

# Identification and molecular characterization of a single-stranded circular DNA virus with similarities to *Sclerotinia sclerotiorum* hypovirulence-associated DNA virus 1

Zhenguo Du · Yafei Tang · Songbai Zhang ·  
Xiaoman She · Guobing Lan · Arvind Varsani ·  
Zifu He

Received: 29 July 2013 / Accepted: 8 October 2013 / Published online: 7 December 2013  
© Springer-Verlag Wien 2013

**Abstract** A putative circular single-stranded DNA (ssDNA) virus was recovered from *Hypericum japonicum* collected in Vietnam. The viral isolate was tentatively named *Hypericum japonicum*-associated circular DNA virus (HJasCV). HJasCV shares 58.7–65.4% nucleotide sequence identity with *Sclerotinia sclerotiorum* hypovirulence-associated DNA virus 1 (SsHADV-1) and SsHADV-1-like viruses. Like this group of viruses, the genome of HJasCV (2 200 nt) has two large ORFs, one in the virion-sense and the other in the complementary-sense DNA. The proteins encoded in the virion-sense and complementary-sense ORFs share 39–46 % and 45–67 % amino acid

sequence identity with the putative capsid and replication-associated proteins (Reps), respectively, of SsHADV-1 and SsHADV-1-like viruses. The putative Rep of HJasCV contains all of the motifs related to rolling-circle replication. Its 111-bp intergenic region (IR) contains a hairpin structure with a geminivirus-like nonanucleotide sequence, TAATGTTAT, at the apex of the loop. Phylogenetic analysis revealed that HJasCV forms a monophyletic clade with SsHADV-1 and SsHADV-1-like viruses.

Recent studies have unveiled unexpected ubiquity and diversity of single-stranded DNA (ssDNA) viruses. Many ssDNA viruses with novel genome architectures and biological characteristics have been discovered in the last few years [3, 19]. Recently, a novel circular ssDNA mycovirus, *Sclerotinia sclerotiorum* hypovirulence-associated DNA virus 1 (SsHADV-1), was discovered and characterized [26]. SsHADV-1 has a genome organization similar to that of circoviruses and geminiviruses, i.e. it is bidirectionally transcribed, and its replication protein (Rep) is closely related to that of geminiviruses [26]. In less than three years, six viruses (or virus-like ssDNA molecules) with similarities to SsHADV-1 have been recovered from cassava leaf samples (cassava-associated circular DNA virus; CasCV), mosquitoes (mosquito VEM SDBVL-G virus; MvemV), dragonflies (dragonfly-associated circular viruses-1, 2, and 3; DfasCV-1, -2 and -3) and European badger faecal matter (Meles meles fecal virus; MmFV) [2, 16, 18, 23]. Recently, Rosario et al. [18] suggested that these viruses are similar and should belong to a new genus, which they have tentatively named *Gemycocircularvirus* (Gemini-myc-like circular DNA virus). Interestingly, these viruses also share some similarities to three geminivirus-like ssDNA viruses recently recovered from sewage

---

Z. Du · Y. Tang · X. She · G. Lan · Z. He (✉)  
Plant Protection Research Institute, Guangdong Academy of  
Agricultural Sciences, Guangzhou 510640, China  
e-mail: hezf@gdppri.com

