8-2009

Triticum mosaic virus: A Distinct Member of the Family *Potyviridae* with an Unusually Long Leader Sequence

Satyanarayana Tatineni
USDA-ARS, Satyanarayana.Tatineni@ars.usda.gov

Amy D. Ziems
University of Nebraska-Lincoln

Stephen N. Wegulo
University of Nebraska-Lincoln, swegulo2@unl.edu

Roy French
USDA-ARS

Phytopathology Aug 2009, Volume 99, Number 8

Virology

Triticum mosaic virus: A Distinct Member of the Family *Potyviridae* with an Unusually Long Leader Sequence

Satyanarayana Tatineni, Amy D. Ziems, Stephen N. Wegulo, and Roy French

First and fourth authors: United States Department of Agriculture–Agriculture Research Service and Department of Plant Pathology; and second and third authors: Department of Plant Pathology, University of Nebraska, Lincoln 68583.
Accepted for publication 4 April 2009.

ABSTRACT

Tatineni, S., Ziems, A. D., Wegulo, S. N., and French, R. 2009. *Triticum mosaic virus*: A distinct member of the family *Potyviridae* with an unusually long leader sequence. Phytopathology 99:943-950.

The complete genome sequence of *Triticum mosaic virus* (TriMV), a member in the family *Potyviridae*, has been determined to be 10,266 nucleotides (nt) excluding the 3′ polyadenylated tail. The genome encodes a large polyprotein of 3,112 amino acids with the "hall-mark proteins" of potyviruses, including a small overlapping gene, PIPO, in the P3 cistron. The genome of TriMV has an unusually long 5′ nontranslated region of 739 nt with 12 translation initiation codons and three small open reading frames, which resemble those of the internal ribosome entry site containing 5′ leader sequences of the members of *Picornaviridae*. Pairwise comparison of 10 putative mature proteins of TriMV with those of representative members of genera in the family *Potyviridae* revealed 33 to 44% amino acid identity within the highly conserved NIb protein sequence and 15 to 29% amino acid identity within the least conserved P1 protein, suggesting that TriMV is a distinct member in the family *Potyviridae*. In contrast, TriMV displayed 47 to 65% amino acid sequence identity with available sequences of mature proteins of *Sugarcane streak mosaic virus* (SCSMV), an unassigned member of the *Potyviridae*. Phylogenetic analyses of the complete polyprotein, NIa-Pro, NIb, and coat protein sequences of representative species of six genera and unassigned members of the family *Potyviridae* suggested that TriMV and SCSMV are sister taxa and share a most recent common ancestor with tritimoviruses or ipomoviruses. These results suggest that TriMV and SCSMV should be classified in a new genus, and we propose the genus *Poacevirus* in the family *Potyviridae*, with TriMV as the type member.

Additional keyword: wheat.

Viruses from several different families infect wheat (*Triticum aestivum* L.) in the Great Plains and other parts of the United States. These viruses include *Agropyron mosaic virus* (AgMV), *Barley yellow dwarf virus*, S*oil-borne wheat mosaic virus*, *Triticum mosaic virus* (TriMV), *Wheat American striate mosaic virus*, *Wheat mosaic virus*, and *Wheat streak mosaic virus* (WSMV) (6,20,21). Among these viruses, WSMV is an economically important virus causing significant yield losses in the United States (6). TriMV was recently reported from Kansas, naturally infecting WSMV-resistant wheat cultivars (20); however, the impact of this virus on yield losses in wheat remains to be known.

The *Potyviridae* is the largest family of positive-stranded RNA viruses infecting plants, divided into six genera based on their genetic relatedness, vector transmission, and genome organization (2,5). The genus *Potyvirus*, with *Potato virus* Y (PVY) as the type member, contains numerous economically important aphid-transmitted virus species and is the most thoroughly characterized genus among the family *Potyviridae*. Other genera include *Rymovirus*, with *Ryegrass mosaic virus* (RGMV) as the type species, transmitted by *Abacarus* mites; *Tritimovirus*, with WSMV as the type member, vectored by wheat curl mites (*Aceria tosichella*); *Ipomovirus*, with *Sweet potato mild mottle virus* (SPMMV) as the type species, transmitted by whiteflies; and *Macluravirus*, with *Maclura mosaic virus* (MacMV) as the type species, with characteristic short virus particles transmitted by aphids. These five genera all contain monopartite viruses, whereas the genus *Bymovirus*, with *Barley yellow mosaic virus* (BaYMV) as the type member, contains bipartite viruses transmitted by plasmodiophorids. Additionally, there are several unassigned members which may warrant separate genera in the family *Potyviridae*. *Blackberry virus* Y (BlVY) (23) and *Spartina mottle virus* (9) do not belong to any of the existing genera. *Sugarcane streak mosaic virus* (SCSMV) was initially reported as a possible member of the genus *Tritimovirus* (10,11), but subsequent phylogenetic analyses concluded that SCSMV is an additional unassigned member in the family *Potyviridae* (12,17).

During 2006, temperature-sensitive WSMV-resistant wheat cultivars produced virus-like symptoms in Kansas, and the causal agent was characterized and identified as TriMV in the family *Potyviridae* (20). The coat protein (CP) of TriMV showed 49% amino acid sequence identity with that of SCSMV (20). Recently, Seifers et al. (19) reported the wheat curl mite, which also transmits WSMV, as the vector of TriMV. During the 2008 growing season, ≈30% of wheat samples submitted to the University of Nebraska-Lincoln Plant and Pest Diagnostic Clinic tested positive for TriMV (28), which warranted the molecular characterization of this new virus. As a first step in this direction, we report the complete genome sequence of TriMV, the amino acid sequence comparison of its polyprotein and predicted mature proteins, and phylogenetic analysis with representative members in six genera and unassigned members of the family *Potyviridae*. Our results provide evidence that TriMV is a distinct species in the family *Potyviridae*, with an unusually long 5′ nontranslated region (NTR) (739 nucleotides [nt]), and is proposed as the type member of a new genus *Poacevirus* in the family *Potyviridae*.

Corresponding author: S. Tatineni
E-mail address: Satyanarayana.tatineni@ars.usda.gov

doi:10.1094/PHYTO-99-8-0943

MATERIALS AND METHODS

Virus source. Wheat plants with mosaic and chlorotic streak symptoms collected from Red Willow County, Nebraska that failed to react with WSMV antiserum but positively reacted with TriMV antiserum in double-antibody sandwich enzyme-linked immunosorbent assay (DAS-ELISA) (Agdia, Elkhart, IN) provided the virus source for the present investigation. A pure culture of TriMV was obtained by inoculating crude sap from this virus isolate at high dilutions (1:6,600) to single-leaf stage wheat seedlings (cv. Tomahawk) for three successive passages. Briefly, virus inoculum was prepared by grinding infected wheat leaves in 20 mM sodium phosphate buffer, pH 7.0 (inoculation buffer), and further diluted to 1:6,600 in inoculation buffer. At each passage, individual infected wheat plants that reacted positively with TriMV antiserum in DAS-ELISA were used as the source of inoculum for the next passage.

