MODULE 1: **TITLE, AUTHORS, etc**

Code assigned:	*2014.008aP*	(to be completed by ICTV officers)

Short title: One new sequence-only species, *Trailing lespedeza virus 1,* in the family *Tombusviridae*
(e.g. 6 new species in the genus *Zetavirus*)

Modules attached (modules 1 and 9 are required)	1 ☒ 2 ☒ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☒

Author(s) with e-mail address(es) of the proposer:

Kay Scheets (kay.scheets@okstate.edu) and Ulrich Melcher (ulrich.melcher@okstate.edu)

List the ICTV study group(s) that have seen this proposal:

A list of study groups and contacts is provided at http://www.ictvonline.org/subcommittees.asp . If in doubt, contact the appropriate subcommittee chair (fungal, invertebrate, plant, prokaryote or vertebrate viruses)	Tombusviridae and Umbravirus Study Group

ICTV-EC or Study Group comments and response of the proposer:

SG comment: The decision to accept this proposal was unanimous.

EC comment: This proposal, which is related to the proposal suggesting the creation of the genus Pelarspovirus, was coded "Ud" because of concerns associated with the creation of the genus (see comments on the related proposal). However, the proposal for creation of the species could be approved this year if the proposal was modified to propose that the species remains unassigned in the family Tombusviridae or possibly be assigned to the current genus Carmovirus.

Response to EC comment: The proposal has been revised for inclusion in the family *Tombusviridae* but unassigned to a genus.

Date first submitted to ICTV:	June 18, 2014
Date of this revision (if different to above):	August 25, 2014

creating and naming one or more new species.
If more than one, they should be a group of related species belonging to the same genus. All new species must be placed in a higher taxon. This is usually a genus although it is also permissible for species to be "unassigned" within a subfamily or family. Wherever possible, provide sequence accession number(s) for one isolate of each new species proposed.

Code	*2014.008aP*	(assigned by ICTV officers)

To create 1 new species within:

		Fill in all that apply.
Genus:	**unassigned**	• If the higher taxon has yet to be created (in a later module, below) write "**(new)**" after its proposed name.
Subfamily:		
Family:	*Tombusviridae*	• If no genus is specified, enter "**unassigned**" in the genus box.
Order:		

Name of new species:	Representative isolate:	GenBank sequence accession number(s)
Trailing lespedeza virus 1	06TGP01091 (TGP 1)	HM640935

Reasons to justify the creation and assignment of the new species:
- Explain how the proposed species differ(s) from all existing species.
  - If species demarcation criteria (see module 3) have previously been defined for the genus, **explain how the new species meet these criteria**.
  - If criteria for demarcating species need to be defined (because there will now be more than one species in the genus), please state the proposed criteria.
- Further material in support of this proposal may be presented in the Appendix, Module 9

The following information shows that the sequence is of viral origin, it is complete, and the assembly has been verified. Additionally, the sequence has characteristics of viruses in the family *Tombusviridae.*

Source of metagenomic sequence data: This viral sequence was determined through the Plant Virus Biodiversity and Ecology project which used a metagenomic survey approach to identify plant viruses at The Nature Conservancy's Tallgrass Prairie Preserve (TPP), Osage County, OK (Wren et al., 2006). Collection information included GPS location, identification of plant species, and symptoms if present. The complete sequence was obtained from contig 30 of a 2006 sample of symptomless *Lespedeza procumbens* (trailing lespedeza), 06TGP01091 (Scheets et al., 2011). The sample was processed to isolate dsRNA, sequenced via Roche 454 pyrosequencing, assembled into contigs, and was identified as being viral in origin via BLASTn, tBLASTx, and BLASTx searches as described (Roossinck et al., 2010). Assembled virus-like contigs were used to search unassembled sequences for additional viral fragments (Scheets et al., 2011). Plant samples 05TGP00121 (*Pellaea atropurpurea*, symptomless) and 06TGP01062 (*Melilotus officinalis*, yellowing leaves) also contained contigs closely related to 06TGP01091 contig 30. Closely related sequences were also derived from a mixture of 8 species collected in 2006 from all aboveground plant tissue in a 0.1 m X 0.1 m plot, and *Lespedeza procumbens* was one of the species. We propose the name Trailing lespedeza virus TGP 1 (TLV TGP 1) for this virus which was initially identified (Scheets et al., 2011) as TGP carmovirus 1 (GenBank: HM640935).

