Virus Research 167 (2012) 34–42

Contents lists available at SciVerse ScienceDirect

Virus Research

journal homepage: www.elsevier.com/locate/virusres

Detection of members of the *Secoviridae* in the Tallgrass Prairie Preserve, Osage County, Oklahoma, USA

Vaskar Thapa[a,1], Ulrich Melcher[b,*], Graham B. Wiley[c,2], Andrew Doust[a], Michael W. Palmer[a], Kimberly Roewe[b,3], Bruce A. Roe[c], Guoan Shen[d], Marilyn J. Roossinck[d,4], Ye Margaret Wang[b,5], Nitin Kamath[b,6]

[a] *Department of Botany, Oklahoma State University, Stillwater, OK 74078, USA*
[b] *Department of Biochemistry & Molecular Biology, Oklahoma State University, Stillwater, OK 74078, USA*
[c] *Department of Chemistry & Biochemistry, University of Oklahoma, Norman, OK 73019, USA*
[d] *Plant Bioogy Division, Samuel Roberts Noble Foundation, Ardmore, OK 73401, USA*

ARTICLE INFO

Article history:
Received 12 January 2012
Received in revised form 21 March 2012
Accepted 24 March 2012
Available online 2 April 2012

Keywords:
Plant virus
Comovirinae
Survey
Polymorphism
Phylogenetics

ABSTRACT

Viruses are most frequently discovered because they cause disease. To expand knowledge of plant-associated viruses beyond these narrow constraints, non-cultivated plants of the Tallgrass Prairie of the United States were systematically surveyed for evidence of viruses. This report discusses putative viruses of the family *Secoviridae* identified by the survey. Sequence analysis suggests the presence of at least six viruses in the study site, including *Bean pod mottle virus*, *Maize chlorotic dwarf virus*, three previously undescribed viruses within the subfamily *Comovirinae* and one unclassifiable virus.

1. Introduction

The family *Secoviridae* has been established recently to include genera from the previously recognized families *Comoviridae* and *Sequiviridae* and unassigned genera *Cheravirus*, *Sadwavirus* and *Torradovirus*, that share small icosahedral virions, positive-sense RNA genomes with a polyprotein expression strategy and type I RNA dependent RNA polymerase (RdRp) (Sanfaçon et al., 2011). In phylogenetic analyses based on the replication protein, three genera belonging to previous *Comoviridae*: *Comovirus*, *Fabavirus* and *Nepovirus*, form a tight grouping, that is now assigned to a new subfamily, *Comovirinae*, within the *Secoviridae* by the International Committee on Taxonomy of Viruses (ICTV) (Sanfaçon et al., 2011). The members of *Comovirinae* possess bipartite genomes encoding two polyproteins, one from each positive-sense single-stranded RNA (RNA 1 and RNA 2). RNA 1 encodes a polyprotein with a set of signature sequences characteristic of helicase-proteases and polymerases. The polyproteins are post-translationally cleaved to give mature forms of each polypeptide. The genera *Comovirus*, *Fabavirus* and *Nepovirus* are mainly distinguished on the basis of the vectors involved in transmission (Mayo and Fritsch, 1994). Previously described members of *Comovirus* are transmitted by beetles and infect mostly cultivated members of Fabaceae as well as some cultivated Solanaceae and Brassicaceae (Petrzik et al., 2005). The

Abbreviations: ICTV, International Committee for the Taxonomy of Viruses; PVBE, Plant Virus Biodiversity and Ecology; TPP, Tallgrass Prairie Preserve; RdRp, RNA-dependent RNA polymerase; LCP, large coat protein; NJ, Neighbor Joining; ML, Maximum Likelihood; LG, Le and Gascuel; BPMV, *Bean pod mottle virus*; MCDV, *Maize chlorotic dwarf virus*.

* Corresponding author. Tel.: +1 405 744 6210, fax: +1 405 744 7799.
E-mail addresses: vaskar.thapa@okstate.edu (V. Thapa), u-melcher-4@alumni.uchicago.edu, ulrich.melcher@okstate.edu (U. Melcher), wileyg@omrf.org (G.B. Wiley), andrew.doust@okstate.edu (A. Doust), mike.palmer@okstate.edu (M.W. Palmer), kimberly.roewe@okstate.edu (K. Roewe), broe@ou.edu (B.A. Roe), gshen@noble.org (G. Shen), mjr25@psu.edu (M.J. Roossinck), wangye@seas.upenn.edu (Y.M. Wang), nmk38@cornell.edu (N. Kamath).
[1] Present address: 2233 Eglinton Ave. E., Toronto, Ontario, Canada M1K2M9.
[2] Present address: Oklahoma Medical Research Foundation, Oklahoma City, OK 73104, USA.
[3] Present address: Department of Chemistry, Oklahoma State University, Stillwater, OK 74078, USA.
[4] Present address: Center for Infectious Disease Dynamics, Pennsylvania State University, University Park, PA 16802 USA.
[5] Present address: Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19104, USA.
[6] Present address: Department of Plant Pathology, Cornell University, Ithaca, NY 14853, USA.

0168-1702/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.virusres.2012.03.016

genus *Fabavirus* is similar to *Comovirus* in its capsid composition, genome size and cytopathology and its members mostly have hosts in vegetables and ornamental plants from the Fabaceae (Lisa and Boccardo, 1996; Koh et al., 2001) but aphids are the primary vector. *Nepovirus* is known from diverse cultivated hosts, mostly woody members from the Vitaceae, Rosaceae, Solanaceae and Malvaceae where nematodes serve as vectors (Sanfaçon et al., 2011).

Currently over 50 (15 *Comovirus*, 4 *Fabavirus* and 34 *Nepovirus* species) recognized members of *Comovirinae* are included in the ICTV's 2009 master species list (Sanfaçon et al., 2011). Among them, few members are reported from natural ecosystems (Sandoval et al., 1995), and the phylogenetic relationships of viruses from native plants and those known to infect cultivated plants have been studied only rarely. Many species of plant viruses remain to be discovered due to relative lack of attention to viruses of plants from non-cultivated areas (Cooper and Jones, 2006; Wren et al., 2006). A project (Plant Virus Biodiversity and Ecology, PVBE) to uncover such viruses in plants of the Tallgrass Prairie Preserve of Osage County, Oklahoma (TPP, managed by the Nature Conservancy) (Hamilton, 2007; Allen et al., 2009) was initiated (Melcher et al., 2008) and resulted in evidence of several novel viruses (Muthukumar et al., 2008; Roossinck et al., 2010; Scheets et al., 2011; Min et al., 2012; Stobbe et al., 2012). We report here the results of searches of the recovered sequences for evidence of members of the family *Secoviridae*.