Virus purification and isolation of viral RNA. Wheat (cv. Tomahawk) leaves infected with TriMV at 14 to 17 days post-inoculation (dpi) showing mosaic and chlorotic streak symptoms were used to partially purify virions as described by Lane (14) with slight modifications. Briefly, frozen infected tissue (20 to 25 g) was ground in an electric blender in 0.1 M sodium citrate buffer, pH 6.5 (SCB) containing 0.1% β-mercaptoethanol (4 ml/g of tissue). The extract was filtered through four layers of muslin cloth and clarified at 7,700 × *g* for 10 min. The supernatant was treated with 2% Triton-X 100 at 4°C for 5 to 10 min and layered on 5 ml of 20% sucrose (wt/vol) in SCB. The virus particles were concentrated by centrifugation at 118,000 × *g* for 1.5 h in a Beckman 50.2 Ti rotor. The virus pellet was resuspended in 18 ml of cold SCB and clarified by low-speed centrifugation at 7,700 × *g* for 10 min. Virions from the supernatant (9 ml) were further concentrated by centrifugation (139,000 × *g* for 1.5 h in a Beckman 70.2 Ti rotor) through a 1-ml cushion of 20% sucrose. Virions were recovered by suspending the pellet in 400 µl of suspension buffer (10 mM Tris-Cl, pH 8.0; 1 mM EDTA; and 100 mM NaCl). All centrifugation steps were carried out at 4°C.

Viral RNA was isolated from partially purified virus by treating with 1% sodium dodecyl sulfate and proteinase K at 200 µg/ml (Ambion, Austin, TX) at 50°C for 20 min, followed by two phenol-chloroform extractions and ethanol precipitation. Polyadenylated viral RNA was purified further by passing viral RNA through an Illustra mRNA purification oligo(dT) column (GE Healthcare, Waukesha, WI). The viral RNAs were analyzed on a 0.9% agarose gel containing 10 mM methyl mercury hydroxide (3).

cDNA synthesis and molecular cloning of TriMV genome. TriMV RNA isolated from partially purified virions, followed by oligo(dT)-cellulose column purification, was used as the template for reverse transcription with the combination of oligo dT and random hexamers as primers. The SuperScript Choice System for cDNA synthesis kit (Invitrogen, Carlsbad, CA) was used to make the first-strand cDNA, followed by second-strand synthesis as per the instructions provided with the kit. *Eco*RI (*Not*I) adaptors were ligated to double-stranded (ds)DNA products of second-strand reaction, followed by passage through column chromatography to select larger dsDNA products. The dsDNA products with *Eco*RI (*Not*I) adaptors were ligated into *Eco*RI-cut dephosphorylated pGEM-7Zf(–) (Promega Corporation, Madison, WI), followed by transformation into *Escherichia coli* JM109. The inserts were characterized by restriction enzyme digestion with *Eco*RI, *Hin*dIII, *Not*I, *Nco*I, and *Pst*I, and cDNA clones with insert size >2.0 kbp were selected for further sequencing.

Reverse-transcription polymerase chain reaction amplification. Sequences obtained from cDNA clones aligned into two contigs and comparison of these sequences with reported genome sequence of WSMV (22) revealed a gap in the genome sequence of TriMV. This gap was amplified by reverse-transcription polymerase chain reaction (RT-PCR), followed by cloning and sequencing. An oligonucleotide, Tr-1, complementary to nucleotides 5457 to 5425, was used for reverse transcription with SuperScript II RT, followed by PCR amplification with oligonucleotides Tr-2, corresponding to nucleotides 2292 to 2324, and Tr-1 using Taq DNA polymerase. PCR conditions used for 5′ rapid amplification of cDNA ends (RACE) were also used for RT-PCR, except for a 3-min extension time at 72°C (see below). The RT-PCR product was ligated into pGEM-Teasy vector (Promega Corporation) and three independent clones were sequenced completely in both directions. The consensus sequence of RT-PCR clones was used to fill the gap in the genome sequence of TriMV.

Determination of the 5′ end sequence of the genomic RNA. The exact 5′ end sequence of TriMV was determined using viral RNA isolated from partially purified virions as the template with the 5′ RACE system (Invitrogen). First-strand cDNA synthesis was performed using the gene-specific primer 1 (GSP1) (5′-CCAAGATACTGCACACAAGAAATTGAACAAAAT-3′, complementary to nucleotides 1240 to 1208), and unincorporated nucleotides and GSP1 primer were removed using sodium iodide, followed by column purification as per the instructions in the kit. The 3′ end of the purified first-strand cDNA was 'C' tailed using terminal deoxynucleotidyl transferase (TdT), and TriMV-specific DNA was amplified using GSP2 (5′-CTGTAAGTATGAGCAGCATATCGTTTCAACATT-3′, complementary to nucleotides 1197 to 1165) and abridged anchor primer (AAP, supplied with kit) using *Taq* DNA polymerase (2.5 U) for 1 cycle at 94°C for 2 min; followed by 35 cycles at 94°C for 20 s, 55°C for 20 s and 72°C for 2 min; and 1 cycle at 72°C for 10 min. The 5′ RACE PCR product was ligated into pGEM-Teasy vector (Promega Corporation) and the inserts were sequenced.

RT-PCR was performed to further examine the authenticity of the length of the 5′ NTR of TriMV. RNA isolated from virions was reverse transcribed with oligonulceotide Tr-26 (complementary to nucleotides 1234 to 1217), followed by PCR amplification with oligonucleotides Tr-16 (corresponding to nucleotides 1 to 33) in combination with Tr-18 (complementary to nucleotides 829 to 812) or Tr-22 (complementary to nucleotides 1009 to 992) or Tr-26. PCR was performed with the same conditions that were used for the 5′ RACE system except for using Herculase II Fusion DNA polymerase (Stratagene, La Jolla, CA). The RT-PCR products were analyzed on a 1.0% agarose gel in 1× Tris-acetate EDTA buffer (40 mM Tris-acetate, pH 8.3, and 1 mM EDTA).

Nucleotide sequencing and sequence analyses. Sequencing of TriMV cDNA, RT-PCR, and 5′ RACE clones was carried out at the University of Florida ICBR Core DNA Sequencing Facility using an Applied Biosystems 3730 model sequencer. Both strands of TriMV clones were sequenced by the primer walking method. Contig sequences were created from individual sequences of cDNA, RT-PCR, and 5′ RACE clones using Sequencher 4.1 (Gene Codes, Ann Arbor, MI). Both strands of at least two to three independent clones covering the entire genome were sequenced to obtain the sequence of genomic RNA. Additionally, sequences of 12 independent cDNA and 5′ RACE clones were used to obtain an unambiguous sequence of the 5′ end of the genomic RNA.

Pairwise sequence comparisons of TriMV mature proteins with selected members of the family *Potyviridae* were carried out using the ALIGN program of online analysis tools (http://molbiol-tools.ca). Multiple sequence alignments were performed with the ClustalW program (25), and the MEGA 4.0 analysis package (24) was used to perform phylogenetic analysis with the neighbor-joining (NJ) method using the JTT matrix and pairwise gap deletion, with 1,000 bootstrap replicates as the test of phylogeny.

RESULTS

Analysis of viral RNA. TriMV RNA extracted from partially purified virions was separated into a single band on 0.9% agarose

gel containing 10 mM methyl mercury hydroxide (Fig. 1, lane 1). The size of TriMV RNA was estimated to be ≈10.0 kb based on RNA molecular size markers (Invitrogen) and WSMV RNA. WSMV virion RNA (9.3 kb) was included in agarose gel as the reference standard, which migrated slightly faster than that of TriMV RNA (Fig. 1, lanes 2 and 3). Taken together, agarose gel electrophoresis of viral RNA and genome sequencing (see below) indicated that the RNA genome of TriMV is ≈1.0 kb longer than that of WSMV.