Further editing of sequence data: The 3950 nt 06TGP01091 contig 30 was assembled from 794 reads and was 97.3% high quality nt. Two independent manual inspections of the sequence were performed by U Melcher and K Scheets. In the first inspection 8 linker bases were removed from the 5' end. The initial sequence did not show the most common type of tombusvirid replicase gene of a long ORF punctuated by a leaky stop codon, but the downstream replicase portion was present as a -1 frameshift. There was no classic XXXYYYZ shifty sequence within the expected region. One A was added to a string of four As at nt 747 to remove the frameshift which was about 8-9 codons upstream of the readthrough stop codon. Incorrectly counting strings of the same base is a common problem with 454 technology, and since this virus replicated to high titre, it was unlikely to be defective in replication. The resultant coding surrounding the readthrough region (FAK*GCL) is similar to sequences of other replicases in the family. When the program Sequencher was used on the sequence with its quality data, 2 low-quality nt from the 5' end and 12 nt (last 10 were low quality) from the 3' end were removed. The twelve 3' nt that were removed were the exact sequence for nt 2289-2300, which is the predicted 5' end of its sgRNA. The trimmed 3' end is CCUUGCCC which is similar to the 3' termini of PCRPV, PLPV, RrLDV and several carmoviruses (See Table 1 for virus abbreviations and Genbank numbers). The trimmed 5' end is also similar to carmoviruses. The final trimmed sequence was 97.8% high quality nt sequence data.

Sequences from a second virus TGP carmovirus 3 (TGP Car 3) with best BLAST hits to coat protein (CP) and/or RNA dependent RNA polymerase (RdRp) of PCRPV and PLPV were found in 7 of 80 samples of *Ambrosia psilostachya* (western ragweed) (Scheets et al., 2011). A consensus sequence of 3233 nt (~80% complete) was assembled for this virus which is missing the 5' UTR, most of ORF1, and the end of the 3'UTR (Scheets et al., 2011). The predicted CP and RdRp readthrough regions of TGP Car 3 show 34% and 45% identities, respectively, with TLV TGP 1 proteins, so it is not an isolate of TLV TGP 1. Since the sequence is incomplete we are not submitting it to ICTV for recognition. Its currently known genome characteristics are similar to TLV TGP 1.

The 3929 nt TLV TGP 1 genome organization is similar to that of carmoviruses, PLPV (Fig. 1), and other proposed members of the family *Tombusviridae* i.e. PCRPV (Scheets et al., 2011), RrLDV (Mollov, Lockhart, and Zlesak, 2013) PelRSV, and ELV with an 88 kDa RdRp produced through readthrough of ORF1 (28 kDa protein) and a 37 kDa CP at the 3' end. Two small overlapping ORFs between the RdRp ORF and CP ORF encode proteins similar to the movement proteins of carmoviruses, PLPV, PCRPV, PelRSV, ELV and RrLDV using a noncanonical CUG start codon for the 2nd ORF (Fig. 2) as shown for PLPV (Castaño, Ruiz, and Hernandez, 2009). Phylogenetic analysis of the replicase clearly shows TLV TGP 1 nested within the family *Tombusviridae* (Fig. 3). RNA structure predictions were performed on the 5' UTR, 5' UTR of a sgRNA predicted to initiate at nt 2289, and the 3' UTR (Scheets et al., 2011). The predicted 3' UTR contains three hairpins and a pseudoknot predicted to form in10 carmoviruses and PLPV (Scheets et al., 2011; Yuan et al., 2010). A long I-shaped stem-loop in the 3' UTR is predicted to act as a cap-independent translational enhancer (CITE) whose terminal loop can base-pair with kissing loops in the 5' UTRs of vRNA and sgRNA (Scheets et al., 2011; Simon and Miller, 2013). All of these characteristics indicate it is a member of the family *Tombusviridae*.

