

Genome sequence of parvovirus from budgerigar (*Melopsittacus undulatus*)

Dominika Kadlečková,¹ Michal Vinkler,² Ruth Tachezy,¹ Martina Saláková¹

AUTHOR AFFILIATIONS See affiliation list on p. 3.

ABSTRACT Here, we report a parvovirus genome identified in *Melopsittacus undulatus*. The genome is 4,547 bp long and codes for two major open reading frames (ORFs): the non-structural replicase protein 1 (NS1) and the structural capsid gene (VP1). Phylogenetic analysis shows that this virus belongs to the genus *Chaphamaparvovirus*.

KEYWORDS parvovirus, budgerigar, *Melopsittacus undulatus*, NGS

Parvoviruses are known to infect the gastrointestinal tracts of domestic fowl as well as urban wild birds (1, 2). So far, there are only a few metagenomics/metatranscriptomic studies exploring the virome in birds (3). Here, we characterize a parvovirus genome found in a healthy female budgerigar (*Melopsittacus undulatus*) collected in September 2021 in Czechia: Přešimasy (Kolín).

Fresh feces were frozen at -80°C . The sample was homogenized in 500 μL of 1 \times phosphate-buffered saline (PBS) and processed according to NetoVir protocol (4). Briefly, the sample was centrifuged, filtered, and treated with nucleases, and nucleic acids enclosed in viral capsids were extracted using QIAamp Viral RNA Mini Kit (Qiagen, Hilden, Germany). Nucleic acids were amplified with a modified WTA2 protocol (Sigma-Aldrich, St. Louis, MO, USA); see NetoVIR protocol (4). The library was prepared with Nextera XT Library Preparation Kit (Illumina, San Diego, CA, USA). The sequencing was performed on the NextSeq 500 Illumina platform (2 \times 150 bp paired-end, Laboratory of Genomics and Bioinformatics, IMG, Czechia). The quality of the reads was checked using FastQC v0.11.9 (Babraham Bioinformatics, Babraham, UK). Trimming was done with Trimmomatic v0.32 (5) with settings to trim adapters from WTA2 and NexteraXT with ILLUMINACLIP, HEADCROP:19, LEADING:15, TRAILING:15, SLIDINGWINDOW:4:20, and MINLEN:50. Trimmed sequences were assembled using SPAdes v3.15.5 (6) with settings `--meta, -k 21, 33, 55, and 77`. Reads were mapped back to contigs with bwa-mem2 v2.2.1 (7), and coverage was extracted with CoverM v0.6.1 (<https://github.com/wwood/CoverM>). Contigs were compared to the nr database (NCBI) downloaded on 8 August 2021 using DIAMOND v2.0.15 (8) with a sensitive setting. Settings not specified were used with default parameters. Most viral reads (83% of all viral reads, 3% of all reads, 386,330 reads out of 13,709,189, and \sim 3% of all reads) were mapped to one contig identified as the parvovirus genome (Fig. 1B). The sequence was blasted with BLASTx (9) against the nr database and showed the highest similarity to avian chapparvovirus (54.04%, [QKX49056.1](https://ncbi.nlm.nih.gov/nuccore/QKX49056.1)). The length of the assembled contig was 4,547 bp, with a G-C content of 45.0%. The first major ORF predicted using Prodigal (10) encoded a non-structural protein 1 (NS1; 650 aa) that plays a role in the initiation of viral replication. The predicted NS1 protein shared 54.6% amino acid identity with NS1 of avian chapparvovirus ([QKX49056.1](https://ncbi.nlm.nih.gov/nuccore/QKX49056.1)). The second major ORF encodes a capsid protein with a length of 546 aa with the highest similarity (48.83%) to the capsid protein of the blue-and-yellow macaw (*Ara ararauna*) Chaphamaparvovirus ([QTE04011.1](https://ncbi.nlm.nih.gov/nuccore/QTE04011.1)). Two short ORFs with hypothetical proteins 1 (106 aa) and 2 (141 aa) were closest to Phoenicopteridae parvo-like hybrid

Editor Jelle Matthijnssens, Katholieke Universiteit Leuven, Leuven, Belgium

Address correspondence to Martina Saláková, martina.salakova@natur.cuni.cz.

The authors declare no conflict of interest.

See the funding table on p. 3.