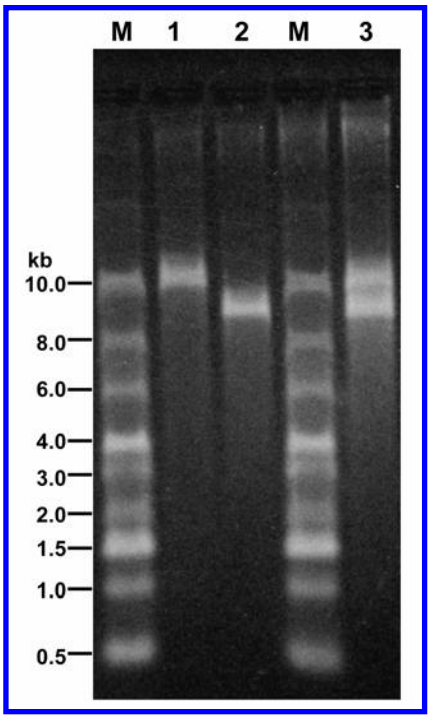

Fig. 1. Analysis of *Triticum mosaic virus* (TriMV) RNA through 1.0% agarose gel in the presence of 10 mM methyl mercury hydroxide. Lane M: RNA size markers (Invitrogen); lane 1: TriMV RNA; lane 2: *Wheat streak mosaic virus* (WSMV) RNA; lane 3: co-migration of TriMV and WSMV RNAs. Note the slow migration of TriMV RNA when compared with 9.3-kb RNA of WSMV (compare lanes 1 and 2; and lane 3).

Genome organization of TriMV. The complete nucleotide sequence of TriMV genomic RNA was assembled from the sequence of multiple overlapping cDNA, RT-PCR, and 5′ RACE clones (Fig. 2). The sequence of TriMV genomic RNA was determined to be 10,266 nt, excluding the variable-length polyadenylated tail at the 3′ end of the genomic RNA. The genome sequence has been deposited in GenBank with the accession number FJ669487. The genome organization of TriMV is typical of monopartite viruses of the family *Potyviridae*, with one large open reading frame (ORF), which initiates (AUG) at nucleotide positions 740 to 742 and terminates with a stop codon (UAG) at nucleotide positions 10076 to 10078 (Fig. 2). The large ORF encodes a polyprotein of 3,112 amino acids with a predicted molecular weight of 352.9 kDa.

Examination of TriMV polyprotein, followed by comparison with those of WSMV, *Brome streak mosaic virus* (BrSMV) (genus *Tritimovirus*), and *Turnip mosaic virus* (TuMV) (genus *Potyvirus*), revealed potential viral-encoded proteinase cleavage sites in the polyprotein (Table 1). Processing of the polyprotein at these cleavage sites would result in the production of mature "hallmark" proteins of potyviruses—P1, HC-Pro, P3, 6K1, CI, 6K2, NIa-VPg, NIa-Pro, NIb, and CP—as in other species of the genera *Potyvirus*, *Rymovirus*, *Tritimovirus*, and *Ipomovirus* (*Sweet potato mild mottle virus*) (Fig. 2). P1 and HC-Pro proteinases of TriMV likely autocatalytically cleave at predicted cleavage sites to produce mature P1 and HC-Pro, respectively. As reported in other potyviruses, NIa-Proteinase is likely responsible for the cleavages in the C-terminal two-thirds of the polyprotein at predicted cleavage sites (Table 1), which would result in release of the remaining mature proteins. The predicted cleavage sites of NIa-Pro are unusual in that they contain a conserved histidine (H) at the –1 position (Table 1) rather than the typical gluta-

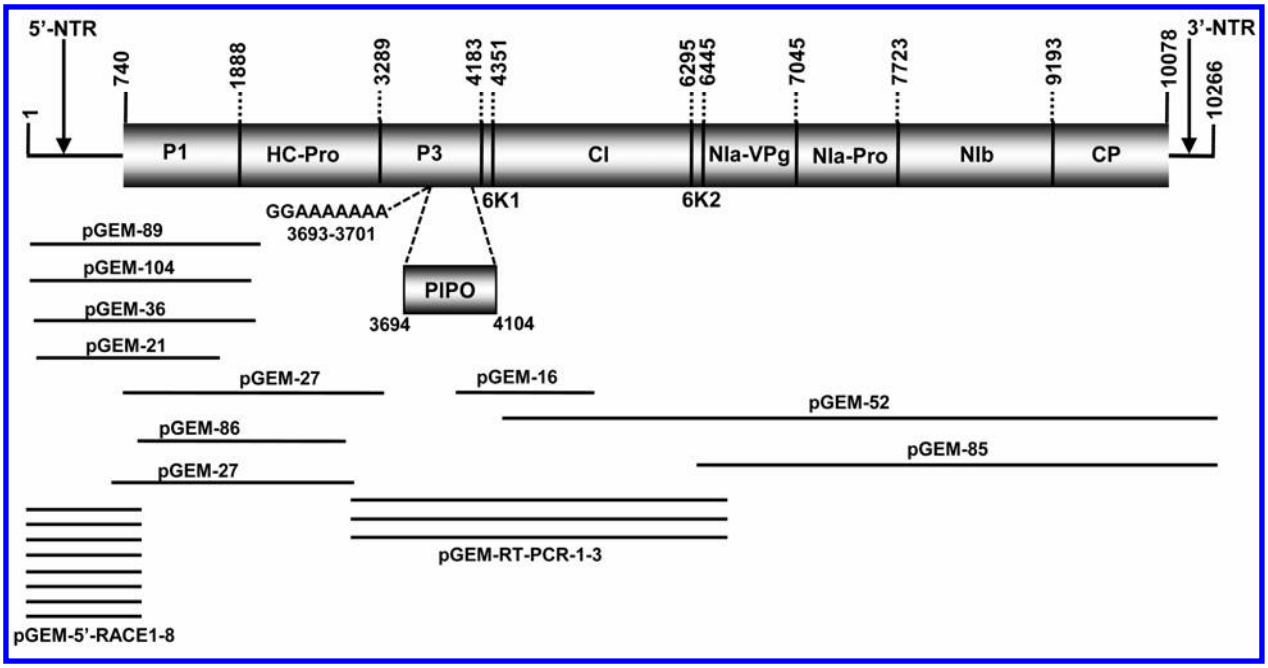

Fig. 2. Schematic representation of the genome organization of *Triticum mosaic virus* (TriMV) depicting coding proteins in cistrons and the location of cleavage sites (dotted lines). The solid lines on either side of the genome organization represent nontranslated regions (NTRs). The location of a small open reading frame (ORF), *pretty interesting Potyvirus ORF* (PIPO) (7), in P3 is shown below the genomic organization as an enlarged view. The location of the conserved motif (G_2A_7) at the 5′ end of PIPO is indicated. Horizontal lines below the genomic organization represent the location of overlapping cDNA, reverse-transcription polymerase chain reaction, and 5′ rapid amplification of cDNA ends clones used to obtain the complete genome sequence of TriMV. Note that several independent clones covering the 5′ end of genomic RNA were sequenced to obtain an unambiguous 5′ end sequence.

mine (Q) or glutamic acid (E) found with almost all other potyviruses (1).

Blast searches revealed potyviral conserved motifs in the TriMV polyprotein, including protease domains in P1, HC-Pro, and NIa; helicase domain in CI; polymerase motifs in NIb; and potyviral coat protein motifs in CP. We also found a conserved tyrosine residue at amino acid position 1979 in TriMV polyprotein, which is present in the same location as in *Tobacco etch virus* (TEV) that forms the phosphodiester linkage of VPg to the 5′ end of the viral RNA (18).