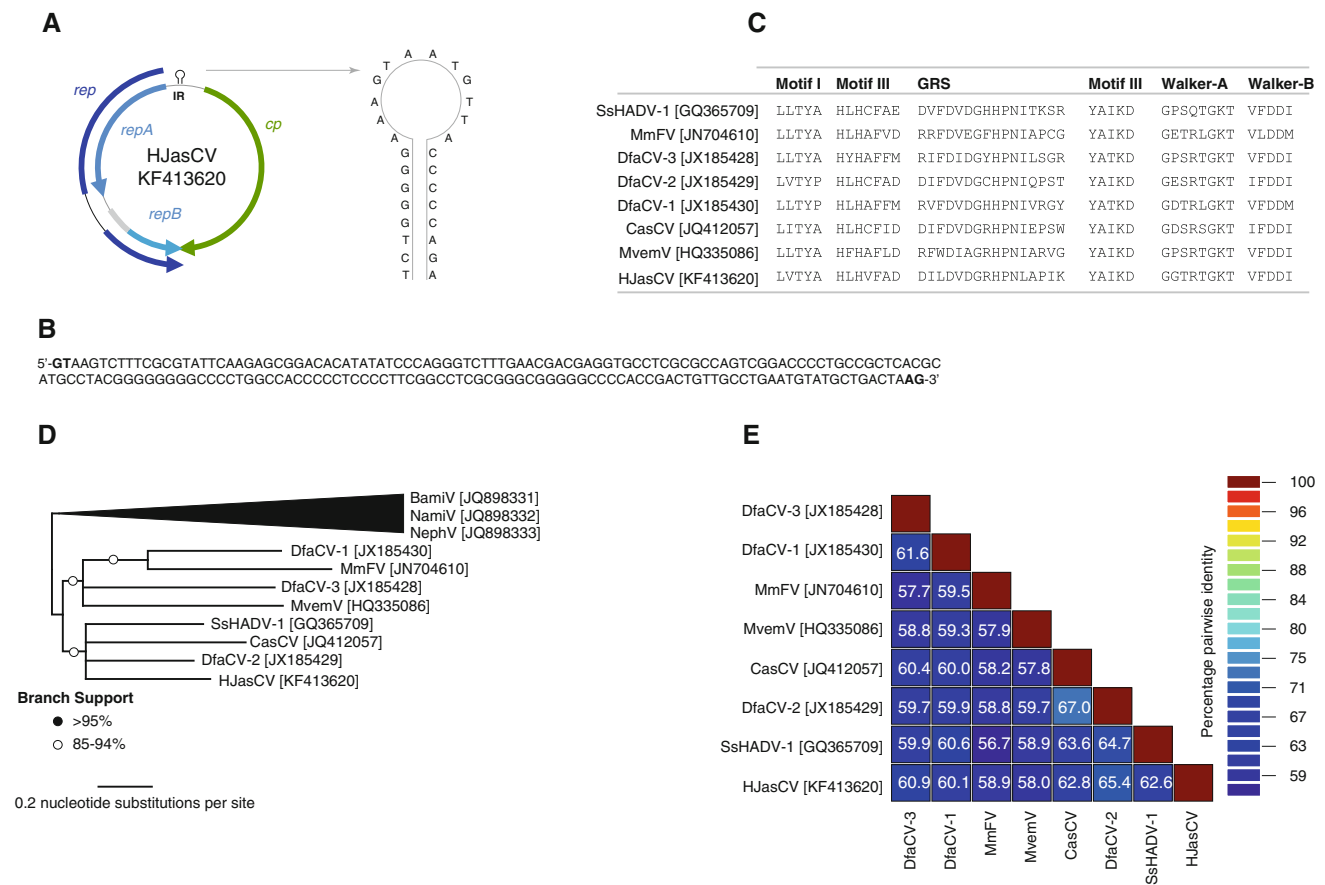
Z. Du · Z. He  
Guangdong Provincial Key Laboratory of High Technology for  
Plant Protection, Guangzhou 510640, China

S. Zhang (✉)  
College of Agriculture, Yangtze University,  
Jingzhou 434025, Hubei, China  
e-mail: yangtze2008@126.com.cn

A. Varsani  
Electron Microscope Unit, Division of Medical Biochemistry,  
Department of Clinical Laboratory Sciences, University of Cape  
Town, Rondebosch, Cape Town 7701, South Africa

A. Varsani  
School of Biological Sciences, University of Canterbury,  
Private Bag 4800, Christchurch 8140, New Zealand

A. Varsani  
Biomolecular Interaction Centre, University of Canterbury,  
Private Bag 4800, Christchurch 8140, New Zealand



