Arch Virol (2016) 161:2311–2316
DOI 10.1007/s00705-016-2912-x

BRIEF REPORT

Analysis of the coding-complete genomic sequence of groundnut ringspot virus suggests a common ancestor with tomato chlorotic spot virus

Soledad de Breuil[1,2] · Joaquín Cañizares[3] · José Miguel Blanca[3] · Nicolás Bejerman[1,2] · Verónica Trucco[1] · Fabián Giolitti[1] · Peio Ziarsolo[3] · Sergio Lenardon[1]

Received: 19 February 2016 / Accepted: 24 May 2016 / Published online: 3 June 2016

Abstract Groundnut ringspot virus (GRSV) and tomato chlorotic spot virus (TCSV) share biological and serological properties, so their identification is carried out by molecular methods. Their genomes consist of three segmented RNAs: L, M and S. The finding of a reassortant between these two viruses may complicate correct virus identification and requires the characterization of the complete genome. Therefore, we present for the first time the complete sequences of all the genes encoded by a GRSV isolate. The high level of sequence similarity between GRSV and TCSV (over 90 % identity) observed in the genes and proteins encoded in the M RNA support previous results indicating that these viruses probably have a common ancestor.

The family *Bunyaviridae* comprises five genera: *Orthobunyavirus*, *Hantavirus*, *Nairovirus*, *Phlebovirus* and *Tospovirus*. Members of the genus *Tospovirus* are the only plant-infecting viruses of the family [18]. The tospovirus genome includes three RNAs, referred to as large (L), medium (M) and small (S). The L RNA is in the negative sense, while the M and S RNAs are ambisense. The L RNA codes for the RNA-dependent RNA polymerase (RdRp), also called L protein, and the M RNA codes for the precursor of two glycoproteins (GPp) and a non-structural protein (NSm). The S RNA codes for the nucleocapsid (N) protein and another non-structural protein (NSs). NSm and NSs have been shown to function as a movement protein and silencing suppressor, respectively [27]. The three genomic RNAs are tightly linked to the N protein to form ribonucleoproteins (RNPs). These RNPs are encapsidated within a lipid envelope consisting of the two virus-encoded glycoproteins (Gn and Gc) and a host-derived membrane.

Tospoviruses are distributed worldwide and include many economically important viruses that infect crops as well as weeds and native plants [17]. The genus takes its name from tomato spotted wilt virus (TSWV), which was discovered in Australia in 1915 and remained the only member of the family until after 1990, when genetic characterization of plant viruses became more common and other distinct members of the genus *Tospovirus* were identified species [9]. Species demarcation criteria for tospoviruses include genome structure and organization, molecular relationships of their N genes, and biological properties such as vector transmission (thrips) and host range. Of these, N protein amino acid (aa) sequence relationships are mainly used as the major descriptor for the establishment of new tospovirus species and for classification within different geographical groups. An identity level of 90 % is currently accepted as the species demarcation criterion [18].

The genetic diversity within and among tospoviruses has always been a subject of intense study. Previous studies suggest the possibility that viruses belonging to the family *Bunyaviridae* are the result of natural reassortments [5],

✉ Soledad de Breuil
debreuil.soledad@inta.gob.ar

[1] Instituto de Patología Vegetal, Centro de Investigaciones Agropecuarias, Instituto Nacional de Tecnología Agropecuaria (IPAVE-CIAP-INTA), Camino 60 Cuadras Km 5,5, X5020ICA Córdoba, Argentina

[2] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Godoy Cruz 2290, C1425FQB CABA, Argentina