2. Materials and methods

2.1. Sample collection and virus sequence

Plant samples for the extraction of plant virus sequences were collected from the TPP. The TPP comprises mostly tallgrass prairie vegetation but also includes wetlands, forests and woodlands (Allen et al., 2009) and includes 763 documented species of vascular plants of which only 12% are non-native (Palmer, 2007). From 2005 to 2008, the PVBE project (Wren et al., 2006) collected samples of 537 plant species as a part of its sample collection from the TPP. Six species (*Asclepias viridis*, *Ambrosia psilostachya*, *Ruellia humilis*, *Panicum virgatum*, *Sorghastrum nutans* and *Vernonia baldwinii*) were sampled multiple times at multiple locations in each of three years. The protocols for virus isolation and genome sequencing were followed as described for other plant viruses in the PVBE project (Melcher et al., 2008; Muthukumar et al., 2008; Roossinck et al., 2010). The sequence of a particularly underrepresented region of RNA 1 of one virus of interest was confirmed by sequencing of an RT-PCR product covering the region using primers 5′ GACCCTGAAGATCCAACTGC 3′ and 5′ CAACCAATGGCATGTTCTCA 3′. This RT-PCR was used also for detection of the virus in a survey of the multiply sampled species during 2009. RNA extracted from young leaves as previously described (Min et al., 2012) was used as template for the reaction and detection was by agarose gel electrophoresis of the amplicons.

2.2. Analysis of sequence polymorphism

By using Clustal W (Larkin et al., 2007) we aligned 37 and 26 contig sequences of RNA 1 and RNA 2, respectively, from different hosts and apparently belonging to a single virus species which we here designate Asclepias virus TGP2. The alignments were manually adjusted according to amino acid expression of codons in RNA 1 and RNA 2. The percentage of nucleotide substitutions in the first, second and third position of codons in polyprotein sequence of RNA 1 and 2 of Asclepias virus TGP2 were calculated by generating a positional nucleotide numerical summary in BioEdit version 7 (Hall, 1999). The ratio of frequencies of non-synonymous (dN) to synonymous substitutions (dS) was calculated for RNA 1 and RNA 2 using *p*-distance (proportion of nucleotide differences to total number of nucleotides) by averaging all sequence pairs, following Nei-Gojobori's (1986) method in MEGA4 (Tamura et al., 2007). In the analysis, all positions containing alignment gaps and missing data were eliminated by applying the pairwise deletion option.

2.3. Sequence alignment and phylogenetic analysis

The RNA 1 and RNA 2 sequences with 1618 and 1038 amino acids respectively of Asclepias virus TGP2 that cover the complete polyprotein regions were used for BLAST search using protein vs. protein (BLASTp) and position-specific iterated (PSI-BLAST) (Altschul et al., 1997; Schäffer et al., 2001). For RNA 1 polyproteins, we used BLASTp with the sequence portion of RNA 1 polyproteins aligned to the conserved domain of the RdRp's of *Comoviridae* (*Comovirinae*, *sensu stricto*) (Marchler-Bauer et al., 2009) because cleavage sites for RdRp are not consistent across the genera of the *Comovirinae* (*Comoviridae*, *sensu stricto*) (Sanfaçon et al., 2011). For RNA 2 polyproteins we used PSI-BLAST because its large coat protein (LCP) conserved domain query in both BLASTp and PSI-BLAST provided hits with very few members. Moreover, the alignments built based on the BLAST hits in M-Coffee (Wallace et al., 2006) showed a poor alignment score (<50%) for phylogenetic analysis. Both BLASTp and PSI-BLAST were performed against the GenBank reference protein database to avoid duplicate hits. The significant viral sequence hits thus obtained from the BLAST searches were aligned using M-Coffee algorithm for multiple sequence alignment (Wallace et al., 2006). Some changes in the alignment were done manually where necessary. We constructed phylogenetic trees for RNA 1 using the conserved domain of RdRp and RNA 2 polyproteins separately using Neighbor-Joining (NJ) and Maximum Likelihood (ML) methods. The NJ trees were developed using the Jukes-Cantor distance and resampled with bootstrap including 1000 replicates (Saitou and Nei, 1987) in Geneious version 4.8.5 (Drummond et al., 2009). For ML tree construction, the protein substitution model (LG model) of Le and Gascuel (2008) was used. LG model was the model best fit to our samples as determined by the protein substitution model analysis in ProtTest version 2.4 (Abascal et al., 2005). The ML trees were then made using PhyML version 3.0 (Guindon and Gascuel, 2003) with bootstrap resampling of 1000 replicates.

3. Results

3.1. BPMV-TGP1

Sequences similar to those of two known members of the *Secoviridae* were identified in the survey. A *Desmodium paniculatum* plant yielded sequences similar to both RNA 1 and RNA 2 of *Bean pod mottle virus* (Table 1; Supplemental files 1 and 2). Five regions encompassing 1302 nt were derived from RNA 1 and 694 nt were similar to two regions of RNA 2. Similarities to the reference BPMV sequence were 80.8 and 79.5% at the nucleotide level and 89.4 and 91.3% at the amino acid levels, respectively for RNAs 1 and 2. No other plants surveyed contained any sequences highly similar to these, and we tentatively named the strain BPMV-TGP1.

3.2. MCDV-TGP1

A strain of MCDV, designated MCDV-TGP1, was identified by 860 nt from five sequence reads that were distributed among four regions of the MCDV reference genome (Table 1; Supplemental files 1 and 2). Similarities between the reference MCDV and MCDV-TGP1 were 79.5 and 87.1% for nucleotide and amino acid sequences,

Table 1
Plant sources of signature sequences of *Secoviridae* viruses from the Tallgrass Prairie Preserve.