Recently, Chung et al. (7) reported an overlapping essential gene in the P3 cistron in the family *Potyviridae*, which was termed the *pretty interesting Potyviridae ORF (*PIPO). PIPO is expressed as a P3-PIPO fusion product via ribosomal frameshifting or transcriptional slippage at a conserved $G_{1-2}A_{6-7}$ motif at the 5′ end of PIPO (7). Analysis of the TriMV genome also revealed a conserved G_2A_7 motif (nucleotides 3693 to 3701) in the P3 cistron similar to other reported members of the *Potyviridae* (Fig. 2). The P3-PIPO fusion product consists of 135 amino acids of P3 (nucleotides 3290 to 3695) and 136 amino acids of PIPO (nucleotides 3694 to 4104) (Fig. 2). Pairwise comparison of the TriMV PIPO protein sequence with those of representative members of *Potyviridae* revealed no significant amino acid identity (data not shown).

The 3′ NTR of TriMV is 188 nt long excluding the polyadenylated tail, which is similar to lengths reported for other potyviruses. The 3′ NTR shares 51% nucleotide sequence identity with that of SCSMV and only 30 to 44% nucleotide identity with the 3′ NTRs of other reported members of *Potyviridae*.

TriMV contains 739-nt leader sequence. The exact 5′ end sequence of TriMV genomic RNA was determined using the 5′ RACE system. Based on the fact that the 5′ NTRs of reported potyviruses range are 112 (RGMV) to 184 nt (PVY), we designed a gene-specific primer ≈500 nt inside the P1 coding region for the 5′ RACE system, expecting a PCR product of ≈650 bp. In contrast, we obtained a single RT-PCR product of ≈1,200 bp (Fig. 3A), which was cloned into pGEM-Teasy vector. Eight independent clones were sequenced and all eight clones possess identical sequences, with the extreme 5′ end sequence as 5′-AAAAUU….. (Fig. 4A). The contig sequence obtained from 5′ RACE and cDNA clones revealed that TriMV contains an unusually long leader sequence of 739 nt (Fig. 4A). The 5′ NTR is A-U rich (25% As and 33% Us), as in other potyviruses with no significant homology with reported potyviruses.

Because the 5′ NTR of TriMV is unusually long relative to other members of the family *Potyviridae*, we further examined its authenticity by RT-PCR amplification using virion RNA as the template, with a forward primer located at the extreme 5′ end of the genomic RNA (Tr-16) and reverse primers ending at nucleotides 829 (Tr-18), 1,009 (Tr-22), and 1,234 (Tr-26). We obtained RT-PCR products with estimated sizes of 850, 1,000, and 1,250 bp with the primer pairs of Tr-16/Tr-18, Tr-16/Tr-22, and Tr-16/Tr-26, respectively (Fig. 3B), indicating that the 5′ NTR sequence obtained by cDNA clones and the 5′ RACE system are, indeed, authentic.

TABLE 1. Putative proteinase cleavage sites in *Triticum mosaic virus* polyprotein

Proteinase	Peptide junction	Amino acid sequence
P1	P1/HC-Pro	EGLTYY/S
HC-Pro	HC-Pro/P3	KDYRIG/G
NIa	P3/6K1	EDYVLH/A
NIa	6K1/CI	DSFVFH/A
NIa	CI/6K2	DCLVFH/G
NIa	6K2/NIa-VPg	EKYVMH/G
NIa	NIa-VPg/NIa-Pro	DNYVPH/S
NIa	NIa-Pro/NIb	SEFIFH/N
NIa	NIb/CP	DEFVFH/S

The 739-nt-long 5′ NTR of TriMV resembles those of species within the family *Picornaviridae*, whose leader sequences range in size from 628 nt (*Human rhinovirus* B; GenBank accession no. NC_001490) to 1,039 nt (*Foot-and-mouth disease virus* type C; GenBank accession no. NC_002554) and contain internal ribosome entry site (IRES) elements (4). Interestingly, the 5′ NTR of TriMV contains 12 translation initiation codons (AUG) in all three reading frames preceding the translation initiation codon of the polyprotein (Fig. 4A), which is unusual for 5′-leader sequences of most RNA plant viruses. Moreover, we observed three small ORFs (sORFs) in the 5′-leader sequence either within the 5′ NTR (sORF2) or overlapping with the 5′ of the polyprotein (sORFs1 and -3) (Fig. 4B). sORF1 encodes 61 amino acids in reading frame 1 from nucleotides 598 to 783; sORF2 encodes 33 amino acids in reading frame 2 from nucleotides 281 to 382; and sORF3 with a start codon at nucleotide 525 and ending at nucleotide 758 in reading frame 3 with 77 amino acids (Fig. 4B). Blast searches revealed that there is no significant homology of proteins encoded by sORFs with other reported protein sequences in GenBank.

Sequence comparison of TriMV mature proteins with those of other members of the family *Potyviridae*. Pairwise comparisons of amino acid identities of 10 mature proteins of TriMV with those of PYV and SCMV (*Potyvirus*), RGMV and AgMV (*Rymovirus*), MacMV (*Macluravirus*), *Cucumber vein yellowing virus* (CVYV) and SPMMV (*Ipomovirus*), WSMV and BrSMV (*Tritimovirus*), BaYMV (*Bymovirus*), and BlVY and SCSMV (unassigned members) clearly indicate that TriMV is a distinct species in the family *Potyviridae* (Table 2). As observed with other potyvirid species, NIb of TriMV shared the highest amino acid identity (33 to 65%) with the corresponding protein of other

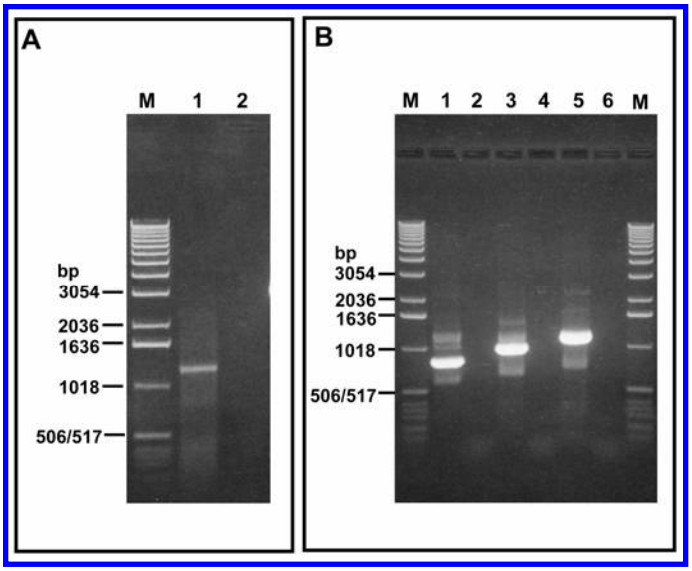

Fig. 3. Analysis of 5′ nontranslated region of *Triticum mosaic virus* (TriMV). **A,** Agarose gel electrophoresis (1% in 1× Tris-acetate EDTA [TAE] buffer) of 5′ rapid amplification of cDNA ends product of TriMV RNA. Gene-specific primer complementary to nucleotides 1240 to 1208 used for reverse transcription, followed by polymerase chain reaction (PCR) amplification with gene-specific primer complementary to nucleotides 1197 to 1165 and an abridged anchor primer (supplied with kit). Lane 1: C-tailed cDNA as the template for PCR; lane 2: cDNA without C-tailing as the template (negative control); lane M: DNA size markers. **B,** Agarose gel electrophoresis (1% in 1× TAE buffer) of reverse-transcription PCR products with extreme 5′ end primer (forward primer, Tr-16) and reverse primers at different locations in the TriMV RNA. Lanes 1, 3, and 5: TriMV cDNA used for PCR; lanes 2, 4, and 6: water control. Lanes 1 and 2: PCR with Tr-16 (5′ end primer) and Tr-18 (complementary to nucleotides 829 to 812) primer pair; lanes 3 and 4: PCR with Tr-16 and Tr-22 (complementary to nucleotides 1009 to 992) primer pair; and lanes 5 and 6: PCR with Tr-16 and Tr-26 (complementary to nucleotides 1234 to 1217) primer pair. cDNA reverse transcribed with primer Tr-26 was used for PCR amplifications.