[3] Instituto de Conservación y Mejora de la Agrodiversidad Valenciana, Universitat Politècnica de Valencia (COMAV-UPV), Camino de Vera s/n, 46022 Valencia, Spain

and several authors have discussed the possibility of natural reassortment events occurring among members of the genus *Tospovirus*. Most studies have focused on reassortments among isolates of TSWV [12, 15, 25, 26], but Webster *et al*. [28] provided evidence of the occurrence of a natural reassortant in Florida (USA) between two tospoviruses classified as members of different species: *Groundnut ringspot tospovirus* and *Tomato chlorotic spot tospovirus*. This reassortant was designated as $L_GM_TS_G$ because its genome is composed of the L and S RNAs coming from GRSV and the M RNA coming from TCSV [28]. Subsequent studies showed that $L_GM_TS_G$ is widespread in vegetable-production regions of South Florida, and it was then detected in South Carolina and Long Island (New York) [29]. The authors point out that this finding may be due to a higher negative selection being imposed on the M genomic RNA of these viruses, or it may be due to a more distant reassortment event between TCSV and GRSV. The presence of GRSV and TCSV has also been reported in other regions of South and Central America and the Caribbean [1, 17]. However, the lack of detection of non-reassorted GRSV isolates and the limited number of fully sequenced genes of GRSV and TCSV available in the GenBank database makes it difficult to determine which of these hypotheses is correct.

In this work, we present the results obtained from the analysis of the coding regions of the genome of an Argentinian GRSV isolate in order to help clarify the mechanism of molecular evolution of tospoviruses, and we discuss aspects involved in its taxonomic classification.

A naturally infected peanut plant displaying symptoms characteristic of a tospovirus infection (plant stunting, smaller and distorted leaves showing severe chlorosis and concentric ringspots) was collected in a commercial peanut field in Córdoba province (Argentina), maintained in a glasshouse at 20–24 °C, and used in the analysis described below. The presence of GRSV was confirmed by enzyme-linked immunosorbent assay (ELISA) and antigen-capture (AC) reverse transcription polymerase chain reaction (RT-PCR). ELISA was done using available commercial antisera for GRSV/TCSV and TSWV following the manufacturer's instructions (Agdia, Elkhart, IN). For AC, a symptomatic peanut leaf was ground in sterilized bicarbonate/carbonate coating buffer, diluted 1:200 w/v, and incubated overnight at 4 °C. RT-PCR amplification was performed with two GRSV-specific primers that anneal in the N gene according to Boari *et al*. [4] using an Access RT-PCR System (Promega Corp. Madison, WI, USA). The amplicon was visualized by gel electrophoresis, ethidium bromide staining and UV transillumination. Positive (Agdia, Elkhart, IN) and negative (healthy peanuts) controls were used in serological and molecular assays. Once the GRSV infection was confirmed, total RNA was isolated from fresh leaf material using the CTAB method [19], including one precipitation with 0.5 vol of 10 M lithium chloride (LiCl) before sample incubation overnight at -80 °C. The amount and quality of the RNA were checked using a spectrophotometer (NanoDrop, Thermo Fisher Scientific, Waltham, MA, USA) and agarose gel electrophoresis. An aliquot of 0.1 vol of 3 M NaAc pH 5.2 and 2.5 vol of EtOH was added to the total RNA extracted, which was sent to Fasteris Life Sciences SA (Switzerland) for processing and sequencing of the small RNA fraction using an Illumina HiSEq 2000 Genome Analyzer. Deep sequencing of small RNAs is a technique that has successfully identified known and unknown viruses with high fidelity using short reads [6, 8, 30]. The raw data were processed to trim adaptor, barcode and low-quality regions using the software clean_reads 0.2.3 [3]. Cleaned reads between 19 and 27 nucleotide (nt) long were assembled *de novo* using the software package Velvet 0.6.04 [31] with the following parameters: *k*-mer (hash) length of 17, cov_cutoff=auto and exp_cov=auto. The origin of individual contigs was identified by BLASTn and BLASTx. The cleaned reads were also mapped against the most closely related genome sequence available from GenBank using the ngs_backbone 1.4.0 software [3]. A whole-genome consensus sequence of the GRSV isolate was determined with the SAMtools mpileup command [11]. The bam files that were obtained were visualized and analyzed using Tablet software [16]. This consensus sequence was used to fill the gaps within the gene sequences in the assembled contigs. Open reading frames (ORFs) were determined using the NCBI ORF Finder program (http://www.ncbi.nlm.nih.gov/gorf/gorf.html), and the aa sequences were predicted from the DNA sequences using the online program Transeq (http://www.ebi.ac.uk/Tools/st/emboss_transeq). All of the genes of the Argentinian GRSV isolate were compared to those of other tospoviruses for which complete sequences are available in the GenBank database. The nt and aa sequences were compared by multiple sequence alignment using the ClustalW option of the MegAlign program (Lasergene software, DNASTAR, Inc., Madison, WI, USA). Dendrograms for the different genes were constructed using the MEGA 6.0 software [24], using bootstrap values for 1,000 replicates to assess the support for clustering of each gene.