Plantid	Easting	Northing	Sampling date	% of total reads	Acc. No. Seg. 1	Acc. No. Seg. 2
***Bean pod mottle virus* – TGP1**						
Desmodium paniculatum (L.) *DC.* (Fabaceae)						
05TGP00183	729279	4080684	21-05-2005	0.83	JN661357–JN661361[a]	JN661362–JN661363
***Maize chlorotic dwarf virus* – TGP1**						
Panicum virgatum L. (Poaceae)						
05TGP00305	730000	4086000	14-06-2005	1.5	JN661364–JN661365	
Erigeron tenuis Torr. & Gray (Asteraceae)						
05TGP00095	727956	4079967	13-05-2005	0.24	[a]	
Solanum virus TGP1						
Solanum dimidiatum Raf. (Solanaceae)						
05TGP00442	733128	4079372	23-06-2005	0.44	JN661369	
Vernonia virus TGP1						
Vernonia baldwinii Torr. (Asteraceae)						
05TGP00396	729000	4078000	22-06-2005	0.77	n.d.	JN661370
Vernonia virus TGP2						
Vernonia baldwinii Torr. (Asteraceae)						
05TGP00448	729000	4077000	23-06-2005	2.1	JN661367–JN661368	JN661366
05TGP00362	731000	4082000	16-06-2005	0.90	[a]	n.d.
Asclepias virus TGP2						
Ambrosia bidentata Michx. (Asteraceae)						
07TGP00023	733007	4081003	09-06-2007	0.05	n.d.	[a]
Ambrosia psilostachya DC. (Asteraceae)						
05TGP00350	734000	4083000	16-06-2005	0.83	[a]	JN661347[a]
06TGP01125	727997	4078017	04-06-2006	0.24	n.d.	JN661308
07TGP00214	734000	4074000	19-06-2007	0.05	JN661260	n.d.
08TGP00003	734004	4078993	10-06-2008	0.11	[a]	n.d.
08TGP00062	729003	4076999	12-06-2008	0.16	n.d.	JN661337
08TGP00136	731003	4078995	20-06-2008	0.04	[a]	n.d.
Arenaria serpyllifolia L. (Caryophyllaceae)						
05TGP00359	733255	4082465	16-06-2005	0.29	n.d.	[a]
Asclepias tuberosa L. (Apocynaceae)						
07TGP00052	731046	4079315	09-06-2007	0.07	[a]	n.d.
Asclepias viridis Walt. (Apocynaceae)						
05TGP00008	728405	4081441	05-05-2005	1.1	JN661218[a]	JN661296
05TGP00294	734000	4074000	08-06-2005	0.23	[a]	n.d.
05TGP00307	730000	4086000	14-06-2005	1.0	JN661215–JN661216	JN661346
05TGP00337	731000	4079000	15-06-2005	2.6	JN661295, JN661217[a]	n.d.
05TGP00360	731000	4079000	16-06-2005	6.5	JN661221–JN661225	JN661352–JN661353[a]
06TGP01122	729007	4078024	04-06-2006	15.3	JN661277–JN661284	JN661303–JN661307[a]
06TGP01212	734005	4083009	13-06-2006	1.5	JN661226–JN661227[a]	JN661309–JN661311
06TGP01258	734002	4071009	14-06-2006	3.4		JN661328[a]
07TGP00026	733009	4081002	09-06-2007	1.5	JN661228–JN661233[a]	JN661354-JN661356[a]
07TGP00045	731012	4078998	09-06-2007	2.6	JN661250–JN661251[a]	JN661314[a]
07TGP00057	731009	4085994	09-06-2007	0.25	JN661258–JN661259	JN661315–JN661317
07TGP00071	734016	4083002	10-06-2007	1.3	JN661220	JN661318–JN661319
07TGP00131	730000	4086000	17-06-2007	4.5	JN661252–JN661254[a]	JN661320–JN661323[a]
07TGP00183	730000	4076000	18-06-2007	15.6	JN661273–JN661276[a]	JN661350–JN661351
07TGP00215	734000	4074000	19-06-2007	1.2	JN661261[a]	JN661327, JN661348–JN661349[a]
07TGP00221	727000	4080000	19-06-2007	0.12	[a]	n.d.
08TGP00006	733003	4078991	10-06-2008	0.13	JN661263	n.d.
08TGP00016	733001	4080998	10-06-2008	0.21	JN661264	[a]
08TGP00024	731001	4081999	11-06-2008	0.68	n.d.	JN661333
08TGP00052	728004	4078002	12-06-2008	54.3	JN661234–JN661237	JN661344[a]
08TGP00060	729003	4076999	12-06-2008	0.07	n.d.	JN661336
08TGP00172	733995	4073990	21-06-2008	10.9	JN661244–JN661249	JN661345
Commelina erecta L. (Commelinaceae)						
05TGP00501	729282	4080023	02-07-2005	0.32	n.d.	JN661302
Dichanthelium acuminatum (Sw.) Gould & C.A. Clark (Poaceae)						
07TGP00115	729000	4078000	17-06-2007	0.06	[a]	n.d.
Dichanthelium oligosanthes (J.A. Schultes) Gould (Poaceae)						
07TGP00111	729000	4078000	17-06-2007	0.12	JN661219	n.d.
Diodia teres Michx. (Rubiaceae)						
07TGP00065	730971	4086003	09-06-2007	0.37	[a]	n.d.
Geranium pusillum L. (Geraniaceae)						
07TGP00148	730057	4080644	17-06-2007	0.10	n.d.	JN661324
Gymnocladus dioicus (L.) K. Koch (Fabaceae)						
05TGP00187	729516	4080711	21-05-2005	0.02	n.d.	JN661297
Juncus torreyi Coville (Juncaceae)						
05TGP00404	728953	4077991	22-06-2005	13.9	JN661238–JN661243	JN661299–JN661301[a]
Panicum anceps Michx. (Poaceae)						
07TGP00019	729019	4077009	08-06-2007	0.06	[a]	n.d.
Panicum virgatum L. (Poaceae)						
06TGP01128	728004	4078019	04-06-2006	0.14	[a]	n.d.

Table 1 (*Continued*)

Plantid	Easting	Northing	Sampling date	% of total reads	Acc. No. Seg. 1	Acc. No. Seg. 2
06TGP01217	734006	4083013	13-06-2006	0.08	n.d.	JN661312
07TGP00106	729000	4078000	17-06-2007	0.12	n.d.	[a]
07TGP00169	734000	4071000	18-06-2007	0.10	JN661255	n.d.
08TGP00007	733003	4078991	10-06-2008	0.47	JN661256	n.d.
08TGP00025	731001	4081999	11-06-2008	0.45	n.d.	JN661334
08TGP00102	731003	4074000	18-06-2008	0.06	n.d.	JN661339
08TGP00143	730011	4085975	20-06-2008	0.07	n.d.	[a]
Prunus angustifolia Marsh. (Rosaceae)						
05TGP00188	729454	4080660	21-05-2005	0.02	[a]	n.d.
Quercus marilandica Muenchh. (Fagaceae)						
06TGP01014	739078	4066635	25-05-2006	0.05	n.d.	JN661343[a]
06TGP01182	738029	4069054	08-06-2006	0.15	[a]	n.d.
Ruellia humilis Nult. (Acanthaceae)						
06TGP01199	734001	4074008	08-06-2006	0.08	[a]	n.d.
07TGP00033	732662	4080907	09-06-2007	0.04	[a]	n.d.
08TGP00019	733001	4080998	10-06-2008	0.39	n.d.	JN661329
08TGP00051	729002	4077994	12-06-2008	0.33	JN661266	n.d.
08TGP00063	729003	4076999	12-06-2008	0.12	n.d.	JN661338
Sorghastrum nutans (L.) Nash (Poaceae)						
06TGP01274	728000	4080000	14-06-2006	0.05	n.d.	JN661313
08TGP00080	730030	4076000	16-06-2008	0.05	JN661267	n.d.
08TGP00103	731005	4073997	18-06-2008	0.08	n.d.	JN661340
Vernonia baldwinii Torr. (Asteraceae)						
07TGP00175	734000	4071000	18-06-2007	1.05	n.d.	JN661325–JN661326[a]
08TGP00021	733001	4080998	10-06-2008	1.94	JN661262	JN661330–JN661332
08TGP00046	729002	4077994	12-06-2008	0.06	n.d.	JN661335
08TGP00105	730955	4073958	18-06-2008	0.08	JN661268	n.d.
08TGP00139	731003	4078995	20-06-2008	0.60	n.d.	JN661342
08TGP00177	733995	4073990	21-06-2008	0.03	[a]	n.d.

n.d.: not detected.

[a] Short sequences available in Supplemental files.

respectively. All but one of the sequence reads came from one of 82 samples of *Panicum virgatum.*

3.3. Solanum virus TGP1

A sample of *Solanum dimidiatum* yielded a polyadenylated 205 nt sequence fragment, tentatively named as belonging to Solanum virus TGP1, whose pre-polyA sequence was about 90% identical to 65–70 nt near the 3′ termini of the genomes of viruses of the *Comovirinae* (Table 1; Supplemental files 1 and 2). The first 130 nt had no detectable similarity to sequences in the nucleotide databases.

