been deposited in the Mammal Collection of the Hungarian Natural History Museum (HNHM) (www.nhmus.hu), under the accession number HNHM-MAMM-28811.

Frozen reference tissue material of muscle is available for HNHM-MAMM-28811 at the Hungarian Natural History Museum (HNHM) Collection of Genetic Resources under voucher ID HNHM-CGR-33863.

Data availability

Cricetus cricetus and the related genomic study were assigned to Tree of Life ID (ToLID) 'mCriCri2', and all sample, sequence, and assembly information are available under the umbrella BioProject PRJEB81695. The sample information is available at the following BioSample accessions: SAMEA115676028, SAMEA115676030, and SAMEA116317919. The genome assembly is accessible from ENA under accession number GCA_964304595.1, and the annotated genome is available through the Ensembl Website (https://projects.ensembl.org/erga-bge/). Sequencing data produced as part of this project are available from ENA at the following accessions: ERX13300601, ERX13300602, ERX13300603, ERX13300604, and ERX13300605. Documentation related to the genome assembly and curation can be found in the ERGA Assembly Report (EAR) document available at https://github.com/ERGA-consortium/EARs/blob/main/Assembly_Reports/Cricetus_cricetus/mCriCri2/. Further details and data about the project are hosted on the ERGA portal at https://portal.erga-biodiversity.eu/data_portal/10034.

Genetic information

The estimated genome size, based on flow cytometry, is between 3.36 Gb and 3.49 Gb. This corresponds to a diploid genome with a haploid number of 11 chromosomes (2n = 22).

DNA/RNA processing

DNA was extracted from liver tissue using the Blood & Cell Culture DNA Midi Kit (Qiagen) following the manufacturer's instructions. DNA quantification was performed using a Qubit dsDNA BR Assay Kit (Thermo Fisher Scientific), and DNA integrity was assessed using a Genomic DNA 165 Kb Kit (Agilent) on the Femto Pulse system (Agilent). The DNA was stored at +4 °C until used. RNA was extracted using a RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. RNA was extracted from four different specimen parts: liver, spleen, testicle and brain. RNA quantification was performed using the Qubit RNA BR kit, and RNA integrity was assessed using a Bioanalyzer 2100 system (Agilent) RNA 6000 Nano Kit (Agilent). The RNA was pooled equimolarly for the library preparation and stored at-80 °C until used.

Library preparation and sequencing

For long-read whole genome sequencing, a library was prepared using the SQK-LSK114 Kit (Oxford Nanopore Technologies, ONT), which was then sequenced on a PromethION 24 A Series instrument (ONT). A short-read whole-genome sequencing library was prepared using the KAPA Hyper Prep Kit (Roche). A Hi-C library was prepared from spleen tissue using the Dovetail Omni-C Kit (Cantata Bio), followed by the KAPA Hyper Prep Kit for Illumina sequencing (Roche). The RNA library was prepared from the pooled sample using the KAPA mRNA Hyper prep kit (Roche). The short-read libraries were sequenced on a NovaSeq 6000 instrument (Illumina). In total, 105x Oxford Nanopore, 84x Illumina WGS shotgun, and 78x HiC data were sequenced to generate the assembly.

