

Potato Virus T RNAの3'末端領域の塩基配列

誌名	日本植物病理學會報 = Annals of the Phytopathological Society of Japan
ISSN	00319473
著者名	越智,資泰 柏崎,哲 平塚,和之 難波,成任 土崎,常男
発行元	日本植物病理學會
巻/号	58巻3号
巻号補足	
掲載ページ	p. 416-425
発行年月	1992年7月

農林水産省 農林水産技術会議事務局筑波産学連携支援センター
Tsukuba Business-Academia Cooperation Support Center, Agriculture, Forestry and Fisheries Research Council
Secretariat



Nucleotide Sequence of the 3'-Terminal Region of Potato Virus T RNA*

Motoyasu OCHI**, Satoshi KASHIWAZAKI***, Kazuyuki HIRATSUKA†,
Shigetou NAMBA**,†† and Tsuneo TSUCHIZAKI**

Abstract

The partial sequence of the capillovirus potato virus T (PVT) RNA containing the 3'-terminal 2392 nucleotides, excluding the poly(A) tail, was determined from cloned cDNA. The sequence contains three open reading frames (ORFs) which encode putative proteins (in the 5'→3' direction) of M_r > 29,000 (>29K; ORF 1), M_r 40,451 (40K; ORF 2) and M_r 23,596 (24K; ORF 3), followed by an untranslated region of 188 nucleotides upstream of the 3' poly(A) tail. The 5'-proximal ORF 1 encodes a product with significant homologies to the C-terminal portions of the putative polymerase proteins of "alpha-like" supergroup of plant RNA viruses, including apple chlorotic leaf spot virus (ACLSV; closterovirus subgroup A), carlavirus, potexvirus and tymovirus. The PVT 24K protein shared several blocks of conserved amino acids with the coat proteins of filamentous viruses, *i.e.* ACLSV, beet yellows virus (BYV; closterovirus subgroup B), citrus tristeza virus (CTV; closterovirus subgroup C), potexvirus and carlavirus. The PVT 40K protein had homologies with putative movement proteins encoded by several plant viruses (e.g. ACLSV 50K protein). PVT and ACLSV showed a similar arrangement of three ORFs in the 3'-terminal region of their genomes. However, this arrangement differs significantly from that of BYV. The sequence represented here is the first report on the group capillovirus which was established recently. These results support a close relationship of PVT with ACLSV which is currently classified in the closterovirus subgroup A.

(Received December 27, 1991)

Key words: potato virus T, RNA, nucleotide sequence, capillovirus, closterovirus, "alpha-like" supergroup.

INTRODUCTION

The potato virus T (PVT) particle is a flexuous filament 640 nm long and 12 nm in diameter. The virus is composed of a single-stranded RNA of M_r 2.2×10^6 (ca. 6.5 kb) and a single coat protein of M_r 27K³¹. PVT, apple stem grooving virus (ASGV) and citrus tatter leaf virus²⁸ have been excluded from the closterovirus group and placed in the new capillovirus group based on the accumulating evidence of differences between these viruses and definitive closteroviruses^{5,26,28,29,31,37}. Closteroviruses still constitute a diverse virus group which is divided into three subgroups, A (e.g., apple chlorotic leaf spot virus

* A preliminary report of this study was presented at the Meeting of the Phytopathological Society of Japan, Tokyo, April 3, 1991. Sequence data from this article have been deposited with the DDBJ/EMBL/GenBank Data Libraries under Accession No. D 10172.

** Laboratory of Plant Pathology, Faculty of Agriculture, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan
東京大学農学部

*** National Agriculture Research Center, Tsukuba, Ibaraki 305, Japan 農業研究センター

† Laboratory of Plant Molecular Biology, The Rockefeller University, New York, N.Y. 10021-6399, USA ロックフェラー大学

†† Author to whom correspondence should be addressed.

(ACLSV)), B (e.g., beet yellows virus (BYV)) and C (e.g., citrus tristeza virus (CTV)) based on the modal particle lengths and biological characteristics⁹. Sequence analysis of the ACLSV¹⁴, BYV³ and CTV³²) RNAs have revealed considerable diversity in their genomic organization.

At present, very little is known about the molecular organization of PVT and other capilloviruses. To assess more precisely the taxonomic position of capilloviruses, we determined the nucleotide sequence of the 3'-terminal portion of PVT RNA, and compared the deduced translational products and the genomic organization with those of other plant viruses. The sequence represented here is the first report on the capillovirus group.

MATERIALS AND METHODS

Virus purification and RNA preparation. A PVT isolate, kindly provided by Dr. T. Nishio, Yokohama Plant Protection Station, Yokohama, Japan, was propagated in *Chenopodium quinoa* by mechanical inoculation. Systemically infected leaf tissue was homogenized in a blender with liquid nitrogen and mixed with 5 volumes of 0.06 M phosphate buffer, pH 7.2, containing 0.1% 2-mercaptoethanol, 0.1% ascorbic acid and 0.01 M EDTA. The sap was centrifuged at $9,000 \times g$ for 5 min. The supernatant was mixed with 1/6 volume of carbon tetrachloride and centrifuged at $9,000 \times g$ for 15 min. The clarified extract was stirred with 6% (W/V) polyethylene glycol (M_r 6,000), 0.1 M NaCl and 1% Triton X-100 at 4°C for 1.5 hr, and then centrifuged at $17,000 \times g$ for 15 min. The pellets were suspended in 0.01 M phosphate buffer, pH 7.2, and centrifuged at $10,000 \times g$ for 10 min. The supernatant was concentrated by a 20% sucrose cushion centrifugation at $170,000 \times g$ for 1.5 hr. Virus suspension was loaded on top of a 20–50% cesium sulfate density gradient and centrifuged at $150,000 \times g$ for 5 hr. The opalescent virus band was collected, recentrifuged at $170,000 \times g$ for 2 hr, and the pellet was resuspended in distilled water. Viral RNA was isolated from purified virus preparation by incubation with 0.4% SDS and 1 mg/ml proteinase K, followed by two phenol-chloroform (1:1 v/v) extractions and subsequent ethanol precipitation.

cDNA synthesis and cloning. Double-stranded cDNA was synthesized using reverse transcriptase and an oligo (dT) primer for the first strand synthesis and *Escherichia coli* DNA polymerase I, RNase H and T4 DNA polymerase for the second strand synthesis. Synthesis of the first strand was monitored by incorporation of [³²P]dCTP, followed by electrophoresis in a 1.0% alkaline agarose gel. The cDNA product was electrophoresed in a 1.0% agarose gel. The fragments greater than 1.3 kb were recovered from excised gel pieces using NaI and glass powder, and finally cloned into *Sma*I-cleaved Bluescript II KS+.

