A novel flavivirus in the soybean cyst nematode

Sadie Bekal,1 Leslie L. Domier,2,3 Biruk Gonfa,2 Nancy K. McCoppin,2,3 Kris N. Lambert2 and Kaustubh Bhalerao1

1Department of Agricultural and Biological Engineering, University of Illinois, Urbana, IL 61810, USA
2Department of Crop Sciences, University of Illinois, Urbana, IL 61810, USA
3US Department of Agriculture, Agricultural Research Service, University of Illinois, Urbana, IL 61801, USA

Heterodera glycines, the soybean cyst nematode (SCN), is a subterranean root pathogen that causes the most damaging disease of soybean in the USA. A novel nematode virus genome, soybean cyst nematode virus 5 (SbCNV-5), was identified in RNA sequencing data from SCN eggs and second-stage juveniles. The SbCNV-5 RNA-dependent RNA polymerase and RNA helicase domains had homology to pestiviruses in the family Flaviviridae, suggesting that SbCNV-5 is a positive-polarity ssRNA virus. SbCNV-5 RNA was present in all nematode developmental stages, indicating a transovarial mode of transmission, but is also potentially sexually transmitted via the male. SbCNV-5 was common in SCN laboratory cultures and in nematode populations isolated from the field. Transmission electron microscopy of sections from a female SCN showed virus particles budding from the endoplasmic reticulum and in endosomes. The size of the viral genome was 19 191 nt, which makes it much larger than other known pestiviruses. Additionally, the presence of a methyltransferase in the SbCNV-5 genome is atypical for a pestivirus. When cDNA sequences were mapped to the genome of SbCNV-5, a disproportionate number aligned to the 3’ NTR, suggesting that SbCNV-5 produces a subgenomic RNA, which was confirmed by RNA blot analysis. As subgenomic RNAs and methyltransferases do not occur in pestiviruses, we conclude that SbCNV-5 is a new flavivirus infecting SCNs.

INTRODUCTION

The soybean cyst nematode (SCN), Heterodera glycines, is a microscopic sedentary obligate plant parasite that feeds on soybean roots. It is considered the most destructive pathogen of soybean in the USA (Wrather & Koenning, 2006). As SCNs cause severe plant damage, a great deal of attention has been paid to identifying natural pathogens of this nematode. Whilst bacterial and fungal pathogens of SCNs have been identified and studied (Kerry, 1987), little is known about viral pathogens of SCNs or any other nematode. Plant-parasitic nematodes from the genera Longidorus, Trichodorus and Xiphinema have been shown to vector viruses to plants; however, these viruses do not infect the nematode but simply adhere to the oesophagus and are sloughed off each time the nematode moults (Hooper, 1974).

The first viral disease to be reported in a nematode was by Loewenberg et al. (1959), who demonstrated that a root-knot nematode disorder was caused by a filterable agent. However, the agent in their study was never visualized or further characterized. Whilst some later studies examined nematode viruses associated with a ‘swarming’ behaviour observed in a number of nematode species (McBride & Hollis, 1966; Ibrahim & Hollis, 1973; Ibrahim et al., 1973), transmission electron microscopy (TEM) of putatively virus-infected nematodes were inconclusive (Mankau, 1981). Other observations in TEM studies of nematodes also suggested that nematodes could be infected by viruses; however, because of the lack of follow-up research, the uncertainty of these assertions persisted (Foor, 1972; Zuckerman et al., 1973).

Poinar and colleagues published the first comprehensive analysis of TEM images of viruses infecting insect parasitic nematodes (Poinar & Hess, 1977; Poinar et al., 1980), but no molecular or biochemical research was conducted on the nematode viruses. It was only in 2011 when the genomes of nematode viruses were first sequenced that molecular details about the types of viruses that replicate in Caenorhabditis elegans (Félix et al., 2011) and SCNs (Bekal et al., 2011) were clarified and it became possible to study nematode viruses in more detail. To date, all reported nematode viruses have had negative-sense ssRNA genomes, with SCNs infected by viruses in the families Bunyaviridae and Rhabdoviridae (Bekal et al., 2011) and in the recently proposed family Nyamiviridae (Kuhn

The GenBank/EMBL/DDBJ accession number for the sequence of SbCNV-5 determined in this study is KF726084.

One supplementary figure is available with the online version of this paper.
et al., 2013), whilst C. elegans has been shown to be infected by viruses in the family Nodaviridae (Félix et al., 2011).

