

Complete genome sequence of tobacco virus 1, a closterovirus from *Nicotiana tabacum*

Fang Wang¹ · Shuishui Qi² · Zhengliang Gao¹ · Ibukun A. Akinyemi² · Dafeng Xu¹ · Benguo Zhou¹

Received: 17 August 2015 / Accepted: 19 December 2015 / Published online: 21 January 2016
© Springer-Verlag Wien 2016

Abstract The complete genome sequence of a novel virus, provisionally named tobacco virus 1 (TV1), was determined, and this virus was identified in leaves of tobacco (*Nicotiana tabacum*) exhibiting leaf mosaic and yellowing symptoms in Anhui Province, China. The genome sequence of TV1 consists of 15,395 nucleotides with 61.6 % nucleotide sequence identity to mint virus 1 (MV1). Its genome organization is similar to that of MV1, containing nine open reading frames (ORFs) that potentially encode proteins with putative functions in virion assembly, cell-to-cell movement and suppression of RNA silencing. Phylogenetic analysis of the heat shock protein 70 homolog (HSP70h) placed TV1 alongside members of the genus *Closterovirus* in the family *Closteroviridae*. To our knowledge, this study is the first report of the complete genome sequence of a closterovirus identified in tobacco.

Introduction

Tobacco (*Nicotiana tabacum*) is an economically important crop worldwide, particularly in China, and its production is affected by several viruses [14]. Currently, more than 30 viruses belonging to 13 genera of seven families have been reported in tobacco, and infected plants can exhibit symptoms such as mosaic, severe stunting, vein banding, and mottle [7, 10, 22]. Early detection of known

viruses and the identification and characterization of novel viruses in tobacco are necessary to develop control measures.

The family *Closteroviridae* includes plant viruses characterized by filamentous particles varying in length from 650 nanometers (nm) to over 2000 nm and a single-stranded positive-sense RNA genome whose size varies from 13 to 19 kilobases. The family *Closteroviridae* currently has four genera: *Closterovirus*, *Ampelovirus*, *Crinivirus* and *Velarivirus* [11]. The genus *Closterovirus* contains well-studied viruses such as beet yellows virus and citrus tristeza virus, among others [11]. Chatzivassiliou et al. reported the first case of a closterovirus isolated from tobacco in Macedonia, Northern Greece. In this study, we identified and characterized the complete genome of a novel virus belonging to a new species in the genus *Closterovirus* by sequencing small RNAs isolated from symptomatic tobacco plants collected in Anhui Province of China. To our knowledge, this is the first report of the complete genome sequence of a closterovirus identified in tobacco [6].

Plant material and extraction of total RNA from tobacco leaves

Leaf samples were collected from several tobacco plants (*Nicotiana tabacum*) of cultivar Yunyan 87 that were grown in China and showed leaf mosaic and yellowing. Leaf samples with similar symptoms were immediately frozen in liquid nitrogen and stored at -80 °C until RNA extraction. Symptomatic samples were pooled at random for sequencing of sRNAs. Total RNA was extracted from the frozen diseased leaves by homogenization with lysis buffer (50 mM Tris-HCl, pH 8.0; 150 mM LiCl; 5 mM

✉ Benguo Zhou
bgzhou@163.com

¹ Tobacco Research Institute, Anhui Academy of Agricultural Sciences, Hefei 230031, Anhui, China

² School of Life Sciences, University of Science and Technology of China, Hefei 230027, Anhui, China

EDTA, pH 8.0; 5 % SDS). The supernatant was treated twice with chloroform and precipitated with isopropanol, followed by re-suspension in nuclease-free water. The integrity of the RNA was verified in ethidium-bromide-stained 1.2 % agarose gels after electrophoresis, and purity was assessed by measuring the absorbance ratio (1.8–2.0) at 260/280 nm using a Eppendorf BioPhotometer Plus (Germany).

