Virus Research 260 (2019) 49–52

Contents lists available at ScienceDirect

Virus Research

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

Short communication

Complete genome sequence and phylogenetic analysis of a novel dicistrovirus associated with the whitefly *Bemisia tabaci*

Erich Y.T. Nakasu[a,1], Marcio Hedil[a,1], Tatsuya Nagata[b], Miguel Michereff-Filho[a], Vivian S. Lucena[a], Alice K. Inoue-Nagata[a,*]

[a] Embrapa Vegetables, Brasília, Brazil
[b] Department of Cell Biology, University of Brasilia, Brasília, Brazil

ARTICLE INFO

Keywords:
High-throughput sequencing
Picornavirales
Dicistroviridae
Next generation sequencing
Insect virus

ABSTRACT

A novel single-stranded RNA virus was detected in a whitefly (*Bemisia tabaci*) sample subjected to high-throughput sequencing. The 8293 nt-long genome presents a polyadenylated 3' end, and contains two ORFs encoding putative 1596 and 849 aa-long proteins. These putative proteins display significant similarity to replicase and capsid polyproteins, respectively, of discitroviruses. Its complete genome sequence shared the highest nucleotide identity (59.8%) with cricket paralysis virus (family *Dicistroviridae*, genus *Cripavirus*). Phylogenetic analyses showed that this new virus putative protein sequences clustered with those from members of *Dicistroviridae*. However, the replicase and capsid polyprotein sequences clustered with those of members of different genera, respectively to *Aparavirus* and *Cripavirus*. RT-PCR using newly collected adult and nymph whitefly samples confirmed the presence of this virus in field populations of *B. tabaci*. Genome sequence and organization, and polyproteins comparison indicate that this virus is a new species of the family *Dicistroviridae*. The name Bemisia-associated dicistrovirus 1 is proposed for this virus.

Repositories: The GenBank accession number for Bemisia-associated dicistrovirus 1 genomic sequence is MH459180.

Viruses assigned to the family *Dicistroviridae* present small, icosahedral particles (approximately 30 nm) carrying a linear, positive-sense ssRNA genome (Valles et al., 2017). Their genomes typically range from 9 to 10 kb, presenting a covalently-linked VPg at the 5′ end and a poly (A) tail at the 3′ end. Their genomes contain two open reading frames (ORFs) and these two genes are expressed by a cap-independent manner via internal ribosome entry sites (IRESs). ORF1 (at the 5′ region) encodes a replicase polyprotein, with RdRP, protease, helicase and VPg domains, whereas ORF2 encodes a structural polyprotein which is cleaved into three to four capsid proteins. Currently, this family is divided into three genera, named *Aparavirus*, *Cripavirus* and *Triatovirus*, based on sequence diversity of replicase and capsid polyproteins, and topology of the second IRES structure, located in the intergenic region (IGR) (Valles et al., 2017).

Whiteflies of the *Bemisia tabaci* complex (Hemiptera: Aleyrodidae) are among the main insect pests of agricultural crops worldwide, as it is highly polyphagous, presenting high reproduction rates and ability to disperse to new environments (Brown et al., 1995). The yield reduction caused by this hemipteran is enormous, mainly derived from the transmission of plant viruses, such as begomoviruses and criniviruses (Navas-Castillo et al., 2011; Perring, 2001). In addition, this insect is able to rapidly develop insecticide resistance, decreasing the effectiveness of chemical control (Horowitz et al., 2005). In this context, the development of virus-based bioinsecticides is an interesting approach for this insect control (Bonning and Miller, 2010). Considering that seven of the 15 official members of the family *Dicistroviridae* infect hemipterans (Bonning and Miller, 2010), we hypothesized that whiteflies could also serve as hosts for dicistroviruses. In the present study, a high-throughput sequencing (HTS) approach led to the identification of a novel dicistrovirus associated with *B. tabaci*. Furthermore, we describe the virus complete genome sequence, analysis of its sequence and phylogeny inference.