members of the family *Potyviridae* (Table 2). P1 was the most diverse protein, with amino acid identity ranging from 15% (SPMMV and BlVY) to 29% (BrSMV) (Table 2). HC-Pro, P3, 6K1, CI, 6K2, NIa-VPg, NIa-Pro, and CP of TriMV shared low levels of amino acid identities (13 to 33%) with the corresponding proteins of members of the family *Potyviridae*, except with SCSMV (Table 2). Thus, among the members of *Potyviridae* that were used for sequence comparisons, TriMV proteins showed the highest amino acid identities with the corresponding proteins of SCSMV (Table 2). However, the degree of amino acid identity was relatively low (47 to 65%). Only the nucleotide sequence encompassing 6K2 through the 3′ end is available for SCSMV; thus, amino acid identities of P1 through CI between TriMV and SCSMV are not known.

Phylogenetic relationships within the family *Potyviridae*. Phylogenetic analysis was performed using the NJ method with the complete polyprotein, CI, NIa-Pro, NIa-VPg, NIb, and CP sequences of representative virus species from different genera and unassigned members within the family *Potyviridae*. Phylogenetic trees with polyprotein and CI sequences resulted in similar topologies, whereas NIb and NIa-Pro sequences produced similar tree topologies but were slightly different from those of the polyprotein and CI sequences (data not shown). Thus, representative trees based on polyprotein and NIb sequences, in addition to NIa-VPg and CP, are presented in Figure 5. Similar phylogenetic trees were observed depicting the genetic relatedness among *Potyvirus* taxa with maximum likelihood analysis based on complete nucleotide sequence and NIb protein sequence (data not shown).

Phylogenetic analyses based on complete polyprotein, NIb, and CP sequences placed TriMV (and SCSMV) in a separate branch between the members of *Tritimovirus* and *Ipomovirus*, such that three groups of viruses form a distinct clade within the family *Potyviridae* (Fig. 5A, B, and D). Trees representing the phylogenetic relationship based on mature protein sequences placed TriMV and SCSMV as sister taxa which share a most recent common ancestor with the members of *Ipomovirus* (with NIb and NIa-Pro) or *Tritimovirus* (with polyprotein and CP). Analysis with NIa-VPg protein yielded a poorly resolved phylogenetic tree which, nevertheless, showed that TriMV and SCSMV are sister taxa (Fig. 5C). In summary, TriMV and SCSMV formed a clade with tritimo- and ipomoviruses in phylograms generated with polyprotein and analyzed mature proteins of potyviruses, except with NIa-VPg (Fig. 5).

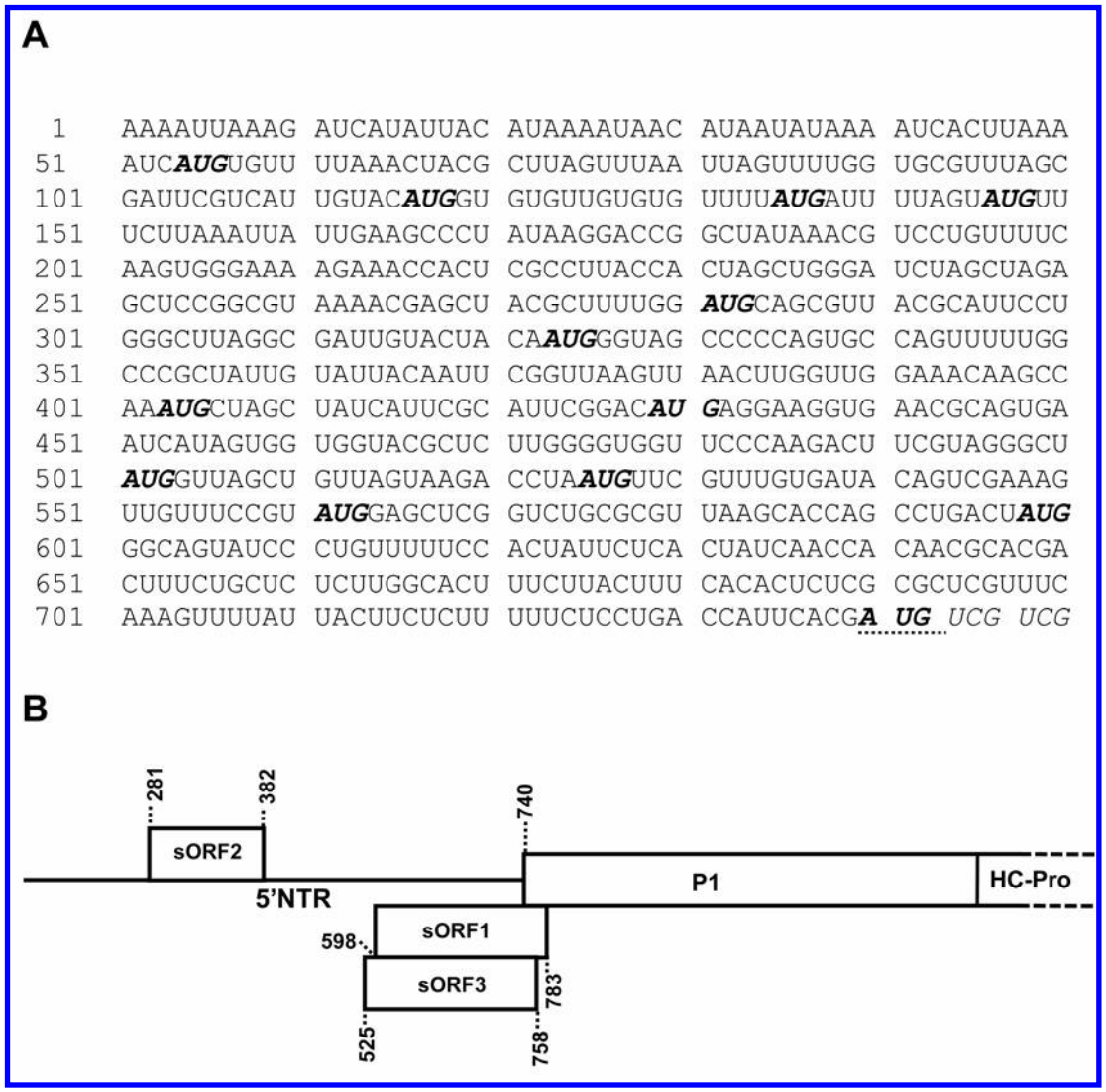

Fig. 4. A, 5′ Nontranslated region (NTR) sequence of *Triticum mosaic virus* (TriMV). The translation initiation codons in the 5′ NTR sequence are indicated with bold italics. Note that there are 12 AUG codons in the 5′ NTR of TriMV genomic RNA. The translation initiation codon of TriMV polyprotein is indicated with bold italics and underlined, followed by two codons of polyprotein indicated in italics. **B,** Enlarged view of schematic representation of 5′ NTR and part of large polyprotein open reading frame (ORF). Location of small ORFs (sORFs) in three reading frames in the 5′ NTR or overlapping with the polyprotein is indicated with rectangle boxes. sORFs 1 and 3 overlap with the 5′ end of the polyprotein ORF, whereas sORF2 is located in the 5′ NTR. The nucleotide coordinates of sORFs are indicated.