Genome assembly methods

The genome was assembled using the CNAG CLAWS pipeline (Gomez-Garrido, 2024). Briefly, ONT reads were preprocessed for quality and length using Trim Galore v0.6.7 and Filtlong v0.2.1 and were filtered to 216 Gb (40 Kb read N50) with a median read quality of 20.8. Initial contigs were assembled using NextDenovo v2.5.0 (Hu et al., 2024), followed by polishing of the assembled contigs using HyPo v1.0.3 (Kundu et al., 2019), removal of retained haplotigs using purge-dups v1.2.6 (Guan et al., 2020) and scaffolding with YaHS v1.2a (Zhou et al., 2023). Finally, assembled scaffolds were curated via manual inspection using Pretext v0.2.5 (Harry, 2022) with the Rapid Curation Toolkit (https://gitlab.com/wtsi-grit/rapid-curation) to remove any false joins and incorporate any sequences not automatically scaffolded into their respective locations in the chromosomal pseudomolecules (or super-scaffolds). Finally, the mitochondrial genome was assembled as a single circular contig of 16,265 bp using the FOAM pipeline v0.5 (https://github.com/cnag-aat/FOAM) and included in the released assembly (GCA_964304595.1). Summary analysis of the released assembly was performed using the ERGA-BGE Genome Report ASM Galaxy workflow (https://doi.org/10.48546/workflowhub.workflow.1103.2).

Genome annotation methods

A gene set was generated using the Ensembl Gene Annotation system (Aken et al., 2016), primarily by aligning publicly available short-read RNA-seq data from BioSample SAMEA116317919 to the genome. Gaps in the annotation were filled via protein-to-genome alignments of a select set of vertebrate proteins from UniProt (Consortium, 2019), which had experimental evidence at the protein or transcript level. At each locus, data were aggregated and consolidated, prioritising models derived from RNA-seq data, resulting in a final set of gene models and associated non-redundant transcript sets. To distinguish true isoforms from fragments, the likelihood of each open reading frame (ORF) was evaluated against known vertebrate proteins. Low-quality transcript models, such as those showing evidence of fragmented ORFs, were removed. In cases where RNA-seq data were fragmented or absent, homology data were prioritised, favouring longer transcripts with strong intron support from short-read data. The resulting gene models were classified into three categories: protein-coding, pseudogene, and long non-coding. Models with hits to known proteins and few structural

