Archives of Virology (2019) 164:1943–1947
https://doi.org/10.1007/s00705-019-04272-3

ANNOTATED SEQUENCE RECORD

Complete genome sequence of a highly divergent carrot torradovirus 1 strain from *Apium graveolens*

Yahya Z. A. Gaafar[1] · **Heiko Ziebell**[1]

Received: 26 November 2018 / Accepted: 2 April 2019 / Published online: 10 May 2019
© Springer-Verlag GmbH Austria, part of Springer Nature 2019

Abstract

A new virus was identified in a celery plant showing chlorotic rings, mosaic and strong yellowing symptoms, and its complete genome sequence was determined. The genomic organization of this novel virus is analogous to that of known members of the genus *Torradovirus*, consisting of two single-stranded RNAs of 6,823 (RNA1) and 4,263 nucleotides (RNA2), excluding the poly(A) tails. BLAST searches against the nucleotide and protein databases showed that this virus is closely related to but different from carrot torradovirus 1 (CaTV1). Comparisons between the two viruses demonstrated relatively low levels of nucleotide and amino acid similarity in different parts of their genomes, as well as considerable differences in the sizes of their two genomic RNAs. However, the protease-polymerase (Pro-Pol) and capsid protein (CP) regions of this virus share >80% amino acid identity with the corresponding regions of CaTV1. Therefore, based on the current ICTV species demarcation criteria for the family *Secoviridae*, the virus from celery is a divergent strain of CaTV1, named "CaTV1-celery". Nevertheless, differences between CaTV1 and CaTV1-celery in genome size, as well as in biological and epidemiological features, may warrant their separation into two distinct species in the future.

A celery plant (*Apium graveolens* variant *graveolens*) was collected in August 2017 in the state of Hesse in Germany. The sample displayed mosaic symptoms with chlorotic rings and strong yellowing. To identify the possible cause of the disease, leaf material was examined by electron microscopy, but no virus particles were observed. However, mechanical inoculation of *Nicotiana benthamiana, N. clevelandii*, *N. occidentalis*-P1 and *Coriandrum sativum* with the sap of the celery plant resulted in systemic infections three weeks after inoculation, with symptoms consisting of chlorosis and necrosis. No virus particles could be observed in samples from symptomatic test plants. Attempts to transmit the pathogen mechanically to the original host species *A. graveolens* or to *Ammi majus*, *Anethum graveolens*, *Daucus carota* and *Petroselinum crispum* were not successful.

To characterise the genome of the suspected virus, total RNA was extracted from symptomatic *N. benthamiana* leaves using an innuPREP RNA Mini Kit (Analytik Jena AG). The ribosomal RNAs were depleted using a RiboMinus Plant Kit (Invitrogen) and the ribo-depleted RNA was used for library preparation using a Nextera XT Library Kit (Illumina). The library was subjected to high-throughput sequencing (HTS) on a MiSeq v3 platform (2x301). The reads were quality trimmed and filtered using Geneious software (version 11.1.3) (Biomatters Limited). The high-quality reads were assembled using the Geneious *de novo* assembly tool. In BLASTn searches, two assembled contigs of 6,727 and 4,106 nt shared 71.7% and 70.5% identical nucleotides with the two genomic RNAs of carrot torradovirus 1 (CaTV1) (NC_025479 and NC_025480).

The 5′ ends of both RNA segments were confirmed using RNA-ligase-mediated amplification of cDNA ends (RLM-RACE) [1]. The 3′ ends of the two RNA segments were determined via RT-PCR using a virus-specific primer and an oligo(d)T primer. The PCR products were cloned and sequenced, and the resulting sequences were assembled. The assembled full-length sequences of the two RNA segments were 6,823 nt (RNA1) and 4,263 nt (RNA2) in

Handling Editor: Sead Sabanadzovic.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00705-019-04272-3) contains supplementary material, which is available to authorized users.