Provenance of virus material

A small RNA library was prepared from the total RNA and sequenced using an Illumina HiSeq-2000 (BGI-ShenZhen, China), yielding 20,372,426 clean reads with sizes between 18 and 28 nucleotides (nt). The small RNAs were assembled into 9,943 contigs (length: 33–1484 nt) with a *k*-mer value of 17 using Velvet [12, 24]. These contigs were compared with the non-redundant nucleotide and protein database of GenBank by BLASTn and BLASTx, respectively. At the nucleotide level, 4,768 contigs showed ≥ 90 % identity and ≥ 90 % coverage with the tobacco and known viral sequences in GenBank by BLASTn, including 513, 246 and 122 sequences that were closely similar to sequences present in cucumber mosaic virus, potato virus Y and tobacco vein banding mosaic virus, respectively. Exactly 44 out of the remaining 5,175 contigs with lengths between 81 and 479 nt were identified by BLASTx (e-value cutoff of 10^{-3}) to have distant similarities to members of the family *Closteroviridae*. Forty-three of these contigs showed highest similarity to mint virus 1 (NC_006944.1), and one contig showed highest similarity to carrot yellow leaf virus (NC_013007.1). In order to further characterize these 44 contigs, the genome of mint virus 1 (MV1) was used as a reference to determine their relative position and orientation. Reverse transcription polymerase chain reaction (RT-PCR) and Sanger sequencing were performed to join these contigs together and confirm ambiguous nucleotides. The 5' and 3' ends of the viral genome were obtained by RACE-PCR (TaKaRa Biotechnology Dalian Co., Ltd) and Sanger sequencing. ORFs were predicted using ORF Finder, while the conserved domains/motifs were analyzed by SMART [13]. A phylogenetic tree was constructed using MEGA5 based on the neighbor-joining algorithm with 1000 replicates [20].

Sequence properties

The complete genome of TV1 (KT203917) is 15,395 nt in size and shares the highest nucleotide sequence identity of 61.6 % with MV1. It contains nine putative ORFs that are similar to the analogously positioned ORFs in MV1

(Fig. 1A). Neither the 193-nt-long 5' UTR nor 356-nt-long 3' UTR of TV1 showed significant sequence similarities to any entries in the GenBank database.

ORF1a and ORF1b code for replication-associated proteins [9]. ORF1a encodes a multifunctional 280-kDa polyprotein, which contains conserved domains of papain-like leader protease (L-PRO, pfam05533), methyltransferase (MTR, pfam01660), and helicase (HEL, pfam01443) [13]. The RNA-dependent RNA polymerase (RdRp) is encoded by ORF 1b and shares 83 % identity with that of MV1. This ORF1b-encoding protein is potentially expressed via a +1 ribosomal frameshift from the stop codon UAG of ORF1a, as occurs in some other members of the family *Closteroviridae* [1, 16]. The quintuple gene block (QGB), which is conserved in members of the family *Closteroviridae*, comprises of ORFs 2–6. One hydrophobic protein of 65 aa (7.3 kDa; p7) that is encoded by ORF2 (nt 9,010–9,207) is predicted to contain a transmembrane domain (aa residues 7–29) and plays an important role in cell-to-cell movement as both a signal and an anchor for insertion into the membrane [18]. ORF3 (nt 9,211–11,028) putatively encodes a 605-aa (67.1-kDa) HSP70 homologue protein (HSP70h) that is predicted to play significant roles in both function of tail assembly and cell-to-cell movement [4]. ORF4 (nt 11,029–12,672) of TV1 encodes a putative CP homologue protein of 547 aa (62.3 kDa; CPh). The CPh may function together with hsp70h and CPm in tail assembly and cell-to-cell movement [17]. ORF5 (nt 12,635–13,285) and ORF6 (nt 13,336–13,962) encode a putative minor coat protein of 216-aa (24 kDa; CPm) and a putative major coat protein of 208 aa (22.9 kDa; CP), respectively. Based on analogy to studied closteroviruses, the CP is predicted to encapsidate most of the helical nucleocapsid, while the CPm together with hsp70h and CPh putatively encapsidate a small portion of the 5' end (the viral “tail”) [2, 5, 19, 21], which is involved in virion assembly and cell-to-cell movement [3, 8]. The remaining genes in TV1 are ORF7 (nt 13,971–14,489) and ORF8 (nt 14,486–15,040). ORF7 encodes a putative systemic transport protein of 172 aa (19.4 kDa; p19) [8], while ORF8 encodes a putative protein of 184 aa (21.4 kDa; p21) that contains a P21-like domain (pfam11757), which may function in suppression of RNA silencing [9, 23].