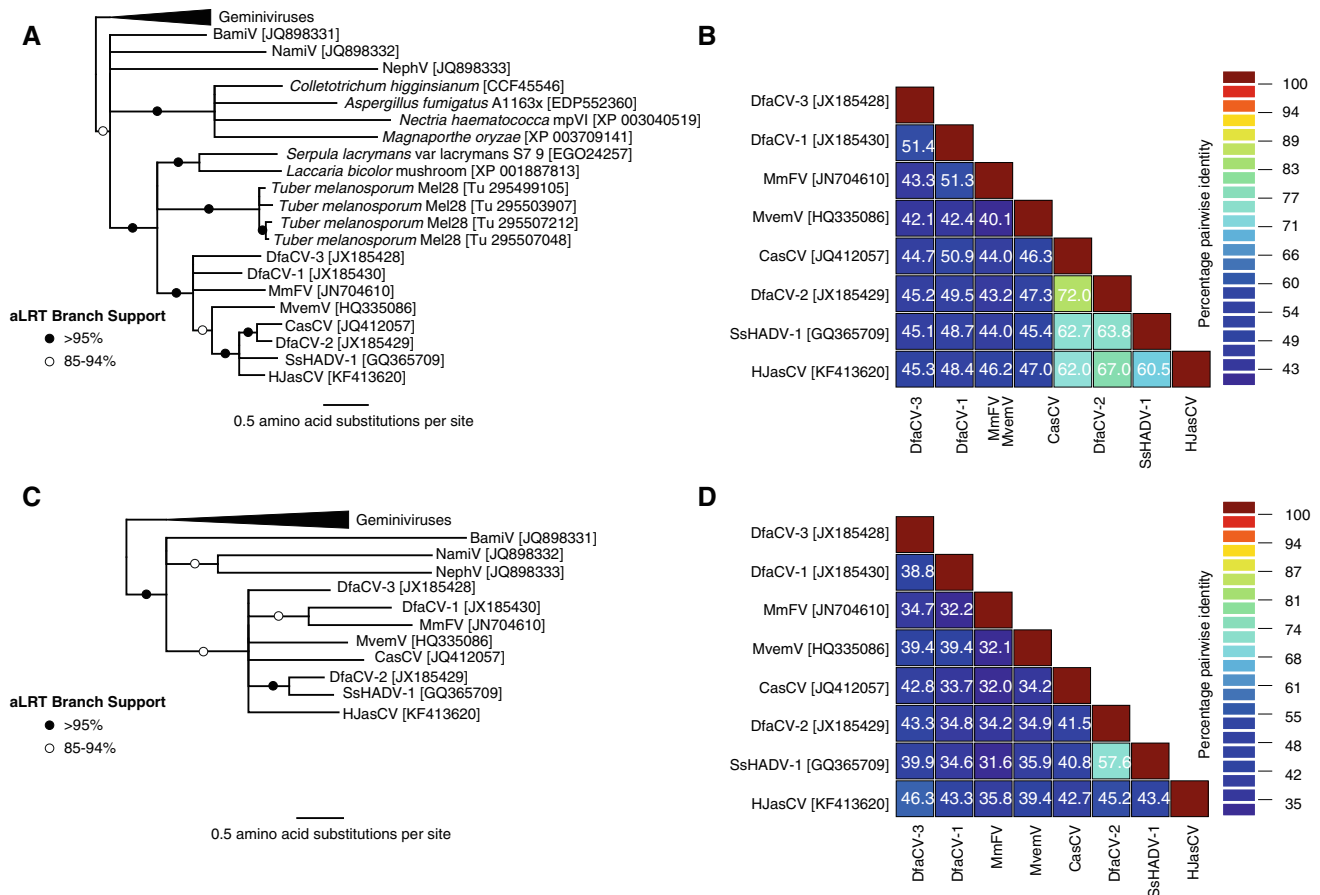
**Fig. 1** **a** Genome organisation of *Hypericum japonicum*-associated circular DNA virus (HJasCV) and the putative stem-loop structure identified in the long intergenic region. **b** Sequence of putative intron identified within the ORF encoding for the putative replication protein. **c** Conserved motifs identified in the Rep of SsHADV-1 and

SsHADV-1-like viruses. **d** Maximum-likelihood phylogenetic tree of SsHADV-1 and SsHADV-1-like viruses rooted with BamiV, NamiV and NephV. **e** Percentage pairwise identities (nucleotide) between SsHADV-1 and SsHADV-1-like viruses

samples, namely baminivirus (BamiV), nimirivirus (Nimiv) and nephavirus (NephV) [17]. The biological properties of SsHADV-1 like viruses have yet to be determined. Nonetheless, SsHADV-1 infects and confers hypovirulence to the plant pathogenic fungus *S. sclerotiorum* [26]. The fact that this virus can be transmitted extracellularly raises the possibility that it has the potential to be used for virus-mediated attenuation of fungal pathogenesis [27]. Furthermore, the recent discovery of SsHADV-1 in river bank and benthic sediments in New Zealand suggested that this virus may be more common in environmental samples than previously thought [10]. In this paper, we report the identification and molecular characterization of a novel ssDNA virus that is similar to SsHADV-1 and may be a new member of the proposed genus *Gemycocircularvirus*.

*Hypericum japonicum* plants exhibiting yellow mosaic symptoms were collected from Phu My, Binh Dinh Province, Vietnam (14° 11' 11" N; 109° 3' 53" E), on February 4, 2013. Total DNA was extracted from the sample using a CTAB method as described by Harrison et al. [9].