✉ Heiko Ziebell
heiko.ziebell@julius-kuehn.de

[1] Julius Kuehn-Institute, Institute for Epidemiology and Pathogen Diagnostics, Messeweg 11-12, 38104 Braunschweig, Germany

length, excluding their poly(A) tails. The complete genome sequences of RNA1 and RNA2 were deposited in the GenBank database (accession nos. MK063924 and MK063925, respectively). Diagnostic primers (HZ-539 5'TGTTAGCAG AGCTACGTCCTC3' and HZ-568 5'CCTGAATCTGCC CACGACTT3') were designed using the Primer3 v. 2.3.7 tool [2] to amplify a partial sequence of RNA2-ORF1 (730 nt) to confirm the presence of this virus in infected plants.

According to the species demarcation criteria proposed by the ICTV *Secoviridae* Study Group, viruses belonging to different species share less than 80% aa sequence identity in the protease-polymerase (Pro-Pol) region of the RNA1 polyprotein and less than 75% aa sequence identity in the coat protein (CP) region [3]. The celery virus shares 86.4% aa sequence identity in the Pro-Pol region and 80.3% aa sequence identity in the CP region with CaTV1 (Table 1). Based on these criteria, the celery virus should be considered a new strain of CaTV1, for which we propose the name "CaTV1-celery" (isolate JKI-29346). However, further comparison of the CaTV1-celery genome with the reference sequences of CaTV1 revealed considerable differences.

The genomic organization of CaTV1-celery is analogous to that of other torradoviruses [4]. Accordingly, RNA1 contains a single open reading frame (ORF1) encoding a polyprotein of 2,076 aa (MW: 237 kDa). A search of the Conserved Domain Database (CDD) of NCBI identified the presence of two conserved domains: the RNA helicase (Hel) (RNA_helicase; pfam00910) and RNA-dependent RNA polymerase (RdRp) (RNA_dep_RNAP; cd01699) in the RNA1 polyprotein. An additional motif, characteristic of cysteine proteases (3C), was identified using the Geneious search tool (Fig. 1a) [5–7]. RNA2 is bicistronic, with RNA2 ORF1 encoding a 201-aa protein (MW: 22.0 kDa) and RNA2 ORF2 encoding a predicted polyprotein of 1,057 aa (MW: 117.1 kDa). The protein encoded by RNA2 ORF1 did not match any domain in the CDD database. The RNA2 ORF2 polyprotein contains two conserved domains: the 3A movement protein (MP) family domain (3A superfamily; cl02970) and picornavirus capsid protein (CP)-like domain (rhv_like superfamily; cl13999). A motif search identified the position of the MP conserved motif (LxxPxL) in the RNA2 ORF2 polyprotein [8, 9]. In addition, the MP and the three mature CP subunit peptides were determined based on homologies to those of other torradoviruses (Fig. 1a).

The percentages of identity in a ClustalW 2.1 pairwise alignment between CaTV1-celery RNAs, ORFs and regions with their cognates in other torradoviruses are listed in Table 1. Comparisons of the CP region sequences showed that CaTV1-celery shares between 36.6 and 80.3% aa sequence identity with torradoviruses. Additionally, the Pro-Pol region of CaTV1-celery shares between 57.3 and 86.4% aa sequence identity. Neighbour-joining trees based on the aa sequence alignments of the Pro-Pol and CP regions showed that the celery virus clusters together with CaTV1 within the genus *Torradovirus* (Fig. 1b and c) [10, 11].