Soybean leaves infested with *B. tabaci* Middle East-Asia Minor 1 (MEAM1) nymphs were collected in the Federal District, Brazil, in 2014 (coordinates: 15°56′97.3″S; 47°37′82.5″W). Nymphs at different developmental stages were carefully removed from the leaves and

Abbreviations: HTS, high-throughput sequencing; RT-PCR, reverse transcription polymerase chain reaction; IRES, internal ribosome entry site; IGR, intergenic region; ORF, open reading frame; BaDV-1, Bemisia-associated dicistrovirus 1

* Corresponding author:
E-mail address: alice.nagata@embrapa.br (A.K. Inoue-Nagata).
[1] E.Y.T. Nakasu and M. Hedil equally contributed to this work.

https://doi.org/10.1016/j.virusres.2018.11.008
Received 17 September 2018; Received in revised form 14 November 2018; Accepted 15 November 2018
Available online 16 November 2018
0168-1702/

transferred to microtubes containing ethanol until further processing. Total RNA of the insects (approximately 100 individuals) was extracted using RNeasy Mini Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. The RNA sample (approximately 2 µg) was treated with Ribo-Zero rRNA Removal Kit (Human/Mouse/Rat) (Illumina, San Diego, CA, USA) prior to the library construction using Illumina TruSeq Stranded Total RNA LT Sample Prep Kit and cDNA sequencing in a HiSeq 2000 platform (2 × 100 bp) at Macrogen Inc. (Seoul, South Korea). Reads were trimmed using Trimommatic 0.35 (Bolger et al., 2014) and the contigs were *de novo* assembled using the Velvet algorithm (Phred = 34, 91 k-mer) (Zerbino and Birney, 2008). Generated contigs were analysed for their identity against a virus RefSeq database (downloaded from NCBI on 03/08/2016) with the tBLASTx algorithm (Altschul et al., 1990) using Geneious 9.1.5 (Biomatters, Auckland, New Zealand).

The library yielded 11,916,943 reads, which generated 1287 contigs. Following tBLASTx analysis, one contig, named as 3634, had hit to Israeli acute paralysis virus of bees (IAPV, EF219380; 53.5% identity, E-value = 2.17e-93). This contig was selected for further studies, and used as reference for extending the contig using the "map-to-reference" command of the Geneious program. Following assembly, 71,241 reads were mapped to contig 3634, generating an approximately 8300 nt-long sequence.

In order to confirm the HTS data, amplicons covering the viral genome were obtained by reverse transcription PCR (RT-PCR) and subsequently Sanger sequenced. All Sanger sequencing reactions were performed using an ABI 3730XL platform at Macrogen, Inc. First, total RNA of approximately 100 nymphs of the original whitefly sample used for HTS was extracted using TRIzol reagent (Thermo Fisher Scientific, Waltham, USA). Four overlapping cDNA fragments covering the genome of the virus were obtained by RT-PCR: fragment 1 from nucleotide position 116 to 1572; fragment 2 from 1432 to 4006; fragment 3 from 3891 to 5937; fragment 4 from 5844 to 7931 (primers in Table 1), cloned to pGEM-T Easy vector (Promega, Madison, USA) and sequenced by Sanger's method. Furthermore, the 5′ and 3′ ends of the virus genome were sequenced following 5′ RACE (version 2.0, Thermo Fisher Scientific) and 3′ RACE (Lucinda et al., 2012), respectively (Table 1). The complete genome sequence obtained by Sanger sequencing is 99.96% identical to the sequence obtained by HTS. The complete viral genome sequence is 8293 nucleotides-long (GC content of 39%) plus a poly(A) tail at the 3′ end. The virus sequence is available on GenBank database under the accession number MH459180.