In silico analysis of deep-sequencing data allowed us to obtain the sequences of every gene coded by the genomic RNAs of the GRSV Argentinian isolate, here called GRSV-AR. *De novo* assembly generated 413 contigs with a median coverage depth of 71.19, using 6,684,792 of 12,376,739 reads. Of these contigs, 38 belonged to tospoviruses, with 19, 14 and 5 corresponding to the L, M and S RNA, respectively. Blastn and Blastx analysis revealed that the most closely related genome was the natural reassortant

$L_GM_TS_G$ (NCBI accession no. HQ644140/1/2), which was used as a reference to match the cleaned reads. The L, M and S segments were obtained by mapping 2,636,174, 403,433 and 1,239,129 reads, respectively, with an average coverage of 10.65. This strategy allowed us to determine the sequence of almost the entire genome, except for some portions of the IGR of the M and S RNAs.

The GRSV-AR isolate showed a genome structure and organization similar to that previously described for tospoviruses [18], with an A/U-rich IGR separating the two ORFs of the M and S RNAs. The consensus terminal 5′UCUCGUUA3′nt sequence at the 3′end and the 5′AGAGCAAU3′ at the 5′end were present in all three genomic segments. The L RNA of GRSV-AR had a single ORF of 8625 nt in length, encoding an RdRp protein of 2874 aa. The M RNA had two ORFs corresponding to the NSm and GPp genes of 912 nt and 3402 nt in length, whose encoded proteins had 303 aa and 1133 aa, respectively. The S RNA contained two ORFs of 1404 nt and 777 nt, encoding the NSs and N proteins of 467 aa and 258 aa in length, respectively. The complete nt sequences of the RdRp, NSm, GPp, NSs and N genes were obtained, and they were deposited in the GenBank database under the accession numbers KT972590–KT972594.

Each coding region of GRSV-AR was compared with those of other GRSV, TCSV and TSWV isolates that have

Table 1 Percentage of nucleotide (nt) and derived amino acid (aa) sequence identity in the GRSV-AR genes compared with other related tospoviruses

Virus isolate	Accession number	Origin/host	RNA L		RNA M				RNA S			
			RdRp		NSm		GPp		NSs		N	
			nt	aa	nt	aa	nt	aa	nt	aa	nt	aa
GRSV-AR	KT972590-94	Argentina peanut	100	100	100	100	100	100	100	100	100	100
$L_GM_TS_G$	HQ644142-40	USA tomato	96.8	98.9	92.5	96.4	91.2	96.4	97.4	99.1	97.3	99.2
GRSV	AF513220/ AF513219	Brazil tabacco			95.4	95.4					96.9	98.5
GRSV-Cb	AF251271	Brazil cubiu									96.8	98.8
GRSV-SA-05	AF213673/AY574055/ JN571117/ S54327	South Africa peanut			95.3	98.4	95.0	97.9	96.7	97.2	95.2	95.7
TCSV-AB8900-2-3	HQ700667	Brazil tomato	82.2	93.6								
TCSV-BR-03	AF213674/ AY574054	Brazil tomato			92.4	95.7	92.0	96.6				
TCSV-V1-3	AF282983	Brazil tomato			93.1	96.4						
TCSV	AF282982	Brazil tomato									83.9	87.6
TCSV- DRF1	KJ399303	Dom. Republic bean									83.5	87.3
TCSV-DRSP1	KJ399304	Dom. Republic chilli pepper									83.5	86.9
TCSV-T36-H	JX244198	USA tomato									83.7	87.6
TCSV-OH13	KM610235	USA tomato									83.0	87.6
TCSV 10-10-14	KP172480	USA annual vinca									84.0	88.0
TSWV-PA01	KT160280-82	USA pepper	78.3	89.5	78.7	86.8	76.5	81.5	78.1	82.1	78.2	79.9

