DATA NOTE

REVISED The genome sequence of the Hungarian meadow viper, *Vipera ursinii rakosiensis* (Méhely, 1893)

[version 2; peer review: 6 approved]

Bálint Halpern[1-3], Judit Vörös[1,4,5], Ann M. Mc Cartney[6], Giulio Formenti[7], Alice Mouton[8],
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

[1]MME Birdlife Hungary, Budapest, Hungary
[2]Department of Systematic Zoology and Ecology, Institute of Biology, ELTE-Eötvös Loránd University, Budapest, Hungary
[3]HUN-REN – ELTE - MTM Integrative Ecology Research Group, Budapest, Hungary
[4]Department of Zoology, Hungarian Natural History Museum, Budapest, Hungary
[5]HUN-REN Balaton Limnological Research Institute, Tihany, Hungary
[6]University of California Santa Cruz, Santa Cruz, California, USA
[7]The Vertebrate Genome Laboratory, The Rockefeller University, New York, New York, USA
[8]SEED Arlon Campus, University of Liège, Arlon, Belgium

v2 **First published:** 26 Jul 2024, **9**:404
https://doi.org/10.12688/wellcomeopenres.22694.1
Latest published: 15 Jan 2025, **9**:404
https://doi.org/10.12688/wellcomeopenres.22694.2

Abstract

We present a genome assembly from an individual female *Vipera ursinii rakosiensis* (the Hungarian meadow viper; Chordata; Lepidosauria; Squamata; Viperidae). The genome sequence is 1,625.0 megabases in span. Most of the assembly is scaffolded into 19 chromosomal pseudomolecules, including the W and Z sex chromosomes. The mitochondrial genome has also been assembled and is 17.38 kilobases in length.

Keywords

Vipera ursinii rakosiensis, Hungarian meadow viper, genome sequence, chromosomal, Squamata

This article is included in the Tree of Life gateway.

Open Peer Review

Approval Status ✓ ✓ ✓ ✓ ✓ ✓

	1	2	3	4	5	6
version 2 (revision) 15 Jan 2025			✓ view	✓ view	✓ view	✓ view
version 1 26 Jul 2024	✓ view	✓ view				

1. **Turk Rhen**, University of North Dakota, Grand Forks, USA
2. **Jia-Tang Li**, University of Chinese Academy of Sciences, Beijing, China
3. **Alastair J. Ludington**, The University of Adelaide, Adelaide, Australia

4. **Divya Tej Sowpati**, CSIR-Centre for Cellular and Molecular Biology, Hyderabad, India

5. **Paul Doniol-Valcroze**, CNRS, Univ. Montpellier, Montpellier, France

6. **Drew Schield**, University of Virginia, Charlottesville, USA

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

Corresponding author: Wellcome Sanger Institute Tree of Life Management, Samples and Laboratory team (Mark.Blaxter@sanger.ac.uk)

Author roles: Halpern B: Investigation, Resources, Writing – Original Draft Preparation, Writing – Review & Editing; **Vörös J**: Investigation, Writing – Original Draft Preparation, Writing – Review & Editing; **Mc Cartney AM**: Project Administration; **Formenti G**: Project Administration; **Mouton A**: Project Administration;

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by Wellcome through core funding to the Wellcome Sanger Institute [206194, https://doi.org/10.35802/206194] and the Darwin Tree of Life Discretionary Award [218328, https://doi.org/10.35802/218328]. The sample collection in Hungary and the lab work was funded by EU LIFE-Fund (HUNVIPHABLIFE18 NAT/HU/000799). Bálint Halpern was supported by the KDP-2021 program of the Ministry of Innovation and Technology from the source of the National Research, Development and Innovation Fund (KDP_2021_ELTE_C1791523).
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: Halpern B, Vörös J, Mc Cartney AM *et al.* **The genome sequence of the Hungarian meadow viper, *Vipera ursinii rakosiensis* (Méhely, 1893) [version 2; peer review: 6 approved]** Wellcome Open Research 2025, **9**:404 https://doi.org/10.12688/wellcomeopenres.22694.2

First published: 26 Jul 2024, **9**:404 https://doi.org/10.12688/wellcomeopenres.22694.1

REVISED **Amendments from Version 1**

The taxonomic authority in the title and taxonomy section has been corrected to (Méhely, 1893), with parentheses, as requested by a peer reviewer.

Any further responses from the reviewers can be found at the end of the article

Species taxonomy

Eukaryota; Opisthokonta; Metazoa; Eumetazoa; Bilateria; Deuterostomia; Chordata; Craniata; Vertebrata; Gnathostomata; Teleostomi; Euteleostomi; Sarcopterygii; Dipnotetrapodomorpha; Tetrapoda; Amniota; Sauropsida; Sauria; Lepidosauria; Squamata; Bifurcata; Unidentata; Episquamata; Toxicofera; Serpentes; Colubroidea; Viperidae; Viperinae; *Vipera*; *Vipera ursinii* (Bonaparte, 1835) (NCBI:txid103942); *Vipera ursinii rakosiensis* (Méhely, 1893) (NCBI:txid394585).

Background

The Hungarian meadow viper (*Vipera ursinii rakosiensis*) – a member of the family *Viperidae* – is a steppe form of the *Vipera ursinii* species group (Ferchaud *et al.*, 2012; Freitas *et al.*, 2020; Nilson & Andrén, 2001; Vörös *et al.*, 2022; Zinenko *et al.*, 2015). This subspecies was originally described by the prominent Hungarian zoologist Lajos Méhely in 1893, type specimens originating from meadows on the banks of the Rákos River, within current boundaries of Budapest (Méhely, 1894).

The total length of this viper is up to 60 cm, typically with dark dorsal zig-zag pattern on grey or yellowish-brown basal colour, while the white throat gradually gets darker towards the belly, with off-white spots. Sexual dimorphism is less prominent than in other species of the family: males and females are distinguishable by total and tail length ratio and subcaudal scale numbers.

Distribution of the Hungarian meadow viper is restricted to the Carpathian-Basin. Only fragmented populations survived in the Hanság and Kiskunság regions of Hungary and Transylvanian region of Romania, while it went extinct in Austria due to cultivation changes destroying its prime habitats. This snake is an inhabitant of steppe remnants. Meadows and pastures that form a mosaic of wet and dry grass habitats is favoured by the species, providing the preferred high microclimatic diversity and abundance of prey (Korsós *et al.*, 2000).

The Hungarian meadow viper is mainly insectivorous, consuming grasshoppers, crickets and small-sized lizards. Adults opportunistically predate on vertebrates such as lizards and rodents. They are diurnal, tending to start each morning basking, usually avoiding being exposed too much as they are predated by many other species.