3.4. Vernonia viruses TGP1 and TGP2

Several sequence reads from three specimens of *Vernonia baldwinii* showed similarity in BLASTx searches to viruses of the genus *Nepovirus* (Table 1; Supplemental files 1 and 2). One specimen, 05TGP00396, yielded sequences with top BLASTx hits to the nepovirus *Cycas necrotic stunt virus* (putative virus designated Vernonia virus TGP1), while the other two, 05TGP00448 and 05TGP00362, had sequences most similar to one or the other genome segment of two other nepoviruses, melon mild mottle virus and *Grapevine fanleaf virus*. Nucleotide similarities of these putative viruses (which we designate Vernonia viruses TGP1 and TGP2, respectively) were not high enough to provide an accurate estimate of percent nucleotide identity. At the amino acid level, identities to the closest GenBank relative were 30.6% and 49.8% for Vernonia virus TGP1 and -TGP2, respectively. The three plant specimens were the only ones of 77 *V. baldwinii* specimens tested that had nepovirus-like sequences.

3.5. Asclepias virus TGP2

Two sets of very closely interrelated sequences from multiple plants were assembled into consensus sequences that had characteristics of RNAs-1 and -2 of members of the *Comovirinae*, but were distinct from BPMV-TGP, Solanum virus TGP1 and Vernonia viruses TGP1 and TGP2. The composite RNA 1 sequence assembled from the available reads was 5107 nt in length and was incomplete relative to other subfamily members at both 5′ and 3′ ends. The 5′ end included the AUG initiation codon of a polyprotein ORF that occupied the remainder of the RNA, encoding 1699 residues. Sequence encoding the C-terminal peptide sequence was missing as was the conserved 3′ structure element. The composite RNA 2 sequence consisted of 3523 nucleotides, and included the complete coding region for a polypeptide of 1070 amino acids. In addition, there was an overlapping ORF potentially encoding a polypeptide of 141 amino acids of which 5 were cysteines. Based on similarity of the 3′ sequence to the polyA-proximal sequences of members of the *Comovirinae*, the 3′ sequence of this putative virus, designated Asclepias virus TGP2, was complete. The designation was chosen because the sequences were found in the highest number in preparations from multiple *Asclepias viridis* specimens. The numeration of TGP2 is used to avoid confusion with recently described Asclepias asymptomatic virus, which was found in abundance in this plant species (Min et al., 2012) and could be designated Asclepias virus TGP1. However, evidence of Asclepias virus TGP2 was also found during analysis of multiple other specimens, including the six frequently sampled species. Plant families represented among those yielding sequence of Asclepias virus TGP2 included Acanthaceae, Asteraceae, Apocynaceae, Caryophyllaceae, Commelinaceae, Fabaceae, Fagaceae, Geraniaceae, Juncaceae, Poaceae, Rosaceae and Rubiaceae. The virus was detected in multiple samples over multiple years in *Asclepias viridis* (2005, 2006, 2007, 2008), *Ambrosia psilostachya* (2005, 2006, 2007, 2008), *P. virgatum* (2006, 2007, 2008), *R. humilis* (2006, 2007, 2008), *S. nutans* (2006, 2008) and *V. baldwinii* (2007, 2008).

To determine whether the RT-PCR amplification used to complete the main part of the Asclepias virus TGP 2 sequence could be used also as a detection tool for presence of the virus in plants, total RNA was prepared from samples of plants harvested in 2009

Table 2
Prevalence[a] of Asclepias virus TGP2 among frequently sampled plant species of the Tallgrass Prairie Preserve over the years 2005–2009.

Species	Years				
	2005	2006	2007	2008	2009[b]
Ambrosia psilostachya	9 (22)[a]	6 (18)	5 (20)	21 (19)	20 (10)
Asclepias viridis	29 (21)	14 (21)	36 (22)	30 (20)	30 (10)
Panicum virgatum	0 (19)	10 (20)	19 (21)	20 (20)	80 (10)
Ruellia humilis	0 (21)	5 (19)	5 (22)	16 (19)	0 (9)
Sorghastrum nutans	0 (20)	5 (19)	0 (20)	10 (20)	n.a.
Vernonia baldwinii	0 (19)	0 (18)	5 (19)	23 (22)	11 (9)

[a] Percent of plant samples assayed (total number in parentheses) that were positive for the presence of Asclepias virus TGP2.
[b] 2009 samples were assayed by Asclepias virus TGP2-specific RT-PCR on total RNA extracts from plant leaf tissue. All other prevalences represent plant samples detected as positive due to the presence of viral sequences in their sequencing results.

and subjected to RT-PCR. In many samples, the amplification reaction produced a single product of the expected size with no other amplicons detected. The percentages of 2009 plant samples showing evidence of the virus in five of the six target sequences (Table 2) were comparable to those observed by sequencing (2005–2008 samples). Strikingly, there was considerable year-to-year and species-to-species variation. When the geographic distribution of positive and negative samples was examined, no clear clustering pattern was observed (data not shown).

3.6. Sequence polymorphism

The overlapping regions of Asclepias virus TGP2 contigs (overlaps ranging from 2 to 15 contig pairs) for RNA 1 and RNA 2 showed higher occurrence of mutation (mainly nucleotide substitution and deletion) in the third position (12% of total third position of RNA 1 and 10% of RNA 2) than the first and second positions (less than 5% of total first and second positions in RNA 1 and RNA 2). The dN/dS ratio for RNA 1 and RNA 2, averaging over all sequence pairs, were 0.215 and 0.271, respectively (Table 3). Overall, RNA 1 and RNA 2 had 20% and 18% of their positions polymorphic, but in both cases only 10% were informatively polymorphic. The distribution of the informatively polymorphic sites among the sequences from the various plants did not support the hypothesis that the polymorphisms were due to several cocirculating isolates.

A BLASTp search using deduced amino acid sequences encoded by RNA 1 and RNA 2 of Asclepias virus TGP2 showed significant hits with the conserved domains of the RdRp and the LCP, respectively, of *Comovirinae* members. Based on the sequence similarity within the conserved RdRp domain of RNA 1, Asclepias virus TGP2 shared 39–47% sequence identity with members of *Comovirus* (five species), *Fabavirus* (four species) and *Nepovirus* (eight species) and 32–40% sequence identity with other members of *Secoviridae* (ten species) (Table 4). The sequence similarity at the RdRp level of Asclepias virus TGP2 is 10–20% more than at the whole polyprotein level of RNA 1 with viruses obtained from BLAST hits (Table 4, implying better conservation of RdRp in the group. A similar BLAST search with the conserved LCP domain of RNA 2 showed 18–25% sequence identity with *Comovirus* (three species) and *Fabavirus* (three species). However, the LCP query of RNA 2 had less coverage of taxa (few members of *Comovirus* and *Fabavirus*) compared with the BLAST search based on the RNA 2 polyprotein (showing hits with *Comovirus*, *Fabavirus*, *Nepovirus* and *Sadwavirus* members) (Table 5).

Table 3
Synonymous and non-synonymous site differences in Asclepias virus TGP2.[a]

	dN	SE	dS	SE	dN/dS
RNA 1 ($n = 37$)[b]	0.070	0.004	0.325	0.012	0.215
RNA 2 ($n = 26$)	0.059	0.012	0.218	0.020	0.271

[a] Estimates of the number of non-synonymous differences per non-synonymous sites (dN), the number of synonymous differences per synonymous sites (dS), their respective standard errors (SE) and dN/dS ratio were performed by using the Nei-Gojobori method (1986) in MEGA4.
[b] n represents the sample size.