Screening of cDNA clones. Recombinant plasmid DNAs were prepared on a small scale for dot blot hybridization with the viral RNA probe labelled with photobiotin (Bresatec, USA). The size of DNA inserts were estimated by agarose gel electrophoresis after digestion with *Pvu*II (*Pvu*II site; 532 and 977 of Bluescript II KS+). A recombinant plasmid containing a 2.4 kb insert was selected and designated pdT11. To confirm that pdT11 had a PVT cDNA insert, the viral RNA was electrophoresed in a 1.0% agarose gel, transferred to a nylon membrane with the capillary blot procedure and hybridized with a photobiotin-labelled pdT11.

Deletion of cDNA. DNA insert of pdT11 was cut out with *Pvu*II and ligated again in inverted orientation into *Sma*I-cleaved Bluescript II KS+ (named pdT11C). DNA insert of each clone (pdT11 and pdT11C) was digested with *Sac*I and *Bam*HI, unidirectionally deleted with exonuclease III followed by mung bean nuclease and Klenow fragment, and re-ligated to get subclones, using a commercial kit (Takara Shuzo, Japan).

Determination of the nucleotide sequence. Single-stranded DNAs were prepared with superinfection of helper phage M13K07 for sequencing. Annealing of DNA templates and fluorescent-dye-labelled primers (Applied Biosystems, USA) was carried out at 60°C for 10 min. The dideoxynucleotide chain termination reaction with *Taq* DNA polymerase (Promega, USA) was made at 70°C for 10 min according to a commercially provided protocol. DNA sequence was determined by an automated DNA sequencer (370A; Applied Biosystems, USA).

UGUAUUGCAAGAGAGGGGACUGAAUUCUGAUUAUUGAAGCUUUUCGUAUGAUCACAGGAUGCAAUUAUUGAGGUCUGAAAAUGAAUUGUCUG 90
V D Q U E E G G T A E S D Y E A F D R S Q D A I I L G L E I E C L 30
AAAUUGUUUGGAUGGACCAAGACCCUGAUCGUAUGAUUACAGAAAGCGUAAACUUUGGAUGGGAUUGAGGUUGGGGGCAAUUGCAAUUG 180
K L F G W D Q D L I D D Y R K L K L W M G C R L G A I A I M 60
AGAUUACACAGGUGAAUUGGAACAUCUUUUUUCAACACAUAUUGCAACACUUGCAUUCACAGUUGUAGGUAUAAUUAUUGACCCAGAGACACU 270
R F T G E F G T F F F F N T I A N I A F T C L R Y N I T R D T 90
GUGAUUGCGUUUUGCAGGAGAUUGACAUUGAUGAUCGCGAAAAUCUGAAUUCAGGAAAGAUAGAGAAGACCUGCUGGCCCAUCUAAACUUUA 360
V I A F A G A G D D M Y A S G K L E I R K A D R E D L L A H L T L 120
AAGCGAAAAGUACAGUUCACAGAAAAGCCAAUGUUUCUGUGGGUGUAUCAAGAAAGAUUGGGAUUGGAGAAGGAGCCAAAGAUUGGUACUU 450
K A K V G U A G T E K P M F C G W Y I K K M G I V K E P R L V L 150
GAGAGGUGUUGAUGCTGAAGAAGAAAGGUGAUAGAUCAAUGCUUACAACUCAAUUGCAGGUUAUCUUAUGGAUACAGAUUAGGU 540
E R W L I A E R K K V I D Q C F I N Y S I E V S Y G Y R L G 180
GAGUACCUUGCGGGAGAUUUUGAUAAUUUGAGGAGAUUUCCAAGCAUUGAGAGAUUGGUGAUAAAGAGAGAGAGCAACUGCCACAGCU 630
E Y L W E Y F D N L E D F C G A I V R L V I K K K Q L P P A 210
AUAAAGAGAUUUUUGAAACCAAGCAUUGAGGUUAUUUACUGGAGAGAGUUCAGGAGACAUUGGAGAGAGAGAGAGCAUCACGGGUCC 720
I R R I F E T S N G V D F S G E V Q E T M G E G E G E H H G S 240
M E L I S V E K A F R R Q W G E E R E S I T G P 22
UGUGGACUCUGGUGCAAUCUACACAAACUCCGCGUUUCACAAUCUGAAGACCAAUUGCGACGUUUACAAGUCUGAGUGCAGUAUUGGUUCU 810
C G L W C N L H 248
V D S G A I Y T N S A F H N L K T K W H V Y K S E C S I G L 52
UGAUUUUGCCGUAUACCGAAAAAUAAUUAUCAAAGGACAUUCCCCUUTUGAUCAAAGGAGGAUUGAUAAACAUCAUGAAAGACGUAUAAACA 900
D L P D N G K I I S K D I L F D Q C E U E I D N I M K D D K Q 82
GGUAUUCGUCUUGCGGAGCCUUGUGUGUUGCGGACUUGUUGCACAUUUCUGUCGAGGAGGAAGUGGAUUGACUUCUUAUUGA 990
V F V H L G A F V F G L V A H F P V D E E V E G L V S I D 112
CAAAUGUCGAGCUGAUUAUAGAGGAGGGCAACCCUAGCUUGCAGGAAAAUUAUAAUUGUGAAUUGGAAGAGUGCGUUUCAGAUUGAAACCAAA 1080
K R R T D L R R A T L A C R K U I K G F V N R G A U F M M K P N 142
CUUUUACAGUCAGGAGGAGGAUUAAGGAGUGGACACUUUUUGUGCCGUUCAAUCAAUUGAAACCAUUGGCUUUGAGAGGUGGUUCU 1170
F S V R K E D L R D G D T F C A A I K I K N L G F E G G F F 172
UCCAUUUUUCAGCAUGUGGCGGAGUGAUUAACAGAACCCUCAAUUGAGUUGUUGACAUUGCUGUUGACAAAACCUUGCUUAGGACAGU 1260
P F S A C G V G I Y R T S N V S F A H A V D K T F A S R T V 202
UCAUGAUUUUGUGGGUACUGAUUAUCUUCUUGGAUUAUCUGGACCAAGCCACUUGGAGGAGCCUUGGAGGAUGGAGAGGAGGCCCAU 1350
H D L V G C T D I L S L D Q Q L D Q A T L E D L E V R R S P I 232
UCUAAGGUUAACGCUUCGAGAGAGAGAUUAGAUAGAAAGAGGAGGAUUGGUUCCAGAGAAACCGGCAUUAAGGAGGAGAGAUUUUG 1440
L R L T A P D E R V M I E R G N W F G K K P A I R R R S F G 262
GAAAAGGAGCCCAAAAAGUUCGUCAGUUAGGUUGCGCGAGUUUGCCUAGUUUUCAGUUCUGAGAGGUUAGAAGGGGCUUGGAUUC 1530
K R R P Q K S S V R S S A S L P R F U S C S E R L E R G F E S 292
AGAGAGUCUGGUGGAUUGUGUGAUGAUGAGAAUUGAGCCCAACAUUUUGUGCAGAUUAGAGACGAGUUCUGGAUUGCAGCAGUGG 1620
E S L A G L V L D Q K Y G P N N F C A D K R R S P E S D S G 322
M D P T T F V Q I R D E V L N L T V 18
CAGCUUACAGCAGUCAUUGGAGCGGUCAGGCCACUCAGGCUUCUGAAGGAGUGGGCAAGGGAACAGAUUGUUGGUAUUCUUCUGGAAGAA 1710
S L Q Q S M G R S H S G S E G K G K T D V A I P L R K N 352
A A Y S S Q W D G Q A T Q A L K D G A K E Q M L R F L F R 48
UCGCAUAUUCUUCAGCGAGUAGAAACCAUUCUGGCCUGACACGAGAUUCGCGUGCGAGGAGUUCUCCAGAUUAGGUUAUGUCAGCGGCUUCAG 1800
R N I F S E 358
I A I S S A S R N T I W P D T E I A S E D L Q I G M S A A S 78
CGGUCUCCACCUCCAGUGGCGUGCGUCUCGUAUAGCCUUAUUAUUGGUGAAUUAUUAUUCUUAUGUCAAAAGUGCUAAUAGCUCUUGUCCA 1890
A G P P P V A A A P I S L T F R V N F N S Y V K M L I A L S 108
ACACUUCACCAAAUUCUUUGUAAAGAAUAAACACUCAGACAAUUGUGUAUUGCCUUUUGCAAUUAUGCUUAUGGAUACUUAUUGAGAA 1980
N T S T S N S F V K N K T L R Q M C M P F A K Y A Y G Y L S E 138
UGGGCUAUGCAACUUGGGCAUUGAGAAAAUUGCCCAAAUCUUGUGCGAAAGGCUAAUUGGGUUGCUUUCGAUUAUUGCAUUGGUGUUGA 2070
M G Y A T T W A Y E K M P K L C R K A K W A F D F A S G L L 168
UUGAUACAACAAUGCAGUUGAUGAUGAUGAAAAAGCCGUCAUUCAAAGGUCUGGGAGCUGAGUUUAUCAAACUCACAGUCUUAUCAA 2160
I D A T T M Q L N D D E K T V I U G L G A R L F K T Q Q S I Q 198
UUGUCUUAUCCAAUGGAGCGGUGAAGCUUAUCAAUCUGUGAAUUAUUGCUGAGAAAGGAGGAAGAGUUGGCGUUUAUUAUUAUCCGCAU 2250
A A U T S T M D G E A I N R E I 213
IAUUGGUCGCGGUCUUGAGGGUUGACUGGCCCUCAUUAUUGAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 2340
CCCCCGAUGUUUUAUUGAUCUAGGGUUGUGAUGGUUUUUCAAGUUUUUUAAGC (A)n 2392

