Arch Virol (2011) 156:1097–1100
DOI 10.1007/s00705-011-0971-6

ANNOTATED SEQUENCE RECORD

Discovery of a genome of a distant relative of chicken anemia virus reveals a new member of the genus *Gyrovirus*

Franciscus A. M. Rijsewijk · Helton F. dos Santos · Thais F. Teixeira · Samuel P. Cibulski · Ana P. M. Varela · Diogenes Dezen · Ana C. Franco · Paulo M. Roehe

Received: 5 December 2010 / Accepted: 5 March 2011 / Published online: 26 March 2011

Abstract A 2.4-kb phi29 polymerase amplification product from serum of a diseased chicken was cloned and sequenced. The 2383-nucleotide sequence showed about 40% identity to a representative genome of chicken anemia virus (CAV), the only member of the genus *Gyrovirus*, family *Circoviridae*. The new genome had an organization similar to that of CAV: a putative 5′ untranscribed region of about 400 nt followed by three partially overlapping open reading frames encoding VP1, VP2 and VP3 homologs. The amino acid identities between these homologs and those of CAV were 38.8%, 40.3%, and 32.2%, respectively. Based on these limited similarities, it is proposed that the newly identified virus is a member of a new species in the genus *Gyrovirus*. For this new species, the name Avian gyrovirus 2 (AGV2) is proposed.

The reported sequence has been submitted to GenBank and has accession number HM590588.

Electronic supplementary material The online version of this article (doi:10.1007/s00705-011-0971-6) contains supplementary material, which is available to authorized users.

F. A. M. Rijsewijk (✉) · H. F. dos Santos · A. C. Franco · P. M. Roehe
Virology Laboratory, Microbiology Department, Institute of Basic Health Sciences, Federal University of Rio Grande do Sul (UFRGS), Av. Sarmento Leite 500, Porto Alegre, Rio Grande do Sul (RS) CEP 90050-170, Brazil
e-mail: frans@rijsewijk.com

H. F. dos Santos · T. F. Teixeira · S. P. Cibulski · A. P. M. Varela · D. Dezen · P. M. Roehe
Institute for Veterinary Research "Desidério Finamor" (IPVDF), Estrada do Conde 6000, Eldorado do Sul, Rio Grande do Sul (RS) CEP 92990-000, Brazil

Chicken anemia virus (CAV) is the sole described member of the genus *Gyrovirus* of the family *Circoviridae* [18]. It was discovered in Japan in 1979, and it causes anemia and immunosuppression in newly hatched chickens [18, 21]. Based on results of genomic sequencing and expression of open reading frames (ORFs) of CAV, a 5′ untranscribed region (5′UTR) [16] and the coding regions for three viral proteins, VP1, VP2 and VP3, were identified [6].

In the middle of 2008, a number of three-week-old chickens in a poultry production unit in the south of Brazil were noticed to display apathy and loss of weight. In an attempt to identify circular DNA viruses that may have contributed to this condition, DNAs from sera of these chickens were subjected to multiply primed rolling-circle amplification (MPRCA) [1]. Here, we report the complete nucleotide sequence of one of the amplified products and show that this sequence is about 40% identical to the genome of CAV. It was inferred that it belongs to a new species in the genus *Gyrovirus*, for which the name Avian gyrovirus 2 (AGV2) is proposed.

On July 14, 2008, four sera were taken from diseased chickens ("Ave-1 to 4") of a poultry production unit in Rio Grande do Sul, Brazil. Isolation of DNA from chicken sera was performed by phenol extraction as described previously [2]. The DNA samples (100 ng) were subjected to MPRCA using Phi 29 DNA polymerase (New England Biolabs, Ipswich, MA, USA) [2]. The amplification products were digested with restriction enzymes, separated on a 0.7% agarose gel, and cloned in plasmid pUC18 (New England Biolabs) using standard cloning methods [17]. AGV2 clones were purified, and about 200 ng of each plasmid was used for nucleotide sequence analysis on a multi-channel automatic sequencing device (ACTGene256-Mega Bace, Uppsala, Sweden) using M13 primers (New England Biolabs) or specific AGV2 primers (primer- walking approach). Before

sequencing, a GC-rich region was PCR amplified in the presence of 50% 7-deaza-GTP and 50% dGTP (New England Biolabs). The sequences that were obtained were aligned using the sequence assembly program SeqMan^{tm}II from DNASTAR Inc., Madison, WI, USA, and analyzed using the Basic Local Alignment Search Tool (BLAST) of the National Center for Biotechnology Information (NCBI). Sequence comparisons and phylogenetic analysis were performed using the MegAlign 500 program of DNASTAR Inc.

