



Novel Circo-Like Virus Detected in a Croatian Bat Population

Ivana Šimić,^a Tomaž Mark Zorec,^b Nina Krešić,^a  Mario Poljak,^b Tomislav Bedeković,^a Ivana Lojkić^a

^aDepartment of Virology, Croatian Veterinary Institute, Zagreb, Croatia

^bInstitute of Microbiology and Immunology, Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia

ABSTRACT Viral metagenomics analysis of samples from bats has been carried out as part of bat rabies surveillance in Croatia. Here, we report the complete genome sequence of a novel circo-like virus isolated from a sample of *Miniopterus schreibersii* bat guano determined by Illumina next-generation sequencing.

The family *Circoviridae* includes two genera, *Circovirus* and *Cyclovirus*, whose circular single-stranded DNA (ssDNA) molecules encode at least two major open reading frames (ORFs), the replication-associated (Rep) and the capsid protein. Members of the *Circoviridae* have been detected in numerous vertebrate and invertebrate organisms (1). Here, we report the complete genome sequence of a novel circo-like virus, Croatia 17_S17, isolated from guano of the *Miniopterus schreibersii* bat, determined by next-generation sequencing.

Nucleic acids were extracted using the iPrep virus kit (Invitrogen) according to the manufacturer's instructions. A sequencing library was constructed using the Nextera XT DNA library preparation kit and sequenced using a MiSeq reagent kit v2 in paired-end mode (2 × 250 nucleotides [nt]).

The complete genome sequence (length, 2,440 bp; GC content, 38%; average depth of coverage, 38.84×) of circo-like virus Croatia 17_S17 was identified among sequence contigs (assembled *de novo* using SPAdes v3.12.0 [2] using the complete raw read data set and default assembler parameters) using DIAMOND blastx (3) (version, GitHub commit c335eda162cee51dad3209e1c0e5241b493fb61) by mapping against the viral (taxid 10239) nonredundant protein data set, obtained from NCBI (11 June 2018), at the cutoff E value of 10⁻⁴ and NCBI blastx (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). The sequence indicated the presence of a *Circoviridae* Rep gene. The completeness and circularity of the sequence were determined by identification of matching 5' and 3' ends, and moreover, the sequence indicated the presence of a cyclovirus-like non-nucleotide motif (TAATACTAT) flanked by a pseudopalindromic sequence representing the stem-loop (4). The sequence residues at the 3' end (matching the 5' end) were manually trimmed, and the sequence was rotated to start with the stem-loop. Remapping to the novel genome sequence using the BWA v0.7.17-r1188 with default parameters and screening with Pilon v1.22 (5) (with the parameter “-fix all”) did not show evidence of misassembly. The average depth of coverage of the novel genome sequence was calculated on the basis of read remapping using SAMtools v1.3 and Awk v4.1.4 using the following command: “samtools depth -a <bam file> | awk {sum += \$3} END {print sum/NR}.”

The novel genome sequence contains four ORFs (identified using ORFfinder; parameters “-s 0 -ml 300”) encoding four putative proteins, including a putative Rep, and three hypothetical proteins (predicted by annotation transfer according to NCBI blastp similarity). The Rep protein was examined for the presence of conserved amino acid motifs, characteristic of circular ssDNA viruses, and showed the presence of all amino acid motifs, characteristic of *Circoviridae* replication-associated proteins. It has three conserved rolling circle replication (RCR) motifs in the N-terminal region, RCR motif I

Citation Šimić I, Zorec TM, Krešić N, Poljak M, Bedeković T, Lojkić I. 2019. Novel circo-like virus detected in a Croatian bat population. *Microbiol Resour Announc* 8:e00280-19. <https://doi.org/10.1128/MRA.00280-19>.

Editor John J. Dennehy, Queens College

Copyright © 2019 Šimić 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/).

Address correspondence to Ivana Lojkić, ilojki@veinst.hr.

I.Š. and T.M.Z. contributed equally to this article.