In August 2014 TLV TGP 1 RdRp and CP were analyzed by BLASTp against GenBank reference proteins. Highest identities for replicases in the top 19 BLASTp scores were PLPV, RrLDV, and NLVCV (45% identity) followed by TCV, AnFBV and HoRV (43-42% identities). Similarly for CP, within the top 14 BLASTp scores PFBV (40%) and NLVCV (38%) had highest identities followed by PLPV, AnFBV and CarMV (36%). These viruses are

either not yet formally recognized by ICTV (RrLDV), unassigned to a genus (PLPV), or members of the genus *Carmovirus*, which currently has species demarcation limits of <57% for replicase and <52% for CP (Rochon et al 2012). Thus the sequence of TLV TGP 1 indicates it represents a unique species in the family *Tombusviridae*, but it does not fit into any current genus. After acceptance of this proposal the GenBank entries NC_015227 and HM640935 will be updated to reflect the name change and status change.

MODULE 9: **APPENDIX**: supporting material
additional material in support of this proposal

References:

Castaño, A., Ruiz, L., and Hernandez, C. (2009). Insights into the translational regulation of biologically active open reading frames of Pelargonium line pattern virus. *Virology* **386**(2), 417-26.

Mollov, D., Lockhart, B., and Zlesak, D. C. (2013). Complete nucleotide sequence of rosa rugosa leaf distortion virus, a new member of the family *Tombusviridae*. *Arch Virol* **158**(12)**,** 2617-20.

Rochon, D., Rubino, L., Russo, M., Martelli, G.P., Lommel, S., 2012. Tombusviridae, 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. 1111-1138.

Roossinck, M. J., Saha, P., Wiley, G. B., Quan, J., White, J. D., Lai, H., Chavarría, F., Shen, G., and Roe, B. A. (2010). Ecogenomics: using massively parallel pyrosequencing to understand virus ecology. *Molecular Ecology* **19**(s1)**,** 81-88.

Scheets, K., Blinkova, O., Melcher, U., Palmer, M. W., Wiley, G. B., Ding, T., and Roe, B. A. (2011). Detection of members of the *Tombusviridae* in the Tallgrass Prairie Preserve, Osage County, Oklahoma, USA. *Virus Research* **160**(1-2)**,** 256-63.

Simon, A. E., and Miller, W. A. (2013). 3' cap-independent translation enhancers of plant viruses. *Annu Rev Microbiol* **67,** 21-42.

Wren, J. D., Roossinck, M. J., Nelson, R. S., Scheets, K., Palmer, M. W., and Melcher, U. (2006). Plant virus biodiversity and ecology. *PLoS biology* **4**(3)**,** e80.

Yuan, X., Shi, K., Young, M. Y., and Simon, A. E. (2010). The terminal loop of a 3' proximal hairpin plays a critical role in replication and the structure of the 3' region of Turnip crinkle virus. *Virology* **402**(2)**,** 271-80.