Despite the observed similarities, the lengths of the 5′ and 3′ untranslated regions (UTR) of RNA1 are 146 and 449 nt long, respectively, and those of RNA2 are 245 and 294 nt long. These values differ from the 5' and 3' UTRs of CaTV1, which are 126 and 212 nt long, respectively, in RNA1 and 611 nt and 301 nt long in RNA2. Additionally, pairwise alignment of the UTRs with those of other torradoviruses, including CaTV1, showed low nt sequence identity values between 24.7% and 68.4% (Table 1). The predicted protein encoded by ORF1 of RNA2 of CaTV1-celery shares only 66.8% aa sequence identity with its homolog in CaTV1, and the 3′UTRs of both segments share ≤ 68.4% nt sequence identity with those of CaTV1. Furthermore, RNAs 1 and 2 of CaTV1-celery are shorter by 96 and 706 nt, respectively, than their CaTV1 counterparts, making the CaTV1-celery genome the shortest torradovirus genome identified so far, with a total size of 11,086 nt, excluding the poly (A) tail (Table 1).

Another criterion that is considered for species demarcation in the family *Secoviridae* is vector specificity. The close evolutionary relationships between CaTV1, CaTV1-celery and lettuce necrotic leaf curl virus (LNLCV) suggest that CaTV1-celery is also an aphid-borne torradovirus [12, 13]. To test this, *Myzus persicae* aphids from a single laboratory clone were reared on CaTV1-celery-infected *N. benthamiana* for seven days, after which ten aphids were transferred to three groups of healthy plants (10 *N. benthamiana*, 10 *A. graveolens* and 10 *D. carota*). After an inoculation access period of seven days, the plants were treated with the systemic insecticide PIRIMOR (Deutsche ICI) according to the manufacturer's instructions. The plants were incubated under greenhouse conditions (at 22 °C; photoperiod of 16 h light and 8 h dark) for two months, but no symptoms were observed. The absence of virus infections in acceptor plants was additionally confirmed by negative RT-PCR results. Aphid transmission experiments were repeated three times.

Although CaTV1-celery is considered a divergent strain of CaTV1 based on their aa sequence similarity in the Pro-Pol and CP regions, the data suggest that it might also be useful to consider other molecular criteria for species demarcation, i.e., the total genome size and the size and degree of sequence similarity of the 5' and 3' UTR. Taking these criteria into consideration, CaTV1-celery might be accepted in the future as a member of a novel species. Indeed, Sanfaçon and colleagues have already suggested that the current demarcation criteria could be revisited and modified as more viruses become characterized [3]. Furthermore, Verbeek and colleagues have proposed additional criteria, i.e., that the aa sequence identity of the RNA2 ORF1 should be less than 75% and that the conservation level in the 3′UTR of both RNAs should be less than 85% [14].

Table 1 Nucleotide (nt) and amino acid (aa) sequence comparisons between carrot torradovirus 1-celery (CaTV1-celery) and previously characterized torradoviruses