Received 1 June 2023

Accepted 30 August 2023

Published 9 October 2023

Copyright © 2023 Kadlečková et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

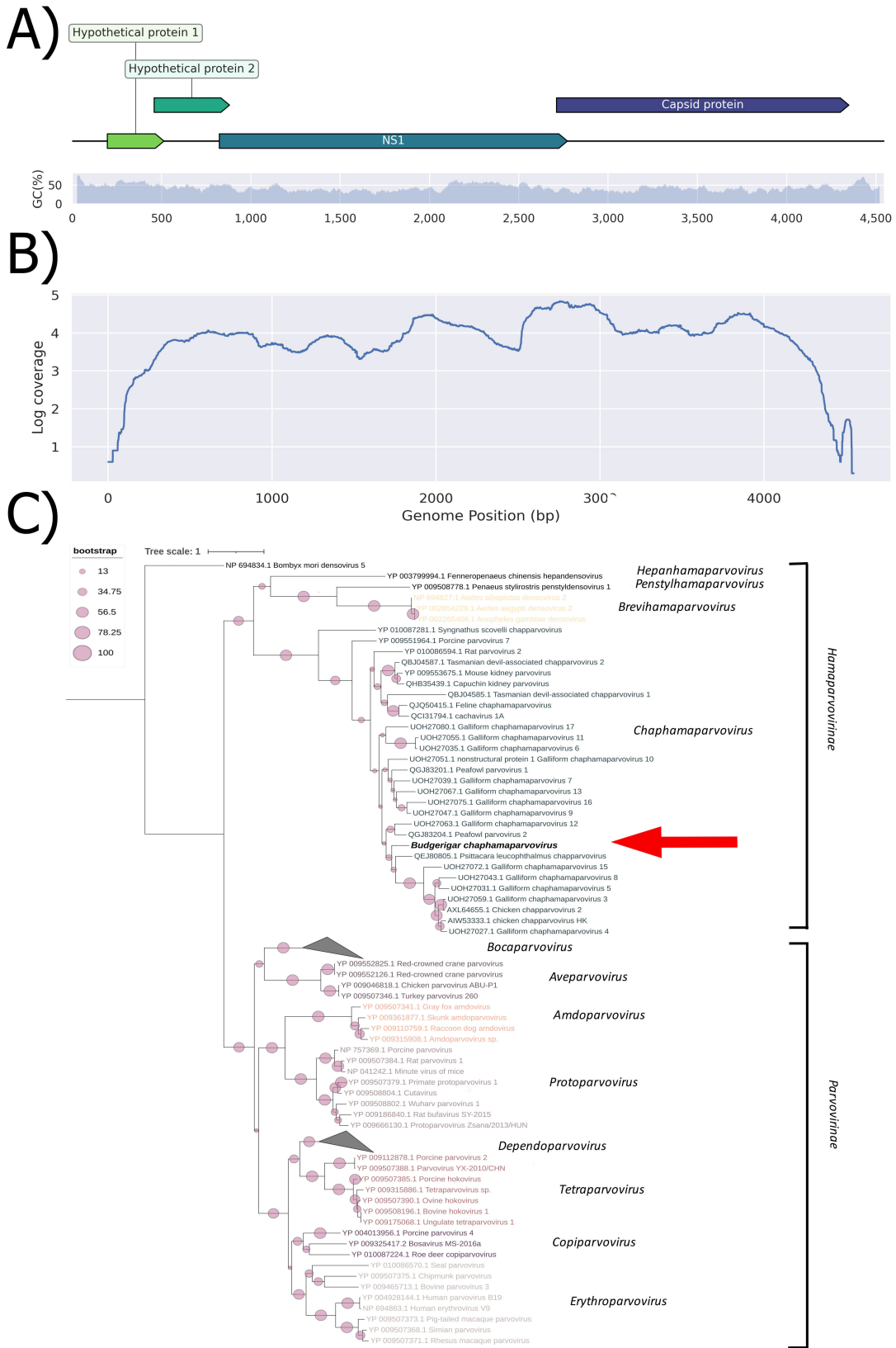


FIG 1 (A) Scheme of the genome of Budgerigar chaphamaparvovirus with GC content displayed below. (B) Depth coverage of the genome. (C) Phylogenetic tree, amino acid sequences of complete NS1 aligned with MAFFT v7.520 (11), trimmed with trimAl v1.4.1 (12), the best model was determined with ModelTest-NG v0.1.7 (13), and a tree with bootstraps (100) was created with PhyML v3.3.20220408 (14) and visualized in iTOL (15).

virus (QTE03742.1) and *Chufflevirus* sp. (QSH48278.1) with an identity of 49.64% and 62.92%, respectively. The genome organization is shown in Fig. 1A.

For the phylogenetic analysis (Fig. 1C), amino acid sequences of NS1 from several genera of *Parvovirinae* and *Hamaparvovirinae* were collected from NCBI. A representative of *Densovirinae* was used as an outgroup. The result shows that the virus, tentatively named Budgerigar chaphamaparvovirus, belongs to the subfamily *Hamaparvovirinae* and genus *Chaphamaparvovirus* (16).