The predicted genome organization of the newly identified virus (Fig. 1) is consistent with the typical organization of viruses of the family *Dicistroviridae* (Valles et al., 2017). The 5′ UTR is 174 nt-long and is expected to contain an IRES for translation of ORF1. The predicted ORF1 encodes a replicase polyprotein, and starts at position 175 and ends at 4965 (4791 nt-long). BLASTp analysis of the deduced 1596 aa-long polyprotein showed a closer relationship to replicase polyproteins of dicistroviruses and picorna-like viruses, with the most significant hit to the replicase protein of Drosophila C virus (DCV, AAC58807, 26% identity, 83% coverage, 2.0e-113). Conserved motifs in the putative proteins were searched through Pfam database using MOTIF search tool (implemented in https://www.genome.jp/tools/motif/) (Finn et al., 2016; "MOTIF, 2018" n.d.). The replicase polyprotein (ORF1) contained motifs related to RdRP_1 (PF00680, E-value: 2.3e-65), RNA_helicase (PF00910, 4.6e-30) and cysteine protease motif Peptidase_C3G (PF12381, 2.2e-03, rice tungro spherical virus peptidase) (Fig. 1), as well as other less significant motifs related to ATPases and peptidases. A putative IGR of 254 nt was detected between ORF1 and ORF2, considering a canonical AUG start codon at position 5220. In order to verify if the IGR had an IRES similar to the ones found in other dicistroviruses, this sequence was further investigated. Initially, Rfam (Nawrocki et al., 2015) was used to search for similar sequences in the non-coding RNA family database. No hits were observed for the putative IGR of the whitefly dicistrovirus. When Rfam analysis was performed for cripavirus and triatovirus IGR sequences, all of them presented hits to the cripavirus internal ribosome entry site (IRES, RF00458). In contrast, no hits were observed for aparavirus IGR sequences. This suggests that the IGR of the whitefly dicistrovirus is not closely related to the IGR of cripaviruses and triatoviruses. Next, an analysis of possible conserved helices between the whitefly dicistrovirus IGR and members of each dicistrovirus genus was performed using Transat (Wiebe and Meyer, 2010), though no evidence of secondary structural conservation was observed between the whitefly dicistrovirus IGR and those of dicistroviruses from different genera (Supplementary Fig. 1). Furthermore, bulge sequences of the typical dicistrovirus IRES (Bonning, 2009; Bonning and Miller, 2010) were not detected in the IGR region, altogether suggesting that this virus is distantly related to known dicistroviruses. Although a non-AUG start codon is expected for ORF2 (Nakashima and Uchiumi, 2009), the starting position of ORF2 could not be identified. Further analysis of the putative ORF2 was performed based on a canonical AUG start codon located 254 nt downstream of the ORF1 stop codon. The putative ORF2 starts at position 5220, is 2550 nt-long and encodes the capsid polyprotein. BLASTp analysis of ORF2 showed closer relationship to capsid polyproteins of dicistroviruses and picorna-like viruses, with most significant hit to the capsid protein precursor of aphid lethal paralysis virus (ALPV; AAN61471, 87% coverage, 26% identity, 5e-58). Three capsid protein motifs were identified in this putative polyprotein, one cricket paralysis virus capsid protein-like motif (CRPV_capsid, PF08762, 2.3e-19) and two picornavirus capsid protein-like Rhv motifs (PF00073) (2.7e-04 and 1,3e-01, respectively). This indicates that ORF2 encodes the capsid polyprotein of the virus (Fig. 1). The complete genome sequence shares maximum

Table 1
Primers used for detection and amplification of the genome of Bemisia-associated dicistrovirus 1 by PCR.

Primer	Sequence (5' → 3')	Annealing region	Genome position	Use
103F	TCGTAGGGTTCTCAGTGTTTCAGTC	5'UTR	116 – 140	Fragment 1
3634R3 (GSP3)	AGTATATGGAGCAAGCCAGC	ORF1	239 - 220	5' RACE
3634R2 (GSP2)	GACCTAATCGCACTCAACAAACG	ORF1	311 - 289	5' RACE
3634R1 (GSP1)	CCGTATTCTTATGCCGATACTTCG	ORF1	384 - 361	5' RACE
1419F	GGCTACATGGGACAGCCTATTGTAG	ORF1	1432 – 1456	Fragment 2
1535R	AGACAATGCAGCACAGTGTAAATGG	ORF1	1572 - 1538	Fragment 1
3878F	AGAAATGGCACGTCATGGTGTC	ORF1	3891 – 3912	Fragment 3
3969R	GACATGCTGAGAACACTCGTGTTTG	ORF1	4006 - 3982	Fragment 2
MB31_3634F	AACTCAGTCAGCCGTTGGAC	ORF2	5588 – 5607	Detection
5831F	CCTTATACAACCGGATGCCCAAAC	ORF2	5844 – 5867	Fragment 4
MB31_3634R	TAAACGACACTGGTCCGGTG	ORF2	5937 - 5918	Fragment 3, detection
3634_3prime	TTTCCTGCGAGTTGGGATGC	3'UTR	7908 - 7927	3' RACE
7898R	ACTAGCATCCCAACTCGCAGG	3'UTR	7931 - 7911	Fragment 4
OLIGOdT50Rz	ATGCCATGCCGACCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	3'UTR	Poly(A) tail	3' RACE