Fig. 1. Nucleotide sequence of the 3'-terminal 2392 nucleotides of PVT RNA deduced from cDNA. The predicted amino acid sequence of the three ORFs is shown below the nucleotide sequence. The nucleotide sequence identical to a putative polyadenylation signal is underlined. Of the sequence, 89.5% was determined on both strands of cDNA and 97.8% of the total sequence was read from at least two independent overlapping subclones.

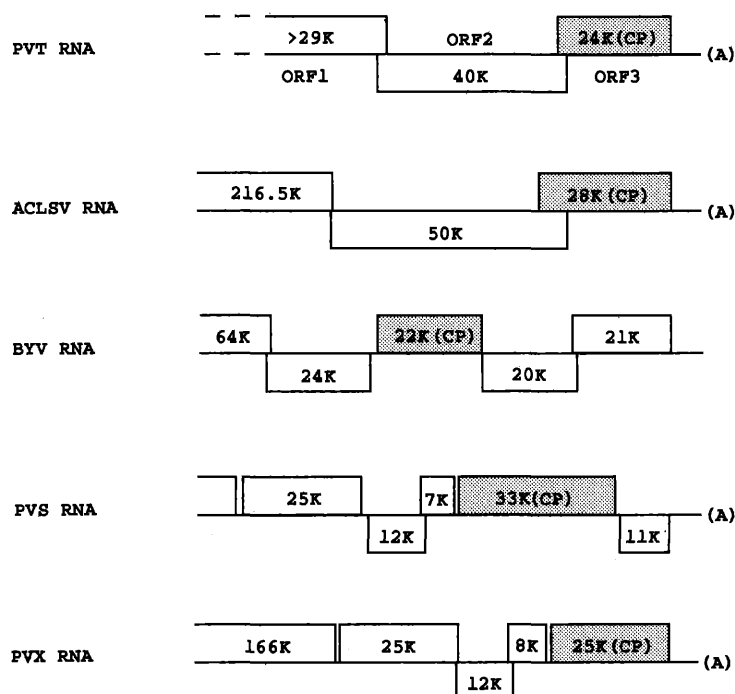


Fig. 2. Comparison of the genome organization of the 3'-terminal region of PVT RNA with those of apple chlorotic leaf spot virus (ACLSV)¹⁴⁾, beet yellows virus (BYV; closterovirus subgroup B)³⁾, potato virus S (PVS; carlavirus)²⁵⁾ and potato virus X (PVX; potexvirus)¹⁷⁾. The positions of ORFs (boxed) and the size of their putative products are shown. The ORFs presumed to encode the coat protein (CP) are hatched.

RESULTS AND DISCUSSION

Nucleotide sequence

Recombinant plasmids, pdT11 and pdT11C, which showed specific hybridization to the viral RNA (data not shown), were used for sequence analysis. The pdT11 contained a poly(A) tract at one end of its cDNA insert, indicating that the cDNA was derived from the 3'-terminal region of PVT RNA. The analysis of both recombinant plasmids provided the sequence of the 3'-terminal 2392 nucleotides upstream of the poly(A) tail of the PVT genome as shown in Fig. 1.

