

SHORT REPORT

Open Access



# The complete genome of the tospovirus *Zucchini lethal chlorosis virus*

R. N. Lima, A. S. De Oliveira, M. O. Leastro, R. Blawid, T. Nagata, R. O. Resende and F. L. Melo\*

## Abstract

**Background:** *Zucchini lethal chlorosis virus* (ZLCV) causes significant losses in the production of cucurbits in Brazil. This virus belongs to the genus *Tospovirus* (family *Bunyaviridae*) and seems to be exclusively transmitted by *Frankliniella zucchini* (Thysanoptera). Tospoviruses have a tripartite and single-stranded RNA genome classified as S (Small), M (Medium) and L (Large) RNA<sub>s</sub>. Although ZLCV was identified as a member of the genus *Tospovirus* in 1999, its complete genome had not been sequenced until now.

**Findings:** We sequenced the full-length genome of two ZLCV isolates named ZLCV-SP and ZLCV-DF. The phylogenetic analysis showed that ZLCV-SP and ZLCV-DF clustered with the previously reported isolate ZLCV-BR09. Their proteins were closely related, except the non-structural protein (NSm), which was highly divergent (approximately 90 % identity). All viral proteins clustered similarly in our phylogenetic analysis, excluding that these ZLCV isolates have originated from reassortment events of different tospovirus species.

**Conclusion:** Here we report for the first time the complete genome of two ZLCV isolates that were found in the field infecting zucchini and cucumber.

**Keywords:** Zucchini, *Tospovirus*, Illumina, Genome, ZLCV

## Body of text

*Zucchini lethal chlorosis virus* (ZLCV) is a member of the genus *Tospovirus*, family *Bunyaviridae* [1]. Although some tospovirus species are notorious for their broad-host range [2], ZLCV mainly infects cucurbits and is only known to be transmitted by *Frankliniella zucchini* (Thysanoptera: Thripidae) [3]. So far ZLCV has just been reported in Brazil, naturally infecting sponge gourd, West Indian gherkin, cucumber, watermelon and several species of squash [4, 5]. Infection by ZLCV abrogate fruit production in zucchini plants and the only resistant cultivar is *Cucurbita maxima* cv. Exposição [6], making ZLCV one of the most economically important viral pathogen for cucurbits in Brazil.

Tospoviruses have a tripartite single-stranded RNA genome and each segment is named according to its size. The L (large) RNA has a negative polarity and encodes a RNA-dependent RNA polymerase (RdRp) [7]. The ambisense M (medium) RNA encodes the precursor of two viral glycoproteins (Gn/Gc) and a non-structural

protein (NSm) involved in viral cell-to-cell movement [8]. The S (small) RNA, which is also ambisense, encodes another non-structural protein (NS<sub>s</sub>) with RNA silencing suppression activity and the nucleocapsid (NP) protein [9].

There are currently 11 approved and 18 tentative tospovirus species, but only a small number of species were completely sequenced. Genome sequence data has the potential to solve key questions in tospovirus evolution, epidemiology and physiology, such as the occurrence and importance of interspecific reassortment [10] and the presence of potential undescribed genes [11]. Even though ZLCV was described in 1999 [1] and some genes of one isolate has been sequenced [12–14], its complete genome is still unknown. Here, we report the complete genome of two ZLCV isolates found infecting zucchini (*Cucurbita pepo* cv. Caserta) and cucumber (*Cucumis sativus* L.) in Brazil and compared them to other tospoviruses.