Virus			CaTV1-celery	CaTV1	LNLCV	MYMoV	SCLSV	ToChSV	ToChV	ToMarV	ToNDV	ToTV
nt length	RNA 1	Accession no.	MK063924	NC_025479	NC_035214	NC_035218	NC_035221	NC_013075	FJ560489	NC_010987	NC_027926	NC_009013
		Whole -poly (A)	6,823	6,917	7,576	7,068	7,045	7,473	7,474	7,221	7,236	7,793
		5′UTR	146	126	170	135	231	138	135	140	150	106
		ORF1	6,228	6,579	6,672	6,576	6,624	6,468	6,468	6,456	6,453	6,477
		3′UTR	449	212	743	357	190	867	871	624	633	1,210
	RNA 2	Accession no.	MK063925	NC_025480	NC_035219	NC_035220	NC_035215	NC_013076	FJ560490	NC_010988	NC_027927	NC_009032
		Whole -poly (A)	4,263	4,969	5,290	4,963	4,730	5,093	5,695	4,898	4,896	5,389
		5′UTR	245	611	326	406	83	171	181	138	138	181
		ORF1	606	609	636	618	702	555	570	573	570	564
		ORF2	3,174	3,504	3,669	3,588	3,165	3,594	3,579	3,576	3,573	3,597
		3′UTR	294	301	766	461	899	814	1,406	652	656	1,105
Pairwise identity %												
nt	RNA 1	5′UTR		67.5	39.1	39.1	30.8	44.4	43.5	41.3	37.4	43.7
		Whole		70.1	53.3	54.4	42.3	43.1	43.1	42.5	42.9	42.5
		ORF1		71.2	55.2	55.9	43.6	45.2	45	44.9	45.3	44.6
		3′UTR		68.4	24.7	26.7	31.4	29.9	30.1	28.6	26.6	24.7
	RNA 2	5′UTR		45	32.3	33.8	46.3	39.8	43.8	47.2	39.6	49.5
		Whole		63.5	46.4	48	39.3	38.2	39.2	38.6	39.1	39.6
		ORF1		65.7	46.9	49.8	36.7	41.1	37.5	37.9	38.7	38.7
		ORF2		65.6	49.1	50.8	43.8	40	40.2	40.2	40.8	40.7
		3′UTR		55.5	30.4	31.3	33.1	34.2	34.2	30.9	33.9	30.5
aa	RNA 1	ORF1		76.1	50.9	51.9	33.5	33.8	33.7	33.9	33.6	33.2
		Hel		99	85.4	85.4	57.3	61.2	61.2	61.2	61.2	59.2
		Pro-Pol		86.4	67.1	71	57.4	58.8	60.1	59.9	59.3	57.3
	RNA 2	ORF1		66.8	40.8	48.5	25.7	27	24.3	27.2	26.7	26.2
		ORF2		71.9	44.9	45.9	30.7	31.1	30.7	30.7	31.2	30.8
		MP		59.1	30.9	31.7	23.3	22.4	22.1	22.1	21.6	22.1
		CP		80.3	54.2	55.9	36.6	40	39.3	39.6	40.4	39.7
		Vp35		73.9	47.7	45.3	31.7	33.1	34.6	33.1	34.3	32.9
		Vp26		83	60.7	68.3	41.9	45.9	42.9	45.5	45.9	45
		Vp23		84.1	54.3	54.4	37.9	41.7	41.3	40.4	41.3	41.7

The viruses used for comparison and their respective abbreviations are as follows: carrot torradovirus 1 (CaTV1), lettuce necrotic leaf curl virus (LNLCV), motherworth yellow mottle virus (MYMoV), squash chlorotic leaf spot virus (SCLSV), tomato chocolate spot virus (ToChSV), tomato chocolate virus (ToChV), tomato marchitez virus (ToMarV), tomato necrotic dwarf virus (ToNDV) and tomato torrado virus (ToTV)