From October to March, snakes spend the winter underground in abandoned burrows. Males emerge from hibernation in the middle of March, 2 to 4 weeks before females. After a few weeks they start to shed their skin, then begin searching for mates. Females shed their skins after the mating season is over. Hungarian meadow vipers are ovoviviparous, females giving birth to average 10 live offspring usually in the end of July or beginning of August. Newborn snakes shed their skin immediately after they leave the transparent sac in which they were born, then they start their individual life, which is filled with challenges (Halpern, 2007; Korsós *et al.*, 2000).

The Hungarian meadow viper is Europe's most endangered venomous snake. It was declared protected in Hungary in 1974 and it is strictly protected since 1988. The subspecies is included in the Bern Convention Appendix II [Council of Europe, 1979 (revised 2002)], and is listed in Appendix I of the Convention on International Trade in Endangered Species of Wild Flora and Fauna (CITES, 1987) and listed on the International Union for Conservation of Nature (IUCN) Red List as Endangered (IUCN, 1996) and also listed on the annex II of the Habitats Directive (Council Directive 92/43/EEC 1992), meaning that its occurrences have to be included in the Natura2000 Network of protected areas (Edgar & Bird, 2005).

As the severe decline of the subspecies was noted by experts by late 1990s, a cooperative conservation effort was initiated in Hungary to save the species from extinction. The Species Conservation Plan targeted to restore populations and habitats with various measures, including captive breeding and reintroduction to reconstructed habitats (Dankovics *et al.*, 2004; Halpern, 2007). The ex-situ populations' genetic screening was a necessity from the very beginning of the programme (Péchy *et al.*, 2015; Vörös *et al.*, 2022), raising various taxonomy or conservation related questions, in which the availability of a precise reference genome opens new horizons. The available reference sequence can also help recent and future efforts, working on more detailed analysis of the phylogeny and phylogeography of the *Vipera ursinii* species-complex, using next generation sequencing methods.

Genome sequence report

The genome was sequenced from one female *Vipera ursinii rakosiensis* (Figure 1) collected from Budapest Zoo, Budapest, Hungary (47.15, 19.31). A total of 28-fold coverage in Pacific Biosciences single-molecule HiFi long reads was generated. Primary assembly contigs were scaffolded with chromosome conformation Hi-C data. Manual assembly curation corrected 204 missing joins or mis-joins and removed 3 haplotypic duplications, reducing the scaffold number by 21.63%, and increasing the scaffold N50 by 0.52%.

The final assembly has a total length of 1,625.0 Mb in 383 sequence scaffolds with a scaffold N50 of 212.8 Mb (Table 1). The snail plot in Figure 2 provides a summary of the assembly statistics, while the distribution of assembly scaffolds on GC proportion and coverage is shown in Figure 3. The cumulative assembly plot in Figure 4 shows curves for subsets of scaffolds assigned to different phyla. Most (99.31%) of the assembly sequence was assigned to 19 chromosomal-level scaffolds, representing 17 autosomes and the W and Z sex chromosomes. Chromosome-scale scaffolds confirmed by the Hi-C

Figure 1. Photographs of *Vipera ursinii rakosiensis:* (**a**) basking female *in situ* in autumnal setting in Kiskunság (**b**) female concealed in dry grass as a defence against predators; (**c**) individual basking exposed; (**d**–**e**) characteristically, this species has vertical pupils and predominantly pale, unmarked labial scales; (**f**) variability in head scale arrangement in Hungarian meadow vipers is substantial, enabling individual identification in conservation efforts – shown is specimen rVipUrs1, utilised for genome sequencing.

data are named in order of size (Figure 5; Table 2). The order and orientation of chromosome W is not determined with full certainty. While not fully phased, the assembly deposited is of one haplotype. Contigs corresponding to the second haplotype have also been deposited. The mitochondrial genome was also assembled and can be found as a contig within the multifasta file of the genome submission.

The estimated Quality Value (QV) of the final assembly is 57.3. The *k*-mer completeness for the primary assembly is 93.12%, for the alternate haplotype is 84.69%, and for the combined assemblies is 99.09%. The assembly has a BUSCO v5.3.2 completeness of 92.6% (single = 91.1%, duplicated = 1.5%), using the sauropsida_odb10 reference set (n = 7,480).

Metadata for specimens, barcode results, spectra estimates, sequencing runs, contaminants and pre-curation assembly statistics are given at https://links.tol.sanger.ac.uk/species/103942.

Methods

Sample acquisition and nucleic acid extraction

Blood sample was collected from an adult female individual (2-gy-02/09, specimen ID ERGA_BP_HU_01, ToLID rVipUrs1), originating from Dabas population in Central Hungary. The

Table 1. Genome data for *Vipera ursinii rakosiensis*, rVipUrs1.1.

Project accession data		
Assembly identifier	rVipUrs1.1	
Species	*Vipera ursinii rakosiensis*	
Specimen	rVipUrs1	
NCBI taxonomy ID	103942	
BioProject	PRJEB55895	
BioSample ID	SAMEA12832258	
Isolate information	rVipUrs1, female: blood sample (PacBio, Hi-C and RNA sequencing)	
Assembly metrics*		***Benchmark***
Consensus quality (QV)	57.3	*≥ 50*
k-mer completeness	99.09% (combined)	*≥ 95%*
BUSCO**	C:92.6%[S:91.1%,D:1.5%], F:1.0%,M:6.3%,n:7,480	*C ≥ 95%*
Percentage of assembly mapped to chromosomes	99.31%	*≥ 90%*
Sex chromosomes	WZ	*localised homologous pairs*
Organelles	Mitochondrial genome: 17.38 kb	*complete single alleles*
Raw data accessions		
PacificBiosciences SEQUEL II	ERR10441188, ERR10441187	
Hi-C Illumina	ERR10177764	
PolyA RNA-Seq Illumina	ERR10908602	
Genome assembly		
Assembly accession	GCA_947247035.1	
Accession of alternate haplotype	GCA_947247025.1	
Span (Mb)	1,625.0	
Number of contigs	2,203	
Contig N50 length (Mb)	2.1	
Number of scaffolds	383	
Scaffold N50 length (Mb)	212.8	
Longest scaffold (Mb)	359.84	

* Assembly metric benchmarks are adapted from column VGP-2020 of "Table 1: Proposed standards and metrics for defining genome assembly quality" from Rhie *et al.* (2021).

** BUSCO scores based on the sauropsida_odb10 BUSCO set using version 5.3.2. C = complete [S = single copy, D = duplicated], F = fragmented, M = missing, n = number of orthologues in comparison. A full set of BUSCO scores is available at https://blobtoolkit.genomehubs.org/view/CAMXUJ01/dataset/CAMXUJ01/busco.

