

**Determination of the taxonomic position and characterization
of yam mosaic virus isolates based on sequence data of the
5'-terminal part of the coat protein cistron***

O. Duterme, D. Colinet, J. Kummert, and P. Lepoivre

Faculté Universitaire des Sciences Agronomiques, Laboratoire de Pathologie Végétale,
Gembloux, Belgium

Accepted January 22, 1996

Summary. The sequences of the N-terminal part of the coat protein cistron from six isolates of yam mosaic virus (YMV-TOG, YMV-COT, YMV-12, YMV-CAR, YMV-BU1 and YMV-BU2) were determined. The analysis of the deduced amino acid sequences revealed the presence of consensus motifs characteristic of the potyvirus genus supporting the classification of YMV as a potyvirus member. The alignment of the N-terminal part of the coat protein of YMV-TOG, YMV-COT, YMV-12 and YMV-CAR showed that they were identical in size (152 aa) while YMV-BU1 and YMV-BU2 were shorter (140 aa) due to a deletion of 12 aa. These amino acid sequences exhibit an overall sequence identity ranging from 70.4% to 97.4% while the identity level with the other potyviruses sequenced in the considered region is below 50%, confirming that YMV is a distinct member of the potyvirus genus. The detailed analysis of the amino acid sequence alignment and of the identity levels observed between the N-terminal part of the coat protein of the six YMV isolates lead us to suggest that they have to be considered as distantly related strains of YMV rather than closely related but distinct viruses.

Introduction

Yam mosaic virus (YMV) was first described by Thouvenel and Fauquet [22–24]. This virus causes economically important losses in yams (*Dioscorea spp.*) and the main symptoms consists of interveinal mosaic or vein banding. YMV has been reported in several African countries: Ivory Coast [22, 23],

* The sequence data reported in this paper have been submitted to the EMBL database under the following accession numbers: YMV-TOG: Z48172, YMV-COT: Z48155, YMV-12: Z48173, YMV-CAR: Z48174, YMV-BU1: Z48156, YMV-BU2: Z48157.

Nigeria [21] and Togo [12]. A morphologically similar virus, serologically related to YMV, has been reported in the Caribbean [8].

YMV particles are flexuous and rod shaped, about 785 nm long and composed of a single-stranded RNA molecule associated with one species of a 34 kDa coat protein (CP). This agent is transmitted by aphids in the non-persistent manner [23]. Due to these biological, microscopical and serological properties, YMV was thus classified in the potyvirus group [24]. Shukla et al. [20] classified YMV as a possible member of the *Potyvirus* genus.

The use of sequence information for the classification of potyviruses is generally considered as more reliable than criteria based on biological and serological properties. Moreover, CP amino acid sequence data can be used to identify and differentiate distinct strains of potyviruses [14, 17–19, 25, 26]. So far, sequence data which would allow the classification of YMV to be confirmed as a potyvirus member has not been published. In order to validate this classification and to more precisely characterize YMV, a RT-PCR assay utilizing degenerate primers was designed to amplify the 5'-terminal part of the CP cistron of six YMV isolates. The implication of the sequence analysis of the N-terminal part of the CP for the taxonomic status of these six YMV isolates are discussed.

Materials and methods

Virus isolates and plant material

The six YMV isolates originated from various countries: Ivory Coast (YMV-12 and YMV-COT), Togo (YMV-TOG), Burkina Faso (YMV-BU1 and YMV-BU2) and Caribbean (YMV-CAR). YMV-TOG, YMV-COT, YMV-CAR, YMV-BU1 and YMV-BU2 were isolated from yam clones (*Dioscorea* spp.). All these isolates were propagated in *Nicotiana benthamiana* by mechanical inoculation from YMV infected yam leaves. The YMV-TOG isolate was provided by the Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH Abteilung Pflanzenviren in Braunschweig, Germany (ref: PV-0106, YMV). YMV-12, a YMV isolate originating from Ivory-Coast, was provided as a purified virus preparation by Dr. J. Dubern (ORSTOM, Montpellier). Host range and symptomatology data of this isolate revealed that this virus was very similar to that studied in the original description of the disease (J. Dubern, pers. comm.). The YMV-COT, YMV-CAR, YMV-BU1 and YMV BU2 isolates were respectively provided by the IIRSDA (Abidjan, Ivory Coast), by Dr. S. Mantell (Wye College, University of London) and by the INERA (Kamboisé, Burkina, Faso). Potato Virus Y (PVY), used as the potyvirus reference strain, was maintained by mechanical transmission in *N. tabacum* cv *Samsun*.