The four obtained Phi 29 amplification products had a unit length of 2.4 kb and contained no *Eco*RI, one *Apa*I site, one *Hin*dIII site and two *Bam* HI sites. One of the amplification products, "Ave-3", was cloned, subcloned, and sequenced. The nucleotide sequence obtained was 2383 nt long and has been submitted to GenBank with accession number HM590588. Figure 1 shows the schematic representation of the genomic organization of AGV2 and Fig. 2 shows the position of AGV2 in relation to other members of the family *Circoviridae*, including a recently proposed genus named *Cyclovirus* [7].

The 5′ untranscribed region (5′ UTR) of AGV2 is by definition located between the polyadenylation site and the start of transcription. Two canonical polyadenylation signals (AATAAA) were found next to each other between nt 27 and 32 and between nt 39 and 44, where in the homologous position in CAV there is only one AATAAA sequence. The actual polyadenylation takes place some 10 to 30 nt downstream of this signal [3]. Using promoter prediction programs [10, 19], a putative TATA box (GATATAAG) was identified between nt 415 and 422, and a transcription start site was postulated at nt 440, which was located 10 nt upstream of the start of ORF1 (nt 450). The estimated size of the 5′ UTR was thus about 400 nt and was similar in size to the 5′ UTR of CAV [16]. This region contained six almost perfect direct repeat (DR) regions of 22 nt: 5′GTACAGGGGGGTACGTCAt/cCAt/g3′. Five or four (depending on the CAV strain) highly similar versions of these repeats are found in the 5′ UTR of CAV strains [11]. These repeats are thought to act as enhancers of transcription and may also play a role in viral replication [9].

Using the ORF Finder protocol of the NCBI [13], 14 open reading frames (ORFs) of more than 100 nt were found in the 2383 nt sequence: eight on the sense strand and six on the antisense strand. Only the three largest ORFs showed homology with proteins found in the NCBI database. All three were encoded on the sense strand, but in different reading frames: ORF1 in frame 3 from nt 450 to nt 1145, ORF2 in frame 1 from nt 577 to nt 951 and ORF3 in frame 2 from nt 953 to nt 2335 (Fig. 1). They encode homologs of CAV VP2, VP3 and VP1, respectively.

ORF1 encodes an amino acid sequence of 231 residues, which displayed 40.3% identity with VP2 of CAV. AGV2 VP2 contains a sequence (CX5R, residues [res.] 107-113—the large capitals indicate conserved residues) that is highly conserved in all protein phosphatases (PTPases), including VP2 of CAV. VP2 of CAV is a dual-specificity PTPase that plays a role in viral replication, cytopathology and virulence [15]. Interestingly, the phosphatase motif is part of a larger motif (WX7HX3CX CX5H) that is shared by CAV, torque teno virus (TTV) and TTV-like mini virus (TLMV) [20]. In AGV2, this motif is found in the sequence WLRQCARSHDEICTCGRWRSH (res. 95-115). Only the gyroviruses have the PTPase arginine (R) within this motif.

ORF2 encodes an amino acid sequence of 124 residues, which displayed 32.2% identity with VP3 of CAV. CAV VP3, or apoptin, was shown to induce apoptosis of hematopoietic cells, causing the observed anemia in infected chickens and to trigger apoptosis of tumor cells, but not of non-tumor cells [8, 12]. In the N-terminal half of CAV VP3, a short hydrophobic isoleucine-rich stretch is found

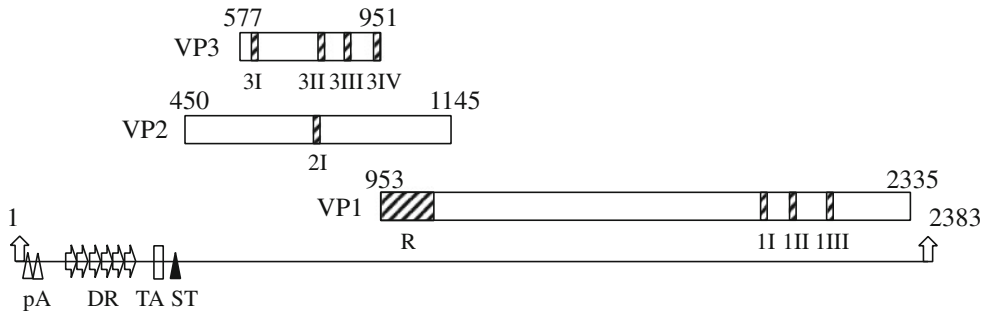

Fig. 1 Schematic representation of the genomic organization of AGV2. The horizontal bars in the upper part represent the sizes and positions of VP3, VP2 and VP1. The numbers above the bars designate their nucleotide positions. The hatched regions within the bars indicate the positions of conserved motifs. Within VP3: 3I = isoleucine stretch, 3II = NLS1, 3III = NES and 3IV = NLS2. Within VP2: 2I = phosphatase motif. Within VP1: R = arginine rich region, 1I, 1II, 1III = rep motifs I, II and III, respectively. The horizontal line in the lower part represents the 2383-nt AGV2 genome. The vertical arrows at the start and end of the genome indicate the positions of the putative stem-loop structures. The open triangles indicate two polyadenylation signals. The six horizontal arrows show the positions of the direct repeats found in the 5' UTR. The open box indicates the position of a putative TATA box, and the closed triangle, the putative transcription start site