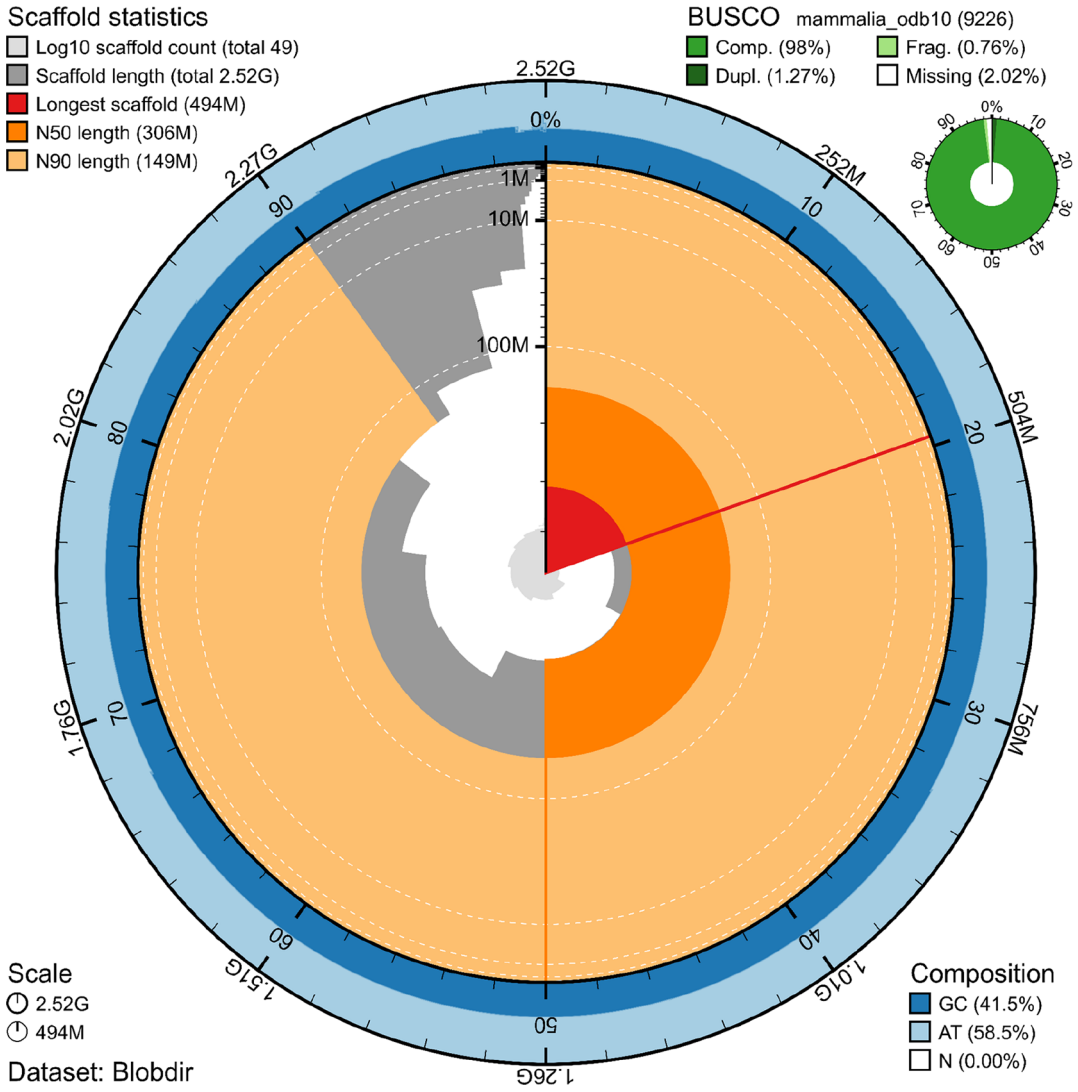

Figure 1. Snail plot summary of assembly statistics. The main plot is divided into 1,000 size-ordered bins around the circumference, with each bin representing 0.1% of the 2,520,178,292 bp assembly including the mitochondrial genome. The distribution of sequence lengths is shown in dark grey, with the plot radius scaled to the longest sequence present in the assembly (494 Mb, shown in red). Orange and pale-orange arcs show the scaffold N50 and N90 sequence lengths (305,542,590 bp and 148,928,470 bp), respectively. The pale grey spiral shows the cumulative sequence count on a logarithmic scale, with white scale lines showing successive orders of magnitude. The blue and pale-blue area around the outside of the plot shows the distribution of GC, AT, and N percentages in the same bins as the inner plot. A summary of complete, fragmented, duplicated, and missing BUSCO genes found in the assembled genome from the Mammalia database (odb10) is shown on the top right.

abnormalities were classified as protein-coding. Models with hits to known proteins but displaying abnormalities, such as the absence of a start codon, non-canonical splicing, unusually small intron structures (<75 bp), or excessive repeat coverage, were reclassified as pseudogenes. Single-exon models with a corresponding multi-exon copy elsewhere in the genome were classified as processed (retrotransposed) pseudogenes. Models that did not fit any of the previously described categories, did not overlap protein-coding genes and were constructed from transcriptomic data were considered potential lncRNAs. Potential lncRNAs were further filtered to remove single-exon loci due to their unreliability. Putative miRNAs were predicted by performing a BLAST search of miRbase (Kozomara et al., 2019) against the genome, followed by RNAfold analysis (Gruber et al., 2008). Other small non-coding loci were identified by scanning the genome with Rfam (Kalvari et al., 2018) and passing the results through Infernal (Nawrocki & Eddy, 2013). Summary analysis of the released annotation was carried out using the ERGA-BGE Genome Report ANNOT Galaxy workflow (10.48546/workflowhub.workflow.1096.1).

