Virus Genes 29:2, 249–255, 2004

Extended Sequence Analysis of Three Danish *Potato Mop-Top Virus* (PMTV) Isolates

TAMARA PEČENKOVÁ, TOMÁŠ MORAVEC, MARIE FILIGAROVÁ, PAVLA ROSECKÁ & NOEMI ČEŘOVSKÁ*

Department of Virology, Institute of Experimental Botany, Czech Academy of Sciences, Na Karlovce la, Prague 6, 160 00 Czech Republic

Received February 17, 2004; Revised and Accepted April 14, 2004

Abstract. The entire nucleotide sequence for the coding regions of a Danish PMTV isolate 54-15 was determined and compared to other known and sequenced isolates of PMTV. Many nucleotide and amino acid changes were found in parts of RNA coding for the triple gene block (TGB) proteins and in the part of the RNA coding for the read-through region of the coat protein (CP). These regions for two other isolates, the mild one 54-10 and the severe one 54-19, were sequenced. Only two amino acid changes were found to correlate with the subdivision of isolates according to symptom development into mild and severe subgroups. In addition, the phylogenetic tree was obtained suggesting the closest relationship between isolates 54-15 and 54-10. Although the sequence comparisons indicate a high genetic stability of PMTV populations, a surprising change was found in the newly sequenced isolates – the replacement of the AUG start codon of the fourth gene of the TGB encoding RNA, coding for a cystein-rich protein, by the less efficient GUG start codon.

Key words: Cystein-rich protein, genome comparison, nucleotide sequence, phylogenetic analysis, PMTV

Introduction

Potato mop-top virus (PMTV) is the type member of the genus *Pomovirus* [1]. PMTV occurs in potato growing regions in Europe, North and South America and Asia in cool wet climate causing a wide range of symptoms in haulms and tubers which vary depending on the potato cultivar and environmental conditions, thus complicating the identification of the virus disease [2,3]. In field conditions the virus is transmitted by motile zoospores of the plasmodiophoromycete fungus *Spongospora subterranea* (Wallr.) Lagerh., which causes powdery scab on tubers [4,5].

PMTV has tubular and rigid particles, measuring 18–22 × 100–150 nm or 250–300 nm [6], encapsidating three RNA components [7]. RNA 1, 6.0 kb long, encodes the viral RNA-dependent RNA polymerase (RdRp) and its read-through domain [8]. The second RNA 3.0 kb long, has a single open reading frame (ORF) encoding the 20 kDa coat protein (CP) and 67 kDa protein produced by read-through of an amber termination codon of CP, which may be involved in virion assembly and virus transmission by fungal vector [9]. The third RNA, 2.9 kb long, contains a triple gene block (TGB) encoding three proteins involved in cell-to-cell movement, and an additional ORF for a predicted 'cystein-rich' protein with unknown function [10].

Previously, molecular data on PMTV was based mostly on the analysis of a single Swedish isolate (PMTV-Sw) whose complete nucleotide sequence was determined [8,10]. Recently, the region encoding part of read-through coat protein was sequenced and compared for nine different Danish PMTV isolates [11]. These isolates could be

*Author for all Correspondence:
E-mail: cerovska@ueb.cas.cz

grouped according to symptom development into three groups – mild, medium and severe but no correlation between symptom grouping and nucleotide sequences was observed [11]. Previous analysis of coat proteins sequence for different PMTV isolates proved that these viruses are highly conserved in this part of genome [10,12,13].

While performing preliminary comparisons of sequence data for isolate 54-15 along with other sequences for isolates available on the sequence databases, we noticed that the distribution of nucleotide changes was not uniform. We sequenced the less conserved regions (including parts of the TGB and the read-through region of CP) for isolates 54-10 and 54-19. The paper describes the findings of this further sequence analysis.

Materials and Methods

Virus Source and 1C-RT-PCR

The isolates of PMTV were kindly provided by Dr. Nielsen from the Danish Institute of Agricultural Sciences, Flakkebjerg, Denmark. Viruses were propagated in *N. debneyi* by mechanical inoculation with the sap extracted from symptomatic leaves.