Gene organization of the 3'-terminal region

Three potential open reading frames (ORFs) were identified in three different reading frames of the positive strand of pdT11 in the 3'-terminal region of the PVT genome (Fig. 2). The deduced amino acid sequences of the proteins encoded by these ORFs are shown in Fig. 1. No other ORFs encoding products of more than 11K were found in the complementary strand. The 2392 nucleotide sequence starts with an incomplete ORF which terminates with an UAA codon at position 745. This ORF potentially encodes a protein of at least 248 amino acids (M_r >29K). The ORF 2 starts from an AUG codon at position 656 and ends with an UAG codon at position 1730. The ORF 3 starts from an AUG codon at position 1566 and ends with an UAG codon at position 2205. ORF 2 overlaps by 89 nucleotides with ORF 1 and by 164 nucleotides with ORF 3. The putative translation products of ORF 2 and ORF 3 have a M_r 40,451 (358 amino acids; M_r 40K) and M_r 23,596 (213 amino acids; M_r 24K), respectively.

These data suggest that PVT greatly differs from carlavirus, potexvirus and BYV (the closterovirus subgroup B) in the gene organization (Fig. 2). The PVT RNA has a 3' poly(A) tail as has been shown for the capillovirus ASGV RNA based on its chromatographic behaviour on oligo (dT) cellulose³⁷⁾. The

PVT	9-101	<u>E</u> <u>S</u> <u>D</u> <u>Y</u> <u>E</u> <u>A</u> <u>F</u> <u>D</u> <u>R</u> <u>S</u> <u>Q</u>	(43aa) <u>F</u> <u>T</u> <u>G</u> <u>E</u> <u>F</u> <u>G</u> <u>T</u> <u>F</u> <u>F</u> <u>F</u> <u>N</u> <u>T</u>	(18aa) <u>I</u> <u>A</u> <u>F</u> <u>A</u> <u>G</u> <u>D</u> <u>D</u> <u>M</u> <u>Y</u> <u>A</u>
ACLSV	1637-1730	<u>E</u> <u>S</u> <u>D</u> <u>Y</u> <u>T</u> <u>A</u> <u>F</u> <u>D</u> <u>V</u> <u>S</u> <u>Q</u>	(43aa) <u>F</u> <u>T</u> <u>G</u> <u>E</u> <u>F</u> <u>S</u> <u>T</u> <u>F</u> <u>L</u> <u>F</u> <u>N</u> <u>T</u>	(18aa) <u>I</u> <u>C</u> <u>F</u> <u>A</u> <u>G</u> <u>D</u> <u>D</u> <u>M</u> <u>C</u> <u>A</u>
PVS	141-234	<u>E</u> <u>S</u> <u>D</u> <u>Y</u> <u>E</u> <u>A</u> <u>F</u> <u>D</u> <u>A</u> <u>S</u> <u>Q</u>	(43aa) <u>F</u> <u>S</u> <u>G</u> <u>E</u> <u>A</u> <u>S</u> <u>T</u> <u>F</u> <u>L</u> <u>F</u> <u>N</u> <u>T</u>	(18aa) <u>I</u> <u>C</u> <u>F</u> <u>A</u> <u>G</u> <u>D</u> <u>D</u> <u>M</u> <u>C</u> <u>A</u>
PVX X3	1240-1333	A N D <u>Y</u> <u>T</u> <u>A</u> <u>F</u> D Q S Q	(43aa) L T G E G P T F D A N T	(18aa) Q V Y A G D D S A L
WC1MV	1075-1167	S N D <u>F</u> <u>T</u> <u>A</u> <u>F</u> D Q S Q	(43aa) L S G E G P T F D A N T	(18aa) Q V Y A G D D M - S
TYMV-type	1576-1668	A N D <u>Y</u> <u>T</u> <u>A</u> <u>F</u> D Q S Q	(43aa) L T G E P G T Y D D N T	(17aa) I M V S G D D S L I
BYV U	5-44	- - - - - - - - - -	K S G A S N T W I G N S	(18aa) L F V S G D D S L I
TMV vulgare	1384-1483	<u>E</u> L D I S K Y D K S Q	(49aa) K S G D V T T E I G N T	(18aa) G A E C G D D S L L
BMV	463-562	<u>E</u> A D L S K F D K S Q	(49aa) R T G D A F T Y E G N T	(18aa) A I E S G D D S L I
A1MV	528-629	<u>E</u> I D F S K Y D K S Q	(49aa) R T G D A L T Y L G N T	(20aa) V V A S G D D S L I
Sindbis	2270-2371	<u>E</u> T D I A S F D K S Q	(49aa) K S G M F L T L E V N T	(20aa) A A E I G D D N I I

Fig. 3. Alignment of the conserved sequences of PVT ORF1 protein with the replication-associated proteins of RNA viruses belonging to the "alpha-like" supergroup: the 216.5K protein of ACLSV¹⁴, partial replication-associated protein of PVS²⁵, the 166K protein of PVX X3¹⁷, the 147K protein of white clover mosaic virus (WC1MV)¹¹, the 206K protein of turnip yellow mosaic virus-type (TYMV)²⁷, the ORF1 polypeptide of BYV-U³, the 183K protein of tobacco mosaic virus vulgare (TMV)¹⁵, the 2A protein of brome mosaic virus (BMV)⁴, the 2A protein of alfalfa mosaic virus (A1MV)⁹ and the nsP4 protein of Sindbis virus³⁶. The numbers in front of the sequence indicate the position of the amino acids in each ORF. The numbers (in parentheses) within the sequences indicate the number of amino acid residues. Highly conserved amino acids are boxed; the other amino acids identical with the sequence of PVT are underlined.

arrangement of the three ORFs found in the 3'-terminal region of the PVT genome closely resembles that in the genome of ACLSV¹⁴ (Fig. 2). The ACLSV genome also has a 3' poly(A) tail and is almost similar in size to the PVT genome¹⁴. The genome of BYV, in contrast, is devoid of a poly(A) tail^{3,19} and is quite different from PVT and ACLSV in their size^{3,14} (Fig. 2). In addition, BYV differs greatly from PVT and ACLSV in its genomic organization³ (Fig. 2).