In 2010, a virus isolate (hereafter ZLCV-SP) from zucchini was found in a commercial field in São Paulo state and transmitted to *Datura stramonium* L. by *F. zucchini* as previously described [3]. Then, virus particles

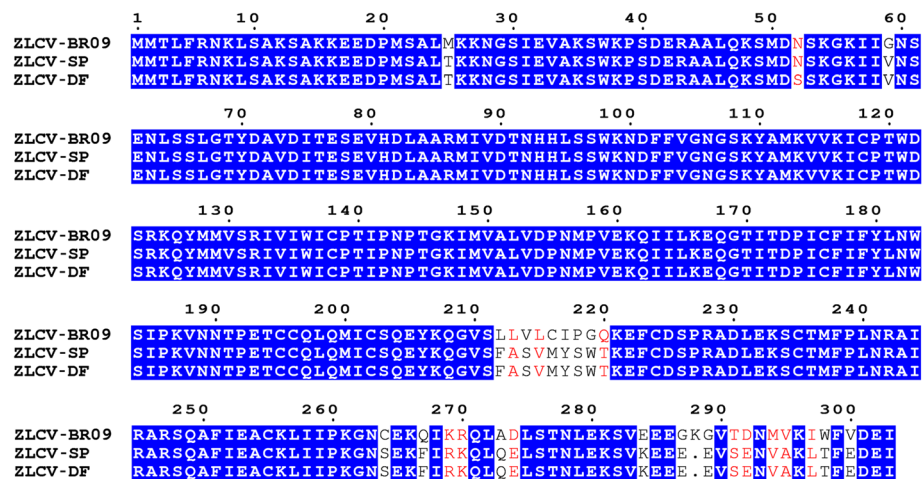
\* Correspondence: flmelo@unb.br  
Laboratory of Virology, Department of Cell Biology, University of Brasília, Brasília, DF 70910-900, Brazil