been completely sequenced (Table 1). For the RdRp gene, the GRSV-AR isolate exhibited 96.8 %, 82.2 % and 78.3 % nt sequence identity to available sequences of GRSV, TCSV and TSWV, respectively. However, the deduced aa sequence identities were higher than 90 % when compared with TCSV isolates. The percent identities

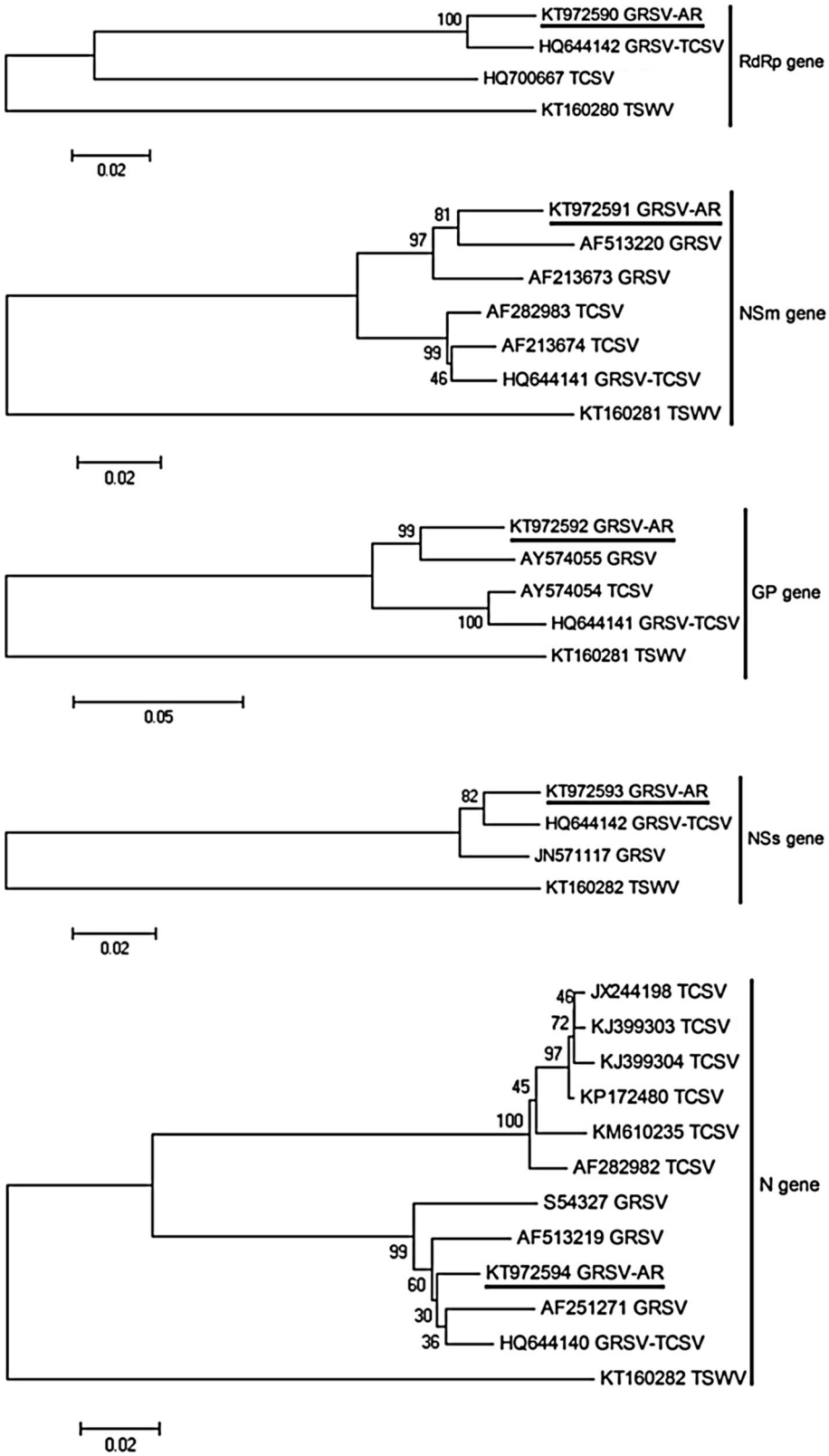

Fig. 1 Phylogenetic reconstruction of GRSV, TCSV and TSWV genes, using the neighbor-joining method in MEGA 6.0. Accession numbers of sequences from GenBank and isolate designations are included for each gene, with the isolate studied in this work (GRSV-AR) underlined in each tree. GRSV-TCSV corresponds to the reassortant $L_GM_TS_G$. The bar represents the number of substitutions per base

of the GRSV-AR genes present in the M RNA were higher than 90 % when compared with other GRSV and TCSV isolates for both nt and aa sequences. The NSs gene had nt and deduced aa sequence identities above 96 % when compared to other GRSV isolates. No comparisons to the NSs gene of TCSV could be made due to the lack of sequence information available in GenBank. Finally, when N nt sequences were compared, identity values with GRSV isolates ranged between 95.2 % and 97.3 %, whereas with TCSV isolates, they ranged between 83.0 % and 84.0 %, and they were less than 80 % with the TSWV isolate. For the N deduced aa sequence, values ranged from 79.9 % to 88.0 % when compared with TSWV and TCSV isolates, sharing the highest identities of 95.7 %-99.2 % with other GRSV isolates. The nt sequences of the GRSV-AR genes were analyzed in DNA neighbor-joining phylogenetic trees in which sequences of GRSV, TCSV and TSWV isolates from different hosts and geographical locations were included (Fig. 1). Genes from the GRSV-AR isolate were closely related to previously characterized GRSV genomes. Furthermore, phylogenetic analysis of nt sequences showed that American GRSV isolates are more closely related to each other than to the South African GRSV isolate. Sequences from TCSV and TSWV isolates clustered in a separate clade.