Rolling-circle amplification (RCA) was conducted using a TempliPhi™ kit (GE Healthcare, USA) to recover the genome of a suspected begomovirus [8]. Digestion of RCA products with the restriction enzyme *PstI* yielded a ~2.2-kb DNA fragment, which was cloned and sequenced by primer walking (Invitrogen Co., Shanghai, China). Unexpectedly, BLASTx [1] analysis of the assembled sequence contigs revealed similarities (E value <  $3 \times 10^{-88}$ ; 77% coverage) to the sequence of DfasCV-2, an SsHADV-1-like virus. Two sets of primers with opposite orientations (forward: 5'- nt2030-AGTCCTCCATCCTCGTGATG-nt2049 3', 5'- nt1822-ATCAGCAGACGTCCCATTTC-nt1841 3'; reverse: 5'- nt1822-GAAATGGGACGTCTGCTGAT-nt1841 -3', 5'- nt1645-GAATTGAGTTTGTTCGGGA-nt1664-3') were designed to amplify the genome of the virus. The resulting sequences were identical to the one recovered by *PstI* digestion of the RCA product. We propose the name *Hypericum japonicum*-associated circular DNA virus for the putative viral isolate (HJasCV; GenBank accession no. KF413620).



**Fig. 2** **a** Maximum-likelihood phylogenetic tree of the Rep amino acid sequences of SsHADV-1 and SsHADV-1-like viruses, fungal introns and other closely related viruses. **b** Two-dimensional pairwise identity plot comparisons of Rep amino acid sequences of SsHADV-1 and SsHADV-1-like viruses. **c** Maximum-likelihood phylogenetic

tree of the CP amino acid sequences of SsHADV-1 and SsHADV-1-like viruses, BamiV, NamiV, NephV. **d** Two-dimensional pairwise identity plot comparisons of CP amino acid sequences of SsHADV-1 and SsHADV-1-like viruses

The putative ORFs of HJasCV were detected using ORF Finder (<http://www.ncbi.nlm.nih.gov/projects/gorf/>). Like SsHADV-1, HJasCV employs an ambisense coding strategy and has two large ORFs, one in the virion-sense and the other in the complementary-sense strand (Fig. 1a). The proteins encoded by the two ORFs share significant amino acid sequence identity with putative capsid and Rep proteins of SsHADV-1 and SsHADV-1-like viruses, respectively. The identification of a putative Rep has been critical in the discovery of novel ssDNA viruses [3, 19].

All known SsHADV-1-like viruses have been found to contain an intron in the Rep coding region (in the complementary-sense DNA), similar to that found in mastreviruses [25]. An intron was identified in the Rep coding region of HJasCV through sequence alignment and manual inspection (Fig. 1b). When the intron was removed, HJasCV was found to encode a Rep protein that is 336 aa long. It contains the rolling-circle replication (RCR)-related motifs (Fig. 1c), including RCR motif I (amino acid

residues 12-16), which is implicated in recognition and binding of the *ori*-associated iterative sequences in geminiviruses [11], RCR motif II (50-56), which is believed to be involved in divalent metal coordination by binding  $Mg^{2+}$  or  $Mn^{2+}$  ions required for the cleavage reaction in circoviruses [20], nanoviruses [7, 13, 22, 24], and geminiviruses [12] through invariant histidine residues, and RCR motif III (91-95), which contains a tyrosine residue that is essential for cleavage and covalently attaches to the 5' end of the cleaved DNA, and an invariant lysine that may be essential for binding and positioning (Fig. 1c) [11, 21, 22]. These three RCR motifs have been found in geminiviruses, nanoviruses, and circoviruses at structurally equivalent positions in the Rep N-terminus [2, 19]. It is noteworthy that whereas the motif II sequences found in circoviruses and nanoviruses contain one histidine residue, those found in geminiviruses and SsHADV-1 or SsHADV-1-like viruses contain two. Additionally, SsHADV-1 and SsHADV-1-like viruses have two other motifs commonly found in the

Rep proteins of geminiviruses: an NTP-binding domain (Walker A, 210–217 and Walker B, 248–252) and a GRS (geminivirus Rep sequence, 67–82) domain. The NTP-binding domain, characterized by a phosphate-binding fold (P-loop), has been proposed to confer ATPase and helicase activities to geminiviral Rep proteins. The GRS domain, which is located between motifs II and III, has been shown to be required for the initiation of geminiviral replication [4, 15].

