Identification of a novel single-stranded, circular DNA virus from bovine stool

Hye Kwon Kim,1 Seong Jun Park,1 Van Giap Nguyen,1 Dae Sub Song,2 Hyoung Joon Moon,3 Bo Kyu Kang3 and Bong Kyun Park1

1Department of Veterinary Virology, College of Veterinary Medicine and BK21 Program for Veterinary Science, Seoul National University, Seoul 151-742, Republic of Korea
2Viral Infectious Disease Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon 305-806, Republic of Korea
3Research Unit, Green Cross Veterinary Products, Yong-in 449-903, Republic of Korea

We report the identification of a novel single-stranded, circular DNA virus isolated from bovine stool. The virus, named bovine stool-associated circular DNA virus (BoSCV), has a genome comprising 2600 bases of circular ssDNA, with two putative ORFs encoding replicase and capsid proteins, arranged inversely. The stem–loop structure was located between the 3′ ends of the two putative ORFs, as in chimpanzee stool-associated circular virus (ChimpSCV) and unlike other circular DNA viruses, including members of the families Circoviridae, Nanoviridae and Geminiviridae. BoSCV was also genetically similar to ChimpSCV, with approximately 30% identity in the replicase and capsid proteins. A phylogenetic analysis based on the replicase protein showed that BoSCV and ChimpSCV are in the same clade. A field survey using BoSCV-specific PCRs targeting ORF1 detected BoSCV and BoSCV-like sequences in bovine and porcine stool samples. BoSCV appears to belong to a new genus of circular DNA viruses.

Viruses with circular DNA genomes are found in humans, animals and plants. They usually consist of a small-sized genome (<10 kb) and capsid proteins without a lipid envelope. The Circoviridae are a family of single-stranded, circular DNA viruses infecting vertebrates, including the genera Circovirus, Gyrovirus and Cyclovirus (proposed) (Todd et al., 2001; Todd, 2005; Blinkova et al., 2010; Li et al., 2010). The families Geminiviridae and Nanoviridae comprise single-stranded, circular DNA viruses known to infect plants (Blinkova et al., 2010). Members of the family Anelloviridae, including torque teno virus, have also been found in humans and other vertebrates (Okamoto, 2009). Cyclovirus is a diverse group of viruses commonly found in chimpanzees, humans and farm animals; others have also been reported in dragonflies and bats (Li et al., 2010; Ge et al., 2011; Rosario et al., 2011) and a new circovirus was also found recently in fish (Lörincz et al., 2011). The potential for cross-species transmission of these broadly occurring viruses was suggested by Li et al. (2011).

Circoviruses and cycloviruses share several characteristics. They contain two major, inversely arranged ORFs encoding the putative replication-associated protein (Rep) and capsid protein (Cap) (Li et al., 2011). A stem–loop structure with nonamer sequences is located between the 5′ ends of the two ORFs (Todd, 2005; Li et al., 2011). However, cycloviruses do not have an intergenic region between the 3′ ends of the two ORFs, whereas circoviruses do (Li et al., 2011).

Blinkova et al. (2010) identified a novel circular DNA virus – neither a circovirus nor a cyclovirus – from chimpanzee stool. The new virus, named chimpanzee stool-associated circular virus (ChimpSCV), has a genome comprising approximately 2600 bases, is phylogenetically different from known animal circular DNA viruses and is related more closely to members of the family Nanoviridae, which contains plant viruses (Blinkova et al., 2010). Additionally, the stem–loop of ChimpSCV was located between the 3′ ends of the two ORFs, which further differentiated it from circoviruses and cycloviruses. Whether ChimpSCV is an animal or plant virus remains unconfirmed.

In this study, another novel circular DNA virus was found in bovine stool, and its genomic structure and phylogeny were analysed.

Five stool samples from highly febrile and anorexic calves were collected from a commercial farm in Korea in April 2011. The stool specimens were negative for bovine viral diarrhea virus, group A rotavirus (ProSpecT Rotavirus microplate assay kit; Oxoid), and bovine coronavirus by PCR-based detection methods (Letellier et al., 1999; Takiuchi...
et al., 2006). Therefore, the potential unknown viral agent was screened by using a random PCR approach.