The nt and deduced aa sequence similarities, together with the results of phylogenetic analysis, indicate that the virus studied in this work (GRSV-AR) is an isolate of GRSV. However, it is interesting to note the high level of sequence similarity between the genes of the M RNA of GRSV and TCSV. GRSV-AR showed aa identity values for the NSm protein between 95.4 % and 98.4 % when compared to other GRSV isolates. When the protein sequence was compared to TCSV isolates, these values ranged between 95.7 % and 96.4 %. The same was observed for the glycoproteins (GP), whose sequences were 96.4 % to 96.6 % identical to those of TCSV isolates and 97.9 % identical to those of other GRSV isolates. These values, above 90 %, are consistent with results obtained by Silva *et al.* [20] and Lovato *et al.* [14], who reported 96.7 % and 92 % identity between NSm and GP aa sequences of TCSV and GRSV, respectively. Londoño *et al.* [13] also found high percentages of identity (94 %-95 %) between the nt sequences of the M RNA of GRSV and TCSV isolates. The low genetic diversity present in the M segment of both viruses was also brought to light with the discovery of a viral isolate that was considered a reassortant between GRSV and TCSV [28]. In the phylogenetic analysis, both GRSV and TCSV genes clustered in separate clades.

Our results confirm that GRSV and TCSV belong to different species and support the hypothesis that the GRSV and TCSV M segment likely comes from a common ancestor, and it may be subjected to a negative selection pressure, probably because such aa sequences are crucial to maintain protein functions.