Immunocapture reverse transcription polymerase chain reaction (IC-RT-PCR) was used to generate cDNA of PMTV RNAs. The tubes were coated with 100 μl anti PMTV IgG (1 μg/ml) (Adgen) diluted in coating buffer for 3 h at 37°C. The wells were then washed (3 μl × 150 μl PBS + T) and 100 μl of the homogenate from PMTV infected leaves in conjugate buffer (1:10) was added. The samples were incubated overnight at 4°C and washed again three times with PBS + T buffer. The reverse transcription and PCR amplification with Superscript II (Gibco) and Taq polymerase following manufacturer's recommendations were done, using PMTV specific primers, whose sequence was based on of the sequence of Sw isolate available in GenBank database (accession number AJ238607, AJ277556 and AJ243719) [8,10]. Primers have introduced restrictions sites that will be relevant for further experimenting, mostly *Nco*I and *Bgl*II. The PCR profile (30 cycles) was: 30 s denaturation at 94°C, 30 s annealing at 55°C and 1 min elongation at 72°C.

Cloning and Sequencing

The sequence was obtained by both sequencing PCR products as well as fragments cloned into pUC57T/A (Fermentas). Sequencing was performed using an ALFexpressII sequencer with the autoRead sequencing kit (AP Life Science). Sequence analyses were carried out using the Genescan software [14], and the similarity search with sequences available in GenBank was performed using BLAST [15]. Multiple sequence alignments were done using Clustal W [16]. The obtained sequences were deposited in the GenBank (Table 1). The sequences of other isolates retrieved from GenBank used for comparison are also shown in this Table 1 [17,18]. The phylogenetic analysis was performed using the programs included in the PHYLIP package [19].

Results

The Genomic Organization of PMTV Isolates 54-10, 54-15 and 54-19

The nucleotide sequences for the three Danish PMTV isolates, namely 54-10, 54-15 and 54-19 were obtained and the preliminary comparison with the isolate Sw was performed (Fig. 1). The sequences were shown to have a high level of similarity and the genomic organization of three Danish isolates was shown to be the same as in the case of the Swedish isolate Sw (Fig. 1 and Table 1).

The Nucleotide Sequence Analysis of PMTV Isolates 54-10, 54-15 and 54-19 and the Comparison with Isolates S, Sw and T

When the 5,629 nt fragment obtained for the RNA 1 of isolate 54-15 was aligned with the sequence of the Sw isolate, the identity score was 99.5%, resulting from 27 nucleotide changes; most of them were A/G (15) and C/U (8) transitions. Only one G/C and three A/U transversions were found. The changes were not evenly distributed along the sequence (Fig. 1). The partial nucleotide sequence of RNA 1 for the isolate 54-19 (Fig. 1, and Table 2) when compared with that of 54-15 turned out to be almost identical. The only nucleotide

Table 1. The nucleotide sequences of newly obtained PMTV isolates as well as other PMTV isolates available in databases used in the sequence comparisons

	Origin	Symptoms	RNA	Length (nt)	Acc No	Citation
54-10	Danish	Mild	2 (CP)	1386	AY277633	–
			3 (TGB)	2286	AY426745	–
54-15	Danish	Severe	1 (RdRp)	5589	AY196959	–
			2 (CP)	2614	AY196094	–
			3 (TGB)	2852	AY187010	–
54-19	Danish	Severe	1	513	AY424977	–
				473	AY424978	–
				1020	AY424979	–
			2 (CP)	1714	AF487407	–
				461	AF508253	[11]
				617	AY436586	–
			3 (TGB)	2458	AY353719	–
S	Scottish		2 (CP)	2046	AJ224991	[18]
Sw	Swedish		1	6043	AJ238607	[8]
			2 (CP)	3134	AJ243719	[10]
			3 (TGB)	2964	AJ277556	–
Todd	Scottish		3 (TGB)	2962	D30753	[17]

change found in isolate 54-19 was also present in the isolate Sw.