Annex:
Include as much information as necessary to support the proposal, including diagrams comparing the old and new taxonomic orders. The use of Figures and Tables is strongly recommended but direct pasting of content from publications will require permission from the copyright holder together with appropriate acknowledgement as this proposal will be placed on a public web site. For phylogenetic analysis, try to provide a tree where branch length is related to genetic distance.
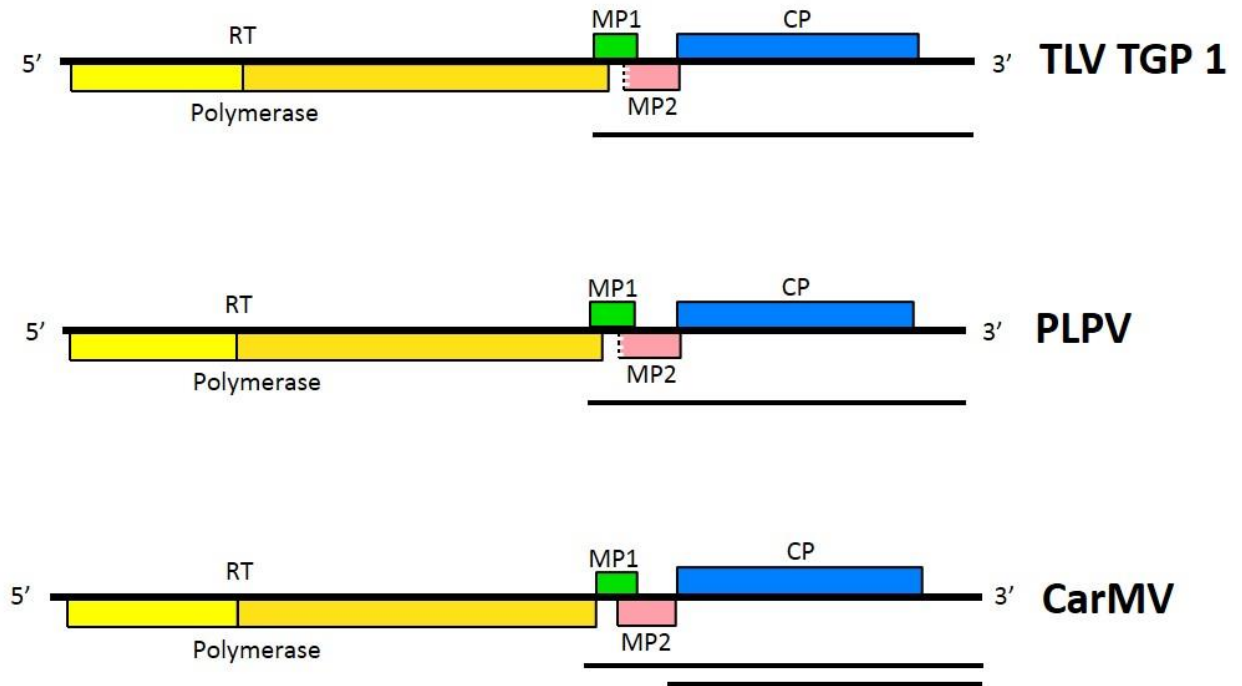
Figure 1. Genome organization of Trailing lespedeza virus TGP 1 (TLV TGP 1), *Pelargonium line pattern virus* (PLPV), and *Carnation mottle virus* (CarMV). PLPV is an unassigned virus in the family *Tombusviridae*. Thin lines under genomes indicate sgRNAs. Dashed line of MP2 indicates noncanonical start codon. Virus abbreviations and accession numbers are listed at the end of this annex.

```
          R   R   N   D   K   P
       AGCGUCGAAAUGAUAAACCAC                      CPMoV            Carmovirus
                M   I   N   H


          A   H   H   A   V   D
       UUGCCCACCAUGCCGUCGACA                      SYMMV
                M   P   S   T



       A   H   S   G   V   S
     GCUCACAGUGGAGUAUCCA                          PLPV
              M   E   Y   P


       A   H   S   G   V   S
     GCCCACAGUGGAGUCAGCA                          PCRPV
              M   E   S   A


       A   H   S   G   V   A
     GCACACUCUGGAGUUGCCA                          RrLDV          unassigned
              M   E   L   P


       A   H   S   G   V   S
     GCUCAUUCUGGCGUCAGCA                          TGPCar3
              M   A   S   A


   T   K   N   G   R   A   V   A   L
 ACAAAGAACGGACGAGCUGUCGCACUCA                     TLV TGP1
          M   D   E   L   S   H   S
                      M   S   H   S


   F   N   F   N
 UUCAAUUUCAACUGAGCUGGAGUGUGUG                     MCMV           Machlomovirus
                  M   E   C   V


   N   F   N   F
 AACUUCAAUUUCUAGUGGCGACCGGC                       PMV
                M   A   T   G


   N   F   N   F
 AACUUCAACUUCUAGCUGGCAACAGGC                      TPAV           Panicovirus
                 M   A   T   G


   F   N   F   G
 UUCAACUUCGGAUAAACCUGGCUACCGGC                    CMMV
                M   A   T   G
```