Although little is known about nematode viruses, a great deal is known about invertebrate viruses, especially arthropod-borne viruses (arboviruses), as these insect-vectored viruses cause significant human and animal diseases, such as yellow fever, dengue and West Nile fever (Alatoom & Payne, 2009). Not all invertebrate viruses are vectored to other animals, but many are arthropod specific (Hobson-Peters et al., 2011). Like nematode viruses, insect-specific viruses are found in many viral taxa, such as mosquito-only flaviviruses (Hoshino et al., 2007), sigmaviruses (Longdon et al., 2010), bunyaviruses (Marklewitz et al., 2011), alphaviruses (Nasar et al., 2012), nidoviruses (Zirkel et al., 2011), reoviruses (Noda et al., 1991) and negeviruses (Vasilakis et al., 2013). In addition to taxonomic diversity, nematode viruses share the tendency to be transmitted vertically. The use of molecular biology techniques has improved virus detection, particularly in nematodes (Bekal et al., 2011). This study reports the discovery of a new SCN nematode virus that is related to viruses in the family Flaviviridae, making this, to our knowledge, the first positive-polarity ssRNA virus identified in a nematode.

RESULTS

Identification of a new virus genome in the H. glycines transcriptome

The transcriptome of H. glycines from eggs and second-stage juveniles was assembled into contigs. The cDNA contigs were compared with known proteins in the National Center for Biotechnology Information (NCBI) databases using BLASTX. A 6565 bp contig was similar to the polyprotein from classical swine fever virus (E-value = 1.29 × 10⁻⁴). Further investigation of the list of BLAST hits revealed other contigs with homology to the closely related bovine viral diarrhea virus 2 or border disease virus. All these viruses are in the family Flaviviridae, which have positive-polarity ssRNA genomes. Typically, viruses in this family have genome sizes between 9.6 and 12.3 kb and have spherical enveloped virions of 40-60 nm in diameter (Meyers et al., 1989; Simmonds et al., 2012). Months after the initial sequence similarity analysis of the nematode viral genome was conducted, a new BLAST search revealed a strong similarity (E-value = 1.2 × 10⁻⁵) to the newly discovered Gentian Kobu-sho-associated virus (GKaV), which has a 23 kb RNA genome. Like the nematode virus, the GKaV is thought to be related to pestiviruses (Kobayashi et al., 2013).

To estimate the size of the nematode virus genome, contig ends were used to find other overlapping contigs until a single 19 191 nt contig was formed. The genome assembly was confirmed by mapping Illumina paired-end reads to the putative viral genome (Fig. 1a). The continuous and overlapping paired-end sequences indicated that the contig was correctly assembled. The fact that the 19 191 nt sequence encoded a single 6004 aa polyprotein also suggested that the genome assembly was correct. This new SCN virus was named SbCNV-5 as it is the fifth virus identified from SCNs. SbCNV-5 contained a 48-base 5’ non-coding region (NCR) but also had a second AUG at position 139. In contrast to the short 5’ NCR, the 3’ NCR was 1131 nt. The mean depth of Illumina sequence reads mapped to the viral genome was 282-fold; however, when the number of mapped reads was plotted over the length of the SbCNV-5 genome, the depth of coverage varied from 20- to 880-fold throughout most of the virus genome but increased dramatically in the 3’ NCR to 3218-fold coverage (Fig. 1a). A second mapping of SCN cDNA sequences, derived from a different RNA sample but using the 454 sequencing platform was conducted. This mapping showed the same read distribution as the Illumina sequence (Fig. 1a), indicating that the pattern of sequences aligned to the SbCNV-5 genome is reproducible. As the increase in sequence depth in the 3’ NCR suggested that a subgenomic RNA is produced by SbCNV-5, an RNA gel blot was conducted on RNA extracted from SCNs harbouring SbCNV-5. The blot was hybridized with a probe produced from the 3’ NCR, which detected subgenomic RNAs of 1.2 and 0.75 kb (Fig. 1b). The two subgenomic RNAs shown in the RNA gel blot are consistent in size with the two peaks in the 3’ NCR fold coverage graph shown in Fig. 1(a).