The sequence of the CP read-through region of RNA 2 for isolate 54-15 was obtained, as well as chosen regions of this RNA for isolates 54-10 and 54-19 (Fig. 1). The 1,386 nt sequence for the read-through domain of the CP displayed six nucleotide changes among the three Danish isolates. The sequence identity was always higher than 99% when each of the three Danish isolates was compared with the Sw isolate. When compared with the Scottish isolate S the sequence identity was lower (97%) due to the fact that the isolate S has a deletion in the read-through region (Table 2) [18, 20]. The distribution of nucleotide changes is shown in Fig. 1.

When comparisons were made for a 400 nt region of the middle part of CP read-through gene (where the sequence was determined for all the Danish and Sw and S isolates), 11 nucleotide changes were found, mostly A/G and C/U substitution, and only one G/U and one C/G substitution.

The nucleotide sequences for the RNA encoding the TGB (RNA 3) of the three isolates 54-10, 54-15 and 54-19 were also determined (Fig. 1). The sequences were adjusted to the length of the shortest isolate (54-10, 2286 nt) and aligned together with those of isolate Sw and isolate Todd (Table 1). The Danish isolates were found to be 99.9% identical when isolates 54-10 and 54-15 were compared and 99.7% when the 54-15 and 54-19 were compared. The nucleotide identity was approximately 97% and 96% when Danish isolates were compared to Sw and Todd, respectively. The total number of changes was 87 with only seven of them being detected among Danish isolates (Table 3). Most of 87 changes were C/U and A/G substitutions. Unexpectedly, the sequence near to the 3′ end of RNA 3 of the Danish isolates contained three neighboring nucleotides identical to the same region of isolate Todd instead of the otherwise more similar isolate Sw. The distribution of nucleotide substitutions in the RNA 3 for these isolates is shown in Fig. 1.

The Comparisons of Deduced Amino Acid Sequences of PMTV 54-10, 54-15 and 54-19 ORFs

When the deduced amino acid sequence from 54-15 RNA1 was compared with the amino acid sequence available for the isolate Sw the similarity was above 99%. The amino acid sequence deduced from the three small parts of RNA 1 for isolate 54-19 were found to be 100% congruent with the isolate 54-15 (Table 3).