**Table 1** Genome comparison of ZLCV isolates

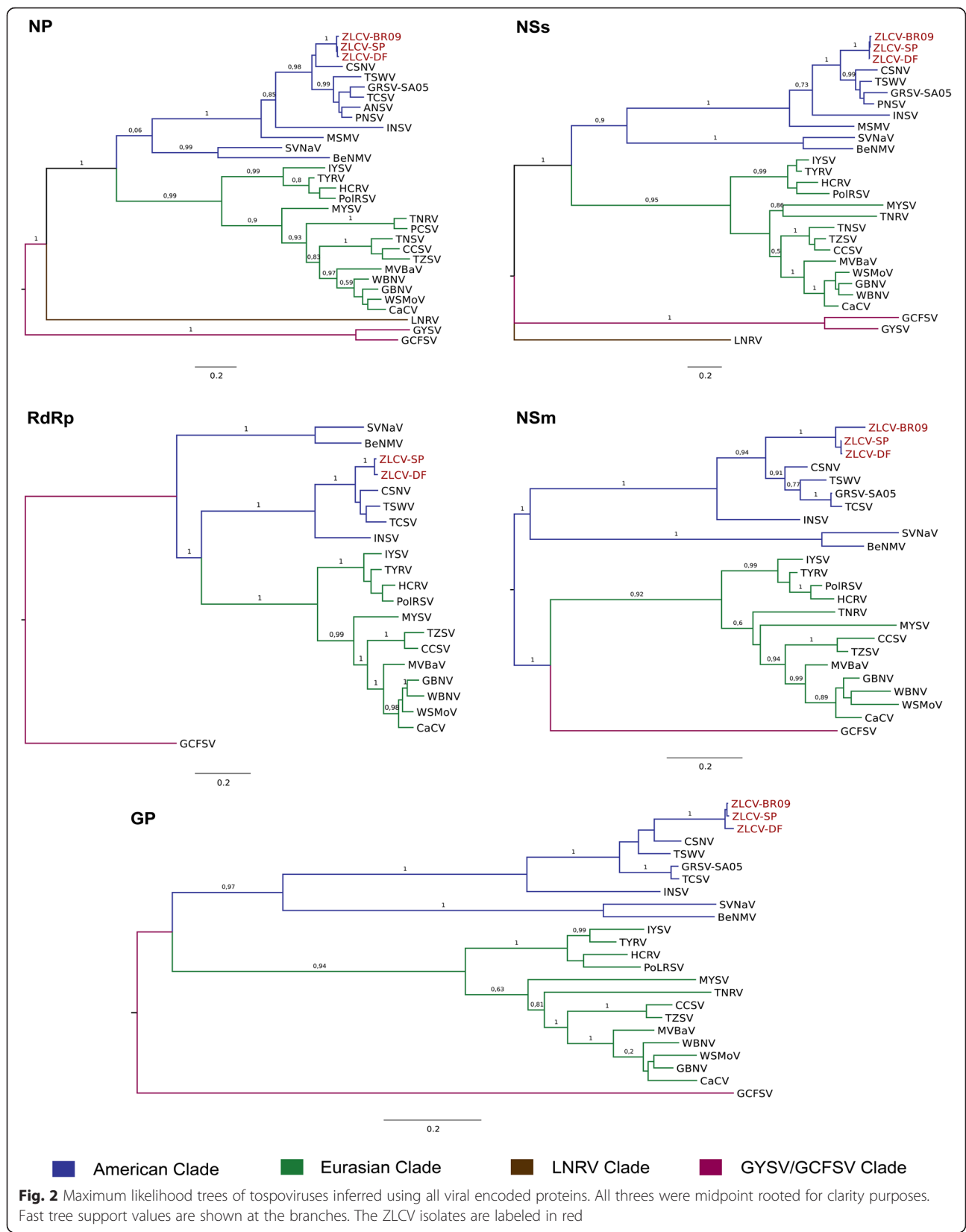
|                        | ZLCV-SP                  | ZLCV-DF           |
|------------------------|--------------------------|-------------------|
| <b>L RNA</b>           |                          |                   |
| L RNA full length (nt) | 8885                     | 8886              |
| 5'UTR (nt)             | 222 (1–222) <sup>a</sup> | 223 (1–223)       |
| L gene ORF (nt)        | 8631 (223–8853)          | 8631 (224–8854)   |
| L protein (aa)         | 2877 (330.85 kDa)        | 2877 (331.16 kDa) |
| 3'UTR (nt)             | 32 (8854–8885)           | 32 (8855–8886)    |
| <b>M RNA</b>           |                          |                   |
| M RNA full length (nt) | 4860                     | 4829              |
| 5'UTR (nt)             | 113 (1–113)              | 113 (1–113)       |
| NSm gene ORF (nt)      | 909 (114–1022)           | 909 (114–1022)    |
| NSm protein (aa)       | 303 (34.4 kDa)           | 303 (34.37 kDa)   |
| IGR (nt)               | 344 (1023–1366)          | 313 (1–1335)      |
| GPs gene ORF (nt)      | 3408 (1367–4774)         | 3408 (1336–4743)  |
| GP protein (aa)        | 1136 (127.58 kDa)        | 1136 (127.73 kDa) |
| 3'UTR (nt)             | 86 (4775–4860)           | 86 (4744–4829)    |
| <b>S RNA</b>           |                          |                   |
| S RNA full length (nt) | 3524                     | 3524              |
| 5'UTR (nt)             | 87 (1–87)                | 87 (1–87)         |
| NSs gene ORF (nt)      | 1404 (88–1491)           | 1404 (88–1491)    |
| NSs protein (aa)       | 468 (53.06 kDa)          | 468 (53.01 kDa)   |
| IGR (nt)               | 1009 (1492–2500)         | 1009 (1492–2500)  |
| NP gene ORF (nt)       | 783 (2501–3283)          | 783 (2501–3283)   |
| NP protein (aa)        | 261 (29.22 kDa)          | 261 (29.24 kDa)   |
| 3'UTR (nt)             | 241 (3284–3524)          | 241 (3284–3524)   |

nt = nucleotides  
<sup>a</sup> = position in the genome  
 aa = amino acids  
 kDa = kilodaltons  
 IGR = Intergenic region  
 GP = glycoproteins precursor