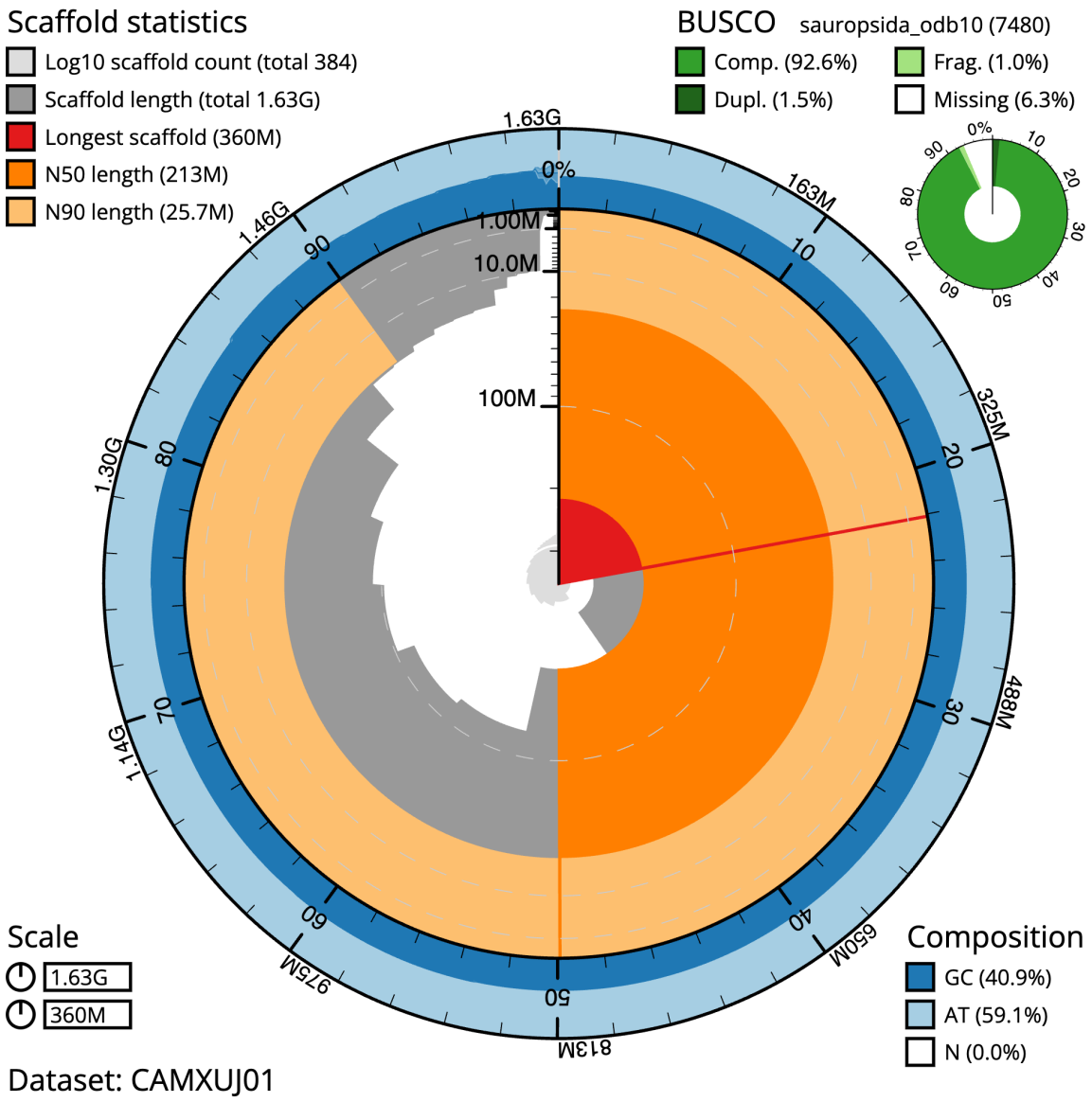

Figure 2. Genome assembly of *Vipera ursinii rakosiensis*, rVipUrs1.1: metrics. The BlobToolKit snail plot shows N50 metrics and BUSCO gene completeness. The main plot is divided into 1,000 size-ordered bins around the circumference with each bin representing 0.1% of the 1,625,023,540 bp assembly. The distribution of scaffold lengths is shown in dark grey with the plot radius scaled to the longest scaffold present in the assembly (359,753,992 bp, shown in red). Orange and pale-orange arcs show the N50 and N90 scaffold lengths (212,821,320 and 25,739,966 bp), respectively. The pale grey spiral shows the cumulative scaffold count on a log scale with white scale lines showing successive orders of magnitude. The blue and pale-blue area around the outside of the plot shows the distribution of GC, AT and N percentages in the same bins as the inner plot. A summary of complete, fragmented, duplicated and missing BUSCO genes in the sauropsida_odb10 set is shown in the top right. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/CAMXUJ01/dataset/CAMXUJ01/snail.

female is a founder of the ex-situ captive population in the Hungarian Meadow Viper Conservation Centre. The blood sampling took place in the Clinic of Budapest Zoo, drawing blood from the caudal vein of the live animal. The sample was placed in liquid nitrogen and later transferred to the Laboratory of Molecular Taxonomy of the Hungarian Natural History Museum where it was stored at –80 °C until cold chain shipping in dry ice to Wellcome Sanger Institute for sequencing.