Comparisons between the amino acid sequences encoded by the TV1 genome and four other closteroviruses showed that each protein of TV1 has the highest predicted amino acid sequence similarity to its counterpart MV1 (Table 1). In addition, a phylogenetic tree using HSP70h amino acid sequences was constructed using MEGA5 based and the neighbor-joining algorithm to investigate the relationship between TV1 and other members of the family *Closteroviridae*. The tree placed TV1 alongside members of the genus *Closterovirus* and closest to mint virus 1

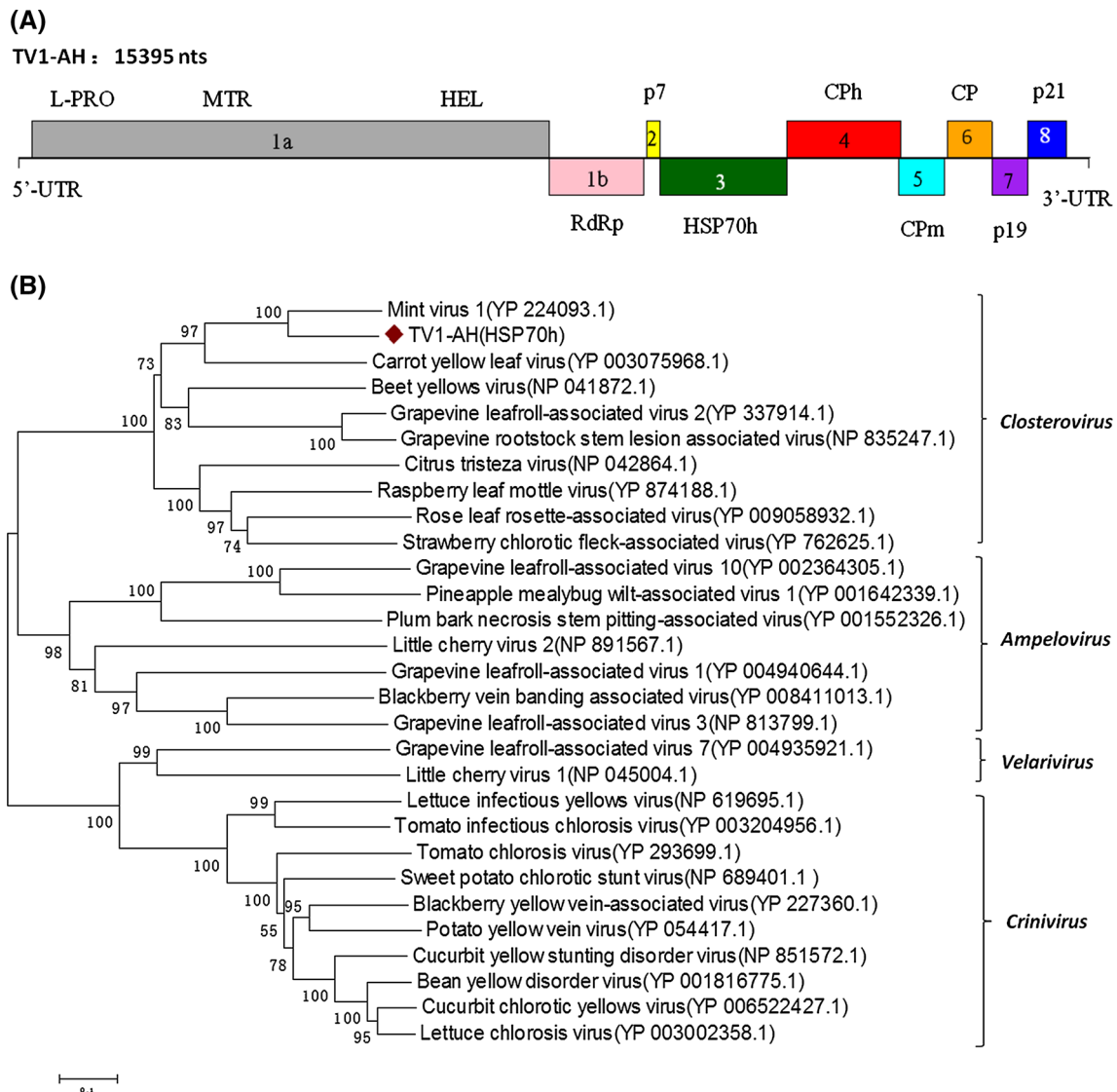


Fig. 1 (A) Schematic representation of the genome organization of tobacco virus 1 (TV1). L-PRO, proteinase; MTR, methyl transferase; HEL, helicase; RdRp, RNA-dependent RNA polymerase; p4, 4-kDa protein; HSP70h, heat shock protein 70 homolog; CPh, coat protein homolog; CPm, minor coat protein; CP, coat protein; p19, 19.4-kDa protein; p21, 21.4-kDa protein. (B) A phylogenetic tree, constructed

by the neighbour-joining method using MEGA 5, showing the relationship between TV1 and some members of the family *Closteroviridae* based on HSP70h. Accession numbers and virus names are given directly in the phylogenetic tree. Values at the nodes show the bootstrap values from 1000 replicates, and the bars represent the evolutionary distances