The 3' non-coding region of PVT consisted of 188 nucleotides upstream of the poly(A) tail (Figs. 1 and 2). It contained a putative polyadenylation signal AAUAAA²⁸ at 97 nucleotides upstream from the poly(A) tail. The signal is typical for eukaryotic mRNAs and was also found in ACLSV¹⁴, white clover mosaic virus (WC1MV)¹¹, papaya mosaic virus (PMV)¹¹, potato virus M³⁰ and *Helenium* virus S¹².

ORF 1

To assess functions of the putative ORF products of PVT, we searched for amino acid sequence homologies between these proteins and other viral proteins in databases. The putative amino acid sequence of ORF 1 (Fig. 1) is apparently homologous to the core of the RNA polymerases (replication-associated proteins) of positive strand RNA viruses (Fig. 3). These proteins are, the 216.5K protein of ACLSV¹⁴, the ORF1 polypeptide (at least 41K) of potato virus S (PVS)²⁵, the 165K protein of potato virus X (PVX)¹⁷, the 147K protein of WC1MV¹¹, the 206K protein of TYMV²⁷, the ORF1 polypeptide (at least 25K) of BYV³, the 180K protein of tobacco mosaic virus (TMV)¹⁵, the 2A protein of brome mosaic virus (BMV)⁴, the 90K protein of alfalfa mosaic virus (A1MV)⁹ and the nsP4 protein of Sindbis virus³⁶ (Fig. 3). The sequence homologies between the PVT >29K protein and the corresponding regions of other viruses shown in Fig. 3 were 46% (ACLSV), 49% (PVS), 34% (PVX), 35% (WC1MV), 32% (TYMV), 28% (BYV; lacking N-terminal sequence), 30% (TMV), 30% (BMV), 30% (A1MV) and 32% (Sindbis), respectively. The homologous regions of these viral proteins contain the sequence DxxxxD (48-54 amino acids) GxxxTxxxNT/S (21-24 amino acids) GDD. This signature sequence is thought to form the core of the RNA-dependent RNA polymerases of positive strand RNA viruses and designated polymerase domain^{7,10,16,18}.

From this analysis, the product of the incomplete ORF 1 of PVT seems to represent the C-terminal portion of the viral polymerase. This putative PVT polymerase was also found to be a close relative of the polymerases of Candresse's "group I" viruses (containing ACLSV (closterovirus subgroup B), carla-, potex- and tymo-viruses) of the "alpha-like" supergroup of plant RNA viruses^{7,16,21,38}. As shown in Fig.

PVT	40K	62-115	SKD <u>L</u> IPLFDQ (13aa)	FY-HLGA <u>F</u> VFGVLV (10aa)	QL-VS-- <u>I</u> IKRR
ACLSV	50K	63-118	LAS <u>L</u> IPVIPS (14aa)	YV-HWGAL <u>S</u> ISID (9aa)	GW-CY-- <u>V</u> YDNEW
BYV	64K	263-329	SLYL <u>L</u> PLLSE (25aa)	LLKVPPTINMHDS (7aa)	LRYLESYFEDDSN
TMV	30K	44-90	S-EVNL <u>L</u> -K (10aa)	CL-- <u>A</u> GLVVTGEW (9aa)	VS-VC-- <u>L</u> VDKRM
TRV-SYM	29K	49-98	H-KL <u>D</u> VFDQ (12aa)	KI-- <u>V</u> GVQLVVTs (9aa)	IQ-ID-- <u>L</u> IDSRL
CPMV-VGNM	58K	206-265	L-DIPVTKD (20aa)	AI-HIGAIETIIQ (10aa)	GG-FL-- <u>L</u> VDLSLH
CaMV-C	ORF1	92-150	V-YLPLITR (19aa)	MV-HLGA <u>V</u> KILLK (10aa)	IK-IA-- <u>L</u> IDRI

Fig. 4. Alignment of conserved domains of putative virus movement proteins. Amino acid sequences of movement proteins of PVT, ACLSV¹⁴, and BYV^{2,3}, TMV¹⁵, TRV-SYM⁶, CPMV-VGNM²⁴ and CaMV-C¹³ are aligned. The three conserved motifs, "L/IPL/V", "G" and "D" are boxed. Identical residues between PVT and ACLSV are shown by underlines. The numbers are as in Fig. 3.

Table 1. Percent amino acid sequence homology between coat proteins of several filamentous viruses^{a)}

	PVT	ACLSV	BYV	CTV	PVS	PVX	WC1MV	PMV
PVT	—	30^{b)}	27	28	25	28	28	30
ACLSV		—	28	25	24	28	25	27
BYV			—	30	22	27	26	29
CTV				—	23	26	25	26
PVS					—	30	36	27
PVX						—	36	43
WC1MV							—	34
PMV								—

a) The sources of sequence data were: PVT (this paper), ACLSV¹⁴, BYV³, CTV³², PVS²⁵, PVX¹⁷, WC1MV¹¹ and PMV¹. Sequence identities were analyzed by the MAXIMUM MATCHING program of DNASIS as described in MATERIALS AND METHODS.

b) Bold letters show homology above 30%.

3, the putative PVT polymerase was similar to those of ACLSV, PVS, PVX, WC1MV and TYMV which belongs to Candresse's "group I" viruses⁷) in the distribution of common amino acids (AF in the position 14 and 15 of PVT; E in the position 66 of PVT) and in the amino acid number between consensus amino acid residues (43 aa) (Fig. 3). However, comparing to this similarity, the putative PVT polymerase was not closely related to respective proteins of "group II" viruses^{2,7}) (containing BYV, TMV, BMV and A1MV in Fig. 3) belonging to the same supergroup, *i.e.* in the distribution of common amino acids (SK in the position 467 and 468 of BMV; K in the position 471 of BMV; KS or RT in the position 523 and 524 of BMV; G in the position 532 of BMV) and in the amino acid number between consensus amino acid residues (49 aa) (Fig. 3). The putative BYV polymerase was reported to be a close relative of the polymerases of tricornaviruses which belongs to Candresse's "group II" (containing tricornaviruses, tobamoviruses and hordei-viruses) of "alpha-like" supergroup²). According to Koonin's phylogeny, Candresse's both "group I" and "group II" are included in the polymerase supergroup III (tobamoviruses, tobamoviruses, hordei-, tricornaviruses, alpha-, rubi-, furo-, potex-, carla- and tymoviruses and BYV, ACLSV and hepatitis E virus). However each of them was on different branches of the same phylogenetic tree^{7,22}).