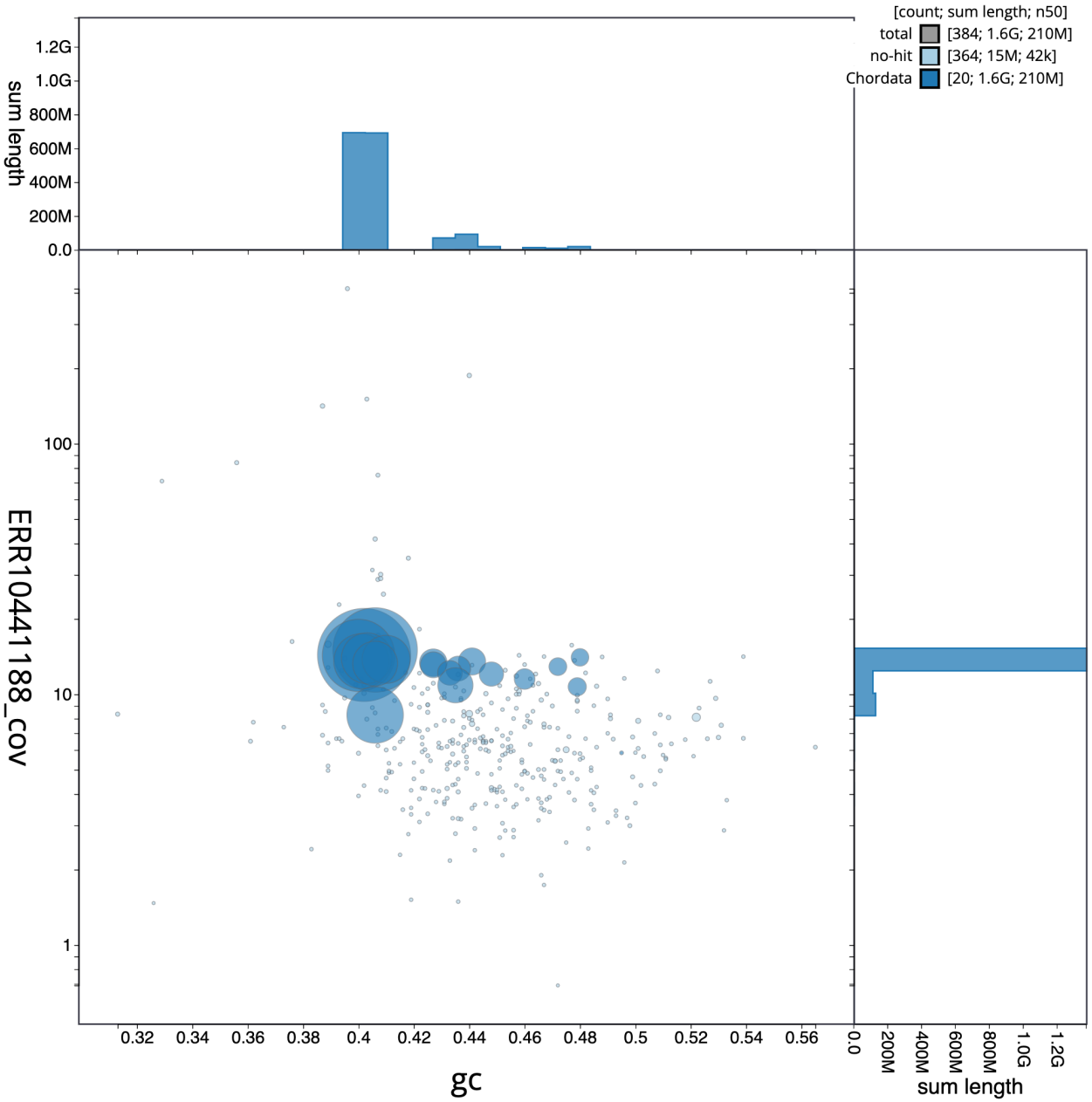

Figure 3. Genome assembly of *Vipera ursinii rakosiensis*, rVipUrs1.1: BlobToolKit GC-coverage plot. Sequences are coloured by phylum. Circles are sized in proportion to sequence length. Histograms show the distribution of sequence length sum along each axis. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/CAMXUJ01/dataset/CAMXUJ01/blob.

The workflow for high molecular weight (HMW) DNA extraction at the Wellcome Sanger Institute (WSI) includes a sequence of core procedures: sample preparation; sample homogenisation, DNA extraction, fragmentation, and clean-up. In sample preparation, the rVipUrs1 blood sample was weighed and dissected on dry ice (Jay *et al.*, 2023). For sample homogenisation, the blood was cryogenically disrupted using the Covaris cryoPREP® Automated Dry Pulverizer (Narváez-Gómez *et al.*, 2023).

HMW DNA was extracted using the Automated MagAttract v1 protocol (Sheerin *et al.*, 2023). DNA was sheared into an average fragment size of 12–20 kb in a Megaruptor 3 system with speed setting 30 (Todorovic *et al.*, 2023). Sheared DNA was