The typical genome organization for a virus in the family Flaviviridae is a single long ORF encoding a polyprotein, flanked by NCRs. In the pestiviruses, the 5’ NCR contains an internal ribosome entry sequence, but flaviviruses also have conserved secondary structure in their 5’ and 3’ NCRs (Brinton & Dispoto, 1988). Similarly, secondary structure predictions indicate that the 3’ NCR for SbCNV-5 contains multiple stable hairpins (Fig. S1, available in the online Supplementary Material). In addition to the secondary structure predictions, the 5’ and 3’ NCRs were analysed for the presence of sequence repeats. The 3’ NCR contained three long repeats that may be part of predicted stem–loop structures (Table 1).

In members of the family Flaviviridae, the N-terminal half of the polyprotein encodes the structural proteins: a small basic capsid protein and two or three envelope proteins. The C-terminal half encodes non-structural proteins: a serine protease, an RNA helicase, a methyltransferase (in flaviviruses) and a RNA-dependent RNA polymerase (RdRp) (Simmonds et al., 2012). SbCNV-5 had no sequence homology to structural proteins in the N-terminal region of the polyprotein but had significant similarity to trypsin-like serine proteases, RNA helicase, S-adenosyl-L-methionine-dependent RNA methyltransferase and an RdRp (Table 2). The trypsin-like serine protease is important for the proteolytic processing of the polyprotein, whilst the RNA helicase is critical for virus replication, as is the positive-polarity ssRNA-type RdRp. The methyltransferase probably plays a role in capping the viral RNA.

In order to gain a greater understanding of the relationship of SbCNV-5 to known viruses, we conducted a phylogenetic
analysis of the conserved region of the RdRp from SbCNV-5 and representative viruses in the family Flaviviridae. The phylogenetic analysis indicated that SbCNV-5 was most closely related to viruses in the genus Pestivirus and was not significantly related to flaviviruses or hepaciviruses. However, the relationship of SbCNV-5 to pestiviruses was distant, as indicated by the deep branch of the phylogenetic tree (Fig. 2). A phylogenetic tree was also reconstructed using the conserved region of the RNA helicase. The resulting tree also suggested that SbCNV-5 was most closely related to the pestiviruses (Fig. 3).

TEM revealed a presence of virus-like particles in female nematodes (Fig. 4). The virus particles were enveloped and measured 70–80 nm in diameter and some exhibited a spherical shape. Virions were also seen clustered inside membrane-bound structures. Virus particles budding from the endoplasmic reticulum were observed in some sections (Fig. 4).

Detection of SbCNV-5 in greenhouse and field populations of SCNs

To determine the prevalence of SbCNV-5 in laboratory strains of SCNs grown on soybean in a greenhouse and in field populations of SCNs, we developed a TaqMan quantitative reverse transcription-PCR (qRT-PCR) assay to detect SbCNV-5 RNA and quantified the SbCNV-5 RNA relative to an abundantly expressed nematode transcript, *H. glycines* fatty acid and retinol binding protein-1 (HgFAR1) (Prior et al., 2001). Detecting the HgFAR1 transcript also served as a positive control for the qRT-PCR assay. The TaqMan assay detected SbCNV-5 in four of the ten SCN populations tested (Table 3), indicating the virus was common in different greenhouse populations. As it is possible that the viruses accumulated to higher levels in laboratory cultures, we tested wild populations of SCNs extracted from soybean field soil samples taken from 23 different locations for the presence of SbCNV-5. In this field experiment, six out of 23 samples had measurable levels of the virus (Table 4).

Detection of SbCNV-5 in different developmental stages of SCNs

To confirm that SbCNV-5 was present in different nematode developmental stages, TaqMan assays were used to test RNA from SCN eggs, second-stage juveniles (J2s), adult males and female nematodes by qRT-PCR as above. SbCNV-5 was detected in all four nematode stages (Table 5). The RNA samples analysed in Table 5 were extracted from SCNs over a period of 4 years, which indicated that the SbCNV-5 infections are long-lived and stable in the nematode populations.