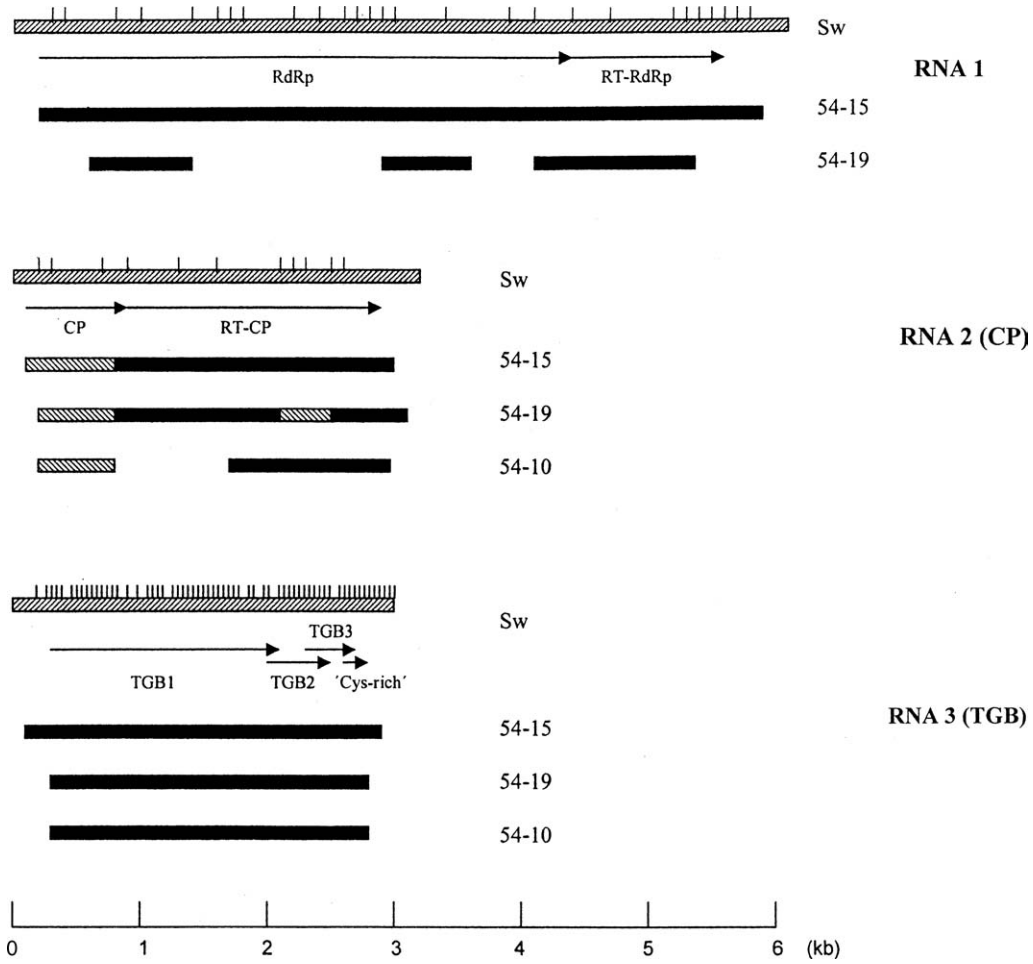

Fig. 1. The obtained nucleotide sequences for three Danish PMTV isolates aligned with the known complete sequence of Sw isolate. The newly obtained sequences are presented with black boxes; the regions previously obtained and published are presented in shadowed boxes. The positions with nucleotide changes found between the Sw and Danish isolate 54-15 for each of RNAs are marked with vertical lines. The translation products are represented by arrows.

The deduced amino acid sequences for 400 nt region of the middle part of the CP read-through gene of the Danish isolates were aligned with the same region from isolates S and Sw (Table 3). The mutation found in this region among Danish isolates was shown to be silent: the amino acid sequences were 100% identical.

Only two amino acid positions for the CP read-through affected by nucleotide changes were found in the case of Danish isolates: one of these two changes was the substitution of amino acid residue Asp (position 261 after the amber stop codon of coat protein gene) in 54-10 instead of the similar amino acid Glu present in 54-15 and 54-19.

Based on the genome organization of isolate Sw RNA 3, three possible overlapping gene components of the TGB were found in isolates 54-10, 54-15 and 54-19 and the deduced amino acid sequences were used for analysis. The similarity in the case of the first protein, 51K ranged from 97% (Danish isolates versus Todd) to 100% (among Danish isolates; Table 3). The one amino acid substitution was found in 54-10 TGB1 when compared to all other isolates: the change was Lys176 (the numbering was based on the Sw sequences) mutated to Arg176 in 54-10 isolate. In the case of the second protein (13K), the similarity among different isolates ranged from 98% to 100%

Table 2. The number of nucleotide changes found when the chosen region of newly obtained sequences of Danish PMTV isolates were aligned and compared (regions of 513 nt, 473 nt, and 1020 nt from RNA 1; 1386 nt from RNA 2 and 2286 nt from RNA 3). The RNA 1 of 54-10 was not sequenced. The number of amino acid changes in deduced proteins arising from these mutation are presented in paranthesis

Isolate RNA		54-10	54-15	54-19
RNA 1	54-10	–	–	–
	54-15	–	–	1 (0)
	54-19	–	1 (0)	–
RNA 2	54-10	–	4 (1)	2 (2)
	54-15	4 (1)	–	4 (1)
	54-19	2 (2)	4 (1)	–
RNA 3	54-10	–	2 (1)	7 (5)
	54-15	2 (1)	–	5 (4)
	54-19	7 (5)	5 (4)	–

Table 3. The similarity of deduced amino acid sequences between different isolates of PMTV