3.7. Phylogenetic inference from RNA 1

The bootstrap trees from both NJ and Maximum Likelihood (ML) methods showed similar topology and strong support for three genera (*Comovirus*, *Fabavirus* and *Nepovirus*) of *Comovirinae* (Fig. 1). The trees provide support for other genera within *Secoviridae*. The newly identified Asclepias virus TGP2 appeared within the clade of *Comovirinae* but had no significant support for placement in any of the three existing lineages at the generic level (Fig. 1). The tree's branching topology showed Asclepias virus TGP2 had an association with *Fabavirus* and *Comovirus* clusters but the lineage had no significant bootstrap support (<70%). A similar clustering of nepoviruses, comoviruses and fabaviruses separate from Asclepias virus TGP2 was obtained when the polyprotein 1 region between the CG 3C-peptidase and the GDD of the RdRp (Sanfaçon et al., 2011) was used for tree construction (data not shown). A separate sister branch for Asclepias virus TGP2 within the clade of *Comovirinae* may have evolved based on its association with non-cultivated hosts when compared to the other members derived from cultivated settings.

3.8. Phylogenetic inference from RNA 2

The consensus bootstrap trees of RNA 2 polyprotein based on ML and NJ showed similar topologies (Fig. 2). As with the RNA 1 polyprotein trees, all three genera of *Comovirinae* had bootstrap support over 70% in RNA 2 trees. Asclepias virus TGP2 appeared as a sister branch to the *Nepovirus* cluster in both ML and NJ trees. However, the Asclepias virus TGP2 branch in the ML tree had no significant statistical support. In both NJ and ML trees from RNA 2, we did not find support for a *Comovirinae* clade as provided by similar trees from RNA 1.

4. Discussion

4.1. Distribution patterns

MCDV-TGP1, Vernonia virus TGP1 and Vernonia virus TGP2 sequences were identified from species that were frequently sampled during the four year course of sampling, yet each putative virus was found in only a single specimen. The rarity suggests that the presence of the virus in these species resulted from spillover from a plant source not included in the study. The *D. paniculatum* and *S. dimidatum* specimens were not sampled frequently, so that the possibility of a close relationship of these hosts with BPMV-TGP1 and Solanum virus TGP1 requires further sampling. Species of the genus *Desmodium* are known hosts of BPMV. In stark contrast to these five viruses, Asclepias virus TGP2 appeared widely distributed in the TPP, evidence of it being encountered in 64 specimens, representing 18 plant species from 12 plant families. This pattern is comparable to that observed with the tymovirus Asclepias asymptomatic virus (Min et al., 2012) and was noted also by direct RT-PCR assay in 2009 samples (Table 2). The apparent dichotomy in virus distributions supports the commonly used classification of plant viruses as generalists, such as the milkweed viruses, and specialists, such as the other members of the *Secoviridae* detected in the study.

Table 4
Plant viruses with significant hits to the polyprotein and the conserved domain of RNA dependent RNA polymerase (RdRp) encoded by RNA 1 of Asclepias virus TGP2.

Virus name	Acronym	Accession number	Genus	Sequence identity (%)	
				Polyprotein	RdRp
Bean pod mottle virus	BPMV	NP_734070.1	*Comovirus*	29	45
Cowpea mosaic virus	CPMV	NP_734057.1	*Comovirus*	32	44
Cowpea severe mosaic virus	CPSMV	NP_734062.1	*Comovirus*	30	44
Red clover mottle virus	RCMV	NP_734030.1	*Comovirus*	28	42
Squash mosaic virus	SqMV	NP_734012.1	*Comovirus*	30	47
Broad bean wilt virus 1	BBWV-1	NP_951030.1	*Fabavirus*	28	41
Broad bean wilt virus 2	BBWV-2	NP_733962.1	*Fabavirus*	29	42
Mikania micrantha wilt virus	MMWV	YP002158831.1	*Fabavirus*	30	44
Patchouli mild mosaic virus	PatMMV	NP_733967.1	*Fabavirus*	29	42
Arabis mosaic virus	ArMV	YP054443.1	*Nepovirus*	29	46
Beet ringspot virus	BRSV	NP_734035.1	*Nepovirus*	28	39
Blackcurrant reversion virus	BRV	NP_734045.1	*Nepovirus*	34	43
Cycas necrotic stunt virus	CNSV	NP_734017.2	*Nepovirus*	35	41
Grapevine chrome mosaic virus	GCMV	YP002000610.1	*Nepovirus*	28	42
Grapevine fanleaf virus	GFLV	NP_734038.1	*Nepovirus*	29	44
Tomato black ring virus	TBRV	NP_958841.1	*Nepovirus*	27	39
Tomato ringspot virus	ToRSV	NP_734007.2	*Nepovirus*	28	41
Apple latent spherical virus	ALSV	NP_734022.1	*Cheravirus*	37	40
Cherry rasp leaf virus	CRLV	YP081454.1	*Cheravirus*	37	41
Satsuma dwarf virus	SDV	NP_734025.1	*Sadwavirus*	29	32
Strawberry latent ringspot virus	SLRSV	YP227373.1	*Sadwavirus*	26	38
Strawberry mottle virus	SMV	NP_733954.1	*Sadwavirus*	32	33
Parsnip yellow fleck virus	PYFV	NP_734447.1	*Sequivirus*	36	37
Tomato marchitez virus	ToMarV	YP001976153.1	*Torradovirus*	28	34
Black raspberry necrosis virus	BRNV	YP654561.1	*Unassigned*	31	34
Maize chlorotic dwarf virus	MCDV	NP_734456.1	*Waikavirus*	31	32
Rice tungo spherical virus	RTSV	NP_734463.1	*Waikavirus*	30	32

4.2. Diversity patterns

The accumulated data presented here suggest the presence of only six members of the *Secoviridae* in the plants sampled and analyzed from the TPP. In the recently completed analysis of sequences from the family *Tombusviridae* in the same survey (Scheets et al., 2011), evidence consistent with the presence of five distinct viruses was discovered. Sequences related to those of the genera *Tobamovirus* (Stobbe et al., 2012) and *Tymovirus* (Min et al., 2012) suggested the presence of only one virus of each genus in the surveyed plants. Yet, other virus families were represented in the TPP by a diversity of viruses (Roossinck et al., 2010 and unpublished observations).

The dN/dS value (<1) for both RNA 1 and RNA 2 of Asclepias virus TGP2 indicates negative selection among the sequences analyzed. We could not analyze population level divergence in the virus sequences from different plant hosts because the Asclepias virus TGP2 sequence used in our analysis is a population estimate (pooled from all samples) and there are few sequence samples from each host. Nevertheless, some comparison is possible with other TPP viruses since similar methods were used with these. The level of polymorphism seen for Asclepias virus TGP2 (19% polymorphic and 10% informatively polymorphic) was somewhat higher than for TGP carmovirus 1 (7.5% polymorphic) and TGP carmovirus 3 (9.2% polymorphic and 2.3% informatively polymorphic) (Scheets et al., 2011). Analysis of the accumulation of variants during passage from cloned sequences also suggests that polymorphism levels are host and virus species specific. That the substantial degree of polymorphism observed is not due to sequencing errors is demonstrated by the 0.5% polymorphism rate observed in the analysis that led to the identification of Passion fruit mosaic virus-TGP (PafMV-TGP) where, consistent with other results from the PVBE project, sequences were obtained from just two plants (Stobbe et al., 2012).