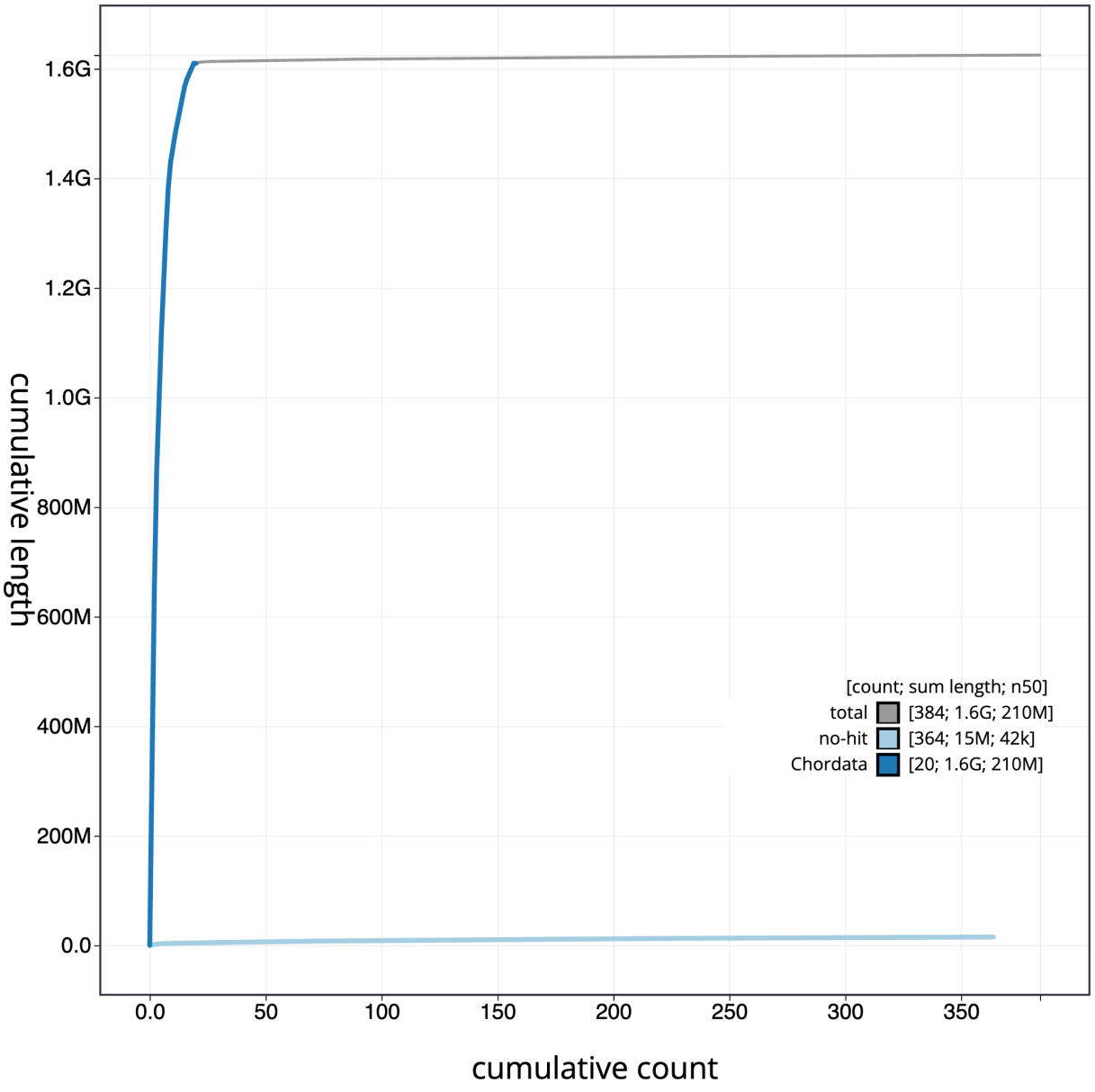

Figure 4. Genome assembly of *Vipera ursinii rakosiensis*, rVipUrs1.1: BlobToolKit cumulative sequence plot. The grey line shows cumulative length for all sequences. Coloured lines show cumulative lengths of sequences assigned to each phylum using the buscogenes taxrule. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/CAMXUJ01/dataset/CAMXUJ01/cumulative.

purified by solid-phase reversible immobilisation (Strickland *et al.*, 2023): in brief, the method employs a 1.8X ratio of AMPure PB beads to sample to eliminate shorter fragments and concentrate the DNA. The concentration of the sheared and purified DNA was assessed using a Nanodrop spectrophotometer and Qubit Fluorometer and Qubit dsDNA High Sensitivity Assay kit. Fragment size distribution was evaluated by running the sample on the FemtoPulse system.

RNA was extracted from an aliquot of the rVipUrs1 blood sample in the Tree of Life Laboratory at the WSI using the RNA Extraction: Automated MagMax™ *mir*Vana protocol (do Amaral *et al.*, 2023). The RNA concentration was assessed using a Nanodrop spectrophotometer 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.

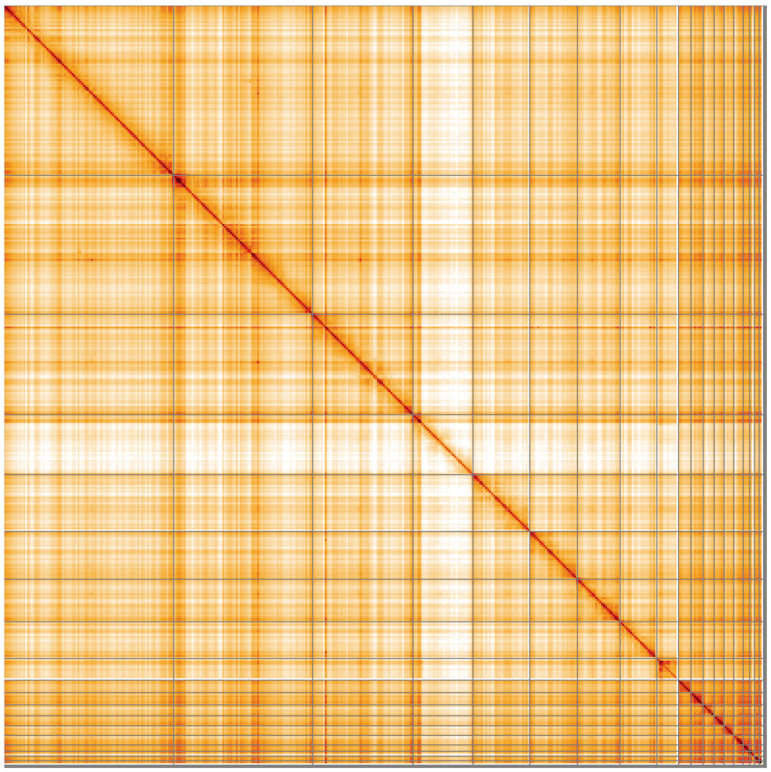

Figure 5. Genome assembly of *Vipera ursinii rakosiensis*, rVipUrs1.1: Hi-C contact map of the rVipUrs1.1 assembly, visualised using HiGlass. Chromosomes are shown in order of size from left to right and top to bottom. An interactive version of this figure may be viewed at https://genome-note-higlass.tol.sanger.ac.uk/l/?d=JtI56nHeRdeUI_b0Pv8Kow.

Table 2. Chromosomal pseudomolecules in the genome assembly of *Vipera ursinii rakosiensis*, rVipUrs1.

INSDC accession	Chromosome	Length (Mb)	GC%
OX365964.1	1	359.75	40.0
OX365965.1	2	294.45	40.5
OX365966.1	3	212.82	40.0
OX365968.1	4	120.14	40.0
OX365969.1	5	101.37	40.5
OX365970.1	6	90.32	41.0
OX365971.1	7	77.67	40.5
OX365973.1	8	26.54	42.5
OX365974.1	9	25.74	44.0
OX365975.1	10	23.23	42.5
OX365976.1	11	21.06	43.5
OX365977.1	12	20.68	43.5
OX365978.1	13	19.98	45.0
OX365979.1	14	13.74	46.0
OX365980.1	15	10.18	48.0
OX365981.1	16	9.66	47.0
OX365982.1	17	9.35	48.0
OX365972.1	W	45.96	43.5
OX365967.1	Z	127.14	40.5
OX365983.1	MT	0.02	41.0

Protocols developed by the WSI Tree of Life laboratory are publicly available on protocols.io (Denton *et al.*, 2023).

Sequencing

Pacific Biosciences HiFi circular consensus DNA sequencing libraries were constructed according to the manufacturers' instructions. Poly(A) RNA-Seq libraries were constructed using the NEB Ultra II RNA Library Prep kit. DNA and RNA sequencing was performed by the Scientific Operations core at the WSI on Pacific Biosciences SEQUEL II (HiFi) and Illumina NovaSeq 6000 (RNA-Seq) instruments. Hi-C data were also generated from the rVipUrs1 sample, using the Arima2 kit and sequenced on the Illumina NovaSeq 6000 instrument.

Genome assembly, curation and evaluation

Assembly was carried out with Hifiasm (Cheng *et al.*, 2021) and haplotypic duplication was identified and removed with purge_dups (Guan *et al.*, 2020). The assembly was then scaffolded with Hi-C data (Rao *et al.*, 2014) using YaHS (Zhou *et al.*, 2023). The assembly was checked for contamination and corrected as described previously (Howe *et al.*, 2021). Manual curation was performed using HiGlass (Kerpedjiev *et al.*, 2018) and PretextView (Harry, 2022). The mitochondrial genome was assembled using MitoHiFi (Uliano-Silva *et al.*, 2023), which runs MitoFinder (Allio *et al.*, 2020) or MITOS (Bernt *et al.*, 2013) and uses these annotations to select the final mitochondrial contig and to ensure the general quality of the sequence.