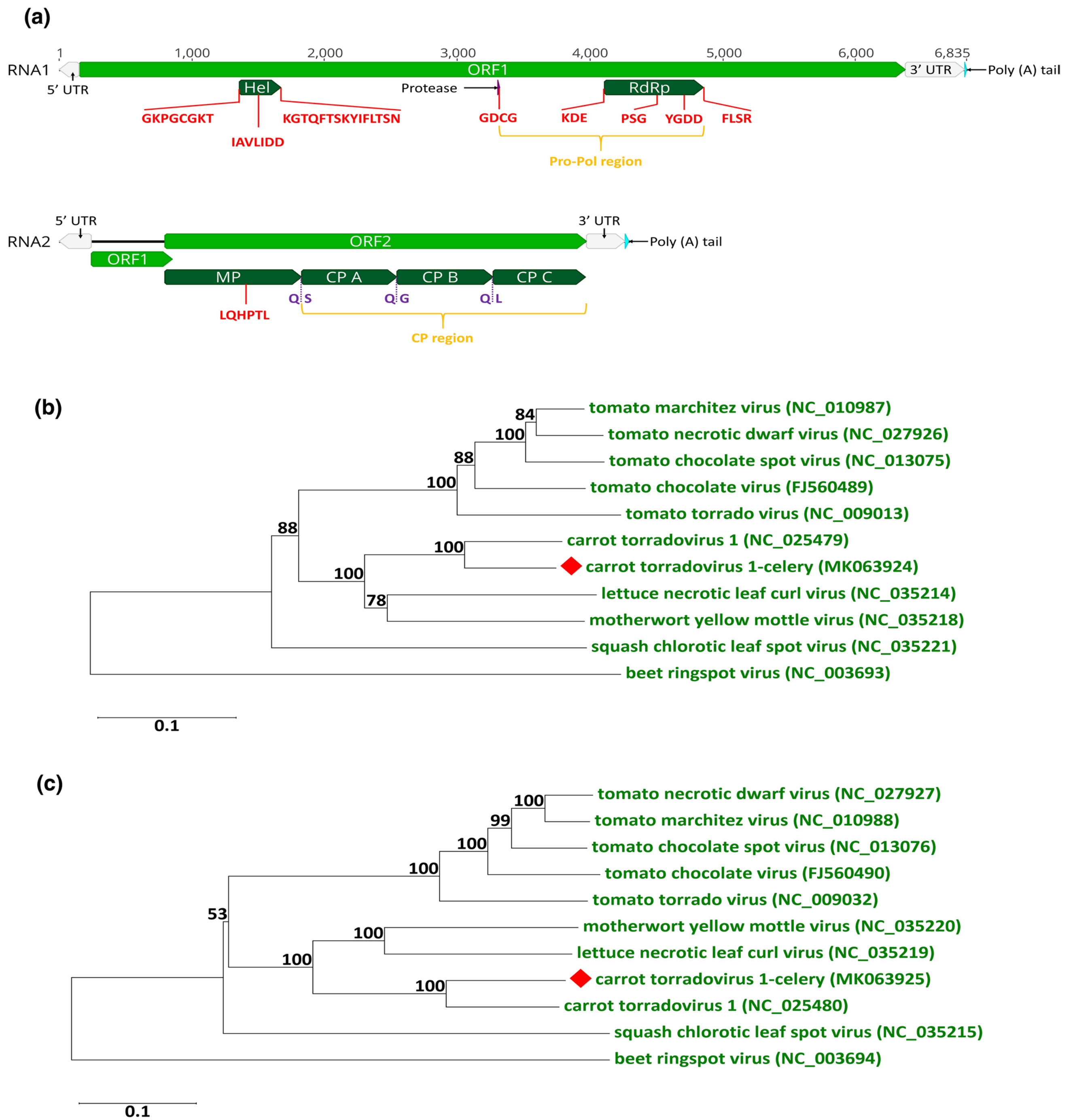

Fig. 1 (a) Schematic representation of the carrot torradovirus 1-celery (isolate JKI 29346) genome. RNA1 encodes a polyprotein containing the helicase (Hel), protease and RNA-dependent RNA polymerase (RdRp). RNA2 has an ORF1 with unknown function and ORF2 encoding a polyprotein containing the movement protein (MP) and three coat protein subunits (CP A = Vp35, CP B = Vp26 and CP C = Vp23). Both strands are flanked by 5′ and 3′ untranslated regions (UTR) and are polyadenylated at the 3’ end. The aa sequences of the conserved motifs in Hel, RdRp and MP are shown in red. The predicted cleavage sites are shown in purple. Taxonomically relevant protein segments are highlighted in yellow **(b and c)**. Neighbour-joining trees based on amino acid sequence alignments of the Pro-Pol **(b)** and CP **(c)** of CaTV1-celery (red diamond) with those of members of the genus *Torradovirus*, with beet ringspot virus (genus *Nepovirus*) as an outgroup. The GenBank accession numbers are shown in brackets. Bootstrap values above 50% (1000 replicates) are indicated for each node, and the scale bar represents a genetic distance of 0.1

Further studies are needed to investigate possible vectors of CaTV1-celery and its potential impact on celery production. Currently, we are developing antibodies against CaTV1-celery as an additional tool for future diagnostic tests and the determination of serological relationships of different torradoviruses.