It is important to point out that the NSm protein participates in processes related to cell-to-cell and long-distance movement in the host [10, 22], whereas the Gn and Gc glycoproteins encoded by the GPp gene are involved in the infection of insect vectors and in subsequent transmission by thrips [7]. This aa sequence so important that a single point mutation can abolish virus transmission by insects [21]. These facts, combined with the coordinated co-evolution between thrips vectors, tospoviruses and their plant hosts [2, 23], could explain both the high level of homology between GRSV and TCSV in the M RNA and the relative prevalence observed for each virus when their geographical distribution was studied. To date, except for $L_GM_TS_G$ in Florida, South Carolina and New York (USA) [29], no reassortants between these two viruses or other different tospoviruses have been reported. This may be attributed to the fact that previous studies have not focused on that aim, and surveys were conducted on the basis of serological tests or amplification of a specific genomic segment. In order to clarify this point, an extensive survey should be carried out that is focused on the detection of mixed infections in plants under natural conditions and on the complete genome analysis of tospoviruses reported in Argentina and other cultivated areas of the world. This work adds to our understanding of the mechanisms of evolution of the members of the genus *Tospovirus*.

Compliance with ethical standards

I have read and have abided by the statement of ethical standards for manuscripts submitted to Archives of Virology.

Funding This work was supported by Fundación Maní Argentino and the PNIND PE 1108072 project of Instituto Nacional de Tecnología Agropecuaria (INTA).

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Almeida MMS, Orílio AF, Melo FL, Rodriguez R, Feliz A, Cayetano X, Martínez RT, Resende RO (2014) The first report of tomato chlorotic spot virus (TCSV) infecting long beans and chili peppers in the dominican republic. Plant Dis 98:1285
2. Ananthakrishnan TN, Annadurai RS (2007) Thrips–tospovirus interactions: biological and molecular implications. Curr Sci 92:1083–1086
3. Blanca JM, Pascual L, Ziarsolo P, Nuez F, Cañizares J (2011) ngs_backbone: a pipeline for read cleaning, mapping and snp calling using next generation sequence. BMC Genom 12:285