Figure 2. Known or predicted MP2 noncanonical start codon (green) contexts for several *Tombusviridae* members and unassigned viruses with similar MP1 and MP2 sequences. MP1 sequences are above the nt sequences.
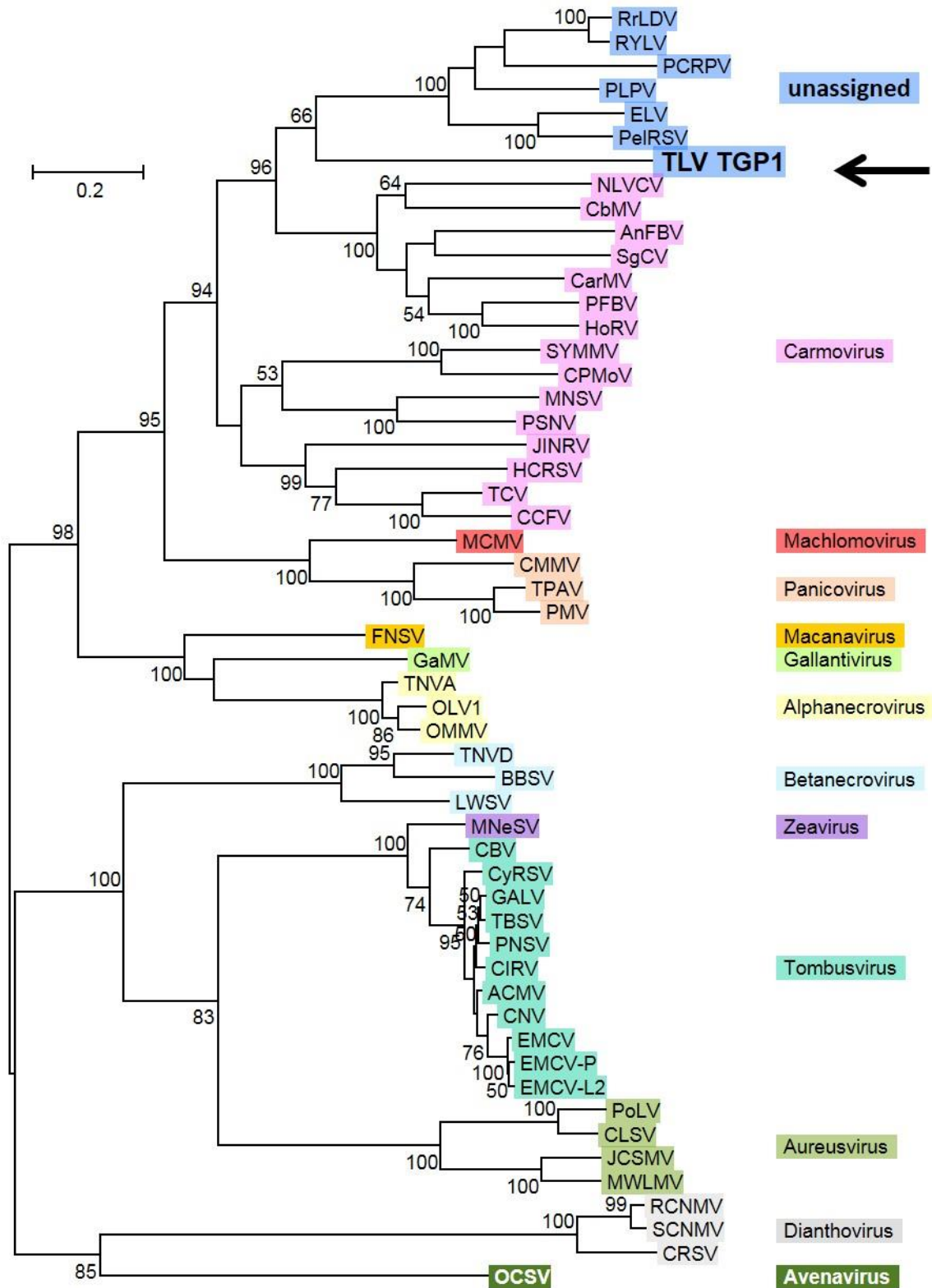
Figure 3. Phylogenetic (distance) analysis of the RdRps of *Tombusviridae* members. **TLV TGP1** clusters with unassigned tombusvirid or unassigned viruses (mid-blue). Alignments were made using Clustal Omega while trees were generated with the Maximum Likelihood algorithm using 1000 bootstrap replicates (showing values >50%).