A Hi-C map for the final assembly was produced using bwa-mem2 (Vasimuddin *et al.*, 2019) in the Cooler file format (Abdennur & Mirny, 2020). To assess the assembly metrics, the *k*-mer completeness and QV consensus quality values were calculated in Merqury (Rhie *et al.*, 2020). This work was done using Nextflow (Di Tommaso *et al.*, 2017) DSL2 pipelines "sanger-tol/readmapping" (Surana *et al.*, 2023a) and "sanger-tol/genomenote" (Surana *et al.*, 2023b). The genome was analysed within the BlobToolKit environment (Challis *et al.*, 2020) and BUSCO scores (Manni *et al.*, 2021; Simão *et al.*, 2015) were calculated.

Table 3 contains a list of relevant software tool versions and sources.

Table 3. Software tools: versions and sources.

Software tool	Version	Source
BlobToolKit	4.2.1	https://github.com/blobtoolkit/blobtoolkit
BUSCO	5.3.2	https://gitlab.com/ezlab/busco
Hifiasm	0.16.1-r375	https://github.com/chhylp123/hifiasm
HiGlass	1.11.6	https://github.com/higlass/higlass
Merqury	MerquryFK	https://github.com/thegenemyers/MERQURY.FK
MitoHiFi	2	https://github.com/marcelauliano/MitoHiFi
PretextView	0.2	https://github.com/wtsi-hpag/PretextView
purge_dups	1.2.3	https://github.com/dfguan/purge_dups
sanger-tol/genomenote	v1.0	https://github.com/sanger-tol/genomenote
sanger-tol/readmapping	1.1.0	https://github.com/sanger-tol/readmapping/tree/1.1.0
YaHS	yahs-1.1.91eebc2	https://github.com/c-zhou/yahs

Wellcome Sanger Institute – Legal and Governance

The materials that have contributed to this genome note have been supplied by a Tree of Life collaborator. 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 undertaken according to a Research Collaboration Agreement or Material Transfer Agreement entered into by the Tree of Life collaborator, Genome Research Limited (operating as the Wellcome Sanger Institute) and in some circumstances other Tree of Life collaborators.

Data availability

European Nucleotide Archive: *Vipera ursinii rakosiensis* (Hungarian meadow viper). Accession number PRJEB55895; https://identifiers.org/ena.embl/PRJEB55895 (Wellcome Sanger Institute, 2023). The genome sequence is released openly for reuse. The *Vipera ursinii rakosiensis* genome sequencing initiative is part of the European Reference Genome Atlas Pilot Project (https://www.erga-biodiversity.eu/pilot-project). All raw sequence data and the assembly have been deposited in INSDC databases. The genome will be annotated using available RNA-Seq data and presented through the Ensembl pipeline at the European Bioinformatics Institute. Raw data and assembly accession identifiers are reported in Table 1.

Author information

Members of the Wellcome Sanger Institute Tree of Life Management, Samples and Laboratory team are listed here: https://doi.org/10.5281/zenodo.10066175.

Members of Wellcome Sanger Institute Scientific Operations: Sequencing Operations are listed here: https://doi.org/10.5281/zenodo.10043364.

Members of the Wellcome Sanger Institute Tree of Life Core Informatics team are listed here: https://doi.org/10.5281/zenodo.10066637.

Members of the Tree of Life Core Informatics collective are listed here: https://doi.org/10.5281/zenodo.5013541.

Acknowledgements

Dr. Endre Sós for providing veterinary assistance in sampling. Judit Bereczki for helping in sample storage and mailing.

References

Abdennur N, Mirny LA: **Cooler: scalable storage for Hi-C data and other genomically labeled arrays.** *Bioinformatics.* 2020; **36**(1): 311–316.
PubMed Abstract | Publisher Full Text | Free Full Text

Allio R, Schomaker-Bastos A, Romiguier J, *et al.*: **MitoFinder: efficient automated large-scale extraction of mitogenomic data in target enrichment phylogenomics.** *Mol Ecol Resour.* 2020; **20**(4): 892–905.
PubMed Abstract | Publisher Full Text | Free Full Text

Bernt M, Donath A, Jühling F, *et al.*: **MITOS: improved *de novo* metazoan mitochondrial genome annotation.** *Mol Phylogenet Evol.* 2013; **69**(2): 313–319.
PubMed Abstract | Publisher Full Text

Challis R, Richards E, Rajan J, *et al.*: **BlobToolKit - interactive quality assessment of genome assemblies.** *G3 (Bethesda).* 2020; **10**(4): 1361–1374.
PubMed Abstract | Publisher Full Text | Free Full Text

Cheng H, Concepcion GT, Feng X, *et al.*: **Haplotype-resolved *de novo* assembly using phased assembly graphs with hifiasm.** *Nat Methods.* 2021; **18**(2): 170–175.
PubMed Abstract | Publisher Full Text | Free Full Text

CITES: **Amendements to appendices I and II of the convention on international trade in endangered species of wild flora and fauna, adopted at the sixth meeting of the conference of the parties.** Ottawa, Canada, 12 to 24 July 1987, 1987.

Dankovics R, Halpern B, Pellinger A, *et al.*: **Rákosi vipera - fajmegőrzési terv. (Hungarian meadow viper – species conservation plan).** Ministry of Enviroment and Water Affairs, 2004.

Denton A, Yatsenko H, Jay J, *et al.*: **Sanger Tree of Life wet laboratory protocol collection V.1.** *protocols.io.* 2023.
Publisher Full Text

Di Tommaso P, Chatzou M, Floden EW, *et al.*: **Nextflow enables reproducible computational workflows.** *Nat Biotechnol.* 2017; **35**(4): 316–319.
PubMed Abstract | Publisher Full Text

do Amaral RJV, Bates A, Denton A, *et al.*: **Sanger Tree of Life RNA extraction: automated MagMax™ mirVana.** *protocols.io.* 2023.
Publisher Full Text

Edgar P, Bird D: **Action plan for the conservation of the meadow viper (*Vipera ursinii*) in Europe.** Bern Convention European Action Plan, 2005.
Reference Source

Ferchaud AL, Ursenbacher S, Cheylan M, *et al.*: **Phylogeography of the *Vipera ursinii* complex (Viperidae): mitochondrial markers reveal an east-west disjunction in the Palaearctic region.** *J Biogeogr.* 2012; **39**: 1836–1847.
Reference Source

Freitas I, Ursenbacher S, Mebert K, *et al.*: **Evaluating taxonomic inflation: towards evidence-based species delimitation in Eurasian vipers (Serpentes: Viperinae).** *Amphibia-Reptilia.* 2020; **41**(3): 285–311.
Publisher Full Text