DISCUSSION

The taxonomic position of TriMV in the family *Potyviridae* was examined by sequencing its genome completely, and it was found that the mature proteins of TriMV possess 13 to 65% amino acid identity with the corresponding proteins of representative members in the family *Potyviridae*. Based on substantial amino acid sequence divergence between TriMV and representative members in other genera of the *Potyviridae*, a separate genus is proposed for TriMV: the genus *Poacevirus*, with TriMV as the type species and SCSMV as a distinct member. TriMV and SCSMV should be considered as two distinct species in the genus *Poacevirus*, because the mature proteins (6K2, NIa, NIb, and CP) of these two viruses differ by 35 to 53% in amino acid identity, which greatly exceeds the species demarcation threshold (<80% amino acid identity) to establish new virus species (2,5). Recently, Viswanathan et al. (27) reported CP sequences of several isolates of SCSMV and proposed a new genus, *Susmovirus*, with SCSMV as the sole member. However, the complete genome sequence of a tentative member species is required before a new genus classification may be considered (2). Thus, we propose the name *Poacevirus* for a new genus that would encompass TriMV and SCSMV. The genus *Poacevirus* was proposed because TriMV and SCSMV naturally infect the *Poaceae* members wheat and sugarcane, respectively.

Even though TriMV has a genome organization similar to monopartite viruses in the genera *Potyvirus*, *Rymovirus*, and *Tritimovirus*, the 5′ NTR of TriMV is unusually long, with 739 nt, which is four to seven times longer than other reported members of the family *Potyviridae*. Interestingly, the family *Picornaviridae*, which contains viruses infecting humans, animals, and birds, possesses long 5′ NTR sequences (628 to 1039 nt), which is highly structured to form an IRES element for protein translation as well as high-order structures responsible for genome replication, both of which are critical for virulence (4,8). The families *Potyviridae* and *Picornaviridae* belong to an evolutionarily similar lineage of a Picorna-like virus superfamily with conserved genomic organization and expression strategy (13). Additionally, the presence of 12 translation initiation codons and three sORFs with 33 to 77 amino acids in the leader sequence of TriMV is a unique feature for a potyvirid species. Two of these sORFs (sORF1 and -3; 61 and 77 amino acids) located in the 5′ NTR are larger than the 6K1 and 6K2 cistrons (50 to 56 amino acids) of TriMV, and the 3′ ends of these sORFs overlap slightly with the ORF encoding a large polyprotein. The role of these sORFs in virus biology is not known and can be examined after the availability of an infectious cDNA clone. The 5′-leader sequences of *Picornaviridae* members also possess several AUG codons preceding the polyprotein translation initiation codon. The 712-nt leader sequence of *Aichi virus* (GenBank accession no. NC_001918) contains 6 AUGs, and a 628-nt leader sequence of *Human rhinovirus* B (GenBank accession no. NC_001490) possesses 13 AUG codons preceding the translation initiation of polyprotein. The 5′ NTR of *Picornaviridae* members also encodes one to three sORFs with 36 to 165 amino acids, although the function of these small ORFs in virus biology is not known.

Multiple initiation codons upstream of the translation initiation codon of a polyprotein would substantially downregulate the ability of ribosomes to recognize the translation initiation codon of the polyprotein. Then, how would TriMV overcome these translation "road blocks" in the eukaryotic translation system? It is possible that the long 5′-leader sequence of TriMV might play an important role in translation of the polyprotein, like an IRES element as reported for the members of *Picornaviridae* (15). Although the 5′ NTR of TriMV folded into several stem-and-loop structures by MFOLD program (data not shown), we do not know whether the TriMV leader sequence contains an IRES element similar to the *Picornaviridae* members. Nevertheless, resemblances between the leader sequences of TriMV (a potyvirid) and the *Picornaviridae* members, both belonging to the Picorna-like superfamily, suggests that the 5′ NTR of TriMV might contain an IRES element or equivalent structure, which would involve cap-independent translation of TriMV mRNA. Because the complete genome sequence of SCSMV is not available, it is not known whether the long 5′ NTR is the characteristic feature of the proposed new genus.

The putative cleavage sites of TriMV and SCSMV NIa-proteinase in the polyprotein were identified between histidine (–1 position) and alanine, glycine, serine, or asparagine at the +1 position, which is not found in the vast majority of other members of the family *Potyviridae*. The conserved histidine residue at the –1 position of NIa-Proteinase cleavage sites in TriMV and SCSMV, together with relatively high amino acid sequence identity of mature proteins (6K2, NIa, NIb, and CP), strongly support the idea that these two viruses form sister taxa in the new genus. Adams et al. (1) inferred that the NIa proteinase cleavage sites of SPMMV also have a histidine residue at the –1 (substrate P1) position. Structural features of the NIb proteinase of TEV have been identified including two residues, threonine-146 and histidine-167, which are part of the binding pocket (S1) that

TABLE 2. Percent amino acid sequence identities of *Triticum mosaic virus* (TriMV) proteins compared with selective members of other genera and unassigned members in the family *Potyviridea*[a]

	Potyvirus		*Rymovirus*		*Macluravirus*	*Ipomovirus*		*Tritimovirus*		*Bymovirus*	Unassigned	
Cistron	PYV	SCMV	RGMV	AgMV	MacMV	CVYV	SPMMV	WSMV	BrSMV	BaYMV	BlVY	SCSMV
P1	17.6	16.5	17.1	16.3	–	18.5[b]	15.0	26.9	28.7	17.4	15.0	–
HC-Pro	20.4	19.0	21.8	20.5	–	20.9[c]	20.4	21.1	19.5	np[d]	20.4	–
P3	17.3	19.5	19.3	19.2	–	17.6	18.4	19.7	19.9	20.1	20.1	–
6K1	19.3	14.9	21.1	19.6	–	19.3	12.5	17.9	22.8	17.9	17.5	–
CI	25.5	26.1	29.0	27.6	–	24.8	24.3	29.2	25.6	24.5	23.4	–
6K2	13.5	22.6	18.3	17.0	–	21.8	26.4	21.6	33.3	10.7	20.6	54.0
NIa-VPg	22.7	22.0	21.4	21.3	–	24.8	24.5	22.9	17.7	21.6	20.9	49.0
NIa-Pro	20.2	22.1	22.4	23.0	–	31.1	24.4	25.4	24.2	19.2	23.0	46.9
NIb	37.9	38.7	39.4	38.7	–	43.7	40.7	41.3	43.1	32.6	38.4	65.2
CP	22.9	22.0	23.6	20.4	19.0	24.7	25.8	25.6	29.6	21.0	18.7	48.3

[a] PVY = *Potato virus* Y (NC_001616), SCMV = *Sugarcane mosaic virus* (NC_003398), RGMV = *Ryegrass mosaic virus* (NC_001814), AgMV = *Agropyron mosaic virus* (NC_005903), MacMV = *Maclura mosaic virus* (U58771), CVYV= *Cucumber vein yellowing virus* (AY578085), SPMMV = *Sweet potato mild mottle virus* (NC_003797), WSMV = *Wheat streak mosaic virus* (NC_001886), BrSMV = *Brome streak mosaic virus* (NC_003501), BaYMV = *Barley yellow mosaic virus* (NC_002990), BlVY = *Blackberry virus* Y (NC_008558), and SCSMV = *Sugarcane streak mosaic virus* (Y17738); – indicates sequence data not available.