4. Boari AJ, Maciel-Zambolim E, Lau DD, Lima GSA, Kitajima EW, Brommonschenkel SH, Zerbini FM (2002) Detection and partial characterization of an isolate of groundnut ringspot virus in *Solanum sessiliflorum*. Fitopatol Bras 27:249–253
5. Briese T, Calisher CH, Higgs S (2013) Viruses of the family Bunyaviridae: are all available isolates reassortants? Virology 446:207–216
6. Hagen C, Frizzi A, Kao J, Jia L, Huang M, Zhang Y, Huang S (2011) Using small RNA sequences to diagnose, sequence, and investigate the infectivity characteristics of vegetable-infecting viruses. Arch Virol 156:1209–1216
7. Hogenhout SA, Ammar E-D, Whitfield AE, Redinbaugh MG (2008) Insect vector interactions with persistently transmitted viruses. Annu Rev Phytopathol 46:327–359
8. Kreuze JF, Perez A, Untiveros M, Quispe D, Fuentes S, Barker I, Simon R (2009) Complete viral genome sequence and discovery of novel viruses by deep sequencing of small RNAs: a generic method for diagnosis, discovery and sequencing of viruses. Virology 388:1–7
9. Law MD, Moyer JW (1990) A tomato spotted wilt-like virus with a serologically distinct N protein. J Gen Virol 71:933–938
10. Lewandowskia DJ, Adkins S (2005) The tubule-forming NSm protein from tomato spotted wilt virus complements cell-to-cell and long-distance movement of Tobacco mosaic virus hybrids. Virology 342:26–37
11. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth GT, Abecasis GR, Durbin R (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079
12. Lian S, Lee JS, Cho WK, Yu J, Kim MK, Choi HS, Kim KH (2013) Phylogeneticand recombination analysis of tomato spotted wilt virus. PLOS ONE 8:e63380. doi:10.1371/journal.pone.0063380
13. Londoño A, Capobianco H, Zhang S, Polston JE (2012) First record of tomato chlorotic spot virus in the USA. Trop Plant Pathol 37:333–338
14. Lovato FA, Nagata T, de Oliveira Resende R, de Avila AC, Inoue-Nagata AK (2004) Sequence analysis of the glycoproteins of tomato chlorotic spot virus and groundnut ringspot virus and comparison with other tospoviruses. Virus Genes 29:321–328
15. Margaria P, Ciuffo M, Rosa C, Turina M (2015) Evidence of a tomato spotted wilt virus resistance-breaking strain originated through natural reassortment between two evolutionary-distinct isolates. Virus Res 196:157–161
16. Milne I, Stephen G, Bayer M, Cock PJA, Pritchard L, Cardle L, Shaw PD, Marshall D (2013) Using tablet for visual exploration of second-generation sequencing data. Brief Bioinform 14:193–202
17. Pappu HR, Jones RAC, Jain RK (2009) Global status of tospovirus epidemics in diverse cropping systems: successes achieved and challenges ahead. Virus Res 141:219–236
18. Plyusnin A, Beaty BJ, Elliott RM, Goldbach R, Kormelink R, Lundkvist A, Schmaljohn CS, Tesh RB (2012) Family Bunyaviridae. In: King AMQ, Adams MJ, Carstens EB, Lefkowitz (eds) Virus taxonomy: ninth report of the international committee on taxonomy of viruses. Elsevier Inc, London, pp 725–741
19. Sambrook J, Russell D (2001) Molecular cloning: a laboratory manual, 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
20. Silva MS, Martins CRF, Bezerra IC, Nagata T, de Ávila AC, Resende RO (2001) Sequence diversity of NSm movement proteins of tospoviruses. Arch Virol 146:1267–1281
21. Sin SH, McNulty BC, Kennedy GG, Moyer JW (2005) Viral genetic determinants for thrips transmission of Tomato spotted wilt virus. Proc Natl Acad Sci USA 102:5168–5173
22. Soellick T-R, Uhrig JF, Bucher GL, Kellmann J-W, Schreier PH (2000) The movement protein NSm of tomato spotted wilt topovirus (TSWV): RNA binding, interaction with the TSWV N protein, and identification of interacting proteins. Proc Natl Acad Sci USA 97:2373–2378
23. Sundaraj S, Srinivasan R, Culbreath AK, Riley DG, Pappu HR (2014) Host plant resistance against Tomato spotted wilt virus in peanut (*Arachis hypogaea*) and its impact on susceptibility to the virus, virus population genetics, and vector feeding behavior and survival. Phytopathology 104:202–210
24. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729
25. Tentchev D, Verdin E, Marchal C, Jacquet M, Aguilar JM, Moury B (2011) Evolution and structure of tomato spotted wilt virus populations: evidence of extensive reassortment and insights into emergence processes. J Gen Virol 92:961–973
26. Timmerman-Vaughan GM, Lister R, Cooper R, Tang J (2014) Phylogenetic analysis of New Zealand tomato spotted wilt virus isolates suggests likely incursion history scenarios and mechanisms for population evolution. Arch Virol 159:993–1003
27. Tsompana M, Moyer JW (2008) Tospoviruses. In: Mahy BWJ, Van Regenmortel MHV (eds) Encyclopedia of virology, vol 5, 3rd edn. Elsevier Ltd, Oxford, pp 157–162
28. Webster CG, Reitz SR, Perry KL, Adkins S (2011) 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 413:216–225
29. Webster CG, Frantz G, Reitz SR, Funderburk JE, Mellinger HC, McAvoy E, Turechek WW, Marshall SH, Tantiwanich Y, McGrath MT, Daughtrey ML, Adkins S (2015) Emergence of groundnut ringspot virus and tomato chlorotic spot virus in vegetables in Florida and the Southeastern United States. Phytopatology 105:388–398
30. Wu Q, Luo Y, Lu R, Lau N, Lai EC, Li W-X, Ding S-W (2010) Virus discovery by deep sequencing and assembly of virus-derived small silencing RNAs. Proc Natl Acad Sci USA 107:1606–1611
31. Zerbino DR, Birney E (2008) Velvet: algorithms for de novo read assembly using de Bruijn graphs. Genome Res 18:821–829