

Complete Genome Sequence of the First Canine Circovirus

Amit Kapoor,^a Edward J. Dubovi,^b Jose Angel Henriquez-Rivera,^a and W. Ian Lipkin^a

Center for Infection and Immunity, Columbia University, New York, New York, USA,^a and Animal Health Diagnostic Center, College of Veterinary Medicine at Cornell, Ithaca, New York, USA^b

We found a highly divergent circovirus in serum samples from several dogs. Phylogenetic analysis indicates that canine circovirus genotype 1 (CaCV-1) represents the first circovirus reported in dogs and is genetically most closely related to the only known mammalian circovirus, porcine circovirus. Here we report the complete genome sequence of the CaCV-1 strain NY214, which will help toward understanding the evolutionary and pathogenic characteristics of mammalian circoviruses.

Circoviruses are genetically diverse nonenveloped viruses with a small monomeric single-strand circular DNA genome and belong to the family *Circoviridae*, which comprises two genera, *Circovirus* and *Gyrovirus* (3). The genus *Circovirus* includes six recognized species that infect mammals and birds, namely, porcine circovirus 1 (8), porcine circovirus 2 (5), canary circovirus (9), goose circovirus (10), pigeon circovirus (10), and beak and feather disease virus (6). Several new circoviruses have been recently identified in animal feces and environmental samples; however the natural hosts of most of these viruses remain unidentified (2, 4, 7). Canine circovirus genotype 1 (CaCV-1) was found in serum samples from several dogs (6 of 205 animals tested) and thus represents the first nonporcine circovirus confirmed to infect mammals.

The complete genome of CaCV-1 (strain NY214) comprises 2,063 nucleotides (nt) as a covalently closed circular DNA with a GC content of 51.7%. All PCR products were sequenced using classical dideoxy Sanger sequencing (4× genome coverage) and assembled using Sequencher 4.7 (Gene Codes Corp.). TG and GG were the most abundant dinucleotides, with an observed/expected frequency ratio of >1.34. The genome contains two putative open reading frames (ORFs), on complementary strands in opposite orientation, that encode the viral replicase (303 amino acids [aa]) and capsid protein (270 aa). Similar to other animal circoviruses, CaCV has two intergenic noncoding regions that are 135 and 203 nt long. The noncoding region between the two major ORFs contains a thermodynamically stable stem-loop for initiation of rolling-circle replication and a characteristic 9-nt sequence, TAGTATTAC (1). In CaCV-1, the palindrome sequence at the origin of replication site is comprised of 12 nucleotide pairs (stem) and an open loop of 10 nt (CATAGTATTA). Similar to other animal circoviruses, the amino terminus of putative CaCV capsid protein contains a 30-aa-long arginine (R)-rich stretch. The capsid and replicase proteins of CaCV-1 share <25% and <50% identities, respectively, with the known animal circoviruses. According to the criteria employed by the International Committee on Taxonomy of Viruses (ICTV) (www.ictvdb.org), circoviruses of the same species should share >75% and >70% nucleotide identity in their complete genome and capsid protein sequences, respectively. Based on these criteria, we tentatively propose that CaCV-1 be classified as a prototype of new species in the genus *Circovirus* of the family *Circoviridae*.

Our intent in submitting this report is to enable others in the virology community to investigate the epidemiology, evolutionary biology, and pathobiology of mammalian circovirus infection. Moreover, the availability of CaCV-1 genome sequence will allow development of molecular reagents that can be used to identify more novel circoviruses that infect other mammalian species. We

note that the 203-nt-long intergenic noncoding region of CaCV-1 has 91% nucleotide identity over 150 nt of sequence with pine marten torque teno virus (11), providing the first direct evidence of an evolutionary relationship between two distinct virus families that includes genetically diverse viruses with single-stranded DNA (ssDNA) circular genomes, *Circoviridae* and *Anelloviridae*.

Nucleotide sequence accession number. The GenBank accession number of the CaCV-1 strain NY214 complete genome sequence is [JQ821392](https://www.ncbi.nlm.nih.gov/nuclot/JQ821392).

ACKNOWLEDGMENTS

We thank Jan L. Medina for excellent technical assistance.

This work is supported by National Institutes of Health grant awards AI090196, AI081132, AI079231, AI57158 (Northeast Biodefense Center-Lipkin), and AI070411 and the Defense Threat Reduction Agency.

REFERENCES

- Cheung AK. 2012. Porcine circovirus: transcription and DNA replication. *Virus Res.* **164**:46–53.
- Delwart E, Li L. 2012. Rapidly expanding genetic diversity and host range of the *Circoviridae* viral family and other Rep encoding small circular ssDNA genomes. *Virus Res.* **164**:114–121.
- King AMQ, Lefkowitz E, Adams MJ, Carstens EB (ed). 2011. *Virus taxonomy: ninth report of the International Committee on Taxonomy of Viruses*. Elsevier, Philadelphia, PA.
- Li L, et al. 2010. Multiple diverse circoviruses infect farm animals and are commonly found in human and chimpanzee feces. *J. Virol.* **84**:1674–1682.
- Meehan BM, et al. 1998. Characterization of novel circovirus DNAs associated with wasting syndromes in pigs. *J. Gen. Virol.* **79**:2171–2179.
- Ritchie BW, Niagro FD, Lukert PD, Steffens WL III, Latimer KS. 1989. Characterization of a new virus from cockatoos with psittacine beak and feather disease. *Virology* **171**:83–88.
- Rosario K, Duffy S, Breitbart M. 2009. Diverse circovirus-like genome architectures revealed by environmental metagenomics. *J. Gen. Virol.* **90**:2418–2424.
- Tischer I, Rasch R, Tochtermann G. 1974. Characterization of papovavirus- and picornavirus-like particles in permanent pig kidney cell lines. *Zentralbl. Bakteriol. Orig. A* **226**:153–167.
- Todd D, et al. 2001. Nucleotide sequence-based identification of a novel circovirus of canaries. *Avian Pathol.* **30**:321–325.
- Todd D, Weston JH, Soike D, Smyth JA. 2001. Genome sequence determinations and analyses of novel circoviruses from goose and pigeon. *Virology* **286**:354–362.
- van den Brand JM, et al. 2012. Metagenomic analysis of the viral flora of pine marten and European badger feces. *J. Virol.* **86**:2360–2365.

Received 29 March 2012 Accepted 2 April 2012

Address correspondence to Amit Kapoor, ak3117@columbia.edu.

Copyright © 2012, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JVI.00791-12