Acknowledgements The authors thank Ms. Roswitha Ulrich for providing the original sample. The authors are grateful to Ms. Angelika Sieg-Müller, Ms. Petra Lüddecke, Ms. Kerstin Herz and Mr. Jonas Hartrick for their outstanding technical assistance, and Mr. Christopher Ziebell for proofreading.

Funding This research was financed by EUPHRESCO (2015-F-172) project "The Application of Next-Generation Sequencing Technology for the Detection and Diagnosis of Non-culturable Organisms: Viruses and viroids". Yahya Z. A. Gaafar was supported by a German Egyptian Research Long Term Scholarship.

Compliance with ethical standards

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

Ethical approval This research did not involve any studies on human or animal participants.

References

1. Coutts RHA, Livieratos IC (2003) A rapid method for sequencing the 5'- and 3'-termini of double-stranded RNA viral templates using RLM-RACE. J Phytopathol 151(9):525–527. https://doi.org/10.1046/j.1439-0434.2003.00755.x
2. Untergasser A, Cutcutache I, Koressaar T et al (2012) Primer3–new capabilities and interfaces. Nucleic Acids Res 40(15):e115. https://doi.org/10.1093/nar/gks596
3. Sanfaçon H, Wellink J, Le Gall O et al (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. Arch Virol 154(5):899–907. https://doi.org/10.1007/s00705-009-0367-z
4. Thompson JR, Dasgupta I, Fuchs M et al (2017) ICTV virus taxonomy profile: *Secoviridae*. J Gen Virol 98(4):529–531. https://doi.org/10.1099/jgv.0.000779
5. Gorbalenya AE, Koonin EV, Donchenko AP et al (1989) Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. Nucleic Acids Res 17(12):4713–4730
6. Gorbalenya AE, Snijder EJ (1996) Viral cysteine proteinases. Perspect Drug Discov Des 6(1):64–86. https://doi.org/10.1007/BF02174046
7. Argos P, Kamer G, Nicklin MJ et al (1984) Similarity in gene organization and homology between proteins of animal picornaviruses and a plant comovirus suggest common ancestry of these virus families. Nucleic Acids Res 12(18):7251–7267
8. Mushegian AR (1994) The putative movement domain encoded by nepovirus RNA-2 is conserved in all sequenced nepoviruses. Arch Virol 135(3–4):437–441. https://doi.org/10.1007/BF01310028
9. Verbeek M, Dullemans AM, van den Heuvel JFJM et al (2007) Identification and characterisation of tomato torrado virus, a new plant picorna-like virus from tomato. Arch Virol 152(5):881–890. https://doi.org/10.1007/s00705-006-0917-6
10. Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33(7):1870–1874. https://doi.org/10.1093/molbev/msw054
11. Larkin MA, Blackshields G, Brown NP et al (2007) Clustal W and clustal X version 2.0. Bioinformatics 23(21):2947–2948. https://doi.org/10.1093/bioinformatics/btm404
12. Verbeek M, Dullemans AM, van der Vlugt RAA (2017) Aphid transmission of *Lettuce necrotic leaf curl virus*, a member of a tentative new subgroup within the genus *Torradovirus*. Virus Res 241:125–130. https://doi.org/10.1016/j.virusres.2017.02.008
13. Rozado-Aguirre Z, Adams I, Collins L et al (2016) Detection and transmission of carrot torrado virus, a novel putative member of the *Torradovirus* genus. J Virol Methods 235:119–124. https://doi.org/10.1016/j.jviromet.2016.05.018
14. Verbeek M, Dullemans A, van den Heuvel H et al (2010) Tomato chocolàte virus: a new plant virus infecting tomato and a proposed member of the genus *Torradovirus*. Arch Virol 155(5):751–755. https://doi.org/10.1007/s00705-010-0640-1

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.