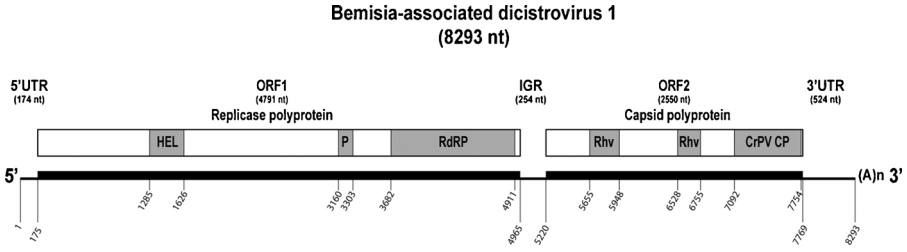

Fig. 1. Schematic representation of the putative genome structure of the Bemisia-associated dicistrovirus 1. The two predicted ORFs are shown, with conserved motifs indicated by shaded boxes. Nucleotide positions indicate start and end of ORFs as well as motifs. Abbreviations: HEL = RNA helicase, P = tungro spherical virus-type peptidase C3G, RdRP = RNA-dependent RNA polymerase, Rhv = picornavirus (Rhinovirus) capsid protein-like, CrPV CP = cricket paralysis virus (CrPV) capsid protein like.

nucleotide identity of 59.8% with that of cricket paralysis virus (AF218039), a cripavirus. Pairwise comparison was performed with SDT v.1.2 (Muhire et al., 2014).

Phylogenetic analyses were conducted using the deduced amino acid sequences of ORF1 and ORF2. Amino acid sequences of the whitefly dicistrovirus were aligned with correspondent protein sequences of the 15 currently accepted dicistrovirus species and five other tentative ones using PROMALS3D (Pei et al., 2008) and trimmed with TrimAI v.1.3 (gappyout method) (Capella-Gutiérrez et al., 2009). Phylogenetic analysis was performed by Bayesian inference using MrBayes v.3.2.6 (10,000 generations) on the Phylogeny.fr platform (Dereeper et al., 2008; Huelsenbeck and Ronquist, 2001). The viruses in the three genera are clearly separated either when analysing replicase (Fig. 2a) or capsid (Fig. 2b) polyprotein sequences. Although mud crab virus (MCV), Taura syndrome virus (TSV) and Macrobrachium rosenbergii Taihu virus (MrTV) appear to be more distantly related to other aparaviruses, they are consistently grouped together with high confidence levels, particularly for the capsid polyprotein. In the case of the whitefly dicistrovirus, however, its replicase polyprotein is more closely related to the replicase of aparaviruses, whereas the capsid polyprotein groups more closely with the capsid polyprotein of cripaviruses. Therefore, the phylogenetic analyses do not support the classification of this virus in an established dicistrovirus genus.

In order to confirm the association of the novel virus with *B. tabaci*, insect collections were carried out in two different locations of the Federal District between February and March of 2017 and of 2018. Non-target insects, such as honeybees (*Apis* sp.), and whitefly predators (the ladybirds *Eriopis* sp. and *Hippodamia* sp.) were also collected. After collection, insects were stored in absolute ethanol until the samples were processed. Total RNA was extracted from the samples using TRIzol reagent and the cDNA was synthesized using random primers and M-MLV reverse transcriptase (Thermo Fischer Scientific). The PCR for virus detection was carried out using a specific detection primer pair (MB31_3634 F and MB31_3634R) (Table 1) which amplified a 350 bp fragment of the virus ORF2. Parameters used for this PCR are: 3 min. at 95 °C; 35 cycles as follows: 95 °C, 45 s; 55 °C, 30 s; 72 °C, 60 s; a final extension step of 5 min at 72 °C. The virus was detected in both *B. tabaci* nymphs and adults (Supplementary Fig. 2). In samples collected in 2017, this virus was detected with relatively high frequency, being present in seven out of 11 adult samples and in one of 11 nymph samples. In contrast, the virus was not detected in honeybee and predator samples, which is indicative of its specific relationship with whiteflies. In samples collected in 2018, this virus was detected in one out of seven adult samples and in four out of 15 nymph samples.

Despite the importance of *B. tabaci* as agricultural pests and as vectors of plant viruses, whitefly-infecting viruses are still poorly studied. In 2001, an iridovirus was described in association with whitefly cell lines (Funk et al., 2001). More recently, two genomoviruses were detected in *B. tabaci* samples based on HTS analysis (Nakasu et al., 2017). However, no evidence has been given about pathological effects of these viruses on *B. tabaci*. The discovery of a novel dicistrovirus associated with whiteflies adds another example of dicistrovirus/hemipteran interaction, while being the first member of dicistroviruses found in *B. tabaci*. Dicistroviruses infect a range of hosts, including beneficial organisms as well as pests of medical and agricultural importance, causing from latent to lethal infection (Bonning and Miller, 2010; Guo et al., 2013). Whether the virus described here is pathogenic to whiteflies is a question that still needs to be elucidated. Genome