were propagated in *D. stramonium* by mechanical inoculation and infected leaves were used for ribonucleoprotein (RNP) purification following the protocol of De Avila et al. [15]. Moreover, cucumber plants showing typical ZLCV symptoms were collected in Planaltina, Federal District, in 2015. Viral particles were semi-purified from leaves as previously described [16]. Briefly, 40 g of leaf material were homogenized in PBS-EDTA plus 0.2 % 2-mercaptoethanol. The plant extract was then filtered and centrifuged through a sucrose cushion at 33,000 x g for 2 h and the pellet resuspended in PBS. Genomic RNA was extracted from purified RNPs of both isolates as previously described by De Oliveira et al. [17] and sequenced at Macrogen (South Korea) using Illumina HiSeq 2000 platform. The resulting paired-end reads were filtered and assembled *de novo* using CLC Genomics Workbench version 6.0.3. The contigs related to ZLCV were selected using Blastx against a RefSeq virus database. To determine if the entire length of each segment was included in the assembled contigs, the reads were mapped back to the ZLCV related contigs. All contigs from both isolates presented the consensus sequences AGAGCAAU and AUUGCUCU at the 5'- and 3'-terminal ends, but some contigs presented distal terminal bases that were trimmed off. These palindromic sequences are conserved among all tospoviruses. Moreover, the ZLCV segments derived from cucumber samples (hereafter ZLCV-DF) presented some unresolved gaps (L segment: 1 gap of 13 nucleotide and 1 of 2 nucleotide). The genome of both isolates were annotated and submitted to NCBI GenBank under the accession numbers no. KU641378-KU641380 (ZLCV-SP) and KU681010-KU681012 (ZLCV-DF).



**Fig. 1** The NSm amino acid sequence alignment of ZLCV isolates. Residues common to the three sequences are shown in white on blue and similar residues are shown in red. The yellow triangle indicates the single amino acid deletion and the black rectangle indicates the coiled-coil (E<sup>265</sup>-E<sup>285</sup>) predicted structure (Uniprot Accession code: Q91PB4). This figure was made with ESPript (<http://esprict.ibcp.fr/ESPript/ESPript/>)



Despite few variations, both genomes presented the same characteristics as listed in Table 1. The 5' and 3' UTR regions of both viruses presented the same size, except for one nucleotide in the 5'UTR of L segment of ZLCV-DF. The intergenic regions (IGR) presented small indels, for example, a duplication of 20 nucleotides in the M segment IGR of ZLCV-SP. Despite the importance of these regions as transcription termination signal [18], the impact of such deletions is still unclear. The L segments vary one nucleotide (as described above) between the isolates, 8,885 nt for ZLCV-SP and 8,886 nt for ZLCV-DF and encode a RdRp of about 331 kDa (2877 aa). The M segment contains 4,860 nt for ZLCV-SP and 4,829 nt for ZLCV-DF, and presented the ambisense arrangement typical of tospoviruses. For both isolates, an NSm protein of about 34 kDa (303 aa) and glycoprotein precursor (GP) protein of about 127 kDa (1136 aa) are predicted. The S segment is 3,524 nt long and encodes the NS<sub>S</sub> and NP proteins with predicted molecular sizes of about 53 kDa (468 aa) and 29 kDa (261 aa), respectively. Since some genes of other ZLCV isolate (ZLCV-BR09) have been previously sequenced [12–14], their encoded proteins were compared with ZLCV-SP and ZLCV-DF (Additional file 1: Table S1). Pairwise sequence comparison of the NP amino acids indicates that ZLCV-SP is most closely related to the ZLCV-DF than ZLCV-BR09. The identities between the NP of ZLCV-SP/ZLCV-DF and between ZLCV-SP/ZLCV-BR09 are 99.23 % and 98.46 %, respectively, while the identity of ZLCV-DF/ZLCV-BR09 is 97.69 %. The NSm from ZLCV-BR09 displayed a significant variation in comparison with ZLCV-SP and ZLCV-DF (approximately 90 % identity) (Additional file 1: Table S1). Interestingly, most variation was located at the C-terminus with an increased number of non-conservative amino acid changes (Fig. 1), suggesting that this region may be under positive selection. Actually, it was recently demonstrated that this same region is implicated in NSm and plasma-membrane interaction [19]. However, the biological impact of these variations remains to be determined.