4.3. Taxonomic classification

Given that viruses of the *Comovirinae* whose genomes are within 75–100% identical in the coat protein coding regions are assigned

Table 5
Plant viruses with significant hits with polyprotein and conserved domain of Large Coat Protein (LCP) of RNA 2 of Asclepias virus TGP2.

Virus name	Acronym	Accession number	Genus	Sequence identity (%)	
				Polyprotein	LCP
Bean pod mottle virus	BPMV	NP_612348.1	*Comovirus*	19	–
Cowpea severe mosaic virus	CPSMV	NP_619517.1	*Comovirus*	19	20
Radish mosaic virus	RaMV	YP001911127.1	*Comovirus*	21	–
Squash mosaic virus	SqMV	NP_620658.1	*Comovirus*	22	21
Turnip ringspot virus	TuRSV	YP003193666.1	*Comovirus*	19	18
Broad bean wilt virus 1	BBWV-1	NP_945135.1	*Fabavirus*	21	–
Broad bean wilt virus 2	BBWV-2	NP_149013.1	*Fabavirus*	22	25
Mikania micrantha wilt virus	MMWV	YP002158823.1	*Fabavirus*	20	21
Patchouli mild mosaic virus	PatMMV	NP_647591.1	*Fabavirus*	21	23
Arabis mosaic virus	ArMV	YP053924.1	*Nepovirus*	21	–
Grapevine fanleaf virus	GFLV	NP_619706.1	*Nepovirus*	25	–
Tomato ringspot virus	ToRSV	NP_620762.1	*Nepovirus*	25	–
Satsuma dwarf virus	SDV	NP_620567.1	*Sadwavirus*	19	–
Strawberry latent ringspot virus	SLRSV	YP227368.1	*Sadwavirus*	24	–

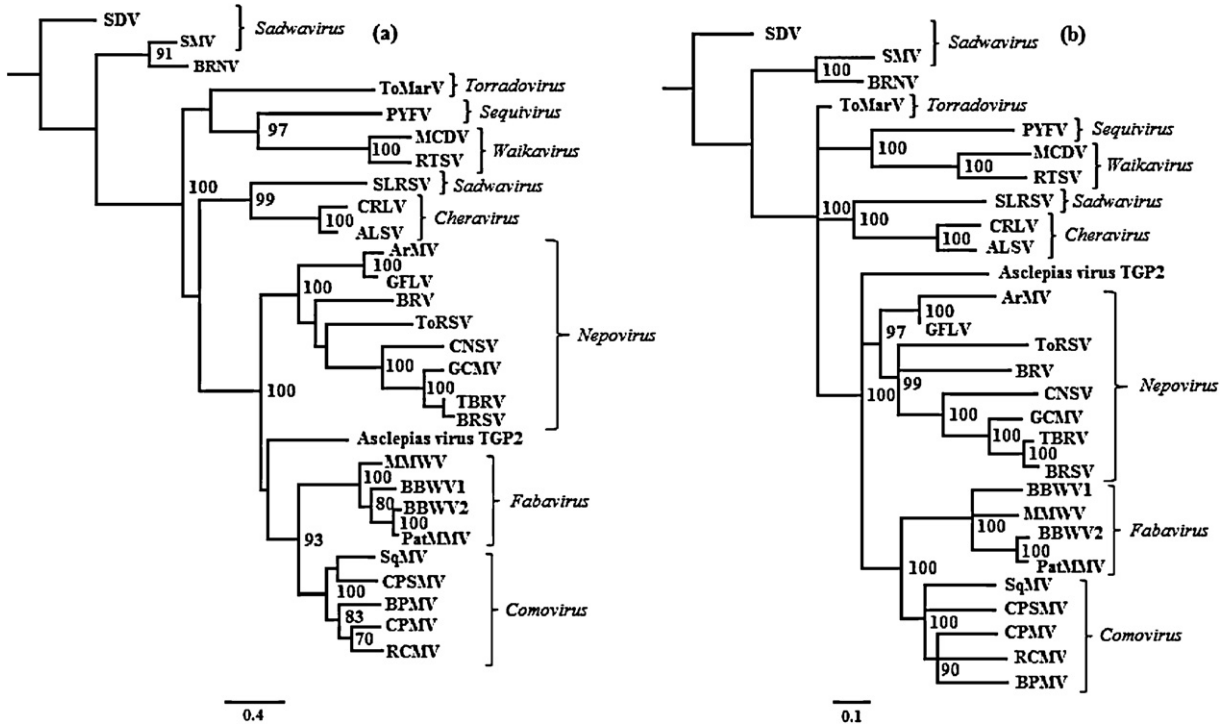

Fig. 1. Maximum Likelihood tree with Le and Gascuel (2008) protein substitution model (a) and Neighbor-Joining tree (b) constructed based on RNA-dependent RNA polymerase (RdRp) of RNA-1 of Asclepias virus TGP2 and members of *Secoviridae* including members of *Comovirinae* (*Comovirus*, *Fabavirus*, *Nepovirus*) that had significant hits in BLAST search. The nodes with bootstrap support >70% are provided in the trees. Refer to Table 3 for the full names of plant viruses abbreviated at the tips of branches.

to the same species, the BPMV-TGP1 is most likely a strain of BPMV since, although the entire coat protein coding region nucleotide sequence is not available, the portions that were well within the range set and the amino acid similarity was above 90%. For the waikavirus MCDV-TGP1, the 79.5% identity with MCDV is within the 70–100% identity range for viruses of the same species. Thus, these are two of less than 20 known viruses that could be identified among the several hundred viruses for which evidence was found (unpublished observations) in the PVBE survey.

Phylogenetic analysis of the conserved region of the RdRp of RNA 1 supports Asclepias virus TGP2 as a member of the *Comovirinae* but in a lineage separate from the three existing genera *Comovirus*, *Fabavirus* and *Nepovirus*. However, the NJ and ML trees based on RNA 2 and rooted with the sequences from *Satsuma dwarf virus*