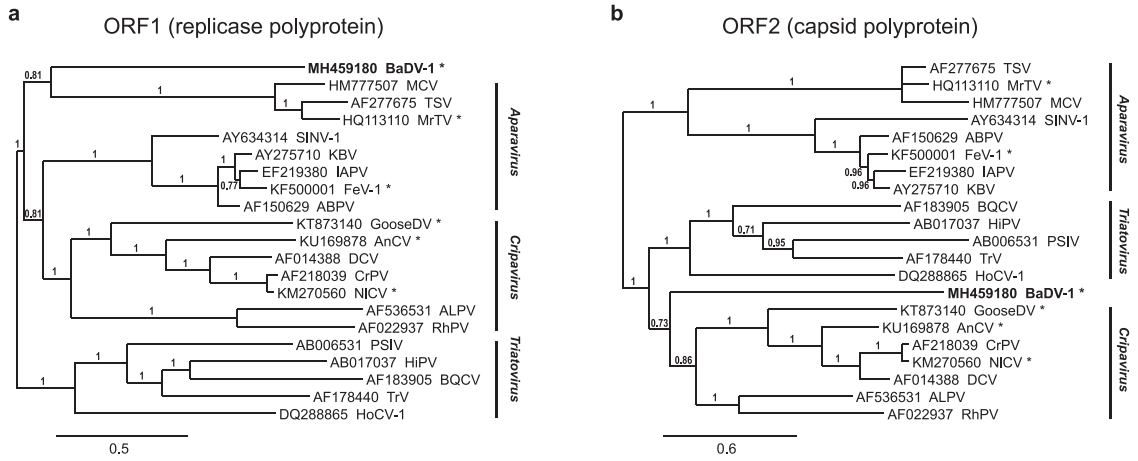

Fig. 2. Phylogenetic analyses of Bemisia-associated dicistrovirus 1 (BaDV-1) polyproteins. a) replicase polyprotein (ORF1); b) capsid polyprotein (ORF2). Phylogenetic trees were constructed by bayesian method using full length polyprotein amino acid sequences of BaDV-1 (in bold), all 15 currently accepted dicistroviruses and still unclassified related viruses (indicated by *). The acronyms given to these unclassified viruses are tentative. Branch support values are indicated above each branch. Scale bars indicate the expected number of changes per site. Acronyms: ABPV, acute bee paralysis virus; ALPV, aphid lethal paralysis virus; AnCV, Anopheles C virus; BQCV, black queen cell virus; CrPV, cricket paralysis virus; DCV, Drosophila C virus; FeV-1, Formica exsecta virus 1; GooseDV, goose dicistrovirus; HiPV, Himetobi P virus; HoCV-1, Homalodisca coagulata virus 1; IAPV, Israeli acute paralysis virus; KBV, Kashmir bee virus; MCV, mud crab virus; MrTV, Macrobrachium rosenbergii Taihu virus; NlCV, Nilaparvata lugens C virus; PSIV, Plautia stali intestine virus; RhPV, Rhopalosiphum padi virus, SINV-1, Solenopsis invicta virus 1; TrV, Triatoma virus; TSV, Taura syndrome virus.

organization and similarities in the polyprotein domains place this virus in the family *Dicistroviridae*, while the distinct IGR-IRES structure and polyprotein phylogenetic analyses suggest that this virus is a member of a new genus in the family *Dicistroviridae*. The low amino acid sequence identities of both polyproteins to other dicistrovirus sequences indicate that it represents a new species. The name Bemisia-associated dicistrovirus 1 (BaDV-1) is proposed for this virus.

Ethical approval

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

Funding

This work was funded by CNPq (403414/2013-0), FAP-DF (0193.001197/2016) and Embrapa (02.15.08.001.00.00). T.N. and A.K.I.N. are CNPq fellows.

Conflict of interest

The authors declare that they have no conflicts of interest.

Acknowledgment

The authors thank José Luiz Pereira for assistance in the collection of whiteflies.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.virusres.2018.11.008.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Bonning, B.C., 2009. The dicistroviridae: an emerging family of invertebrate viruses. Virol. Sin. https://doi.org/10.1007/s12250-009-3044-1.

Bonning, B.C., Miller, W.A., 2010. Dicistroviruses. Annu. Rev. Entomol. 55, 129–150. https://doi.org/10.1146/annurev-ento-112408-085457.