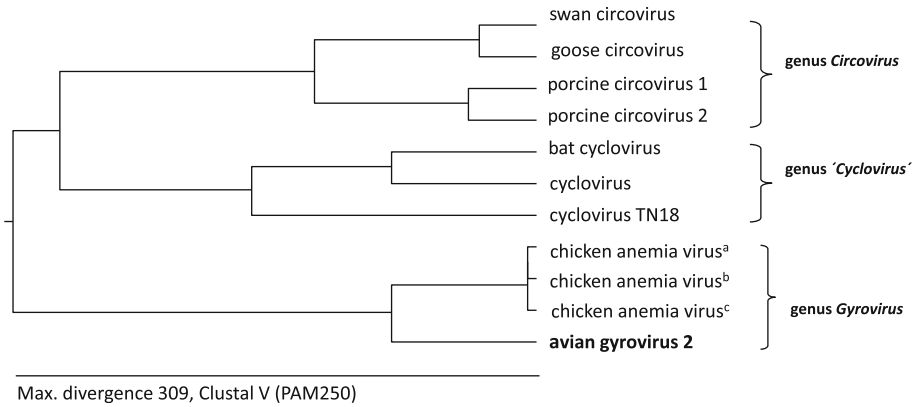

Fig. 2 Phylogenetic tree based on the alignment of amino acid sequences of the VP1 capsid proteins of eight different members of the family *Circoviridae* and the one described in this study. The avian gyrovirus described in this study is in bold. The VP1 sequences used have the following accession numbers: swan circovirus, ABU48446.1; goose circovirus, AAQ03996.1; PCV1, AAN77864.1; PCV2, ABD42929.1; bat cyclovirus, ADU77000.1; cyclovirus, ADU77014.1; cyclovirusTN18, ADD62480.1; chicken anemia virus[a], AAT85595; chicken anemia virus[b], ADJ18278; chicken anemia virus[c], AAT85598. The AGV2 sequence has NCBI accession number HM590588

(CAV res. 33-46) that is required for self-association as well as for binding cellular proteins [4]. This amino acid sequence is well conserved in AGV2 VP3 at res. 38-51. In the C-terminal half of CAV VP3, a bipartite nuclear localization sequence (NLS) was identified. The first part (NLS1) is located at res. 82-88, and the second part (NLS2) at res. 111-121. Between these two domains, a putative nuclear export sequence (NES) was found at res. 97-105. At homologous positions in VP3 of AGV2, NLS-like sequences were found: RRPRR (res. 84-88) and KKLRL (res. 120-124). A NES domain, however, was clearly conserved in AGV2 VP3.

ORF3 encodes an amino acid sequence of 460 residues, which displayed an identity of 38.8% with VP1 of CAV. Like other capsid proteins, it starts with a region that is rich in arginine and lysine residues (27 of the first 60 amino acids). Three conserved replication motifs have been identified in VP1 of CAV: FATLT at res. 313-317, QRWHTLV at res. 351-357, and YALK at res. 402 405 [14]. At the homologous positions, the following sequences were found in VP1 of AGV2: FAALS (res. 325-329), RRWLTLV (res. 363-369) and KAMA (res. 412-415). This comparison suggests that the three motifs are not well conserved and that the third motif is not conserved at all. Another candidate for the third replication motif is a YKFGT sequence that is present in VP1 of both CAV and AGV2, located only 7 residues before the YALK/KAMA sequences.

Immediately downstream of the stop codon of VP1 of AGV2, a 73-nt-long GC-rich sequence (joining nt 2324-2383 to nt 1-23) is found that can form two stem-loop structures: 5'-gggggggggggttttgccccccccacaaaccccccccccccggggggggatcctcccccccg gacccccg-3'. This sequence shows 60% identity to the corresponding GC-rich region of CAV [11].

This is the first time since the discovery of CAV that a genome of a member of a related species has been discovered. Although there are no proposed rules to demarcate species in the genus *Gyrovirus*, there are recognized criteria used for the genus *Circovirus*: (i) complete nucleotide sequence identity of less than 75% and (ii) capsid protein amino acid sequence identity of less than 70% [5]. The application of these criteria to the data described here (identities of 40% with the CAV genomic sequence and 38.8% with the CAV capsid protein) justify the description of this new virus as a member of a separate species in the genus *Gyrovirus*.