Within the 111-bp intergenic region, we identified a stem-loop structure with the nonanucleotide sequence TAATGTTATC positioned in the loop (Fig. 1a). A conserved nonanucleotide sequence has been found in all SsHADV-1 and SsHADV-1-like viruses except Meles meles fecal virus (MmFV) [23]. Similar nonanucleotide sequences are also present in other ssDNA viruses and have been reported to be essential in replication initiation of these viruses [19]. It is noteworthy that the nonanucleotide sequence of HJasCV is different from those of SsHADV-1 and SsHADV-1-like viruses by one nucleotide (the A in the fifth position is substituted by a G). The effects of the difference on the molecular biology of HJasCV are unknown. However, it has been shown that changes in the nonanucleotide motif may not inhibit the cleavage reaction between positions 7 and 8 for geminiviruses, although the reaction efficiency may be affected [12].

A maximum-likelihood (ML) phylogenetic tree was constructed using a MUSCLE [5]-aligned dataset (including complete genome sequences of all SsHADV-1 and SsHADV-1-like viruses, BamiV, NamiV and NephV) using PHYML 3.0 [6] with GTR+I+G4 nucleotide substitution model (Fig. 1). This, coupled with the nucleotide pairwise identities calculated using SDT v1.0 [14], clearly reveals that HJasCV is most closely related to CasCV, SsHADV-1 and DfaCV-2, sharing ~62–68% nucleotide sequence identity (Fig. 1d, e). Nonetheless, it is clearly evident that SsHADV-1 and SsHADV-1-like viruses share >57% pairwise genome-wide identity.

We also constructed ML phylogenetic trees using the aligned Rep and CP amino acid sequences encoded by SsHADV-1 and SsHADV-1-like viruses and related viruses (in the case of the Rep protein, we included related fungal intergenon sequences). The amino acid sequence ML trees were inferred using PHYML with the LG substitution model (Fig. 2). The analyses of the Rep and CP of HJasCV with SsHADV-1 and SsHADV-1 like viruses also indicated that HJasCV is most closely related, sharing >60% amino acid sequence identity in Rep and >40% in CP, to CasCV, SsHADV-1 and DfaCV-2 (Fig. 2).

Altogether, our analysis reveals that HJasCV is a novel SsHADV-1-like virus. Given that the only known host of SsHADV-1 and SsHADV-1-like viruses to date is a fungus [26], and considering the close pathogenic or symbiotic

associations between fungi and plants, it is highly likely that HJasCV infects fungi associating with *H. japonicum*. The identification of the fungi associating with *H. japonicum* and an infectious clone of HJasCV would be required to demonstrate the host and biology of HJasCV.

**Acknowledgments** This work was supported by the International Science & Technology Collaboration Program of China (No. 2011DFB30040), the Science & Technology Program of Guangdong, China (No. 2011B050400003, 2012A020200015), and Foundation of President of GAAS of China (No. 201207).

## References

- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402
- Dayaram A, Opong A, Jaschke A, Hadfield J, Baschiera M, Dobson RC, Offei SK, Shepherd DN, Martin DP, Varsani A (2012) Molecular characterisation of a novel cassava associated circular ssDNA virus. *Virus Res* 166:130–135
- 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
- Desbiez C, David C, Mettouchi A, Laufs J, Gronenborn B (1995) Rep protein of tomato yellow leaf curl geminivirus has an ATPase activity required for viral DNA replication. *Proc Natl Acad Sci USA* 92:5640–5644
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32: 1792–1797
- Guindon S, Dufayard JF, 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
- Hafner GJ, Stafford MR, Wolter LC, Harding RM, Dale JL (1997) Nicking and joining activity of banana bunchy top virus replication protein in vitro. *J Gen Virol* 78(Pt 7):1795–1799
- Haible D, Kober S, Jeske H (2006) Rolling circle amplification revolutionizes diagnosis and genomics of geminiviruses. *J Virol Methods* 135:9–16
- Harrison BD, Liu YL, Khalid S, Hameed S, OtimNape GW, Robinson DJ (1997) Detection and relationships of cotton leaf curl virus and allied whitefly-transmitted geminiviruses occurring in Pakistan. *Annu Appl Biol* 130:61–75
- Kraberger S, Stainton D, Dayaram A, Zawar-Reza P, Gomez C, Harding JS, Varsani A (2013) Discovery of Sclerotinia sclerotiorum hypovirulence-associated virus-1 in Urban River sediments of Heathcote and Styx Rivers in Christchurch city, New Zealand. *Genome Announcements* 1:e00513–e00559
- Laufs J, Schumacher S, Geisler N, Jupin I, Gronenborn B (1995) Identification of the nicking tyrosine of geminivirus Rep protein. *FEBS Lett* 377:258–262
- Laufs J, Traut W, Heyraud F, Matzeit V, Rogers SG, Schell J, Gronenborn B (1995) In vitro cleavage and joining at the viral origin of replication by the replication initiator protein of tomato yellow leaf curl virus. *Proc Natl Acad Sci USA* 92:3879–3883
- Merits A, Fedorkin ON, Guo D, Kalinina NO, Morozov SY (2000) Activities associated with the putative replication initiation protein of coconut foliar decay virus, a tentative member of the genus Nanovirus. *J Gen Virol* 81:3099–3106