ORF 2

The putative PVT 40K protein (ORF2) showed partial homologies with the 50K protein of ACLSV (27%)¹⁴, the 64K protein of BYV (26%)^{2,3}, the 30K protein of TMV (28%)¹⁵, the 29K protein of tobacco rattle virus-SYN (26%, tobamovirus)⁶, the 58K protein of cowpea mosaic virus-VGNM (28%; CPMV, comovirus)²⁴ and the ORF1-encoded protein of cauliflower mosaic virus-C (27%; CaMV, caulimovirus)¹³ (Fig. 4). Recently, putative cell-to-cell movement proteins (MPs) have been shown to possibly be classified into two families (I and II)²³. Alignment of the sequences of PVT 40K protein showed homologies with those of family (I) MPs (MPs of tobamoviruses, tobamoviruses, caulimoviruses and comoviruses, ACLSV and geminiviruses with bipartite genomes). This protein exhibited three conserved motifs centered at conserved L/IPLV, G, and D residues, of which LPL, G and D are also reported as characteristic motifs for the family (I) MPs²³ (Fig. 4).

PVT	39-176	EQM---	LRFLF-GRIAISSAS	(17aa)	GMSAAS	(25aa)	LI-A-----	LSNTSTNS---		
ACLSV	35-159	EQI---	LESIE-ANIAIQGTS	(13aa)	SMEDQS	(12aa)	LIKA-----	FKTSSSDP---		
BYV	24-170	DKLRKNFEEL-	KLKGVPEDN	(11aa)	ATIGTS	(12aa)	KASFGGGKELYL	THGELNSFLG		
CTV	55-191	ALNRDLFLTL	KGKHPNLPDKD	(12aa)	AVKSSS	(12aa)	TREGVEVDLS	DKLWTDVV-FN-		
PVS	114-235	EQM-AQITADI-	AGLGVPTHEH	(12aa)	ASVSSS	(12aa)	TGA-----	VPLDSIIA---		
PVX	57-177	EDL-SEIEAIW-	KNMKVPTDT	(12aa)	ADVGS	(12aa)	NG-----	ISRARLAA---		
WC1MV	41-152	AEI-EAITKTW	AETFKIPNDV	(12aa)	ADVGS	(12aa)	AG-----	VSRQLAQ---		
PMV	39-152	EQL-KSVSTLM-	VAAKVPAAS	(12aa)	YDNGSS	(12aa)	E-----	ISLAQLAS---		
PVT	---	FVKNK-T---	LRQMCMPEAKYAYG-YLSE-MGY-ATWA---	YEKMPKLC--	RKAKWVA	AFDFASGLLIDTTMQLN				
ACLSV	---	NI-NKMT---	FRQCEAFAPERN-GLVK-LKYKGVFTNLFTTMPEVG---	SKYPELM	EDFNKGLNMFIMNKAQ					
BYV	SQKLEGGK-PNK	LRFCRFTQKD-VISLRKE-YRGKLPPIA-RANRHGLPAEDHY---	LA	ADFI	STSTELTDLQQS					
CTV	S-KGIGNR-TNAL	LRVWGRT--NDALY-LAFCRQNRNLSYGG-RPLDAGIPAGYHY--	LC	ADFL	TGAGLTDLECA					
PVS	---	IMKNRAG--LRK	VCRLYAP--VV-WNYM-LVQNRPPSD-WQA-MGFQWNAF--	AF	DFDYVTNGAAIQPV					
PVX	---	AIKEVCT--LRQ	FCMKYAP--VV-WNWM-LTNSPPAN-WQA-QGFKPEHKF--	AF	DFDFNGVTNPAAIMPK					
WC1MV	---	AIKIHCT--LRQ	FCMYFAN--IV-WNIM-LDTKTPPAS-WSK-LGYKESKF--	AF	DFDFGVNHPAALMPA					
PMV	---	IVKASGTS--LRK	FCRYFAP--II-WN-L-RTDKMAPAN-WEA-SGYKPSAKF--	AF	DFDFGVNHPAAMQPP					
SALT BRIDGE										

SALT BRIDGE

Fig. 5. Alignment of amino acid sequences of PVT ORF3 protein with homologous regions of the coat proteins of ACLSV¹⁴, BYV³, CTV³², PVS²⁶, PVX¹⁸, WC1MV¹¹ and PMV^{1,33,35}. Segments potentially involved in salt bridge formation together with three conserved motifs, S, LRQ and AFDF are indicated. These are commonly observed in the coat proteins of rod-shaped and filamentous viruses⁹. Identical residues between PVT and ACLSV are shown by underlines. Gaps (—) were inserted to maximize the alignment. The numbers are as in Fig. 3.

ORF 3

The PVT 24K protein displayed partial homologies with the published coat protein sequences of ACLSV¹⁴, BYV³, CTV³², PVS²⁵, PVX¹⁷, WC1MV¹¹ and PMV^{1,33,35} (Table 1). They share three small blocks of conserved amino acids, S, LRQ and AFDF (Fig. 5), of which R in LRQ and AFDF are also present in the coat proteins of filamentous viruses^{9,20,34}. These conserved amino acids were suggested to be important acting as a salt bridge for maintaining the correct tertiary structure of the coat proteins and/or playing a role in the interaction with the viral RNAs (Fig. 5)^{9,34}. Based on these characteristics, we presumed that ORF 3 of the PVT genome encodes the coat protein. This putative product showed around 30% homology, including the above salt bridges, with the coat proteins of ACLSV, carla-, potex- and poty-viruses. The PVT putative coat protein showed slightly higher homology with ACLSV (Fig. 5) than CTV, BYV, potex- and carla-viruses (Table 1).

Relationship to closterovirus

The sequence represented here is the first report on the group capillovirus which was established recently²⁶. Notwithstanding some level of homology in the coat protein with BYV (27%) and CTV (28%)³², significant differences of PVT to BYV in the genomic organization (Fig. 2) and the amino acid sequence of the putative viral polymerase (Fig. 3) suggest that PVT- capillovirus may be evolutionarily distinct from BYV-closterovirus (subgroup B). This supports the current classification of these viruses in separate groups (capillovirus and closterovirus).