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

Halpern B, (ed.): **A rákosi vipera védelme. Tanulmánygyűjtemény. (Studies on the conservation of the Hungarian Meadow Viper.).** Duna-Ipoly Nemzeti Park Igazgatóság, Rosalia 3, 2007.
Reference Source

Harry E: **PretextView (Paired REad TEXTure Viewer): a desktop application for viewing pretext contact maps.** 2022; [Accessed 19 October 2022].
Reference Source

Howe K, Chow W, Collins J, *et al.*: **Significantly improving the quality of genome assemblies through curation.** *GigaScience.* 2021; **10**(1): giaa153.
PubMed Abstract | Publisher Full Text | Free Full Text

IUCN European Reptile & Amphibian Specialist Group: **Vipera ursinii ssp. rakosiensis.** The IUCN Red List of Threatened Species 1996: e.T23003A9407721, 1996.
Publisher Full Text

Jay J, Yatsenko H, Narváez-Gómez JP, *et al.*: **Sanger Tree of Life sample preparation: triage and dissection.** *protocols.io.* 2023.
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.
PubMed Abstract | Publisher Full Text | Free Full Text

Korsós Z, Újvári B, Péchy T: **Life history, population characteristics and conservation of the Hungarian meadow viper (*Vipera ursinii rakosiensis*).** *Amphib Reptil.* 2000; **21**(3): 267–278.
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

Méhely L: **Eine neue Giftschlange der ungarischen Fauna (*Vipera rákosiensis* My.).** *Mat.Természettud.Ért.* 1894; **12**(2–3): 87–92.

Narváez-Gómez JP, Mbye H, Oatley G, *et al.*: **Sanger Tree of Life sample homogenisation: covaris CryoPREP® automated dry pulverizer V.1.** *protocols.io.* 2023.
Publisher Full Text

Nilson G, Andrén C: **The meadow and steppe vipers of Europe and Asia – the *Vipera* (*Acridophaga*) *ursinii* complex.** *Acta Zoologica Academiae Scientiarum Hungaricae.* 2001; **47**: 87–267.
Reference Source

Péchy T, Halpern B, Sós E, *et al.*: **Conservation of the Hungarian meadow viper *Vipera ursinii rakosiensis*.** *Int Zoo Yearb.* 2015; **49**(1): 89–103.
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.
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

Sheerin E, Sampaio F, Oatley G, *et al.*: **Sanger Tree of Life HMW DNA extraction: automated MagAttract v.1.** *protocols.io.* 2023.
Publisher Full Text

Simão FA, Waterhouse RM, Ioannidis P, *et al.*: **BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs.** *Bioinformatics.* 2015; **31**(19): 3210–3212.
PubMed Abstract | Publisher Full Text

Strickland M, Cornwell C, Howard C: **Sanger Tree of Life fragmented DNA clean up: manual SPRI.** *protocols.io.* 2023.
Publisher Full Text

Surana P, Muffato M, Qi G: **sanger-tol/readmapping: sanger-tol/readmapping v1.1.0 - Hebridean Black (1.1.0).** *Zenodo.* 2023a.
Publisher Full Text

Surana P, Muffato M, Sadasivan Baby C: **sanger-tol/genomenote (v1.0.dev).** *Zenodo.* 2023b.
Publisher Full Text

Todorovic M, Sampaio F, Howard C: **Sanger Tree of Life HMW DNA fragmentation: diagenode Megaruptor®3 for PacBio HiFi.** *protocols.io.* 2023.
Publisher Full Text

Uliano-Silva M, Ferreira JGRN, Krasheninnikova K, *et al.*: **MitoHiFi: a python pipeline for mitochondrial genome assembly from PacBio high fidelity reads.** *BMC Bioinformatics.* 2023; **24**(1): 288.
PubMed Abstract | Publisher Full Text | Free Full Text

Vasimuddin M, Misra S, Li H, *et al.*: **Efficient architecture-aware acceleration of BWA-MEM for multicore systems**. In: *2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS).* IEEE, 2019; 314–324.
Publisher Full Text

Vörös J, Ursenbacher S, Jelić D, *et al.*: **Well-known species, unexpected results: high genetic diversity in declining *Vipera ursinii* in central, eastern and southeastern Europe.** *Amphib Reptil.* 2022; **43**(4): 407–423.
Reference Source

Wellcome Sanger Institute: **The genome sequence of the Hungarian meadow *viper, Vipera ursinii* rakosiensis Méhely, 1893.** European Nucleotide Archive. [dataset], accession number PRJEB55895, 2023.

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

Zinenko O, Stümpel N, Mazanaeva L, *et al.*: **Mitochondrial phylogeny shows multiple independent ecological transitions and northern dispersion despite of Pleistocene glaciations in meadow and steppe vipers (*Vipera ursinii* and *Vipera renardi*).** *Mol Phylogenet Evol.* 2015; **84**: 85–100.
PubMed Abstract | Publisher Full Text

Open Peer Review

Current Peer Review Status: ✔ ✔ ✔ ✔ ✔ ✔

Version 2

Reviewer Report 15 February 2025

https://doi.org/10.21956/wellcomeopenres.26052.r117757

✔ **Drew Schield**
University of Virginia, Charlottesville, VA, USA

The authors have generated a high-quality reference genome for *Vipera ursinii*, which will be a valuable resource for research at various scales, from focused conservation genetics work on the species to broad comparative genomics investigations.

The authors have also taken into account constructive feedback from previous reviewers, which has improved the presentation of the results and summary of the reference genome assembly.

I have no additional concerns about this version of assembly report.

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: Genomics, population genetics, evolution

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 31 January 2025

https://doi.org/10.21956/wellcomeopenres.26052.r117495

Paul Doniol-Valcroze
CNRS, Univ. Montpellier, Montpellier, France

The authors present a chromosome level assembly for the highly endangered *rakosiensis* subspecies of the Meadow viper. Despite some very minor writing details outlined bellow, the rationale to produce such genomic resources is well justified and explained as well as the life history of the species or the origin of the sampled individual/population. I'll be very happy to see this genome note published and this genome made available for future studies.

Background:

- Rewrite the sentence about the ssp desciption: e.g. "was originally described [...] from types specimens originating from [...]"
- Ref for limited sexual dimorphism (end of 2nd paragraph)
- Change the sentence "Meadows and pastures [...] a mosaic of wet and dry grass habitats are favoured"
- I would deleted the mention to *small-sized lizards* when the authors explain that the species is insectivorous especially because they describe the opportunistic lizard consumption in the subsequent sentence: "The Hungarian meadow viper is mainly insectivorous, consuming grasshoppers, crickets and small-sized lizards. Adults opportunistically predate on vertebrates such as lizards and rodents."
- "The available [...] using next generation sequencing methods". Could be worth mentioning here the different and sometimes non concordant phylogenetic studies of the genus, a field which will certainly benefit from such genetic resources.