Received 13 March 2019

Accepted 25 March 2019

Published 18 April 2019

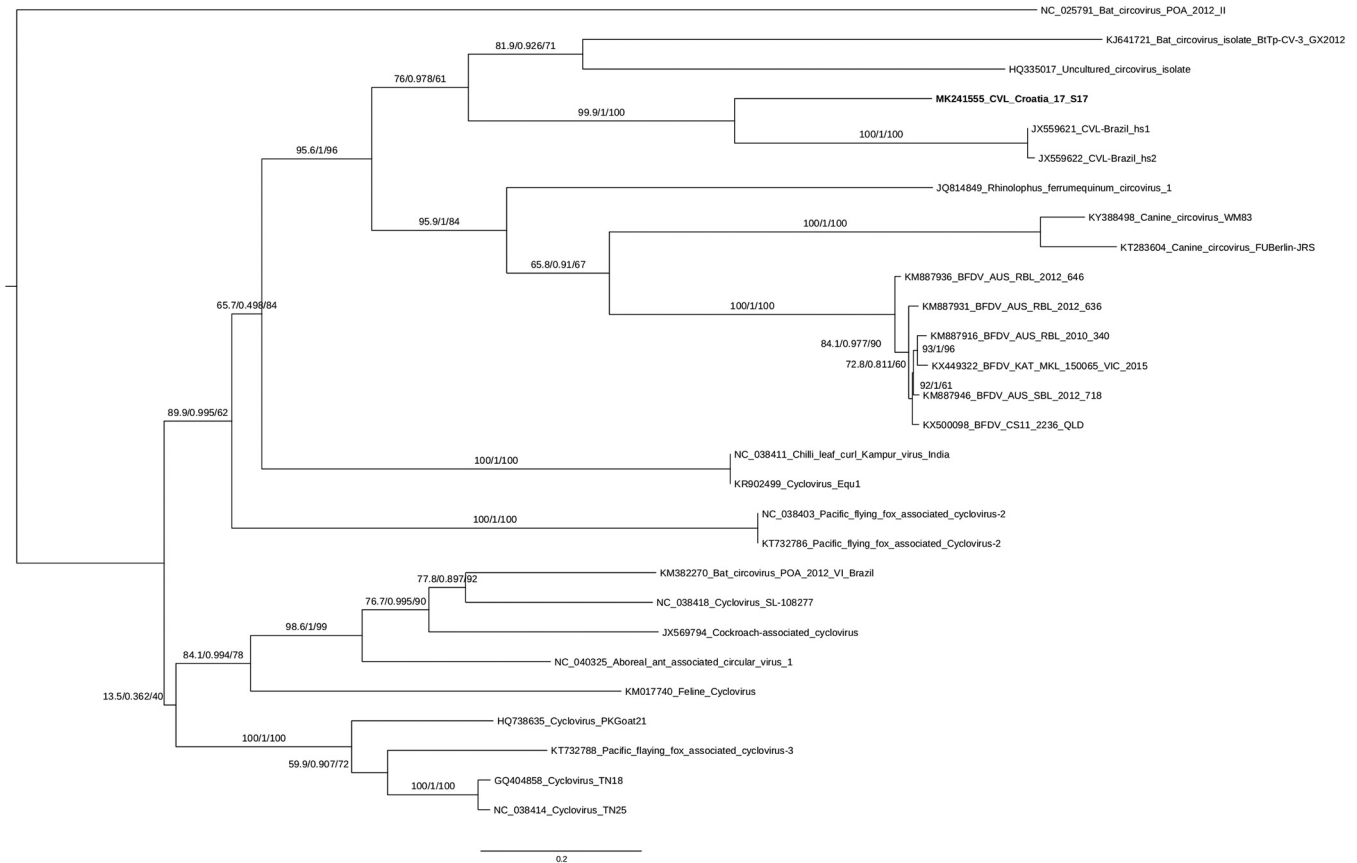


FIG 1 Phylogenetic tree of circo- and circo-like virus Rep sequences from diverse species and samples. The maximum likelihood phylogenetic tree was generated with the TIM3+F+R3 substitution model using IQ-TREE (parameters “-alrt 1000 -abayes -bb 1000”). The sequence identifiers include the NCBI accession number and isolate name. The novel Croatian isolate (Circo-like virus Croatia 17_S17) is indicated in bold. Branch supports are provided in the following format: SH-aLRT (%) / aBayes / ultrafast bootstrap (%).