ACKNOWLEDGMENTS

We are thankful to Balraj Melepat, Daniel Divín, Kateřina Černá, Lucie Schmiedová, and Tao Li for their help with parrot sampling.

This research was funded by the project Durable, ID Project No. 101102733, and by the Czech Science Foundation (Grant No. P502/19-20152Y).

AUTHOR AFFILIATIONS

¹Department of Genetics and Microbiology, Faculty of Science BIOCEV, Charles University, Vestec, Czech Republic

²Department of Zoology, Faculty of Science, Charles University, Prague, Czech Republic

AUTHOR ORCIDS

Martina Saláková  <http://orcid.org/0000-0003-0827-1211>

FUNDING

Funder	Grant(s)	Author(s)
EC European Health and Digital Executive Agency (HaDEA)	101102733	Ruth Tachezy
Czech Science Foundation	P502/19-20152Y	Michal Vinkler

AUTHOR CONTRIBUTIONS

Dominika Kadlečková, Formal analysis, Investigation, Software, Writing – original draft | Michal Vinkler, Conceptualization, Data curation, Resources, Writing – review and editing | Ruth Tachezy, Supervision, Writing – review and editing, Funding acquisition | Martina Saláková, Conceptualization, Data curation, Methodology, Project administration, Supervision, Writing – review and editing

DATA AVAILABILITY

Sequencing reads have been deposited into Sequence Read Archive (SRA) under the accession number [SRR21387183](#) (BioProject [PRJNA875655](#)). The genome of parvovirus was deposited into NCBI GenBank under accession number [OP359044](#).

REFERENCES

1. Reuter G, Boros Á, Delwart E, Pankovics P. 2014. Novel circular single-stranded DNA virus from turkey faeces. *Arch Virol* 159:2161–2164. <https://doi.org/10.1007/s00705-014-2025-3>
2. Chang W-S, Eden J-S, Hall J, Shi M, Rose K, Holmes EC. 2020. Metatranscriptomic analysis of virus diversity in urban wild birds with parvovirus disease. *J Virol* 94:e00606-20. <https://doi.org/10.1128/JVI.00606-20>
3. François S, Pybus OG. 2020. Towards an understanding of the avian virome. *J Gen Virol* 101:785–790. <https://doi.org/10.1099/jgv.0.001447>
4. Conceição-Neto N, Zeller M, Lefrère H, De Bruyn P, Beller L, Deboutte W, Yinda CK, Lavigne R, Maes P, Van Ranst M, Heylen E, Matthijssens J. 2015. Modular approach to customise sample preparation procedures for viral metagenomics: a reproducible protocol for virome analysis. *Sci Rep* 5:16532. <https://doi.org/10.1038/srep16532>
5. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
6. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. *Genome Res* 27:824–834. <https://doi.org/10.1101/gr.213959.116>
7. Vasimuddin M, Misra S, Li H, Aluru S. 2019. Efficient architecture-aware acceleration of BWA-MEM for Multicore systems, p 314–324. In 2019 IEEE International parallel and distributed processing symposium (IPDPS). Rio de Janeiro, Brazil. <https://doi.org/10.1109/IPDPS.2019.00041>
8. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 12:59–60. <https://doi.org/10.1038/nmeth.3176>

9. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. <https://doi.org/10.1186/1471-2105-10-421>
10. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. <https://doi.org/10.1186/1471-2105-11-119>
11. Katoh K, Kuma K, Toh H, Miyata T. 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* 33:511–518. <https://doi.org/10.1093/nar/gki198>
12. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. TrimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25:1972–1973. <https://doi.org/10.1093/bioinformatics/btp348>
13. Darriba D, Posada D, Kozlov AM, Stamatakis A, Morel B, Flouri T. 2020. Modeltest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. *Mol Biol Evol* 37:291–294. <https://doi.org/10.1093/molbev/msz189>
14. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 59:307–321. <https://doi.org/10.1093/sysbio/syq010>
15. Letunic I, Bork P. 2019. Interactive tree of life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* 47:W256–W259. <https://doi.org/10.1093/nar/gkz239>
16. Péntzes JJ, Söderlund-Venermo M, Canuti M, Eis-Hübinger AM, Hughes J, Cotmore SF, Harrach B. 2020. Reorganizing the family parvoviridae: a revised taxonomy independent of the canonical approach based on host association. *Arch Virol* 165:2133–2146. <https://doi.org/10.1007/s00705-020-04632-4>