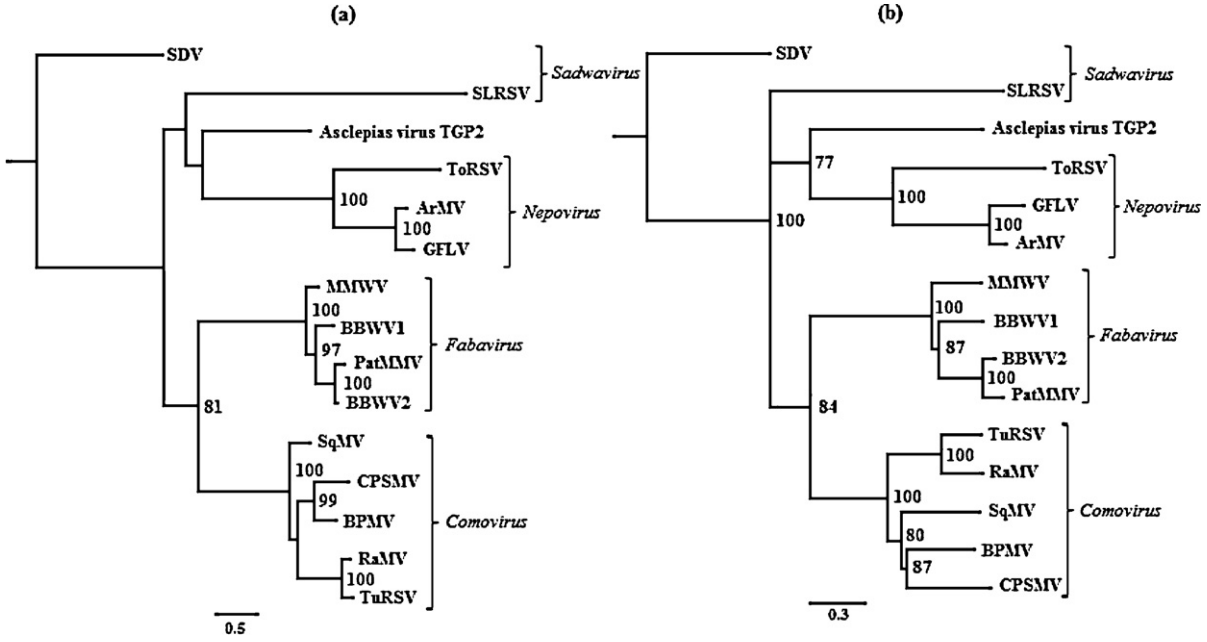

Fig. 2. Maximum Likelihood tree with Le and Gascuel (2008) protein substitution model (a) and Neighbor-Joining tree (b) constructed based on polyprotein of RNA 2 of Asclepias virus TGP2 and members of *Secoviridae* including members of *Comovirinae* (*Comovirus*, *Fabavirus*, *Nepovirus*) that had significant hits in BLAST search. The nodes with bootstrap support >70% are provided in the trees. Refer to Table 4 for the full names of plant viruses abbreviated at the tips of branches.

and *Strawberry latent ringspot virus* did not provide support for unambiguous placement of Asclepias virus TGP2 in the *Comovirinae*. In many phylogenetic studies of *Comovirinae* (*Comoviridae*, *sensu stricto*) members, RdRp sequence from RNA 1 provides convincing evidence of the evolutionary relationships that can also be supported from genome organization, sequence identity and biology of a virus (Li et al., 2000; Koh et al., 2001; Kobayashi et al., 2005; Petrzik et al., 2005; Le Gall et al., 2008). Moreover, BLAST searches and phylogenetic trees based on RdRp among members of *Secoviridae* show that the RdRp is not only well conserved within the group but also has sufficient information to depict resolution at family, subfamily and generic levels in the phylogenetic analysis (Sanfaçon et al., 2009). The findings indicate that the RdRp domain of RNA 1 is better in explaining phylogenetic relationships within the family *Secoviridae* than the information from RNA 2.

In the majority of the literature using RNA 2 of *Comovirinae* (*Comoviridae*, *sensu stricto*) in building phylogenies, the coat protein has been a main target for analysis (Chen and Bruening, 1992; Petrzik and Koloniuk, 2010). However, this literature suggests that the coat protein has better phylogenetic resolution for species within a genus than across genera. For Asclepias virus TGP2, we found it difficult to align its coat protein sequence from the conserved domain with other members, indicating high divergence. However, the alignment of the RNA 2 using the polyprotein of Asclepias virus TGP2 with related members (obtained by PSI-BLAST hit) showed more coverage of taxa and higher sequence identity than did the alignment derived from its coat protein, although the clades in the NJ and ML phylogenetic trees showed no consistent support for these observations. The results indicate that sequence divergence of RNA 2 of Asclepias virus TGP2 with related members was large enough to provide phylogenetic resolution at subfamily and genus level. Since, currently, RdRp phylogeny is the most reliable in explaining the evolutionary relationship of Asclepias virus TGP2 with known *Comovirinae* members from cultivated hosts we are led to predict that Asclepias virus TGP2 will eventually be recognized as a member of the subfamily. That two of the five criteria (Sanfaçon et al., 2011) for classification in a new genus (additional ORFs and phylogenetic separation) are met by this sequence, this virus may also become the first recognized member of a new genus. However, further information about its genome organization and biology will be necessary to assign it definitively to one of the three existing genera or to a new genus.

4.4. Viruses of non-cultivated plants

Although Asclepias virus TGP2 occurs as a separate branch from the three existing genera within the *Comovirinae* clade, the subfamily itself forms a monophyletic group. Therefore the divergent lineage of Asclepias virus TGP2 in the group may be due to its association with non-cultivated plant hosts in natural environments that are different from *Comovirinae* hosts under cultivation. Such divergence has been found in other groups, including a divergent lineage of a geminivirus infecting a native grass, *Eragrostis curvula* in a natural setting from South Africa (Varsani et al., 2009). Varsani et al. (2009) proposed the divergent lineage of the virus from native hosts within *Geminiviridae* as evidence of ancient characteristics of the family. This may be the case with Asclepias virus TGP2, but further support from other *Comovirinae* members from natural settings are necessary to confirm this supposition. This suggestion is supported by the overlapping ORF of RNA 2, an overlap not observed in any other known member of the subfamily. Consistent with this view is that analysis of other TPP viral sequences also yielded unusual genome features. The PafMV-TGP sequence has an additional ORF, partially encoded by extra nucleotides and partly by overprinting of the N-terminus of the movement protein ORF (Stobbe et al., 2012). The encoded polypeptide, such as the one from Asclepias virus TGP2, is rich in cysteine residues. The TPP's Ambrosia asymptomatic virus 1 has a 3′ terminal ORF putatively encoding a nucleic acid binding protein has a unique genomic organization among known members of the *Alphaflexiviridae* (Melcher et al., 2008).

Acknowledgments

The work was supported by funding from NSF-EPSCoR (EPS-0447262), NSF (IOS-0639139) and the Oklahoma Agricultural Experiment Station whose Director has approved the manuscript for publication. The authors thank Plant Virus Biodiversity and Ecology colleagues for gathering of materials and assistance in nucleic acid preparation and Kay Scheets and Jeanmarie Verchot-Lubicz for critiques of a draft of the manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.virusres.2012.03.016.