(FTEFN), RCR motif II (VHVQG), and RCR motif III (YCKK), as well as superfamily 3 helicase Walker A, B, and C motifs (1). The RCR motif I does not unambiguously correspond to those found in circoviruses or cycloviruses but was identified as described by Castri-gnano et al. (6). The RCR motif II, VHVQG, is closer to the described nanovirus xHUQG motif (“U” represents bulky hydrophobic amino acids—I, L, V, M, F, Y, W). The RCR motif III, YCKK, is compatible with the Rep proteins of circovirus, cyclovirus, geminivirus, and nanovirus. Moreover, the helicase superfamily 3 Walker A, B, and C motifs could be identified in the C-terminal region (GPPGTGKS, IIDDF, ITSN) (1).

The virus reported here exhibits a type V circular ssDNA virus genome organization (1).

Phylogenetic analysis conducted using IQ-TREE (7–9) (parameters “-alrt 1000 -abayes -bb 1000”) (Fig. 1), based on the multiple sequence alignment of the Rep proteins (aligned using Muscle v3.8.31), demonstrated that the circo-like virus Croatia 17_S17 clusters close to circo-like virus sequences from Brazil, CLV-BR hs1 and CLV-BR hs2 (GenBank accession no. **JX559621** and **JX559622**), which were detected in human feces (6, 10).

Data availability. The complete genome sequence of circo-like virus Croatia 17_S17 has been deposited at GenBank under the accession no. **MK241555**. The raw reads were deposited in SRA under the accession no. **SRR8759221**.

ACKNOWLEDGMENTS

This research was supported by Croatian Science Foundation grant no. 8513 (BatsRabTrack) and Slovenian Research Agency grant no. P3-0083.

REFERENCES

1. Rosario K, Duffy S, Breitbart M. 2012. A field guide to eukaryotic circular single-stranded DNA viruses: insights gained from metagenomics. *Arch Virol* 157:1851–1871. <https://doi.org/10.1007/s00705-012-1391-y>.
2. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
3. 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>.
4. Li L, Kapoor A, Slikas B, Bamidele OS, Wang C, Shaukat S, Masroor MA, Wilson ML, Ndjango JB, Peeters M, Gross-Camp ND, Muller MN, Hahn BH, Wolfe ND, Triki H, Bartkus J, Zaidi SZ, Delwart E. 2010. Multiple diverse circoviruses infect farm animals and are commonly found in human and chimpanzee feces. *J Virol* 84:1674–1682. <https://doi.org/10.1128/JVI.02109-09>.
5. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:112963. <https://doi.org/10.1371/journal.pone.0112963>.
6. Castrignano SB, Nagasse-Sugahara TK, Kisielius JJ, Ueda-Ito M, Brandão PE, Curti SP. 2013. Two novel circo-like viruses detected in human feces: complete genome sequencing and electron microscopy analysis. *Virus Res* 178:364–373. <https://doi.org/10.1016/j.virusres.2013.09.018>.
7. Nguyen LT, Schmidt H, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32:268–274. <https://doi.org/10.1093/molbev/msu300>.
8. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jeremiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* 14:587–589. <https://doi.org/10.1038/nmeth.4285>.
9. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol* 35:518–522. <https://doi.org/10.1093/molbev/msx281>.
10. Castrignano SB, Nagasse-Sugahara TK, Garrafa P, Monezi TA, Barrella KM, Mehnert DU. 2017. Identification of circo-like virus-Brazil genomic sequences in raw sewage from the metropolitan area of São Paulo: evidence of circulation two and three years after the first detection. *Mem Inst Oswaldo Cruz* 112:175–181. <https://doi.org/10.1590/0074-02760160312>.