14. Muhire B, Martin DP, Brown JK, Navas-Castillo J, Moriones E, Zerbinì FM, Rivera-Bustamante R, Malathi VG, Briddon RW, Varsani A (2013) A genome-wide pairwise-identity-based proposal for the classification of viruses in the genus Mastrevirus (family Geminiviridae). *Arch Virol* 158:1411–1424
15. Nash TE, Dallas MB, Reyes MI, Buhrman GK, Ascencio-Ibanez JT, Hanley-Bowdoin L (2011) Functional analysis of a novel motif conserved across geminivirus Rep proteins. *J Virol* 85: 1182–1192
16. Ng TF, Willner DL, Lim YW, Schmieder R, Chau B, Nilsson C, Anthony S, Ruan Y, Rohwer F, Breitbart M (2011) Broad surveys of DNA viral diversity obtained through viral metagenomics of mosquitoes. *PLoS One* 6:e20579
17. Ng TFF, Marine R, Wang C, Simmonds P, Kapusinszky B, Boddhidatta L, Oderinde BS, Wommack KE, Delwart E (2012) High variety of known and new RNA and DNA viruses of diverse origins in untreated sewage. *J Virol* 86 (22):12161–12175
18. Rosario K, Dayaram A, Marinov M, Ware J, Kraberger S, Stanton D, Breitbart M, Varsani A (2012) Diverse circular ssDNA viruses discovered in dragonflies (Odonata: Epiprocta). *J Gen Virol* 93:2668–2681
19. 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
20. Steinfeldt T, Finsterbusch T, Mankertz A (2006) Demonstration of nicking/joining activity at the origin of DNA replication associated with the Rep and Rep' proteins of porcine circovirus type 1. *J Virol* 80:6225–6234
21. Steinfeldt T, Finsterbusch T, Mankertz A (2007) Functional analysis of cis- and trans-acting replication factors of porcine circovirus type 1. *J Virol* 81:5696–5704
22. Timchenko T, de Kouchkovsky F, Katul L, David C, Vetten HJ, Gronenborn B (1999) A single rep protein initiates replication of multiple genome components of faba bean necrotic yellows virus, a single-stranded DNA virus of plants. *J Virol* 73:10173–10182
23. van den Brand JM, van Leeuwen M, Schapendonk CM, Simon JH, Haagmans BL, Osterhaus AD, Smits SL (2012) Metagenomic analysis of the viral flora of pine marten and European badger feces. *J Virol* 86:2360–2365
24. Vega-Rocha S, Gronenborn B, Gronenborn AM, Campos-Olivas R (2007) Solution structure of the endonuclease domain from the master replication initiator protein of the nanovirus faba bean necrotic yellows virus and comparison with the corresponding geminivirus and circovirus structures. *Biochemistry* 46:6201–6212
25. Wright EA, Heckel T, Groenendijk J, Davies JW, Boulton MI (1997) Splicing features in maize streak virus virion- and complementary-sense gene expression. *Plant J Cell Mol Biol* 12:1285–1297
26. Yu X, Li B, Fu Y, Jiang D, Ghabrial SA, Li G, Peng Y, Xie J, Cheng J, Huang J, Yi X (2010) A geminivirus-related DNA mycovirus that confers hypovirulence to a plant pathogenic fungus. *Proc Natl Acad Sci USA* 107:8387–8392
27. Yu X, Li B, Fu Y, Xie J, Cheng J, Ghabrial SA, Li G, Yi X, Jiang D (2013) Extracellular transmission of a DNA mycovirus and its use as a natural fungicide. *Proc Natl Acad Sci USA* 110: 1452–1457