RNA extraction and cDNA synthesis

For the YMV-COT, YMV-CAR, YMV-BU1 and YMV-BU2 isolates, as well as for PVY, total RNA was extracted from 2 g of symptomatic leaves by the method of Chirgwin et al. [2]. For the YMV-TOG isolate, total RNA was extracted from 100 mg of infected *N. benthamiana* leaves using the TRIzol Reagent (Life technologies) while viral RNA of YMV-12 isolate was extracted from the purified virus preparation received from Dr. Dubern. 200 µg of YMV-12 was diluted in 100 µl of DEPC-treated water. After addition of 25 µl of 25% SDS, the mixture was incubated at 60 °C during 30 min. 75 µl of DEPC-treated water were then added and the

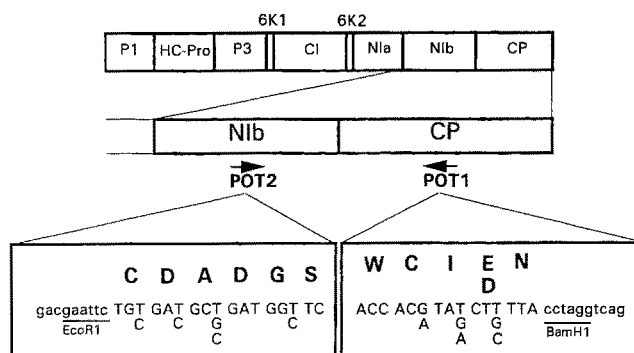


Fig. 1. Amino acid and nucleotide sequence of the degenerate primers POT1 and POT2 and their localization in the potyvirus genome. *P1* First protein; *HC-Pro* helper component-protease; *P3* third protein; *6K1* first 6K peptide; *CI* cytoplasmic inclusion protein; *6K2* second 6K peptide; *NIa* nuclear inclusion a protein; *NIb* nuclear inclusion b protein (RNA polymerase); *CP* coat protein

RNAs were purified by two phenol/chloroform extractions. After ethanol precipitation, the RNA was resuspended in 50 μ l of DEPC treated water.

Single stranded cDNAs were synthesized from 5 μ g of total or viral RNA using an oligo-dT primer according to the Superscript Preamplification System protocol (Gibco-BRL). The resulting cDNAs were diluted 5-fold with sterile water. Five microliter from these preparations were then used for the PCR amplifications.

PCR experiments

Two degenerate primers (POT1 and POT2) were used for the amplification of cDNA corresponding to the 3'-terminal part of the NIb cistron and to the 5'-terminal part of the CP cistron [3]. The sequence of the two degenerate primers and their location in the potyvirus genome are presented in Fig. 1. Amplification of the cDNA was carried out in a 50 μ l volume using the Dynazyme DNA Polymerase kit. Five microliter of 10-fold Dynazyme PCR buffer, 100 ng of each POT1 and POT2 primers, 100 μ M of each dATP, dCTP, dGTP, dTTP and 1 unit of Dynazyme DNA polymerase were added to 5 μ l of the cDNA preparation. Sterile water was added to a final volume of 50 μ l.