For the detection of unknown viral sequences, a particle-associated nucleic acid PCR (PAN-PCR)-based approach was applied, based on the results of earlier studies (Stang et al., 2005; Kim et al., 2011). In this approach, faecal samples were suspended in Dulbecco’s modified Eagle’s medium to create a 10% suspension and centrifuged at 3000 r.p.m. (1750 g) for 20 min at 4 °C. After centrifugation, the supernatant was filtered through a 0.2 μm filter and refiltered through a second filter of the same pore size. Before DNA and RNA extraction, the semi-purified samples were treated with 10 U DNase I (Promega) and RNase A (Biosesang) for 30 min at 37 °C. The enzymic reaction was stopped with EDTA and the DNA and RNA were extracted. The extracted nucleic acids were subjected to PAN-PCR using K-random (5′-GACCATCTAGCGACCTCCACMNNM-3′) primers for cDNA and first-strand PCRs (T4 DNA polymerase), and K (5′-GACCATCTAGCGACCTCCAC-3′) primers for random amplification, as in previous studies (Stang et al., 2005; Kim et al., 2011).

The randomly amplified DNA fragments were cloned into TA vectors and subsequently transformed into competent Escherichia coli cells (DH5α), using a commercial TA cloning kit (Topcloner TA kit; Enzynomics). The selected recombinant plasmids were sequenced using M13F(−20) and M13R universal primers. The sequencing was performed by Macrogen Inc. (Seoul, Korea). The resultant sequences were aligned and processed to remove overlapping sequences using the CAP3 sequence-assembly program (Huang & Madan, 1999). The final sequences were analysed by a BLAST search (http://www.ncbi.nlm.nih.gov).

A total of eight and 13 plasmid inserts (two contigs and four singlets from RNA, and one contig and 10 singlets from DNA) were obtained from RNA and DNA, respectively, by PAN-PCR. When they were analysed in a BLAST search, four of the six RNA-originating sequences revealed similarity to Clostridium spp. sequences; the others were similar to Acetivibrio cellulolyticus and the TonB-dependent receptor family of bacteria. Three DNA-originating sequences were similar to bacteriophage sequences, and five sequences were similar to those from Bacteroides fragilis, Prevotella copri, Symbio bacterium thermophilum, Variorvax paradoxus and Bacillus subtilis. Two of the three remaining sequences were unknown, but one singlet was similar to the capsid protein of ChimpSCV (ABD24798.1, 37% identity by BLASTx).

For genomic sequencing of a suspected novel circular DNA virus, inverse PCR was performed using two primers: primer 1 (5′-AGAATCATCCGAGACCGCT-3′) and primer 2 (5′-GTGGGCTCTATGGCCTTTGT-3′). A Maxime PCR PreMix kit (iNtRON) was used for long PCR, according to the manufacturer’s directions. The PCR was as follows: 95 °C for 5 min; 40 cycles of 95 °C for 30 s, 50 °C for 30 s, 72 °C for 3 min; and 72 °C for 7 min. The amplicon (>2000 bp) was purified using a Qiagen gel extraction kit and submitted to Macrogen for cloning and sequencing. The sequence was obtained using the M13F(−20) and M13R universal primers and primer-walking methods.

The novel circular DNA virus was named bovine stool-associated circular DNA virus (BoSCV, strain CP11-49-3, GenBank accession no. JN634851) and its genome contained 2600 bases (Fig. 1). The possible ORFs were predicted using the GLIMMER program (Delcher et al., 1999); the stem–loop structure with a nonamer sequence was predicted using the Mfold webserver (Zuker, 2003). The predicted ORF region and stem–loop were drawn using the pDRAW32 program (http://www.acaclone.com). GLIMMER predicted two major ORFs (containing >200 bp). ORF1 was 768 bases long, starting with ATG (encoding 255 aa); ORF2 was 1035 bases long, starting with GTG (encoding 344 aa). When ORF1 and ORF2 were analysed

---

**Fig. 1.** Genomic structure of bovine stool-associated circular DNA virus (BoSCV). Two major, inversely arranged ORFs encoding putative replication-associated protein (Rep) and capsid protein (Cap) are indicated on the genome. The stem–loop structure is indicated by the box on the genome; detailed nucleotide sequences are also presented.
using BLASTX, ORF1 was found to be similar to the putative replicase protein of ChimpSCVs (25–32 % amino acid identity) and ORF2 was similar to the putative capsid protein of ChimpSCVs (27–33 % amino acid identity).

The stem–loop structure with high G+C content was predicted to be located in the downstream intergenic region. The stem–loop overlapped with the 3′ end of the ORF1 coding region. The sequence of the stem–loop was CAGTATTACCTTGA, containing a nonamer sequence similar to those of both circoviruses and members of the family Nanoviridae (Table 1).