Brown, J.K., Frohlich, D.R., Rosell, R.C., 1995. The sweetpotato or silverleaf whiteflies: biotypes of Bemisia tabaci or a species complex? Annu. Rev. Entomol. 40, 511–534. https://doi.org/10.1146/annurev.en.40.010195.002455.

Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973. https://doi.org/10.1093/bioinformatics/btp348.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.F., Guindon, S., Lefort, V., Lescot, M., Claverie, J.M., Gascuel, O., 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. https://doi.org/10.1093/nar/gkn180.

Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G.A., Tate, J., Bateman, A., 2016. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. 44, D279–D285. https://doi.org/10.1093/nar/gkv1344.

Funk, C.J., Hunter, W.B., Achor, D.S., 2001. Replication of insect iridescent virus 6 in a whitefly cell line. J. Invertebr. Pathol. 77, 144–146. https://doi.org/10.1006/jipa.2000.5006.

Guo, Z.-X., He, J.-G., Xu, H.-D., Weng, S.-P., 2013. Pathogenicity and complete genome sequence analysis of the mud crab dicistrovirus-1. Virus Res. 171, 8–14. https://doi.org/10.1016/j.virusres.2012.10.002.

Horowitz, A.R., Kontsedalov, S., Khasdan, V., Ishaaya, I., 2005. Biotypes B and Q of Bemisia tabaci and their relevance to neonicotinoid and pyriproxyfen resistance. Arch. Insect Biochem. Physiol. 58, 216–225. https://doi.org/10.1002/arch.20044.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755. https://doi.org/10.1093/bioinformatics/17.8.754.

Lucinda, N., da Rocha, W.B., Inoue-Nagata, A.K., Nagata, T., 2012. Complete genome sequence of pepper yellow mosaic virus, a potyvirus, occurring in Brazil. Arch. Virol. 157, 1397–1401. https://doi.org/10.1007/s00705-012-1313-z.

MOTIF search [WWW Document], n.d. URL https://www.genome.jp/tools/motif/ (Accessed 8.9.2018).

Muhire, B.M., Varsani, A., Martin, D.P., 2014. SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. PLoS One 9, e108277. https://doi.org/10.1371/journal.pone.0108277.

Nakashima, N., Uchiumi, T., 2009. Functional analysis of structural motifs in dicistroviruses. Virus Res. 139, 137–147. https://doi.org/10.1016/j.virusres.2008.06.006.

Nakasu, E.Y.T., Melo, F.L., Michereff-Filho, M., Nagata, T., Ribeiro, B.M., Ribeiro, S.G., Lacorte, C., Inoue-Nagata, A.K., 2017. Erratum to: discovery of two small circular ssDNA viruses associated with the whitefly Bemisia tabaci. Arch. Virol. 162, 3563–3566. https://doi.org/10.1007/s00705-017-3535-6.

Navas-Castillo, J., Fiallo-Olivé, E., Sánchez-Campos, S., 2011. Emerging virus diseases transmitted by whiteflies. Annu. Rev. Phytopathol. https://doi.org/10.1146/annurev-phyto-072910-095235.

Nawrocki, E.P., Burge, S.W., Bateman, A., Daub, J., Eberhardt, R.Y., Eddy, S.R., Floden, E.W., Gardner, P.P., Jones, T.A., Tate, J., Finn, R.D., 2015. Rfam 12.0: updates to the RNA families database. Nucleic Acids Res. 43, D130–D137. https://doi.org/10.1093/nar/gku1063.

Pei, J., Kim, B.H., Grishin, N.V., 2008. PROMALS3D: a tool for multiple protein sequence and structure alignments. Nucleic Acids Res. 36, 2295–2300. https://doi.org/10.1093/nar/gkn072.

Perring, T.M., 2001. The Bemisia tabaci species complex. Crop Prot. 20, 725–737. https://doi.org/10.1016/S0261-2194(01)00109-0.

Valles, S.M., Chen, Y., Firth, A.E., Guérin, D.M.A., Hashimoto, Y., Herrero, S., de Miranda, J.R., Ryabov, E., 2017. ICTV virus taxonomy profile: dicistroviridae. J. Gen. Virol. 98, 355–356. https://doi.org/10.1099/jgv.0.000756.

Wiebe, N.J.P., Meyer, I.M., 2010. Transat — a method for detecting the conserved helices of functional rna structures, including transient, pseudo-knotted and alternative structures. PLoS Comput. Biol. 6, e1000823. https://doi.org/10.1371/journal.pcbi.1000823.

Zerbino, D.R., Birney, E., 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18, 821–829. https://doi.org/10.1101/gr.074492.107.