Acknowledgments We thank Felipe Bardini Rhodenbusch (MV) of the BIOVET Company, São Paulo, Brazil, who supplied sera of chickens. Helton F. dos Santos is supported by the National Council for Scientific and Technological Development (CNPq) grant no.141809/2010-8, and Paulo M. Roehe is a CNPq 1B research fellow. This work was supported by CNPq research grant no.578436/2008-6.

References

1. Dean FB, Nelson JR, Giesler TL, Lasken RS (2001) Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification. Genome Res 11:1095–1099
2. Dezen D, Rijsewijk FA, Teixeira TF, Holz CL, Cibulski SP, Franco AC, Dellagostin AO, Roehe PM (2010) Multiply-primed rolling-circle amplification (MPRCA) of PCV2 genomes: applications on detection, sequencing and virus isolation. Res Vet Sci 88:436–440
3. Edmonds M (2002) A history of poly A sequences: from formation to factors to function. Prog Nucl Acid Res Mol Biol 71:285–389
4. Heilman DW, Teodoro JG, Green MR (2006) Apoptin nucleocytoplasmic shuttling is required for cell type-specific

localization, apoptosis, and recruitment of the anaphase-promoting complex/cyclosome to PML bodies. J Virol 80:7535–7545
5. Index of Viruses—Circoviridae (2006) In: ICTVdB—the universal virus database, version 4. In: Büchen-Osmond C (ed) Columbia University, New York. http://www.ncbi.nlm.nih.gov/ICTVdb/Ictv/fs_index.htm
6. Koch G, Van Roozelaar DJ, Verschueren CAJ, Van Der Eb AJ, Noteborn MHM (1995) Immunogenic and protective properties of chicken anaemia virus proteins expressed by baculovirus. Vaccine 13:763–770
7. 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
8. Los M, Panigrahi S, Rashedi I, Mandal S, Stetefeld J, Essmann F, Schulze-Osthoff K (2009) Apoptin, a tumor-selective killer. Biochim Biophys Acta 1793:1335–1342
9. Miller MM, Jarosinski KW, Schat KA (2005) Positive and negative regulation of chicken anemia virus transcription. J Virol 79:2859–2868
10. Neural Network Promoter Prediction of the Berkeley Drosophila Genome Project. http://www.fruitfly.org/seq_tools/promoter.html
11. Noteborn MH, de Boer GF, van Roozelaar DJ, Karreman C, Kranenburg O, Vos JG, Jeurissen SH, Hoeben RC, Zantema A, Koch G, van Ormondt H, van der Eb AJ (1991) Characterization of cloned chicken anemia virus DNA that contains all elements for the infectious replication cycle. J Virol 65:3131–3139
12. Noteborn MH (2009) Proteins selectively killing tumor cells. Eur J Pharmacol 625:165–173
13. ORF Finder NCBI. http://www.ncbi.nlm.nih.gov/gorf/orfig.cgi
14. Okamoto H, Nishizawa T, Tawara A, Peng Y, Takahashi M, Kishimoto J, Tanaka T, Miyakawa Y, Mayumi M, Peng YH, Nishizawa T, Takahashi M, Ishikawa T, Yoshikawa A, Okamoto H (2002) Analysis of the entire genomes of thirteen TT virus variants classifiable into the fourth and fifth genetic groups, isolated from viremic infants. Arch Virol 147:21–41
15. Peters MA, Crabb BS, Washington EA, Browning GF (2006) Site-directed mutagenesis of the VP2 gene of Chicken anemia virus affects virus replication, cytopathology and host-cell MHC class I expression. J Gen Virol 87:823–831
16. Phenix KV, Meehan BM, Todd D, McNulty MS (1994) Transcriptional analysis and genome expression of chicken anaemia virus. J Gen Virol 75:905–909
17. Sambrook J, Russell DW (2001) Molecular cloning: a laboratory manual, vol 1, 3rd edn. Cold Spring Harbour Laboratory Press, New York
18. Schat KA (2009) Chicken anemia virus. Curr Top Microbiol Immunol 331:151–183
19. SoftBerry FPROM program. http://www.softberry.ru/berry.phtml?group=programs&subgroup=promoter&topic=fprom
20. Takahashi K, Iwasa Y, Hijikata M, Mishiro S (2000) Identification of a new human DNA virus (TTV-like mini virus, TLMV) intermediately related to TT virus and chicken anemia virus. Arch Virol 145:979–993
21. Yuasa N, Taniguchi T, Yoshida I (1979) Isolation and some characteristics of an agent inducing anemia in chicks. Avian Dis 23:366–385