To observe the phylogenetic relationship between BoSCV and other circular DNA viruses, full amino acid sequences from replicase-associated ORFs were compared. Amino acid and other circular DNA viruses, full amino acid sequences were aligned using the COBALT tool (Papadopoulos & Agarwala, 2007). The phylogeny of the aligned amino acid sequences was analysed using MEGA 4.0 (Tamura et al., 2007). The resultant dendrogram was drawn by the neighbour-joining (NJ) method (using the p-distance model and other default parameters) with 1000 bootstrap replications.

A new cluster of BoSCV and ChimpSCV was observed in the NJ tree (Fig. 2a). The new cluster was related more closely to circoviruses and cycloviruses than to viruses in the families Nanoviridae and Geminiviridae. Based on the aligned sequences, including gaps, using COBALT, BoSCV showed 16.9–19.7 % amino acid identity to ChimpSCVs. The BoSCV also had 7.1–12.3, 10.2–11.2, 8.1–10.6 and 6.9–7.0 % identity to circoviruses, cycloviruses and members of the Nanoviridae and Geminiviridae, respectively.

To investigate the presence of this newly identified virus in other field samples, a new primer set was designed: BoSCV rep1 (5′-CTGAGCATCAGTTGGGATGG-3′, nt 2006–1987) and BoSCV rep2 (5′-TCTGGTGGAGTTCCGCAG-3′, nt 1696–1715). For PCRs, i-Star master PCR premix (iNtRON) was used, following the manufacturer’s instructions. PCR was performed as follows: 95 °C for 5 min; 35 cycles of 95 °C for 30 s, 53 °C for 30 s, 72 °C for 30 s; and 72 °C for 7 min. When the target band (311 bp) was amplified, the selected amplicon was purified and sequenced using both primers. PCR was performed on 25 bovine stool samples from four farms and on another 22 porcine stool samples from four additional farms.

In addition, six of the complete feed products from the bovine and porcine farms were investigated. With the feed products, samples were homogenized to a 40 % suspension in PBS containing 3 log10(TCID50) ml−1 of porcine circovirus 2 (PCV2, GenBank accession no. FJ905466) and centrifuged at 1750 g for 20 min. DNA was subsequently extracted from the supernatant. To validate that DNA extraction from feed was adequate and not inhibited by unknown feed ingredients, PCV2-specific PCR was used to detect PCV2 DNA, as a positive control (Yang et al., 2003).

The resultant sequences from the stool samples were aligned using CLUSTAL_X v. 1.83 (Thompson et al., 1997). The aligned data were analysed and converted to an electronic file for phylogenetic analysis using BioEdit (Hall, 1999). The UPGMA tree was constructed using MEGA 4.0 (Tamura et al., 2007), using a p-distance model for distance estimation; percentage frequencies of the groupings were determined after 1000 bootstrap evaluations.

In the case of bovine stool samples, two of the four tested farms (including the farm where BoSCV was first detected) were positive for BoSCV (the 311 bp target band was detected). Among the samples from the BoSCV-positive farms, approximately 60 % of the faecal samples were positive for BoSCV. Only one porcine farm produced samples that tested positive for BoSCV by PCR. Three of the seven porcine faecal samples from that farm showed positive bands. However, the target size was slightly larger than 311 bp, suggesting a sequence different from that of BoSCV. Positive bands were not found in any of the complete feed products; positive control (PCV2-specific) bands were detected.

The three partial ORF1 sequences from the bovine samples from one farm were highly similar to BoSCV, but had only 74.4–87.7 % nucleotide identity to the BoSCV isolated from a different farm. The three partial sequences from the porcine samples matched each other exactly and had 39 % identity to BoSCV. When these partial nucleotide sequences were compared phylogenetically, BoSCV and partial replicate sequences from the field bovine and porcine stools were related to ChimpSCV; all differed from circoviruses and members of the family Nanoviridae (Fig. 2b).

Although a novel circular DNA virus was discovered and named BoSCV, it is not clear whether BoSCV was an infectious agent that caused high fever and anorexia in cattle. The partial nucleotide sequences for Clostridium spp. were also detected by PAN-PCR from RNA extracts. As Clostridium spp. may be pathogenic (Bagge et al., 2010), it is possible that an infecting Clostridium spp. was responsible

Table 1. Comparison of nonamer sequences from BoSCV and other circular DNA viruses

<table>
<thead>
<tr>
<th>Virus</th>
<th>9-mer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>BoSCV</td>
<td>CAGTATTAC</td>
</tr>
<tr>
<td>ChimpSCV*</td>
<td>AATAGTTAC</td>
</tr>
<tr>
<td>PCV2/circovirus*</td>
<td>AATAGTTAC</td>
</tr>
<tr>
<td>Chimp1/circovirus*</td>
<td>TAAATTAC</td>
</tr>
<tr>
<td>FBNYV/Nanoviridae*</td>
<td>GAGTATTAC</td>
</tr>
<tr>
<td>ToCSV/Geminiviridae*</td>
<td>TAAATTAC</td>
</tr>
</tbody>
</table>

*Nonamer sequences from previous papers (Blinkova et al., 2010; Li et al., 2010).
for the high fever and anorexia. Therefore, the aetiology of BoSCV ‘infections’ should be elucidated by further studies.