SbCNV-5 is not integrated into the nematode genome

Sequences homologous to transcripts of RNA viruses have been detected previously integrated into host genomes and are sometimes actively transcribed (Belyi et al., 2010;
Taylor et al., 2010). We therefore assessed the possibility that SbCNV-5 was integrated into the SCN genome using the same genomics approach described by Bekal et al. (2011). SCN genomic sequences generated using a SOLiD sequencer were aligned with the viral sequences. In this experiment, a total of $2.7 \times 10^8$ 50 bp reads of genomic DNA from SCN population TN10 (~100-fold coverage) were compared with the SbCNV-5 genome and with a section of the SCN genome. None of the SOLiD fragments showed a significant match to the SbCNV-5 genome sequence. A similar result was obtained when SOLiD data from SCN population TN20 was used. When the same SOLiD data were compared with a much shorter ~5000 bp single-copy region of the SCN genome (Craig et al., 2008), over 1500 SOLiD reads aligned to the template, confirming that this approach detected SCN genomic sequences. These data indicated that SbCNV-5 was not integrated into the SCN genome in either of the two nematode populations tested.

**DISCUSSION**

To our knowledge, SbCNV-5 is the first nematode virus that is predicted to have a positive-polarity RNA genome. This virus was shown to infect SCN populations in both greenhouse and in soybean fields. However, we cannot presently assign a specific pathology to SbCNV-5 with any confidence due to the fact that SCNs are often co-infected with several other viruses (Bekal et al., 2011). The observation that SbCNV-5 was present in SCN developmental stages including egg, J2 and adult females suggests that the virus is transmitted transovarially to the next nematode generation. Furthermore, the presence of SbCNV-5 in male SCNs indicates that the virus could also be transmitted sexually. SCNs are subterranean obligate plant parasites with an obligate sexual life cycle where SCN males often mate with more than one female; hence, males may spread SbCNV-5 to uninfected females, thus stably maintaining the virus in the nematode population.

Initial BLAST searches of the SbCNV-5 genome in the NCBI databases suggested that this nematode virus had significant similarity to pestiviruses in the family *Flaviviridae*. This similarity was confirmed by phylogenetic analysis of the RdRp and C-terminal helicase domains. Both analyses showed that SbCNV-5 fell in a clade between pestiviruses and flaviviruses, with closer similarity to pestiviruses. However, its distant relationship was indicated by the

<table>
<thead>
<tr>
<th>Domain</th>
<th>E-value</th>
<th>Polyprotein position (aa range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNA helicase</td>
<td>$9.02^{-24}$</td>
<td>3078–3226, 3268–3421</td>
</tr>
<tr>
<td>Trypsin-like serine protease</td>
<td>$3.70^{-8}$</td>
<td>2878–2955, 3000–3030</td>
</tr>
<tr>
<td>S-Adenosyl-l-methionine-dependent RNA methyltransferase</td>
<td>$4.73^{-6}$</td>
<td>4662–4842</td>
</tr>
<tr>
<td>RdRp</td>
<td>$4.20^{-21}$</td>
<td>5249–5431</td>
</tr>
</tbody>
</table>

*Fig. 2.* Phylogram of the conserved RdRp domains from SbCNV-5 and viruses in the family *Flaviviridae*. The RdRp protein regions were aligned with those of hepatitis C virus (GenBank accession no. AF011751), aa 2521–2795. The viruses used in the comparison were: Meaban virus (MV), Gadgets Gully virus (GGV), dengue virus type 1 (DV1), Aroa virus (AV), Gentian Kobu-sho-associated virus (GKaV), classical swine fever virus (CSFV), border disease virus (BDV), bovine viral diarrhea virus 3 (BVDV3), BVDV2, BVDV1, hepatitis C virus 3a (HCV3a), HCV7, HCV1a, GB virus C (GBV-C), GBV-A and GBV-D. Genus designations are shown to the right of the phylogram. Bootstrap values are indicated as the percentage of iterations that supported the node. GenBank accession numbers are given in parentheses. Bar, amino acid substitutions per site.
deep branches on the phylogenetic trees. A closer look at the genome structure showed that SbCNV-5 possesses features typical of flaviviruses that are not shared by pestiviruses. The most notable flavivirus feature is the presence of the S-adenosyl-L-methionine-dependent RNA methyltransferase domain directly upstream from the RdRp. Pestiviruses do not express a methyltransferase as they utilize an internal ribosome entry sequence for translation initiation and do not need the cap-modifying enzyme (Simmonds et al., 2012).