[b] Compared with P1a.

[c] Compared TriMV P1 with CVYV P1b.

[d] HC-Pro cistron is not present.

directly interact with the substrate P1 position (16). These two residues are strictly conserved among members of the *Potyviridae*, where the substrate P1 (–1 position) is either glutamine or glutamic acid. Interestingly, the homologous S1 residues are threonine–asparagine in NIa of SPPMV rather than threonine–histidine. Molecular modeling suggests that the histidine to asparagine substitution would accommodate docking of substrate peptides bearing histidine at substrate P1 position in the cleavage site (1). The homologous S1 residues in the NIas of both TriMV and SCSMV are serine–histidine. It is possible that replacement of threonine with the smaller serine residue may similarly accommodate the presence of histidine at substrate P1 (–1 position).

The fact that phylogenetic analysis of the complete polyproteins or mature proteins of selected members of the family *Potyviridae* consistently placed TriMV and SCSMV together on a separate branch from other reported potyvirid species provides further evidence that these two viruses represent sister taxa in a new genus. TriMV and SCSMV share a most recent common ancestor with the members of *Tritimovirus*, mite-transmitted viruses (based on polyprotein and CP sequences), or with *Ipomovirus*, whitefly-transmitted viruses (based on NIb sequence), suggesting that vector transmission criteria may not be used alone for taxonomic classification of potyviruses, because the mite-transmitted viruses were grouped into phylo-

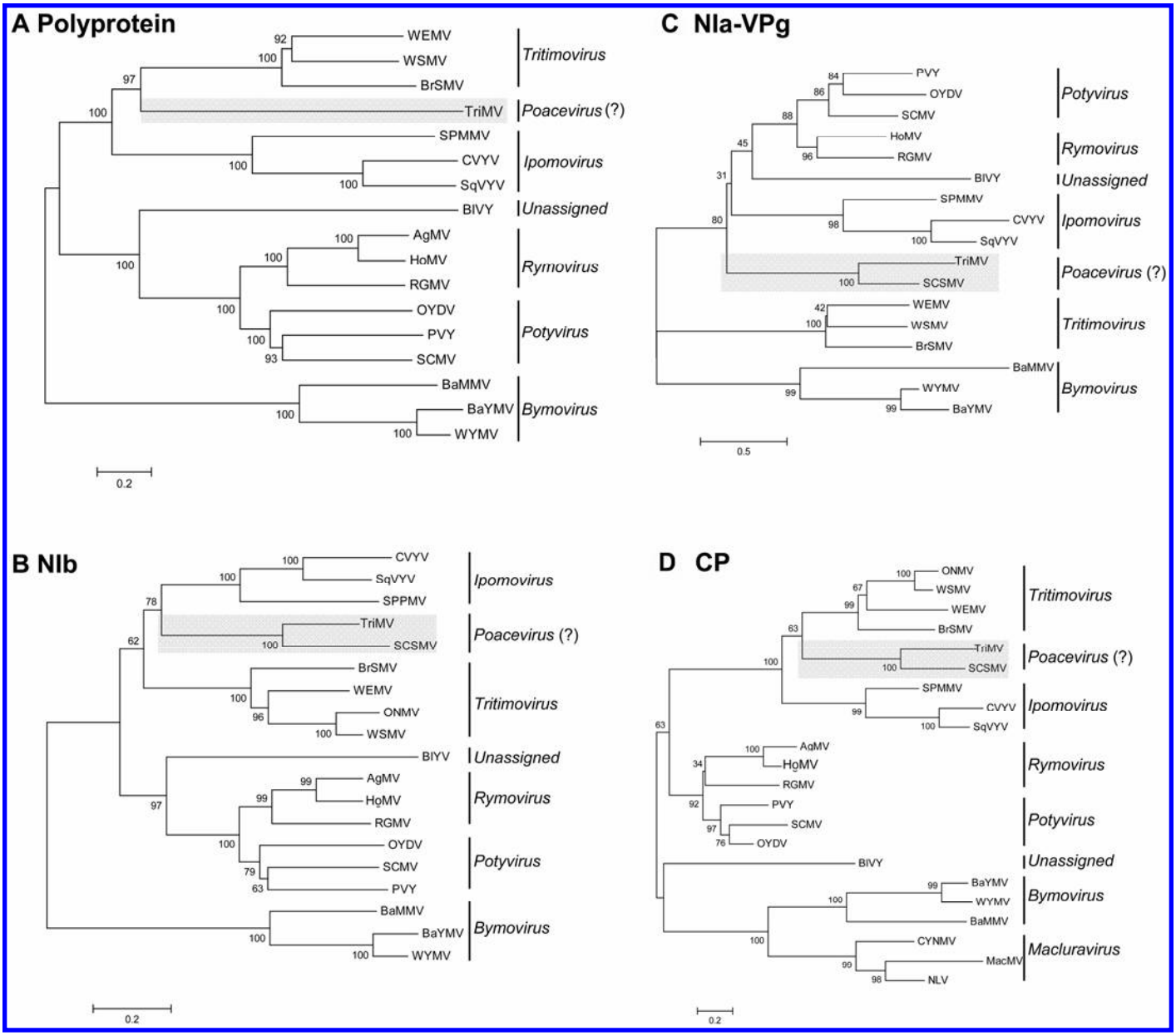

Fig. 5. Unrooted bootstrap consensus phylogenetic trees derived from the amino acid sequences of **A,** polyprotein; **B,** NIb; **C,** NIa-VPg; and **D,** coat protein (CP) of *Triticum mosaic virus* (TriMV), unassigned and representative species of various genera of the family *Potyviridae*. Phylogenetic trees were constructed by the neighbor-joining method using the JTT matrix and pairwise gap deletion with 1,000 bootstrap replicates; bootstrap support is indicated at branch points. The bar represents the number of amino acid replacements per site. The recognized genera with representative members are indicated on the right side of the phylogenetic trees, and TriMV and *Sugarcane streak mosaic virus* (SCSMV) are indicated with shade. GenBank accession numbers of potyvirid species used in phylogenetic tree analysis are *Agropyron mosaic virus* (AgMV, NC_005903), *Barley mild mottle virus* (BMMV, NC_003483), *Barley yellow mosaic virus* (BaYMV, NC_002990), *Blackberry virus* Y (BlVY, NC_008558), *Brome streak mosaic virus* (BrSMV, NC_003501), *Chinese yam necrotic mosaic virus* (CYNMV, AB044386), *Cucumber vein yellowing virus* (CVYV, AY578085), *Hordeum mosaic virus* (HoMV, NC_005904), *Maclura mosaic virus*, MacMV, U58771), *Narcissus latent virus* (NLV, DQ450199), *Oat necrotic mottle virus* (ONMV, NC_005136), *Onion yellow dwarf virus* (OYDV, NC_005029), *Potato virus* Y (PVY, NC_001616), *Ryegrass mosaic virus* (RGMV, NC_001814), *Squash vein yellowing virus* (SqVYV, NC_010521), *Sugarcane mosaic virus* (SCMV, NC_003398), SCSMV (Y17738), *Sweet potato mild mottle virus* (SPMMV, NC_003797), *Triticum mosaic virus* (TriMV, FJ669487), *Wheat eqlid mosaic virus* (WEMV, NC_009805), *Wheat streak mosaic virus* (WSMV, NC_001886), and *Wheat yellow mosaic virus* (WYMV, NC_002350).

genetically distinct genera *Tritimovirus*, *Rymovirus*, and *Poacevirus* (TriMV).