BoSCV was found to be similar to ChimpSCV in many aspects (Blinkova et al., 2010). Not only the genome sizes, but also the putative ORF1 and ORF2 of BoSCV were similar to the replicase (25–32 %) and capsid (27–33 %) proteins of ChimpSCV, respectively, as determined by BLASTX. BoSCV also had a stem-loop structure between the 3′ ends of each ORF, a feature of ChimpSCV that differentiates it from other circular DNA viruses (Blinkova et al., 2010). Although the nonamer sequence of BoSCV was more similar to that of circoviruses or members of the family Nanoviridae, rather than ChimpSCV, BoSCV was shown to be closely related to ChimpSCV by several other common features found in these viruses. In addition, as BoSCV-specific PCR amplicons almost disappeared from the ssDNA-specific nuclease-treated DNA compared with the non-treated DNA, BoSCV may have ssDNA as its genome (Supplementary Fig. S1, available in JGV Online).

The COBALT tool has been reported to be the preferred tool for aligning amino acid sequences (Papadopoulos & Agarwala, 2007); therefore, the replicate protein sequences were aligned using this tool prior to phylogenetic analyses. The phylogenetic analysis based on the replicate protein sequences showed that BoSCV and ChimpSCV aligned in a new clade that was related to, but different from, circoviruses and cycloviruses. This result also showed that BoSCV is closely related phylogenetically to ChimpSCV.

The results of this study did not determine conclusively whether BoSCV infected bovine species or originated from ingested plant viruses. A plant virus, pepper mild mottle virus, has been frequently found in human stool specimens (Zhang et al., 2006; Colson et al., 2010). If BoSCV originated from ingested plant viruses, feed is a plausible source for the virus contamination, because most farms utilize commercial feed for their animals. However, in this study, no evidence of BoSCV or BoSCV-like viruses was identified in the feeds that might be used on the bovine and porcine farms investigated. As a limited number of feed samples were tested, further investigation is needed to eliminate this possibility completely. Furthermore, the possibility that these viruses may be

---

Fig. 2. (a) Phylogenetic analysis of BoSCV and other single-stranded circular DNA viruses, based on full replicase amino acid sequences. (b) Phylogenetic relationships among BoSCV and BoSCV-like partial ORF1 sequences obtained from bovine and porcine stool specimens, respectively. B12, B15 and B19 were from bovine stool samples; P17, P18 and P20 were from porcine stool samples. ChimpSCV, PCV2 and the Nanoviridae BBTV sequence were used for reference and are presented with their GenBank accession numbers.
released from some bovine intestinal parasite should be investigated.

The diversity of cycloviruses in different species suggests that the viruses could infect animals rather than originate from an ingested source (Li et al., 2010). In the present field study, about 65% of the bovine faecal samples from two of the four bovine farms were PCR-positive, and approximately 43% of porcine samples from one of the four porcine farms investigated were PCR-positive. When positive samples were sequenced and analysed genetically, the partial ORF1 sequences from the bovine and porcine samples showed 74.4–87.7 and 39% similarity to the BoSCV ORF1 sequence, respectively. Furthermore, the viral sequences were related phylogenetically to BoSCV and ChimpsCV. Conclusive determination of whether BoSCV is a plant virus and whether it is capable of infecting cattle will require additional studies, including virus isolation and experimental infection.

In conclusion, a novel circular DNA virus (BoSCV) was identified from bovine stool and shown to be similar to ChimpsCV in genome organization, as well as by phylogenetic analysis. BoSCV-like sequences were also detected in porcine stool in a preliminary investigation. Therefore, BoSCV is proposed as a representative of a new genus of circular DNA viruses. Further investigation is warranted to determine the nature of its association with cattle and swine.

Acknowledgements

This study was supported by a grant from the Biogreen 21 Program, Rural Development Administration, Korea (# 20110301-061-544-001-02-00).

References