It is surprising that all the amino acid sequence motifs described for the three kinds of proteins (putative viral polymerase, MP, and coat protein) of ACLSV-closterovirus (subgroup A) are preserved also in PVT sequences. In addition, the extensive homologies in the genomic organization between PVT and ACLSV suggest their close evolutionary relationship. The genome size of CTV-closterovirus (subgroup C)³² was quite different from PVT, and the CTV genome was supposed to be devoid of a poly(A) tail¹⁴. These molecular data indicate the heterogeneous properties of viruses within the closterovirus group. After this work was completed, V.V. Dolja together with E. Koonin independently made a preliminary analysis of our amino acid sequences of the PVT putative polymerase and the movement and coat proteins. Sequence database screening (program SMART, SWISSPROT, release 18) showed, completely in line with our conclusions, that the putative RNA polymerase of PVT belongs to "alpha-like" supergroup of polymerases (or the polymerase supergroup III), and that most close relatives of the PVT 40K and 24K proteins are the 50K protein and the coat protein of ACLSV, respectively

(Dolja, personal communication). The alignments of these proteins (program OPTAL) are statistically significant giving the scores of 11.3 SD for the capsid protein alignment and 7.8 SD for the movement protein alignment (All the methods used for analysis are described appropriately in the paper by Koonin *et al.*²³⁾). These results suggest that ACLSV and PVT should be classified in the same group, although the total sequence of PVT should be compared to that of ACLSV.

Note added in proof

During the processing of this paper, Dr. N. Yoshikawa *et al.* (Iwate University, Japan) reported the sequence data of ASGV genome at the annual meeting of the Phytopathological Society of Japan, Morioka, May 28, 1992. The data showed the complete differences in the genome organization of PVT and ASGV. ASGV genome has two open reading frames that encode a large polyprotein containing polymerase and coat protein and another protein coding viral serine protease. These data and our own data suggest that PVT should be excluded from the capillovirus group and that PVT and ACLSV (which is greatly different in many characteristics from subgroup B and C of closteroviruses) could be classified into a new virus group.

We gratefully acknowledge Dr. A. Kawai, Yokohama Plant Protection Station, Yokohama, Japan, Dr. H. Hibino, National Agriculture Research Center, Tsukuba, Japan, and Dr. Y. Minobe, National Institute of Agrobiological Resources, Tsukuba, Japan, for technical advice. We would like to thank to Dr. I. Havukkala, the National Institute of Agrobiological Resources, Tsukuba, Japan, for critical reading this manuscript. We also would like to thank to Dr. V.V. Dolja, Texas A&M University, College Station, Texas, USA, for reviewing this manuscript and helpful discussions.

Literature cited

1. Abouhaidar, M.G. (1988). Nucleotide sequence of the capsid protein gene and the 3' non-coding region of papaya mosaic virus RNA. *J. gen. Virol.* 69: 219-226.
2. Agranovsky, A.A., Boyko, V.P., Karasev, A.V., Koonin, E.V. and Dolja, V.V. (1991). The putative 65K protein of beet yellows closterovirus is a homologue of HSP70 heat shock proteins. *J. Mol. Biol.* 217: 603-610.
3. Agranovsky, A.A., Boyko, V.P., Karasev, A.V., Lunina, N.A., Koonin, E.V. and Dolja, V.V. (1991). Nucleotide sequence of the 3'-terminal half of beet yellows closterovirus RNA genome: unique arrangement of eight virus genes. *J. gen. Virol.* 72: 15-23.
4. Ahlquist, P., Dasgupta, R. and Kaesberg, P. (1984). Nucleotide sequence of the brome mosaic virus genome and its implications for viral replication. *J. Mol. Biol.* 172: 369-383.
5. Bar-Joseph, M. and Murrant, A.F. (1982). Closterovirus group. CMI/AAB Descriptions of Plant Viruses. No. 260.
6. Boccardo, M., Hamilton, W.D.O. and Baulcombe, D.C. (1986). The organization and interviral homologies of genes at the 3' end of tobacco rattle virus RNA1. *EMBO J.* 5: 223-229.
7. Candresse, T., Morch, M.D. and Dunne, J. (1990). Multiple alignment and hierarchical clustering of conserved amino acid sequences in the replication-associated proteins of plant RNA viruses. *Res. Virol.* 141: 315-329.
8. Cornelissen, B.J.C., Brederode, F. Th., Veeneman, G.H., van Boom, J.H. and Bol, J.F. (1983). Complete nucleotide sequence of alfalfa mosaic virus RNA2. *Nucleic Acids Res.* 11: 3019-3025.
9. Dolja, V.V., Boyko, V.P., Agranovsky, A.A. and Koonin, E.V. (1991). Phylogeny of capsid proteins of rod-shaped and filamentous RNA plant viruses: two families with distinct patterns of sequence and probably structure conservation. *Virology* 184: 79-86.
10. Domier, L.L., Shaw, J.G. and Rhoaws, R.E. (1987). Potyviral proteins share amino acid sequence homology with picorna-, como- and caulimoviral proteins. *Virology* 158: 20-27.
11. Forster, R.L.S., Bevan, M.W., Harbison, S.-A. and Gardner, R.C. (1988). The complete nucleotide sequence of the potyvirus white clover mosaic virus. *Nucleic Acids Res.* 16: 291-303.
12. Forster, G.D., Millar, A.W., Meehan, B.M. and Mills, P.R. (1990). Nucleotide sequence of the 3'-terminal region of *Helium* virus S RNA. *J. gen. Virol.* 71: 1877-1880.
13. Gardner, R.C., Howarth, A.J., Hahn, P., Brown-Luedi, M., Shepherd, R.J. and Messing, J. (1981). The complete nucleotide sequence of an infectious clone of cauliflower mosaic virus by M13mp7 shotgun sequencing. *Nucleic Acids Res.* 9: 2871-2888.