Table 1 Similarities (percentage) between the amino acid sequences encoded by the genes in the TV1 genome and their counterparts in several closely related closteroviruses: beet yellows virus (BYV),

carrot yellow leaf virus (CYLV), mint virus 1 (MV1) and grapevine leafroll-associated virus 2 (GLRaV-2)

Virus	ORF1a	ORF1b	P7	Hsp70	CPh	CPm	CP	P19	P21
MV1	48.89	81.84	57.58	65.62	53.04	53.04	46.15	25.89	39.78
CYLV	33.60	63.81	20.00	48.70	34.70	37.17	31.92	13.13	22.90
BYV	26.48	56.84	26.15	44.72	8.01	—	27.23	10.38	23.66
GLRaV-2	23.94	57.20	15.38	39.04	—	30.94	30.62	13.87	15.38

(Fig. 1B). In conclusion, this study confirms the presence of tobacco virus 1 (TV1) in leaves from a tobacco plant in Anhui province of China. Considering that the sequence similarities of all taxonomically relevant proteins (i.e., RdRp, HSP70h and CP) between the studied virus and recognized closteroviruses are far below the species demarcation threshold proposed by the *Closteroviridae* Study Group [15], we propose this virus to be representative of a new species in the genus, for which we propose the name “*Tobacco virus 1*”.

Acknowledgements This study was supported by Key Program of Anhui Province Tobacco Monopoly Administration (no. 20150551007) and the Chinese National Natural Science Foundation (no. 31272011).