Deduced proteins	PMTV isolates					
	54-10	54-15	54-19	S	Sw	Todd
RdRp*						
54-15	–	–	100	–	99	–
54-19	–	100	–	–	99	–
RT CP*						
54-10	–	100	100	98	100	–
54-15	100	–	99	98	100	–
54-19	100	99	–	98	100	–
TGB 1						
54-10	–	98	99	–	98	97
54-15	98	–	98	–	98	97
54-19	98	98	–	–	96	95
TGB 2						
54-10	–	100	99	–	98	98
54-15	100	–	99	–	98	98
54-19	99	99	–	–	98	99
TGB 3						
54-10	–	100	100	–	97	94
54-15	100	–	100	–	97	94
54-19	100	100	–	–	97	94
Cys rich						
54-10	–	100	98	–	88	88
54-15	100	–	98	–	88	88
54-19	98	98	–	–	90	90

*Partial amino acid sequence was analyzed.
¯Unavailable data.

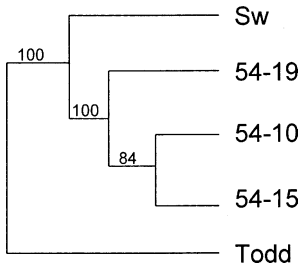

Fig. 2. Phylogenetic tree constructed for PMTV isolates from the multiple alignments of 2286 nt long sequence from RNA 3 (encoding TGB) by DNADIST, KITSCH and CONSENSE. Numbers at branching points are bootstrap values derived from 100 replicons.

(Table 3). The lowest similarity score was found for the third TGB protein (21K). The similarity for the Danish isolates was 94% and 97% with Todd and Sw, respectively (Table 3). No differences were found among Danish isolates in this region.

In the case of Swedish isolate Sw, the fourth ORF on RNA 3 was found, coding the so-called 'cystein-rich' protein. However, this ORF has the first codon for Met mutated to the codon for Val, in all sequenced Danish isolates. The amino acid sequence comparison of this region showed the lowest similarity for isolates 54-10 and 54-15 when compared with Sw and Todd (Table 3).

Phylogenetic Analysis

In order to elucidate the phylogeny of PMTV virus isolates, for each of the analyzed genomic regions or deduced proteins 100 bootstrap replicates were obtained and the corresponding phylogenetic trees constructed, using both parsimony and distant methods. In both approaches, when the alignment of 2286 nt of the RNA 3 was used in the tree construction procedure, the same phylogenetic tree was obtained that supported closer relationship between Danish isolates 54-10 and 54-15 than 54-15 and 54-19 (Fig. 2). However, the relatedness of isolates 54-10, 54-15 and 54-19 varied for other genomic regions when performing different tree construction methods. The topology in this part of the tree was in many cases influenced by input order of sequences during the tree construction which might arise from the fact that these isolates were almost identical. When the multiple amino acid sequence alignments were used for the

evolutionary tree reconstruction, the situation was similar and the trees of the similar topology were obtained (data not shown). Many obtained trees placed the isolate 54-19 as more closely related to Sw and Todd (or S) than to 54-10 and 54-15.

Discussion

The nucleotide sequences covering the most of genomic RNAs of PMTV isolates 54-10, 54-15 and 54-19 were obtained (Fig. 1, Table 1). According to Nielsen and Nicolaisen [11], isolate 54-15 and 54-19 produces severe symptoms while the isolate 54-10 produces mild symptoms in indicator plants. The preliminary comparison of these isolates with previously sequenced PMTV isolates available in GenBank confirmed low variability of PMTV (Table 2) and pointed out regions that could be possible candidates for differentiation of these isolates by symptom severity: namely most of the nucleotide sequence of the RNA 3 and the CP read-through region of RNA 2. However, only few amino acid changes were detected, and these were mainly found for the isolate 54-19 (Table 3). Only two amino acid changes correlated with grouping of isolates based on the symptom development. One of them was found in the middle part of the deduced TGB1 protein for the isolate 54-10 (mild) where amino acid Arg (position 176 according to Sw TGB1 sequence) replaced amino acid Lys present in two severe isolates. This Lys and its neighboring amino acid residue Tyr175 are present and conserved in many viruses belonging to different genera [21]. The other change was found in the read-through region of CP (position 261 after the amber stop-codon of coat protein according to Sw read-through of CP sequence), where the Asp amino acid residue was replaced with Glu in the case of 54-10.