Taken together, the unusually long 5′-leader sequence with several translation initiation codons and sORFs, low amino acid sequence identity with representative members of the genera in the family *Potyviridae*, and relatively high (47 to 65%) amino acid identity with mature proteins of SCSMV suggest that TriMV should be placed in a new genus, *Poacevirus*, as the type species, with SCSMV as a distinct member.

ACKNOWLEDGMENTS

We thank S. Gowda for critical reading of the manuscript. Mention of proprietary or brand names are necessary to report factually on available data; however, the United States Department of Agriculture (USDA) neither guarantees nor warrants the standard of the product, and the use of the name by USDA implies no approval to the exclusion of others that also may be suitable.

LITERATURE CITED

1. Adams, M. J., Antoniw, J. F., and Beaudoin, F. 2005. Overview and analysis of the polyprotein cleavage sites in the family *Potyviridae*. Mol. Plant Pathol. 6:471-487.
2. Adams, M. J., Antoniw, J. F., and Fauquet, C. M. 2005. Molecular criteria for genus and species discrimination within the family *Potyviridae*. Arch. Virol. 150:459-479.
3. Bailey, J. M., and Davidson, N. 1976. Methylmercury as a reversible denaturing agent for agarose gel electrophoresis. Anal. Biochem. 70:75-85.
4. Belsham, G. J. 2009. Divergent picornavirus IRES elements. Virus Res. 139:183-192.
5. Berger, P. H., Adams, M. J., Barnett, O. W., Brunt, A. A., Hammond, J., Hill, J. H., Jordan, R. L., Kashiwazaki, S., Rybicki, E., Spence, N., Stenger, D. C., Ohki, S. T., Uyeda, I., van Zaayen, A., Valkonen, J., and Vetten, H. J. 2005. Pages 819-841 in: Virus Taxonomy–Classification and Nomenclature of Viruses, 8th Report of the ICTV. C. M. Fauquet, M. A. Mayo, J. Maniloff, U. Desselberger, and L. A. Ball, eds. Elsevier Academic Press, San Diego, CA.
6. Brakke, M. K. 1987. Virus disease in wheat. Pages 585-603 in: Wheat and Wheat Improvement, 2nd ed. E. G. Heyne, ed. American Society of Agronomy, Crop Science Society of America, Soil Science Society of America, Madison, WI.
7. Chung, B. Y.-W., Miller, W. A., Atkins, J. F., and Firth, A. E. 2008. An overlapping essential gene in the *Potyviridae*. Proc. Natl. Acad. Sci. USA 105:5897-5902.
8. Fernández-Miragall, O., López de Quinto, S., and Martínez-Salas, E. 2009. Relevance of RNA structure in the activity of picornavirus IRES elements. Virus Res. 139:172-182.
9. Götz, R., Huth, W., Lesemann, D. E., and Maiss, E. 2002. Molecular and serological relationships of *Spartina mottle virus* (SpMV) strains from Spartina spec. and from Cynodon dactylon to other members of the *Potyviridae*. Arch. Virol. 147:379-391.
10. Hall, J. S., Adams, B., Parsons, T. J., French, R., Lane, L. C., and Jensen, S. G. 1998. Molecular cloning, sequencing and phylogenetic relationships of a new potyvirus: sugarcane streak mosaic virus and a reevaluation of the classification of the *Potyviridae*. Mol. Phylogenet. Evol. 10:323-332.
11. Hema, M., Joseph, J., Gopinath, K., Sreenivasulu, P., and Savithri, H. S. 1999. Molecular characterization and interviral relationships of a flexuous filamentous virus causing mosaic disease of sugarcane (*Saccharum officinarum*) in India. Arch. Virol. 144:479-490.
12. Hema, M., Sreenivasulu, P., and Savithri, H. S. 2002. Taxonomic position of sugarcane streak mosaic virus in the family *Potyviridae*. Arch. Virol. 147:1997-2007.
13. Koonin, E. V., Wolf, Y. I., Nagasaki, K., and Dolja, V. V. 2008. The big bang of picorna-like virus evolution antedates the radiation of eukaryotic supergroups. Nat. Rev. 6:925-939.
14. Lane, L. C. 1986. Propagation and purification of RNA plant viruses. Methods Enzymol. 118:687-696.
15. Martínez-Salas, E., Pacheco, A., Serrano, P., and Fernandez, N. 2008. New insights into internal ribosome entry site elements relevant for virus gene expression. J. Gen. Virol. 89:611-626.
16. Phan, J., Zdanov, E., Evdokimov, A. G., Tropea, J. E., Peters, H. K., Kapust, R. B., Li, M., Wlodawer, A., and Waugh, D. S. 2002. Structural basis for the substrate specificity of *Tobacco etch virus* protease. J. Biol. Chem. 277:50564-50572.
17. Rabenstein, F., Seifers, D. L., Schubert, J., French, R., and Stenger, D. C. 2002. Phylogenetic relationships, strain diversity and biogeography of tritimoviruses. J. Gen. Virol. 83:895-906.
18. Schaad, M. C., Haldeman-Cahill, R., Cronin, S., and Carrington, J. C. 1996. Analysis of the VPg-Proteinase (NIa) encoded by tobacco etch potyvirus: effects of mutations on subcellular transport, proteolytic processing, and genome amplification. J. Virol. 70:7039-7048.
19. Seifers, D. L., Martin, T. J., Harvey, T. L., Fellers, J. P., and Michaud, J. P. 2009. Identification of the wheat curl mite as the vector of *Triticum mosaic virus*. Plant Dis. 93:25-29.
20. Seifers, D. L., Martin, T. J., Harvey, T. L., Fellers, J. P., Stack, J. P., Ryba-White, M., Haber, S., Krokhin, O., Spicer, V., Lovat, N., Yamchuk, A., and Standing, K. G. 2008. *Triticum mosaic virus*: a new virus isolated from wheat in Kansas. Plant Dis. 92:808-817.
21. Skare, J. M., Wijkamp, I., Denham, I., Rezende, J. A. M., Kitajima, E. W., Park, J.-W., Desvoyes, B., Rush, C. M., Michels, G., Scholthof, K. B. G., and Scholthof, H. B. 2006. A new eriophyid mite-borne membrane-enveloped virus-like complex isolated from plants. Virology 347:343-353.
22. Stenger, D. C., Hall, J. S., Choi, I.-R., and French, R. 1998. Phylogenetic relationships within the family *Potyviridae*: wheat streak mosaic virus and brome streak mosaic virus are not members of the genus *Rymovirus*. Phytopathology 88:782-787.
23. Susaimuthu, J., Tzanetakis, I. E., Gergerich, R. C., and Martin, R. R. 2008. A member of a new genus in the *Potyviridae* infects *Rubus*. Virus Res. 131:145-151.
24. Tamura, K., Dudley, J., Nei, M., and Kumar, S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24:1596-1599.
25. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 24:4876-4882.
26. Valli, A., López-Moya, J. J., and García, J. A. 2007. Recombination and gene duplication in the evolutionary diversification of P1 proteins in the family *Potyviridae*. J. Gen. Virol. 88:1016-1028.
27. Viswanathan, R., Balamuralikrishnan, M., and Karuppaiah, R. 2008. Characterization and genetic diversity of *Sugarcane streak mosaic virus* causing mosaic in sugarcane. Virus Genes 36:553-564.
28. Wegulo, S., Hein, G., and Ziems, A. 2009. Wheat disease update. Page 133 in: Proc. 2009 Crop Production Clinics. University of Nebraska–Lincoln Extension.