References

- Agranovsky AA, Koonin EV, Boyko VP, Maiss E, Frotschl R, Lunina NA, Atabekov JG (1994) Beet yellows closterovirus: complete genome structure and identification of a leader papain-like thiol protease. *Virology* 198:311–324
- Agranovsky AA, Lesemann DE, Maiss E, Hull R, Atabekov JG (1995) “Rattlesnake” structure of a filamentous plant RNA virus built of two capsid proteins. *Proc Natl Acad Sci USA* 92:2470–2473
- Alzhanova DV, Hagiwara Y, Peremyslov VV, Dolja VV (2000) Genetic analysis of the cell-to-cell movement of beet yellows closterovirus. *Virology* 268:192–200
- Alzhanova DV, Napuli AJ, Creamer R, Dolja VV (2001) Cell-to-cell movement and assembly of a plant closterovirus: roles for the capsid proteins and Hsp70 homolog. *EMBO J* 20:6997–7007
- Alzhanova DV, Prokhnovsky AI, Peremyslov VV, Dolja VV (2007) Virion tails of beet yellows virus: coordinated assembly by three structural proteins. *Virology* 359:220–226
- Chatzivassiliou EK, Katis NI, Tchomguia M, Peters D (1999) A closterovirus (family: *Closteroviridae*) isolated from tobacco crops in Northern Greece (Macedonia). *Plant Dis* 83:198
- Ding M, Yang C, Zhang L, Jiang ZL, Fang Q, Qin XY, Zhang ZK (2011) Occurrence of Chilli veinal mottle virus in *Nicotiana tabacum* in Yunnan, China. *Plant Dis* 95:357
- Dolja VV (2003) Beet yellows virus: the importance of being different. *Mol Plant Pathol* 4:91–98
- Dolja VV, Kreuze JF, Valkonen JPT (2006) Comparative and functional genomics of closteroviruses. *Virus Res* 117:38–51
- Dong JH, Yin YY, Xu XY, Duan YM, Zhang ZK (2010) First report of tomato spotted wilt virus in tomato and tobacco in China. *J Plant Pathol* 92:S121–S121
- Karasev AV (2000) Genetic diversity and evolution of closteroviruses. *Annu Rev Phytopathol* 38:293–324
- 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
- Letunic I, Doerks T, Bork P (2015) SMART: recent updates, new developments and status in 2015. *Nucleic Acids Res* 43:D257–D260
- Lu XP, Gui YJ, Xiao BG, Li YP, Tong ZJ, Liu Y, Bai XF, Wu WR, Xia L, Huttner E, Kilian A, Fan LJ (2013) Development of DArT markers for a linkage map of flue-cured tobacco. *Chin Sci Bull* 58:641–648
- Martelli GP, Agranovsky AA, Bar-Joseph M, Boscia D, Candresse T, Coutts RH, Dolja VV, Hu JS, Jelkmann W, Karasev AV, Martin RR, Minafra A, Namba S, Vetten HJ (2012) Family *Closteroviridae*. In: King A, Adams MJ, Carstens EB, Lefkowitz E (eds) *Virus taxonomy*. Ninth report of the International Committee on Taxonomy of Viruses. Elsevier-Academic Press, San Diego, pp 987–1001
- Menzel W, Goetz R, Lesemann DE, Vetten HJ (2009) Molecular characterization of a closterovirus from carrot and its identification as a German isolate of carrot yellow leaf virus. *Arch Virol* 154:1343–1347
- Napuli AJ, Alzhanova DV, Doneanu CE, Barofsky DF, Koonin EV, Dolja VV (2003) The 64-kilodalton capsid protein homolog of beet yellows virus is required for assembly of virion tails. *J Virol* 77:2377–2384
- Peremyslov VV, Pan YW, Dolja VV (2004) Movement protein of a closterovirus is a type III integral transmembrane protein localized to the endoplasmic reticulum. *J Virol* 78:3704–3709
- Satyanarayana T, Gowda S, Ayllon MA, Dawson WO (2004) Closterovirus bipolar virion: evidence for initiation of assembly by minor coat protein and its restriction to the genomic RNA 5' region. *Proc Natl Acad Sci USA* 101:799–804
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28:2731–2739
- Tian TY, Rubio L, Yeh HH, Crawford B, Falk BW (1999) Lettuce infectious yellows virus: in vitro acquisition analysis using partially purified virions and the whitefly *Bemisia tabaci*. *J Gen Virol* 80:1111–1117
- Tian YP, Liu JL, Zhang CL, Liu YY, Wang B, Li XD, Guo ZK, Valkonen JPT (2011) Genetic diversity of potato virus Y infecting tobacco crops in China. *Phytopathology* 101:377–387
- Tzanetakis IE, Postman JD, Martin RR (2005) Characterization of a novel member of the family *Closteroviridae* from *Mentha* spp. *Phytopathology* 95:1043–1048
- Wu Q, Luo Y, Lu R, Lau N, Lai EC, Li WX, Ding SW (2010) Virus discovery by deep sequencing and assembly of virus-derived small silencing RNAs. *Proc Natl Acad Sci USA* 107:1606–1611