The structure and length of the 3’ NCR for SbCNV-5 is atypical, in that it is long in comparison with flaviviruses that have shorter 3’ NCRs of 384–764 nt (Proutski et al., 1997; Markoff, 2003; Bryant et al., 2005). In flaviviruses, the NCRs contain repeated sequences, hairpin structures and/or pseudoknots critical for viral genome replication, translation and transmission of the virus (Armstrong & Rico-Hesse, 2001; Proutski et al., 1999; Brinton et al., 1986; Bredenbeek et al., 2003). In flaviviruses, the 3’ NCR encodes one or more subgenomic flavivirus RNAs (sfRNAs), but they are not formed in pestiviruses or hepaciviruses (Simmonds et al., 2012); thus the presence of a sfRNA is important evidence for classifying SbCNV-5. Whilst it is difficult to compare SbCNV-5 and flavivirus 3’ NCRs due to their size and sequence divergence, the SbCNV-5 3’ NCR appeared to

Fig. 3. Phylogeny of the conserved C-terminal domain of RNA helicases from SbCNV-5 and viruses in the family Flaviviridae. The protein regions were aligned with BVDV1 at aa 2095–2190. See Fig. 2 for abbreviations of the viruses used in the comparison. Genus designations are shown to the right of the phylogram. Bootstrap values are indicated as the percentage of iterations that support the node. GenBank accession numbers are given in parentheses. Bar, amino acid substitutions per site.

Fig. 4. TEM images of a virus-infected unmated SCN female. (a) A virus budding from the endoplasmic reticulum. (b–d) Enveloped virus in endosomes marked by an arrow. Bars, 100 nm.

Table 3. Detection of SbCNV-5 RNA by qRT-PCR in SCN greenhouse cultures

<table>
<thead>
<tr>
<th>Nematode population</th>
<th>HgFAR-1 (reference)*</th>
<th>SbCNV-5*</th>
</tr>
</thead>
<tbody>
<tr>
<td>TN11</td>
<td>8.9</td>
<td>ND</td>
</tr>
<tr>
<td>TN16</td>
<td>8.8</td>
<td>ND</td>
</tr>
<tr>
<td>TN9</td>
<td>8.9</td>
<td>25.3</td>
</tr>
<tr>
<td>TN10</td>
<td>19.5</td>
<td>22.1</td>
</tr>
<tr>
<td>TN6</td>
<td>28.8</td>
<td>35.4</td>
</tr>
<tr>
<td>TN20</td>
<td>28.4</td>
<td>36.3</td>
</tr>
<tr>
<td>PA2</td>
<td>25.7</td>
<td>ND</td>
</tr>
<tr>
<td>Race 1</td>
<td>21.1</td>
<td>ND</td>
</tr>
<tr>
<td>Race14</td>
<td>29.2</td>
<td>ND</td>
</tr>
<tr>
<td>Ov5</td>
<td>33.8</td>
<td>ND</td>
</tr>
</tbody>
</table>

ND, No signal detected.
*The units for columns 2 and 3 are threshold cycle (Ct) numbers.
**Table 4. Detection of SbCNV-5 RNA by qRT-PCR in field-isolated SCN**

<table>
<thead>
<tr>
<th>Nematode population</th>
<th>HgFAR-1 (reference)*</th>
<th>SbCNV-5*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sagel</td>
<td>16.1</td>
<td>ND</td>
</tr>
<tr>
<td>Boyer</td>
<td>15.2</td>
<td>ND</td>
</tr>
<tr>
<td>Prough</td>
<td>18.4</td>
<td>ND</td>
</tr>
<tr>
<td>Constant</td>
<td>18.1</td>
<td>ND</td>
</tr>
<tr>
<td>Brown</td>
<td>24.2</td>
<td>34.0</td>
</tr>
<tr>
<td>Harrel</td>
<td>22.5</td>
<td>28.2</td>
</tr>
<tr>
<td>Maxheimer</td>
<td>28.4</td>
<td>ND</td>
</tr>
<tr>
<td>Power</td>
<td>21.6</td>
<td>35.4</td>
</tr>
<tr>
<td>Conklin</td>
<td>34.6</td>
<td>27.6</td>
</tr>
<tr>
<td>Rice</td>
<td>22.3</td>
<td>32.4</td>
</tr>
<tr>
<td>Maher</td>
<td>23.5</td>
<td>ND</td>
</tr>
<tr>
<td>Koonce</td>
<td>21.7</td>
<td>35.1</td>
</tr>
<tr>
<td>IL3/7-03-10</td>
<td>23.9</td>
<td>ND</td>
</tr>
<tr>
<td>IL3/7-30-10</td>
<td>29.0</td>
<td>ND</td>
</tr>
<tr>
<td>Larmi</td>
<td>29.8</td>
<td>ND</td>
</tr>
<tr>
<td>IL3/7-15-11</td>
<td>19.81</td>
<td>ND</td>
</tr>
<tr>
<td>Sp10/ Ks</td>
<td>27.9</td>
<td>ND</td>
</tr>
<tr>
<td>Kannawha, IA</td>
<td>30.3</td>
<td>ND</td>
</tr>
<tr>
<td>Graves Chapel, MO</td>
<td>35.8</td>
<td>ND</td>
</tr>
<tr>
<td>Gerbeke, ND</td>
<td>26.0</td>
<td>ND</td>
</tr>
<tr>
<td>Ames, IA</td>
<td>31.1</td>
<td>ND</td>
</tr>
<tr>
<td>Putman, OH</td>
<td>28.1</td>
<td>ND</td>
</tr>
<tr>
<td>GPC, MO</td>
<td>38.3</td>
<td>ND</td>
</tr>
</tbody>
</table>