The start codon of the fourth ORF of RNA 3 found in Sw was found to be mutated from AUG to less efficient GUG in the case of Danish isolates thus being the most remarkable feature concerning these isolates. The disruption or down regulation of this ORF, however, is not relevant to symptom variations, since it was found in both severe (54-15, 54-19) and mild (54-10) symptom inducing isolates.

Phylogeny analysis confirmed that the three Danish isolates are closely related to each other and when the almost whole length RNA3 sequence alignment was employed in the tree construction procedure, the isolate 54-15 (severe) was found to be more closely related to 54-10 (mild) than to 54-19 (severe). Sw isolate was found to be more closely related to Danish isolates than the isolates Todd is. However, when other parts of genomes were used for the genetic analysis, e.g. the 400 nt long region of RNA2 (encoding CP read-through) or different deduced amino acid sequences from TGB genes, the relatedness among isolates varied.

Acknowledgments

This research was supported by the grants No.522/01/1121 and No.522/04/1329 of the Grant Agency of Czech Republic. The authors wish to thank to Dr. S. L. Nielsen, Department of Plant Protection, Danish Institute of agricultural Sciences, Slagelse, Denmark, for providing the PMTV isolates, and to Mmes R. Hadámková and D. Cibochová, Institute of Experimental Botany, Prague, Czech Republic, for their excellent technical assistance. We thank also Dr. Neil Boonham from Central Science Laboratory, Hutton, York, for critical reading of our manuscript.

References

1. Torrance L., and Mayo M.A., Arch Virol *142*, 435–439, 1997.
2. Kurppa A., Annales Agric Fenniae *28*, 285–295, 1989.
3. Sokmen M.A., Barker H., and Torrance L., Ann Appl Biol *133*, 55–63, 1998.
4. Arif M., Torrance L., and Reavy B., Ann Appl Biol *126*, 493–503, 1995.
5. Jones R.A.C., and Harrison B.D., Annals App Biol *63*, 1–17, 1969.
6. White R.F., Kassanis B., and James M., J Gen Virol *15*, 175–177, 1972.
7. Kallender H., Buck K.W., and Brunt A.A., Neth J Plant Pathol *96*, 47–50, 1990.
8. Savenkov E.I., Sandgren M., and Valkonen J.P., J Gen Virol *80*, 2779–2784, 1999.
9. Kashiwazaki S., Scott K.P., Reavy B., and Harrison B.D., Virology *206*, 701–706, 1995.
10. Sandgren M., Savenkov E.I., and Valkonen J.P., Arch. Virol *146*, 467–477, 2001.
11. Nielsen S.L., and Nicolaisen M., Arch Virol *148*, 381–388, 2003.

12. Čeřovská N., Moravec T., Rosecká P., Filigarová M., and Pečenková T., Acta Virol *47*, 41–44, 2003.
13. Mayo M.A., Torrance L., Cowan G., Jolly C.A., Macintosh S.M., Orrega R., Barrera C., and Salazar L.F., Arch Virol *141*, 1115–1121, 1996.
14. Burset M., and Guigo R., Genomics *34*, 353–367, 1996.
15. Altschul S.F., Stephen F., Gish W., Miller W., Myers E.W., and Lipman D.J., J Mol Biol *215*, 403–410, 1990.
16. Thompson J.D., Higgins D.G., and Gibson T.J., NAR *22*, 4673–4680, 1994.
17. Scott K.P., Kashiwazaki S., Reavy B., and Harrison B.D., J Gen Virol *75*, 3561–3568, 1994.
18. Reavy B., Arif M., Cowan G.H., and Torrance L., J Gen Virol *79*, 2343–2437, 1998.
19. Felsenstein J., Cladistics *5*, 164–166, 1989.
20. Torrance L., Cowan G.H., Sokmen MA., and Reavy B., J Gen Virol *80*, 2211–2215, 1999.
21. Morozov S.Y. and Solovyev A.G., J Gen Virol *84*, 1351–1366, 2003.