Results

Genome assembly

The genome assembly has a total length of 2,520,178,292 bp in 12 super-scaffolds (10 autosomes plus the X and Y chromosomes), the mitogenome, five unlocalised sequences and 31 unplaced sequences (Figures 1 and 2), with a GC content of 42%. It features a contig N50 of 47,690,202 bp (L50 = 17) and a scaffold N50 of 305,542,590 bp (L50 = 4). There are 130 gaps, totalling 26.0 kb in cumulative size. The single-copy gene content analysis using the Mammalia database (odb10) with BUSCO v5.8.0 (Manni et al., 2021) resulted in 98.0% completeness (96.7% single-copy and 1.3% duplicated). 94.0% of reads k-mers were present in the assembly, and the assembly has a base accuracy Quality Value (QV) of 54.4 as calculated by Merqury (Rhie et al., 2020).

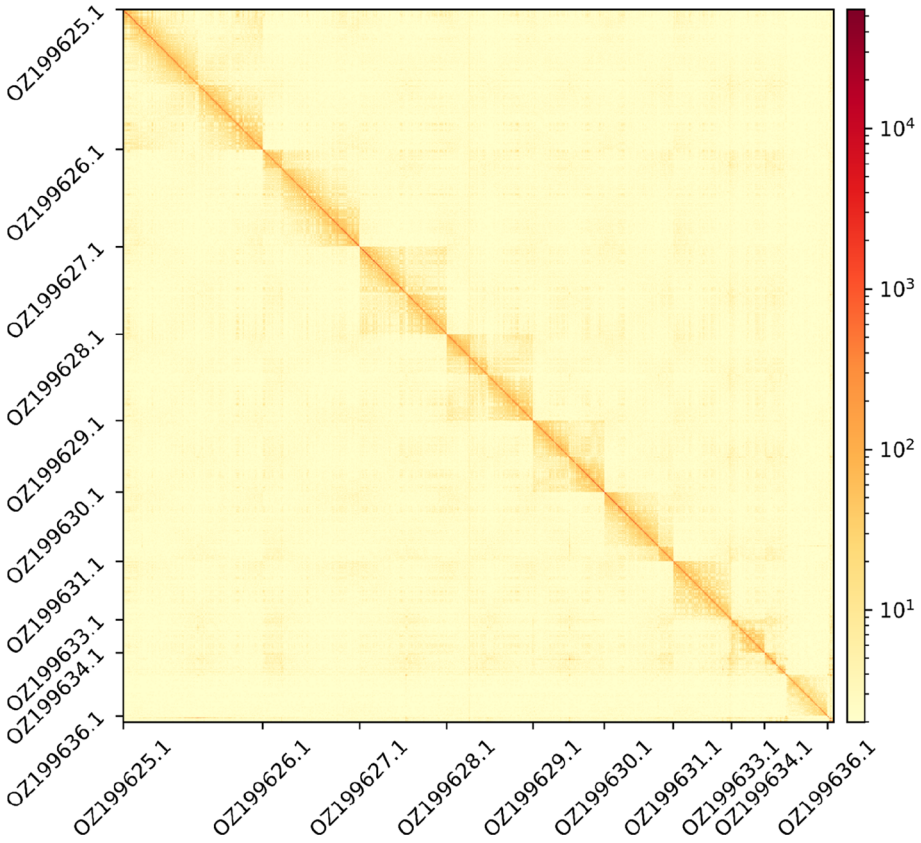

Figure 2. Hi-C contact map showing spatial interactions between regions of the genome. The diagonal corresponds to intra-chromosomal contacts, depicting chromosome boundaries. The frequency of contacts is shown on a logarithmic heatmap scale. Hi-C matrix bins were merged into a 200 kb bin size for plotting. Due to space constraints on both axes, we show the first 10 autosomes (GenBank accessions: OZ199625.1 - OZ199634.1) and both sex chromosomes, the X (GenBank accession: OZ199632.1) and Y chromosome (GenBank accession: OZ199636.1).

Table 1. Statistics from assembled gene models.