To infer evolutionary relationships among tospoviruses, we compiled sequences of each tospovirus protein. The final data sets contain 33, 28, 23, 23 and 20 sequences for the NP, NS<sub>S</sub>, GP, NSm and RdRp proteins, respectively. The GenBank accession numbers are listed in Additional file 1: Table S2. The sequences were aligned using MAFFT [20] and the maximum likelihood trees were inferred using FastTree [21], both implemented in Geneious 9.1. As previously observed, tospoviruses can be divided into two geographic groups, with distinct viral species observed in Europe/Asia (Eurasian clade) and the Americas (American Clade) (Fig. 2). However, the analysis of NP and NS<sub>S</sub> trees indicate that some tospoviruses are clearly distinct from any species of these two groups, such as the

Lisianthus necrotic ringspot virus (LNRV), Groundnut chlorotic fan-spot virus (GCFSV) and the *Groundnut yellow spot virus* (GYSV) (Fig. 2). The ZLCV isolates clustered within the American clade and no reassortment had occurred during the evolution of these ZLCV isolates (Fig. 2).

Considering the high prevalence of tospoviruses worldwide, the complete genome of the ZLCV isolates is important for future surveillance and research. Additional investigations in important plant crops should keep being performed to extend the number of characterized species.

## Additional file

**Additional file 1: Table S1.** Percent identity among ZLCV isolates proteins. **Table S2.** GenBank accession number of the sequences used in this study. (DOCX 61 kb)

## Abbreviations

aa, amino acids; GP, glycoproteins precursor; IR, Intergenic region; kDa, kilodaltons; nt, nucleotides; RdRp, RNA-dependent RNA polymerase; RNP, ribonucleoprotein; ZLCV, *Zucchini lethal chlorosis virus*

## Acknowledgments

This work was supported by CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), and FAPDF (Fundação de Apoio à Pesquisa do Distrito Federal).

## Authors' contribution

FLM, MOL, RNL, ROR and TN conceived and designed the experiments. MOL and TN performed the experiments. ASO, FLM, RB and RNL analyzed the data. FLM, ROR and TN contributed reagents/materials/analysis tools. ASO, RNL and FLM wrote the paper. All authors read and approved the final manuscript.

## Competing interests

The authors declare that they have no competing interests.

Received: 7 March 2016 Accepted: 28 June 2016

Published online: 07 July 2016

## References

1. Bezerra IC, Resende RD, Pozzer L, Nagata T, Kormelink R, De Avila AC. Increase of tospoviral diversity in Brazil with the identification of two new tospovirus species, one from chrysanthemum and one from zucchini. *Phytopathology*. 1999;89:823–30.
2. Pappu HR, Jones RAC, Jain RK. Global status of tospovirus epidemics in diverse cropping systems: Successes achieved and challenges ahead. *Virus Res*. 2009;141:219–36.
3. Nakahara S, Monteiro RC. *Frankliniella zucchini* (Thysanoptera : Thripidae), a new species and vector of tospovirus in Brazil. *Proc Entomol Soc Wash*. 1999;101:290–4.
4. Yuki VA, Rezende JAM, Kitajima EW, Barroso PAV, Kuniyuki H, Groppo GA, Pavan MA. Occurrence, distribution, and relative incidence of five viruses infecting cucurbits in the state of Sao Paulo, Brazil. *Plant Dis*. 2000;84:516–20.
5. Camelo-García VM, Lima EFB, Rezende JAM. Identification of natural hosts of Zucchini lethal chlorosis virus. *Tropical Plant Pathology*. 2015;40:345–9.
6. Giampán JS, Rezende JAM, Piedade SMS. Yield loss caused by Zucchini lethal chlorosis virus (ZLCV) on zucchini squash 'Caserta'. *Summa Phytopathol*. 2009;35:223–5.
7. De Haan P, Kormelink R, Resende RD, Vanpoelwijk F, Peters D, Goldbach R. tomato spotted wilt virus-l rna encodes a putative rna-polymerase. *J Gen Virol*. 1991;72:2207–16.