Table 1. ABBREVIATIONS USED IN THE PROPOSAL

Alphanecrovirus

OLV1	*Olive latent virus 1*	X85989
OMMV	*Olive mild mosaic virus*	AY616760
TNVA	*Tobacco necrosis virus A*	M33002

Aureusvirus

CLSV	*Cucumber leaf spot virus*	EU127904
JCSMV	*Johnsongrass chlorotic stripe mosaic virus*	AJ557804
MWLMV	*Maize white line mosaic virus*	EF589670
PoLV	*Pothos latent virus*	X87115

Avenavirus

OCSV	*Oat chlorotic stunt virus*	X83964

Betanecrovirus

BBSV	*Beet black scorch virus*	AF452884
LWSV	*Leek white stripe virus*	X94560
TNVD	*Tobacco necrosis virus D*	U62546

Carmovirus

AnFBV	*Angelonia flower break virus*	DQ219415
CbMV	*Calibrachoa mottle virus*	GQ244431
CCFV	*Cardamine chlorotic fleck virus*	L16015
CarMV	*Carnation mottle virus*	X02986
CPMoV	*Cowpea mottle virus*	U20976
HCRSV	*Hibiscus chlorotic ringspot virus*	X86448
HoRSV	*Honeysuckle ringspot virus*	HQ677625
JINRV	*Japanese iris necrotic ring virus*	D86123
MNSV	*Melon necrotic spot virus*	M29671
NLVCV	*Nootka lupine vein clearing virus*	EF207438
PSNV	*Pea stem necrosis virus*	AB086951
PFBV	*Pelargonium flower break virus*	AJ514833
SgCV	*Saguaro cactus virus*	U72332
SYMMV	*Soybean yellow mottle mosaic virus*	FJ457015
TCV	*Turnip crinkle virus*	M22445

Dianthovirus

CRSV	*Carnation ringspot virus*	L18870, M8858
RCNMV	*Red clover necrotic mosaic virus*	J04357, X08021
SCNMV	*Sweet clover necrotic mosaic virus*	L07884, S4602

Gallantivirus

GaMV	*Galinsoga mosaic virus*	Y13463

continued

	Macanavirus	
FNSV	*Furcraea necrotic streak virus*	FJ768020
	Machlomovirus	
MCMV	*Maize chlorotic mottle virus*	X14736
	Panicovirus	
CMMV	*Cocksfoot mild mosaic virus*	EU081018
PMV	*Panicum mosaic virus*	U55002
TPAV	*Thin paspalum asymptomatic virus* (proposed)	JX848617
	Tombusvirus	
AMCV	*Artichoke mottled crinkle virus*	X62493
CIRV	*Carnation Italian ringspot virus*	X85215
CBV	*Cucumber Bulgarian virus*	AY163842
CNV	*Cucumber necrosis virus*	M25270
CyRSV	*Cymbidium ringspot virus*	X15511
EMCV	*Eggplant mottled crinkle virus*	JQ864181
EMCV-P	*Eggplant mottled crinkle virus*-P (pear latent virus)	AY100482
EMCV-L	*Eggplant mottled crinkle virus*-L (lisianthus necrosis virus)	DQ011234
GALV	*Grapevine Algerian latent virus*	AY830918
PNSV	*Pelargonium necrotic spot virus*	AJ607402
TBSV	*Tomato bushy stunt virus*	M21958
	Zeavirus	
MNeSV	*Maize necrotic streak virus*	AF266518
	unassigned tombusvirid	
PLPV	*Pelargonium line pattern virus*	AY613852
	unassigned viruses	
ELV	elderberry latent virus	AY038066
PCRPV	pelargonium chlorotic ring pattern virus	AY038069
PelRSV	pelargonium ringspot virus	AY038068
RrLDV	rosa rugosa leaf distortion virus	KC166238
RYLV	rose yellow leaf virus	KC166239
TLV TGP1	*Trailing lespedeza virus* TGP 1 (proposed tombusvirid)	HM640935