	No. genes	No. transcripts	Mean gene length (bp)	No. single-exon genes	Mean exons per transcript
Protein-coding	20,091	29,272	44,713	2,155	11.0
Ig/TCR segments	126	126	77–1,721	91	1.1–1.4
Pseudogenes	434	434	4,164	56	5.1
lncRNA	959	1,635	22,305	40	3.3
snRNA	1,216	1,216	120	1,216	1.0
snoRNA	744	744	110	744	1.0
rRNA	138	138	494	138	1.0
tRNA	22	22	68	22	1.0
miRNA	115	115	85	115	1.0
scRNA	44	44	157	44	1.0
Other non-coding	95	95	97–292	95	1.0–1.0

Table 2. Annotation completeness and consistency scores calculated by BUSCO run in protein mode (mammalia_odb10) and OMArk (Metazoa-v2.0.0.h5).

	Complete	Single-copy	Duplicated	Fragmented	Missing
BUSCO	9,080 (98.4%)	8,997 (97.5%)	83 (0.9%)	29 (0.3%)	117 (1.3%)
OMArk	15,144 (97.7%)	15,139 (95.2%)	405 (2.5%)	-	355 (2.2%)
	Consistent	**Inconsistent**	**Contaminants**	**Unknown**	
OMArk	19,859 (98.2%)	273 (1.3%)	0 (0.0%)	85 (0.4%)	

Genome annotation

The genome annotation consists of 20,091 protein-coding genes with an associated 29,272 transcripts, in addition to 126 immune system receptor segment genes, 434 pseudogenes, and 3,238 non-coding RNA genes of various types (Table 1). Using the longest isoform per transcript, the single-copy gene content analysis using the Mammalia (odb10) database with BUSCO resulted in 98.4% completeness. Using the OMAmer Metazoa-v2.0.0.h5 database (Nevers et al., 2025) resulted in 97.7% completeness and 98.2% consistency (Table 2).

Author contributions

GS coordinated the project; GS, LS and VN collected the species; TC identified the species; VN and LS sampled and preserved biological material and provided metadata; AsB, RM, TM, THS, RAO, JM-F, JP provided sampling and metadata support and management; MG and LA extracted DNA, prepared libraries, and performed sequencing; FC, FCFand JGG performed genome assembly and curation under the supervision of TA; FM, SS, and LH performed genome annotation; CB generated the analysis and report. All authors contributed to the writing, review, and editing of this genome note and read and approved the final version.

Acknowledgements

We would like to express our gratitude to the colleagues of the Bükk, Hortobágy, Körös-Maros, and Fertő-Hanság National Park Directorates for their invaluable assistance during field samplings. We acknowledge the support of the Freiburg Galaxy Team: Saim Momin and Björn Grüning, Bioinformatics, University of Freiburg (Germany), funded by the German Federal Ministry of Education and Research BMBF grant 031 A538A de. NBI-RBC and the Ministry of Science, Research and the Arts Baden-Württemberg (MWK) within the framework of LIBIS/de. NBI Freiburg. We would like to acknowledge the assembly reviewer, Caroline Menguy from Genoscope.

Literature Cited

Aken BL, Ayling S, Barrell D, *et al.*: **The Ensembl gene annotation system.** *Database.* 2016; **2016**: baw093.
Publisher Full Text

Banaszek A, Bogomolov P, Feoktistova N, *et al.*: *Cricetus cricetus. The IUCN Red list of threatened species 2020.* IUCN; 2020.

Consortium, U: **UniProt: A worldwide hub of protein knowledge.** *Nucleic Acids Res.* 2019; **47**(D1): D506–D515.