Genome sequence report: NA

Methods:

- Even is it is implicit here it could be worth mentioning that the specimen is a wild (not born in captivity) animal : "The female is a founder of the ex-situ captive population in the Hungarian Meadow Viper Conservation Centre."

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: My work is interested in describing and understanding various mechanisms at the origin of biodiversity especially speciation, hybridization and migration. I use various approaches such as population genomics, phylogenomics, chemical ecology, deployment of multisensor-loggers or phenotyping. My main study models are Songbirds, Reptiles & Amphibians and Butterflies.

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 30 January 2025

https://doi.org/10.21956/wellcomeopenres.26052.r117491

Divya Tej Sowpati
CSIR-Centre for Cellular and Molecular Biology, Hyderabad, India

I commend the authors for producing a high quality genome assembly of the Hungarian meadow viper. The article is complete in terms of understanding the genome quality.

It would be nice to have the following aspects covered:

1. **Repeat content:** Distribution of interspersed as well as tandem repeats.

2. Any analysis that was done on DNA methylation using the PacBio data.

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: Computational 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.

Reviewer Report 30 January 2025

https://doi.org/10.21956/wellcomeopenres.26052.r117753

Alastair J. Ludington
The University of Adelaide, Adelaide, Australia

Article summary:
The Data Note submitted by Halpern et al. presents the fist high quality genome assembly of the endangered Hungarian meadow viper (*Vipera ursinii rakosiensis*). This genome assembly, along with complementary sequencing data, will serve as a valuable resource for future ecological, conservation and demographic studies on *V. ursinii*, as well as for broader ecological and evolutionary research in herpetology. Using a combination of PacBio HiFi and Hi-C sequencing, the authors have assembled a 1.63Gb reference genome, consisting of 19 pseudomolecules (17 autosomes and the W and Z sex chromosomes), in addition to the mitochondrial genome. Assembly summary metrics, including BUSCO gene completeness (92.6%), Merqury K-mer completeness (99.09%), consensus quality (57.3) and assembly contiguity (N50 = 212.8Mb), indicate that the final assembly accurately represents the underlying data and is of high quality.

Constructive comments:
The assembly metrics presented in the manuscript indicate that the reference genome is of high quality. A nice addition to the text could be a comment on how the quality of this genome assembly compares to other available reference genomes for snakes in the Viperidae family, or even more broadly. Specifically relating to chromosome number, sequence accuracy and relative levels of gene completeness.

Second, there is mention of RNA-sequencing in the methods, though there is no mention of its use in the results. My assumption is that this is to be used for the annotation of protein coding genes, which if so, would be a valuable resource if provided in the future.

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: Evolutionary Biology, comparative 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.

Version 1

Reviewer Report 26 September 2024

https://doi.org/10.21956/wellcomeopenres.24995.r95430

Jia-Tang Li
University of Chinese Academy of Sciences, Beijing, People's Republic of China, China

Peer Review Report

Summary of the Article
This study presents the genome assembly of the Hungarian meadow viper (*Vipera ursinii rakosiensis*), a highly endangered snake species in Europe. The significance of this work lies in its potential contributions to conservation biology and serpentine genomics research. The authors have employed a combination of PacBio HiFi and Illumina sequencing technologies, enhanced by Hi-C, to successfully assemble a 1.63 Gb genome encompassing 19 chromosomal pseudomolecules, including the sex chromosomes and mitochondrial genome. The genome's integrity and quality have been highly appraised by BUSCO and Merqury metrics, indicating an

exceptional assembly quality, which is vital for further genomic studies.

Evaluation Report
1. Rationale for Creating the Dataset
The article articulates a well-defined rationale for the dataset creation, including the species' endangered status, biological significance, and implications for conservation efforts. The detailed context of sample collection also solidifies the justification for the dataset.

2. Appropriateness of Protocols and Technical Soundness of the Work
The study employs cutting-edge sequencing technologies and assembly strategies, which are recognized as state-of-the-art in genomics research.

3. Sufficiency of Details on Methods and Materials
The article provides a comprehensive experimental workflow, including sample processing, DNA and RNA extraction, sequencing platform selection, assembly strategies, and quality control measures, offering ample information for other researchers to replicate the study's methods.

4. Presentation and Accessibility of the Dataset
The dataset is made available through public databases with detailed access links and identifiers. The data is presented in standardized formats that are easy to utilize and analyze.

Constructive Recommendations
1. While the genome assembly is commendable, it is recommended that the authors situate their findings within the broader context of serpentine genomics. By benchmarking against other assemblies, the authors can underscore the precision and accuracy of their work, providing additional layers of validation and confidence in the data presented.

2. Genome annotation is a critical step in understanding the functions of genes and proteins, which is vital for revealing the biological characteristics and evolutionary relationships of species. It is hoped that the authors will promptly complete the genome annotation work and make the results publicly available.

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: Evolutionary Biology, Functional 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.

Reviewer Report 18 September 2024

https://doi.org/10.21956/wellcomeopenres.24995.r95435

Turk Rhen
Department of Biology, University of North Dakota, Grand Forks, North Dakota, USA

Overall, this Data Note is well written and provides the first high-quality genome assembly for an endangered snake, the Hungarian meadow viper (*Vipera ursinii rakosiensis*). This genome will be a valuable resource for current and future studies aimed at conservation and recovery of the species. The methods for DNA and RNA extraction, sequencing, and assembly of the genome are up to date and appropriate. The finished genome of 1.63 Gb has been assembled into 19 chromosomal pseudomolecules, including the sex chromosomes and the mitochondrial genome. Basic descriptive statistics indicate the genome assembly is of good quality: the assembly had a contig N50 of 2.1 Mb and a scaffold N50 of 213 Mb, *k*-mer completeness of 99.99%, and a BUSCO v5.3.2 completeness of 92.6%. The raw data and assembly are available in the link to the project accession in the European Nucleotide Archive.

I have a few minor comments for the authors:

- First, RNA-Seq on blood was done and the raw sequence data has been made available. However, the authors do not provide the RNA integrity number (RIN) of the sample that was sequenced. This is an important basic metric that should be reported to allow readers to judge whether the RNA sequence data is likely to be excellent, good, or poor.

- Second, the contrast and color scheme in Figure 5 showing the Hi-C contact map does not facilitate viewing of the chromosomes. The authors should consider revising this figure. Third, while I understand that Data Notes do not normally include any analyses or conclusions, it would be useful to provide some basic information about the karyotype of this species in the background section (to know if all chromosomes have been assembled). Similarly, a description of the range of assembled genome sizes in other snakes would help readers place this genome assembly in context.

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: Testing, a affiliation

Reviewer Expertise: Genomics, sex determination, sexual differentiation, genetics, development

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.