ND, No signal detected.

*The units for columns 2 and 3 are threshold cycle (Ct) numbers.

Flavivirus sRNAs are formed by an unusual mechanism, where RNA secondary structure prevents the host 5’→3’ exonuclease XRN1 from hydrolysing the 3’ NCR of the viral genome. Typically, a single sRNA is formed from the 3’ NCR, but additional sRNAs have been observed (Liu et al., 2010; Roby et al., 2014; Silva et al., 2010). Additional sRNAs can be formed by mutations that disrupt critical 3’ NCR secondary structures (Funk et al., 2010); however, we have not observed significant RNA sequence variation/mutations that might explain the formation of two sRNAs. It is more likely that the two sRNAs occur naturally in SbSNV-5, presumably by the action of different RNA-degrading enzymes (Silva et al., 2010). In fact, smaller sRNAs are often overlooked and their functional significance is currently unknown (Roby et al., 2014.). SbCNV5 consistently has two sRNAs, making it similar to yellow fever virus and dengue viruses 1–4 (Liu et al., 2010; Silva et al., 2010). As sRNAs are critical for viral pathogenicity via evasion of type 1 IFN responses (Schuessler et al., 2012), it would also be interesting to determine whether a similar mechanism was occurring in the nematode. Collectively, the data suggest that SbCNV-5 has protein homology to pestiviruses but significant structural similarities found only in flaviviruses.

In addition, SbCNV-5 had strong similarity to a plant virus, GKAv, and appears to share some genome features such as the same order of serine proteinase, RNA helicase, methyltransferase and RdRp. One difference is that GKAv comes from an RNA digestion experiment on crude plant extracts (Kobayashi et al., 2013). Thus, it is possible that the dsRNA replicating intermediate was detected from preparations instead of the viral genomic RNA packaged in virions. Consequently, we suggest that both SbCNV-5 and GKAv are positive-polarity ssRNA viruses. Similarities between SbCNV-5 and GKAv, coupled with the fact that some plant-parasitic nematodes can vector plant viruses, led us to test soybean plants infested with SCNs harbouring SbCNV-5 for systemic spread of SbCNV-5 into stems and leaves. Preliminary attempts to detect SbCNV-5 in plant leaves did not detect the virus (data not shown), but more work needs to be carried out to determine whether SbCNV-5 can replicate in plant cells.

The TEM images of SbCNV-5 also revealed a similarity to flaviviruses. A typical flavivirus has an enveloped spherical virion of approximately 50 nm. TEM images of SbCNV-5-infected SCNs showed 80 nm enveloped virus particles that were spherical in shape. The larger size of the SbCNV-5 virion may be necessary to accommodate its larger genome size. Flaviviruses assemble in the rough endoplasmic reticulum and bud from here, after which they quickly move though the Golgi apparatus and accumulate in endosome vesicles before they are released from the cell by exocytosis (Mackenzie & Westaway, 2001; Westaway et al., 2003). In TEM images of SbCNV-5-infected SCNs, both
viral budding and endosomes containing viruses were visible, suggesting that the nematode virus replicates in a similar fashion to flaviviruses. Whilst the morphogenesis of pestiviruses is still poorly understood, it is also assumed that virions are transported via the membrane system of the host-cell secretary pathway to the cell surface and then released by exocytosis (Burack et al., 2012).