Delibes-Mateos M, Smith AT, Slobodchikoff CN, *et al.*: **The paradox of keystone species persecuted as pests: A call for the conservation of abundant small mammals in their native range.** *Biol. Conserv.* 2011; **144**(5): 1335–1346.
Publisher Full Text

Feoktistova NY, Surov AV, Tovpinetz NN, *et al.*: **The common hamster as a synurbist: A history of settlement in European cities.** *Zoologica Poloniae.* 2013; **58**(3–4): 116–129.
Publisher Full Text

Feoktistova NY, Meschersky IG, Shenbrot GI, *et al.*: **Phylogeography of the common hamster (*Cricetus cricetus*): Paleoclimatic reconstructions of Late Pleistocene colonization.** *Integrative Zoology.* 2023; **18**(3): 581–599.
PubMed Abstract | **Publisher Full Text**

Gamperl R, Vistorin G, Rosenkranz W: **A comparative analysis of the karyotypes of Cricetus cricetus and Cricetulus griseus.** *Chromosoma.* 1976; **55**(3): 259–265.
Publisher Full Text

Gomez-Garrido J: **CLAWS (CNAG's Long-read Assembly Workflow in Snakemake) [Computer software].** 2024.

Gregory TR: **Animal genome size database.** 2002.
Reference Source

Gruber AR, Lorenz R, Bernhart SH, *et al.*: **The vienna RNA websuite.** *Nucleic Acids Res.* 2008; **36**(suppl_2): W70–W74.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

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

Harry E: **PretextView (Paired REad TEXTure Viewer): A desktop application for viewing pretext contact maps.** 2022.

Hędrzak MJ, Badach E, Kornaś SA: **Preliminary assumptions for identification of the common hamster (Cricetus cricetus) as a service provider in the agricultural ecosystem.** *Sustainability.* 2021; **13**(12): 6793.
Publisher Full Text

Hu J, Wang Z, Sun Z, *et al.*: **NextDenovo: An efficient error correction and accurate assembly tool for noisy long reads.** *Genome Biol.* 2024; **25**(1): 107.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Kalvari I, Nawrocki EP, Argasinska J, *et al.*: **Non-Coding RNA Analysis Using the Rfam Database.** *Curr. Protoc. Bioinformatics.* 2018; **62**(1): e51.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Kozomara A, Birgaoanu M, Griffiths-Jones S: **miRBase: From microRNA sequences to function.** *Nucleic Acids Res.* 2019; **47**(D1): D155–D162.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Kryštufek B, Hoffmann IE, Nedyalkov N, *et al.*: **Cricetus cricetus (Rodentia: Cricetidae).** *Mamm. Species.* 2020; **52**(988): 10–26.
Publisher Full Text

Kundu R, Casey J, Sung W-K: **HyPo: Super fast & accurate polisher for long read genome assemblies.** *Biorxiv.* 2019; 2019–12.

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**

Mazzoni CJ, Ciofi C, Waterhouse RM: **Biodiversity: An atlas of European reference genomes.** *Nature.* 2023; **619**: 252–252.
Publisher Full Text

Nawrocki EP, Eddy SR: **Infernal 1.1: 100-fold faster RNA homology searches.** *Bioinformatics.* 2013; **29**(22): 2933–2935.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Nevers Y, Warwick Vesztrocy A, Rossier V, *et al.*: **Quality assessment of gene repertoire annotations with OMArk.** *Nat. Biotechnol.* 2025; **43**(1): 124–133.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Rhie A, Walenz BP, Koren S, *et al.*: **Merqury: Reference-free quality, completeness, and phasing assessment for genome assemblies.** *Genome Biol.* 2020; **21**(1): 245.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Steppan SJ, Schenk JJ: **Muroid rodent phylogenetics: 900-species tre reveals increasing diversification rates.** *PLoS One.* 2017; **12**(8): e0183070.
Publisher Full Text

Szatmári L, Sramkó G, Nyíri V, *et al.*: **Under pressure: The Critically Endangered common hamster in Hungary reached an all-time demographic low in 2024.** *Oryx.* 2024; **58**(4): 424–425.
Publisher Full Text

Theissinger K, Fernandes C, Formenti G, *et al.*: **How genomics can help biodiversity conservation.** *Trends Genet.* 2023; **39**: 545–559.
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**