References

Abascal, F., Zardoya, R., Posada, D., 2005. ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21, 2104–2105.
Allen, M.S., Hamilton, R.G., Melcher, U., Palmer, M.W., 2009. Lessons from the Prairie: Research at The Nature Conservancy's Tallgrass Prairie Preserve. Oklahoma Academy of Science, Edmond, OK.
Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research 25, 3389–3402.
Chen, X.J., Bruening, G., 1992. Nucleotide-sequence and genetic-map of Cowpea severe mosaic-virus RNA 2 and comparisons with RNA 2 of other comoviruses. Virology 187, 682–692.
Cooper, I., Jones, R.A.C., 2006. Wild plants and viruses: under-investigated ecosystems. Advances in Virus Research 67, 1–47.
Drummond, A.J., Ashton, B., Cheung, M., Heled, J., Kearse, M., Moir, R., Stones-Havas, S., Thierer, T., Wilson, A., 2009 Geneious v. 4.8.5, http://www.geneious.com.
Guindon, S., Gascuel, O., 2003. A simple fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Systematic Biology 52, 696–704.
Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41, 95–98.
Hamilton, R.G., 2007. Restoring heterogeneity on the Tallgrass Prairie Preserve: applying the fire-grazing interaction model. In: Masters, R.E., Galley, K.E.M. (Eds.), Proceedings of the 23rd Tall Timbers Fire Ecology Conference: Fire in Grassland and Shrubland Ecosystems. Tall Timbers Research Station, Tallahassee, pp. 163–169.
Kobayashi, Y.O., Kobayashi, A., Hagiwara, K., Uga, H., Mikoshiba, Y., Naito, T., Honda, Y., Omura, T., 2005. Gentian mosaic virus: a new species in the genus *Fabavirus*. Phytopathology 95, 192–197.
Koh, L.H., Cooper, J.I., Wong, S.M., 2001. Complete sequences and phylogenetic analyses of a Singapore isolate of broad bean wilt fabavirus. Archives of Virology 146, 135–147.
Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and clustal X version 2.0. Bioinformatics 23, 2947–2948.
Le Gall, O., Christian, P., Fauquet, C.M., King, A.M.Q., Knowles, N.J., Nakashima, N., Stanway, G., Gorbalenya, A.E., 2008. Picornavirales a proposed order of positive-sense single-stranded RNA viruses with a pseudo-T=3 virion architecture. Archives of Virology 153, 715–727.
Le, S.Q., Gascuel, O., 2008. An improved general amino acid replacement matrix. Molecular Biology and Evolution 25, 1307–1320.
Li, C.J., Yoshikawa, N., Takahashi, T., Ito, T., Yoshida, K., Koganezawa, H., 2000. Nucleotide sequence and genome organization of Apple latent spherical virus: a new virus classified into the family Comoviridae. Journal of General Virology 81, 541–547.
Lisa, V., Boccardo, G., 1996. Fabaviruses: broad bean wilt virus and allied viruses. In: Harrison, B.H., Murant, A.F. (Eds.), Polyhedral Virions and Bipartite RNA Genomes: The Plant Viruses, vol. 5. Plenum Press, New York, pp. 229–250.
Marchler-Bauer, A., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., He, S., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Liebert, C.A., Liu, C., Lu, F., Lu, S., Marchler, G.H., Mullokandov, M., Song, J.S., Tasneem, A., Thanki, N., Yamashita, R.A., Zhang, D., Zhang, N., Bryant, S.H., 2009. CDD: specific functional annotation with the Conserved Domain Database. Nucleic Acids Research 37, D205–D210.
Mayo, M.A., Fritsch, C., 1994. A possible consensus sequence for VPg of viruses in the family Comoviridae. FEBS Letters 354, 129–130.

Melcher, U., Muthukumar, V., Wiley, G.B., Min, B.E., Palmer, M.W., Verchot-Lubicz, J., Nelson, R.S., Roe, B.A., Ali, A., Thapa, V., Pierce, M.L., 2008. Evidence for novel viruses by analysis of nucleic acids in virus-like particle fractions from *Ambrosia psilostachya*. Journal of Virological Methods 152, 49–55.

Min, B.-E., Feldman, T., Wiley, G.B., Muthukumar, V., Roe, B.A., Palmer, M.W., Ali, A., Roossinck, M.J., Melcher, U., Nelson, R.S., 2012. Physiological and ecological analysis of a novel tymovirus in *Asclepias viridis* (milkweed) from the Tallgrass Prairie Preserve of Oklahoma. Phytopathology 102, 166–176.

Muthukumar, V., Melcher, U., Pierce, M.L., Wiley, G.B., Roe, B.A., Palmer, M.W., Thapa, V., Ali, A., Ding, T., 2008. Non-cultivated plants of the Tallgrass Prairie Preserve of northeastern Oklahoma frequently contain virus-like sequences in particulate fractions. Virus Research 141, 169–173.

Nei, M., Gojobori, T., 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Molecular Biology and Evolution 3, 418–426.

Palmer, M.W., 2007. The vascular flora of the Tallgrass Prairie Preserve, Osage County, Oklahoma. Castanea 72, 235–246.

Petrzik, K., Hola, M., Spak, J., 2005. Complete nucleotide sequence of Radish mosaic virus RNA polymerase gene and phylogenetic relationships in the genus *Comovirus*. Acta Virologica 49, 271–275.

Petrzik, K., Koloniuk, I., 2010. Emerging viruses in the genus *Comovirus*. Virus Genes 40, 290–292.

Roossinck, M.J., Saha, P., Wiley, G.B., Quan, J., White, J.D., Lai, H., Chavarría, F., Shen, G., Roe, B.A., 2010. Ecogenomics: using massively parallel pyrosequencing to understand virus ecology. Molecular Ecology 19 (Suppl. 1), 18–88.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4, 406–425.

Sandoval, C.R., Ramsdell, D.C., Hancock, J.F., 1995. Infection of wild and cultivated *Vaccinium* spp. with Blueberry leaf mottle nepovirus. Annals of Applied Biology 126, 457–464.

Sanfaçon, H., Wellink, J., Le Gall, O., Karasev, A., Vlugt, R.V., Wetzel, T., 2009. *Secoviridae*: a proposed family of plant viruses within the order Picornavirales that combines the families Sequiviridae and Comoviridae the unassigned genera Cheravirus and Sadwavirus, and the proposed genus Torradovirus. Archives of Virology 154, 899–907.

Sanfaçon, H., Iwanami, T., Karaev, A.V., van der Vlugt, T., Wellink, J., Wetzel, T., Yoshikawa, N., 2011. Secoviridae. In: King, A.M.Q., Adams, M.J., Carstens, E.B., Lefkowitz, E.J. (Eds.), Virus Taxonomy: Classification and Nomenclature of Viruses: Ninth Report of the International Committee on Taxonomy of Viruses. Elsevier, San Diego, pp. 979–1001.

Schäffer, A.A., Aravind, L., Madden, T.L., Shavirin, S., Spouge, J.L., Wolf, Y.I., Koonin, E.V., Altschul, S.F., 2001. Improving the accuracy of PSI-BLAST protein database searches with composition-based statistics and other refinements. Nucleic Acids Research 29, 2994–3005.

Scheets, K., Blinkova, O., Melcher, U., Palmer, M.W., Wiley, G.B., Ding, T., Roe, B.A., 2011. Detection of members of the Tombusviridae in the Tallgrass Prairie Preserve, Osage County, Oklahoma, USA. Virus Research 160, 256–263.

Stobbe, A.H., Melcher, U., Palmer, M.W., Roossinck, M.J., Shen, G., 2012. Co-divergence and host-switching in the evolution of tobamoviruses. Journal of General Virology 93, 408–418.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24, 1596–1599.

Varsani, A., Shepherd, D.N., Dent, K., Monjane, A.L., Rybicki, E.P., Martin, D., 2009. A highly divergent South African geminivirus species illuminates the ancient evolutionary history of this family. Virology Journal 6, 36.

Wallace, I.M., O'Sullivan, O., Higgins, D.G., Notredame, C., 2006. M-Coffee: combining multiple sequence alignment methods with T-Coffee. Nucleic Acids Research 34, 1692–1699.

Wren, J.D., Roossinck, M.J., Nelson, R.S., Scheets, K., Palmer, M.W., Melcher, U., 2006. Plant virus biodiversity and ecology. PLoS Biology 4, e80.