8. Kormelink R, Storms M, Vanlent J, Peters D, Goldbach R. Expression and subcellular location of the nsm protein of tomato spotted wilt virus (tsww), a putative viral movement protein. *Virology*. 1994;200:56–65.
9. Takeda A, Sugiyama K, Nagano H, Mori M, Kaido M, Mise K, Tsuda S, Okuno T. Identification of a novel RNA silencing suppressor, NSs protein of Tomato spotted wilt virus. *Febs Letters*. 2002;532:75–9.
10. Webster CG, Reitz SR, Pery KL, Adkins S. A natural M RNA reassortant arising from two species of plant- and insect-infecting bunyaviruses and comparison of its sequence and biological properties to parental species. *Virology*. 2011;413:216–25.
11. Firth AE. Mapping overlapping functional elements embedded within the protein-coding regions of RNA viruses. *Nucleic Acids Res*. 2014;42:12425–39.
12. Silva MS, Martins CR, Bezerra IC, Nagata T, de Avila AC, Resende RO. Sequence diversity of NS(M) movement protein of tospoviruses. *Arch Virol*. 2001;146:1267–81.
13. Nagata T, Carvalho KR, Sodre RDA, Dutra LS, Oliveira PA, Noronha EF, Lovato FA, Resende RDO, De Avila AC, Inoue-Nagata AK. The glycoprotein gene of Chrysanthemum stem necrosis virus and Zucchini lethal chlorosis virus and molecular relationship with other tospoviruses. *Virus Genes*. 2007;35:785–93.
14. Hallwass M, Leastro MO, Lima MF, Inoue-Nagata AK, Resende RO. Sequence determination and analysis of the NSs genes of two tospoviruses. *Arch Virol*. 2012;157:591–6.
15. Deavila AC, Dehaan P, Smeets MLL, Resende RD, Kormelink R, Kitajima EW, Goldbach RW, Peters D. Distinct levels of relationships between tospovirus isolates. *Arch Virol*. 1993;128:211–27.
16. Silva KN, Melo FL, Orilio AF, Nagata T, Silva MS, Fernandes CD, Fragoso RR, Dessaune SN, Resende RO. Biological and molecular characterization of a highly divergent Johnsongrass mosaic virus isolate from Pennisetum purpureum. *Arch Virol*. 2016; Apr 21. [Epub ahead of print].
17. de Oliveira AS, Melo FL, Inoue-Nagata AK, Nagata T, Kitajima EW, Resende RO. Characterization of bean necrotic mosaic virus: a member of a novel evolutionary lineage within the Genus Tospovirus. *PLoS ONE*. 2012;7:e38634.
18. van Knippenberg I, Goldbach R, Kormelink R. Tomato spotted wilt virus S-segment mRNAs have overlapping 3' ends containing a predicted stem-loop structure and conserved sequence motif. *Virus Res*. 2005;110:125–31.
19. Singh P, Indi SS, Savithri HS. Groundnut Bud Necrosis Virus Encoded NSm Associates with Membranes via Its C-Terminal Domain. *Plos One*. 2014;9: e99370.
20. Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol Biol Evol*. 2013;30:772–80.
21. Price MN, Dehal PS, Arkin AP. FastTree 2 - Approximately maximum-likelihood trees for large alignments. *Plos One*. 2010;5:e9490.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