In conclusion, SbCNV-5 has protein homology to pestiviruses but has a genome structure, sRNAs and viral maturation similar to flaviviruses. Given that flaviviruses are common in insects (Huhtamo et al., 2012) and that nematodes and insects are distant relatives, it might not be surprising that nematodes carry similar viruses. Whilst we do not have evidence to illustrate that SbCNV-5 is vectored by the nematode to any other organism, some flaviviruses only infect insects. More knowledge about the modes of transmission of SbCNV-5 is necessary. If field and greenhouse bioassays determine that SbCNV-5 perturbs SCN growth, then this virus may be an effective biological control agent for the most damaging pathogen of soybeans. Also, because SbCNV-5 is a positive-polarity RNA virus, we believe that it may be more amenable to being converted into a viral vector or used as an infectious RNA for manipulating the expression of SCN genes.

METHODS

Nematode strains. *H. glycines* Ichinohhe, 1952, was inbred by Terry Niblack (TN) and strains TN10 and TN20 were used in this study. All SCN strains were isolated in the USA; nematode strain TN10 was collected in Missouri and strain TN20 was collected in Tennessee. The nematodes were grown by standard methods and the eggs were harvested and purified as described by Niblack et al. (1993). The SCN J2s were surface sterilized as described previously (Craig et al., 2008).

Sequencing of SCN DNA. Total RNA was extracted from eggs and surface-sterilized J2s (Craig et al., 2008) of SCN strains TN10 and TN20. A normalized cDNA library was constructed and sequenced at the W. M. Keck Center for Comparative and Functional Genomics, Roy J. Carver Biotechnology Center, University of Illinois at Urbana-Champaign, USA, as described previously (Bekal et al., 2011). A single illumina sequencing run generated 1.55 × 10^6 reads (75 nt each) for the TN10 egg RNA sample and 1.12 × 10^6 reads (75 nt each) for the TN10 J2 RNA, whilst the 454 sequencing generated 5.79 × 10^6 reads (200–400 nt each) from TN10. SeqWright DNA Technology Services (Houston, TX, USA) collected the genomic DNA sequence from SCN strains TN10 and TN20 using the SOLiD sequencing platform.

Nucleotide sequence analysis. The sequence reads were assembled de novo into contigs using the CLC Genomics Workbench (Cambridge, MA, USA). Several assemblies were conducted with different parameters. The resulting contigs were compared with GenBank sequences using the BLASTX function in the Workbench. The BLASTX results for each contig were sorted by contig size, which revealed large contigs with similarity to viral sequences. End sequences of the virus contigs were used to search for overlapping contig sequences in the lists of contigs assembled by the de novo assembler. Additional stretches of cDNA were added to the ends of the contigs until no further end sequences were detected. The contigs were reassembled to their final size using the CLC Genomics Workbench. The Illumina paired-end reads and SOLiD reads were compared with the contigs using the mapping function in the Workbench. Consistent, overlapping paired-end alignments across the contig were used to verify that the virus genome was correctly assembled. The coverage of Illumina reads at the ends of the full virus genome gradually lowered to one or two read coverage, indicating the end of the viral genome. Coverage for the virus genome was calculated by adding the total number of nucleotides in matching Illumina reads for the virus and then dividing by the length of the virus genome.

The CLC Genomics Workbench was also used for phylogenetic analyses. The conserved regions of the viral RdRp used in the protein sequence alignments and phylogenetic trees were chosen as described by Longdon et al. (2010). The protein alignments and trees were performed using the neighbour-joining and unweighted pair group method with arithmetic mean algorithms with 10 000 bootstrap replications in the CLC Genomics Workbench (Loewenstein et al., 2008). The secondary structures of the SbCNV5 5′ and 3′ NCRs were predicted using pknotsRG (Reeder et al., 2007) on the BiBiServ website (http://bibiserv.techfak.uni-bielefeld.de/pknotsrg/). The conserved protein domains in the SbCNV-5 polypolyprotein were identified by comparing the predicted virus protein with the SUPERFAMILY 1.75 HMM library (Gough et al., 2001). Repeated sequences in the SbCNV-5 genome were identified using the pattern discovery software module in the CLC Genomics Workbench.