14. German, S., Candresse, T., Lanneau, M., Huet, J.C., Pernollet, J.C. and Dunez, J. (1990). Nucleotide sequence and genomic organization of apple chlorotic leaf spot closterovirus. *Virology* 179: 104-112.
15. Goelet, P., Lomonosoff, G.P., Bulter, P.J.G., Akam, M.E., Gait, M.J. and Karn, J. (1982). Nucleotide sequence of tobacco mosaic virus RNA. *Proc. Natl. Acad. Sci. USA* 79: 5818-5822.
16. Goldbach, R. (1987). Genome similarities between plant and animal RNA viruses. *Microbiol. Sci.* 4: 197-202.
17. Huisman, M.J., Linthorst, H.J.M., Bol, J.F. and Cornelissen, B.J.C. (1988). The complete nucleotide sequence of potato virus x and its homologies at the amino acid level with various plus-stranded RNA viruses. *J. gen. Virol.* 69: 1789-1798.
18. Kamer, G. and Argos, P. (1984). Primary structural comparison of RNA-dependent RNA polymerases from plant, animal and bacterial viruses. *Nucleic Acids Res.* 12: 7269-7282.
19. Karasev, A.V., Agranovsky, A.A., Rogov, V.V., Mioshnicenko, N.A., Dolja, V.V. and Atabekov, J.G. (1989). Virion RNA of beet yellows closterovirus: Cell-free translation and some properties. *J. gen. Virol.* 70: 241-245.
20. Kashiwazaki, S., Hayano, Y., Minobe, Y., Omura, T., Hibino, H. and Tsuchizaki, T. (1989). Nucleotide sequence of the capsid protein gene of barley yellow mosaic virus. *J. gen. Virol.* 70: 3015-3023.
21. Kashiwazaki, S., Minobe, Y., Omura, T. and Hibino, H. (1990). Nucleotide sequence of barley yellow mosaic virus RNA1: A close evolutionary relationship with potyviruses. *J. gen. Virol.* 71: 2781-2790.
22. Koonin, E.V. (1991). The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. *J. gen. Virol.* 72: 2197-2206.
23. Koonin, E.V., Mushegian, A.R., Ryabov, E.V. and Dolja, V.V. (1991). Diverse groups of plant RNA and DNA viruses share related movement proteins that may possess chaperone-like activity. *J. gen. Virol.* 72: 2895-2903.
24. Lomonosoff, G.P. and Shanks, M. (1983). The nucleotide sequence of cowpea mosaic virus B RNA. *EMBO J.* 2: 2253-2258.
25. Mackenzie, D.J., Tremaine, J.H., and Stace-Smith, R. (1989). Organization and interviral homologies of the 3'-terminal portion of potato virus S RNA. *J. gen. Virol.* 70: 1053-1063.
26. Milne, R.G. (1988). Taxonomy of the rod-shaped filamentous viruses. *In the Plant Viruses* (Milne, R.G. ed.). Vol. 4. Plenum Press, New York. pp. 3-50.
27. Morch, M.-D., Boyer, J.-C. and Haenni, A.-L. (1988). Overlapping open reading frames revealed by complete sequencing of turnip yellow mosaic virus genomic RNA. *Nucleic Acids Res.* 16: 6157-6173.
28. Nishio, T., Kawai, A., Takahashi, T., Namba, S. and Yamashita, S. (1989). Purification and properties of citrus tatter leaf virus. *Ann. Phytopath. Soc. Japan* 55: 254-258.
29. Ohki, S., Yoshikawa, N., Inoue, N. and Inoue, T. (1989). Comparative electron microscopy of *Chenopodium quinoa* leaves infected with apple chlorotic leaf spot, apple stem grooving, or citrus tatter leaf virus. *Ann. Phytopath. Soc. Japan* 55: 245-249.
30. Rupasov, V.V., Morozov, S. YU., Kanyuka, K.V. and Zavriev, S.K. (1989). Partial nucleotide sequence of potato virus M RNA shows similarities to potexviruses in gene arrangement and the encoded amino acid sequences. *J. gen. Virol.* 70: 1861-1869.
31. Salazar, L.F. and Harrison, B.D. (1978). Potato virus T. CMI/AAB Descriptions of Plant Viruses. No. 187.
32. Sekiya, M.E., Lawrence, S.D., McCaffery, M. and Cline, K. (1991). Molecular cloning and nucleotide sequencing of the coat protein gene of citrus tristeza virus. *J. gen. Virol.* 72: 1013-1020.
33. Short, M.N., Turner, D.S., March, J.F., Pappin, D.J.C., Parente, A. and Davies, J.W. (1986). The primary structure of papaya mosaic virus coat protein. *Virology* 152: 280-283.
34. Shukla, D.D., Strike, P.M., Tracy, S.L., Gough, K.H. and Ward, C.W. (1988). The N and C termini of the coat proteins of potyviruses are surface-located and the N terminus contains the major virus-specific epitopes. *J. gen. Virol.* 69: 1497-1508.
35. Sit, T.L., Abouhaidar, M.G. and Holy, S. (1989). Nucleotide sequence of papaya mosaic virus RNA. *J. gen. Virol.* 70: 2325-2331.
36. Strauss, E.G., Rice, C.M. and Strauss, J.H. (1984). Complete nucleotide sequence of the genomic RNA of Sindbis virus. *Virology* 133: 92-110.
37. Yoshikawa, N. and Takahashi, T. (1988). Properties of RNAs and proteins of apple stem grooving and apple chlorotic leaf spot viruses. *J. gen. Virol.* 69: 241-245.
38. Zimmern, D. (1988). Evolution of RNA viruses. *In "RNA Genetics"* (Holland, J.J., Domingo, E.R. and Ahlquist, P. eds.), CRC Press, Boca Raton. pp. 211-240.

和 文 摘 要

越智資泰・柏崎 哲・平塚和之・難波成任・土崎常男：Potato virus T RNA の 3' 末端領域の塩基配列

Capillovirus グループの potato virus T (PVT) RNA の 3' 末端の 2,392 塩基の配列を決定した。本配列は 5' 末端側より互いに末端の重複した ORF 1 (少なくとも 248 アミノ酸 (29 K) 以上), ORF 2 (358 アミノ酸; 40 K), および ORF 3 (213 アミノ酸; 24 K) をコードし, 3' 末端には 188 塩基よりなる非翻訳領域と poly (A) が認められた。29 K タンパク質は, alpha-like supergroup のうち, ポリメラーゼグループ II に属する数種植物 RNA ウイルス (beet yellows virus (BYV; closterovirus subgroup B), tobacco mosaic virus, tricornavirus) のポリメラーゼタンパク質の C-末端に特有のアミノ酸配列よりも, ポリメラーゼグループ I に属するウイルス (apple chlorotic leaf spot virus (ACLSV; closterovirus subgroup A), carlavirus, potexvirus) のそれに類似していた。24 K タンパク質は, ひも状ウイルス (ACLSV, BYV, citrus tristeza virus (CTV; closterovirus subgroup C), potexvirus, carlavirus) の外被タンパク質と共通のアミノ酸配列ブロックを有していた。40 K タンパク質は, 植物ウイルスの cell-to-cell movement タンパク質と推定される配列と類似していた。また PVT は, ACLSV と 3' 末端側において ORF の配置が酷似しており, poly (A) のない BYV とはまったく異なっていた。Capillovirus グループの遺伝子構造に関しては, 本報告が最初である。以上から, 現在 closterovirus subgroup A に分類されている ACLSV は, 遺伝子構造上むしろ PVT に類似し, subgroup B および C のウイルスとは異なることが明らかになり, ACLSV はむしろ PVT と同一グループを形成すると考えられることから, その分類については再検討の必要があろう。