RNA gel blot detection of SbCNV-5 RNA. Total RNA was extracted from SCNs infected with SbCNV-5 as described above. The cDNA was synthesized using random primers and the Thermoscript RT-PCR system (Life Technologies). The resulting cDNA was PCR amplified using primers that hybridize to the 3′ end of SbCNV-5. Primers SbCV-5-3OF (5′-GTGGTACGATGGACGTAA-3′) and SbCV-5-3OR (5′-GTTTTAGGTCTTACCCAAT-3′) were added to a 50 µl PCR at a concentration of 10 pmol per primer, along with 10 µl Taq polymerase 5× master mix (New England BioLabs). The PCR was run at 95 °C for 30 s, followed by 30 cycles of 95 °C for 30 s, 55 °C for 30 s and 68 °C for 1 min and 30 s, with one cycle of 68 °C for 10 min. The resulting PCR products were purified using ExoSAP-IT (Affymetrix) and a second round of PCR was then conducted as described above using nested primers SbCV-5-3IF (5′-CACAATCCGTTTACCCCAT-3′) and SbCV-5-3IR (5′-CACATGTTGACGACCCCTA-3′). The second amplification produced a 1051 bp PCR product, which was purified using a QiAquick PCR Purification kit (Qiagen). The PCR product was sequenced at the W. M. Keck Center for Comparative and Functional Genomics using SbCV-5-3IF and SbCV-5-3IR as sequencing primers. The resulting sequence matched the 3′ NCR of SbCNV-5, so this PCR product was used to generate a biotinylated probe using a Bright Star Psoralen-Biotin labelling kit (Life Technologies). Total RNA from SCN of infected SCNs was denatured in formamide and run on a 1% RNAse-free TAE/agarose gel using RNA Millennium markers (Life Technologies). The RNA was alkaline blotted onto Hybond-N membrane (Sambrook & Russell, 2001) and the probe was hybridized (0.1 ng ml^-1) in Ultrahybe (Life Technologies). Conjugated streptavidin–alkaline phosphatase (Promega) diluted 1:20 000 in blocking buffer was allowed to bind the probe. Following washing, the hybridized probe was detected using Tropix CDP-STAR chemiluminescence substrate (Life Technologies). The blot was exposed to film for 15 min.

qRT-PCR for virus detection. Primers and probes for the TaqMan assays were synthesized by Life Technologies. The qRT-PCR was conducted using a TaqMan EZ RT-PCR Core Reagents kit (Life Technologies) and thermal cycling was performed using standard conditions on an Applied Biosystems 7900HT sequence detection system. The following probes were used to detect ScCNV-5: ScCNV-5-3′ primer 5′-GGACCACTGTCGGGCGTGGTTG-3′, ScCNV-5-5′ primer 5′-TGAGGCATGTTCACAAAC-3′ and ScCNV-5 probe 5′-FAM-CCTCT-TGAGTCTAGATGG-MGBNFQ-3′. HgFA1 was detected using
TEM analysis of SbCNV-5-infected SCNs. A soybean seedling was inoculated with 500 J2s of SCN strain TN10, which contained SbCNV-5. The nematodes were allowed to infect the plant for 5 days, and they were then transferred to hydroponic culture to allow the development of uninfected SCN females for an additional 2 weeks. Small holes were bored in the cuticles of uninfected adult female nematodes before being soaked in Karnovsky’s fixative (phosphate-buffered 2% glutaraldehyde and 2.5% paraformaldehyde) overnight at 4°C. After primary fixation, a secondary microwave-facilitated Karnovsky’s fixation was used, and the tissue was then washed in 200 mM cacodylate buffer (pH 7.2). Microwave fixation was also used for the secondary 2% osmium tetroxide fixation, followed by the addition of 3% potassium ferricyanide. After washing with water, saturated uranyl acetate was added for en bloc staining. Next, the tissue was dehydrated in a series of increasing concentrations of ethanol. Acetonitrile was used as the transition fluid between ethanol and the epoxy. An infiltration series was carried out with an epoxy mixture using the Epon substitute Lx112 (LADD Research Industries). The resulting blocks were polymerized at 90°C overnight and trimmed, and ultrathin sections were prepared with diamond knives. Sections were stained with uranyl acetate and lead citrate, and examined by TEM (model H600; Hitachi).

ACKNOWLEDGEMENTS

We would like to thank the United Soybean Board and the Soybean Nanotechnology grant program in the College of ACES at the University of Illinois for financial support. We also thank Professor Jason Bond for providing field soil samples containing SCN cysts.

REFERENCES