The following thermal cycle scheme was used for 37 reaction cycles (Biometra cycler): template denaturation at 94 $^{\circ}$ C for 30 sec, primer annealing at 40 $^{\circ}$ C (cycles 1–5) or 50 $^{\circ}$ C (cycles 6–37) for 1 min and DNA synthesis at 72 $^{\circ}$ C for 2 min. After a final 10 min elongation step at 72 $^{\circ}$ C at the end of these 37 cycles, amplification products were analysed by electrophoresis in a 1% agarose gel, in Tris-acetate-EDTA buffer [15]. Bands were visualised by ethidium bromide staining.

Cloning and sequencing of PCR-amplified fragments

After electrophoresis, bands of expected lengths were excised and eluted with the QIAEX Gel Extraction Kit from QIAGEN. For the cloning of the amplified fragments, we used the EcoRI and BamHI restriction sites located respectively in the POT1 and POT2 primers. The YMV-BU2 fragment was cloned in a EcoRI-BamHI digested pBluescript vector (Stratagene). The other amplified fragments were cloned in two separate fragments (due to the presence of an internal EcoRI restriction site for YMV-COT, YMV-BU1, YMV-CAR and

YMV-TOG, and due to the presence of a BamHI restriction site for YMV-12). Those digested fragments were inserted in their corresponding sites in the pBluescript vector.

For the sequencing reactions, the T7 primer, the primer Reverse or an internal synthetic oligonucleotide derived from the sequences were used. The double-stranded DNA sequencing by the dideoxy chain termination method was carried out using T7DNA polymerase (Pharmacia) according to manufacturer's instructions.

Comparison of the sequences

The nucleotides and the derived amino acid sequences were compared to the EMALL and SWALL library from the EMBL using the Fasta program [11]. Multiple alignment and dendrogram were performed with the Clustal V program [5, 6].

Results

RT-PCR amplifications

RT-PCR were carried out using the two degenerate primers, POT1 and POT2, designed to amplify cDNA fragments spanning the 3'-terminal part of the NIB cistron and the 5'-terminal part of the potyvirus CP cistron [3]. RT-PCR assays were carried out on RNA extracted from plants infected with the YMV-TOG, YMV-COT, YMV-CAR, YMV-BU1, YMV-BU2 isolates and on a purified virus preparation of YMV-12. Total RNA isolated from *N. tabacum* cv Samsun infected with PVY was used as a positive control. Amplification with primers POT1 and POT2 yielded a fragment of about 1.3 kb for the six YMV isolates and a fragment of about 1.2 kb for PVY (Fig. 2, lane 2–8) whose length was in agreement with the known sequence of that virus [13]. No fragments were amplified from RNA extracted from healthy *N. benthamiana* (Fig. 2, lane 9).

Cloning and sequence analysis

The various PCR fragments corresponding to the 5'-terminal part of the CP cistron were cloned in the pBluescript vector and their nucleotide sequence were determined from three independent clones obtained from a single amplification.

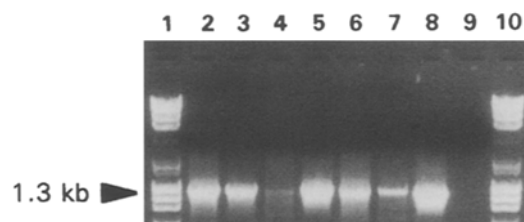


Fig. 2. Electrophoretic gel of the RT-PCR assays performed with primers POT1 and POT2 on RNA extracted from a healthy *N. benthamiana* (9) and from plants infected with: YMV-TOG (2); YMV-COT (3); YMV-12 (4); YMV-CAR (5); YMV-BU1 (6); YMV-BU2 (7) and PVY (8). 1, 10 Molecular weight markers



Fig. 3. Amino acid sequence alignment of the N-terminal part of the coat protein of the six YMV isolates and of PVY-D. /: presumed cleavage site between the NIB and the CP cistron. The trypsin-resistant core region of the CP begins at the amino acid residue equivalent to D₃₃ in PVY [28]

Table 1. Sequence identity (in percent) between the N-terminal part of the coat protein of the six YMV isolates

	YMV-TOG	YMV-COT	YMV-12	YMV-CAR	YMV-BU1	YMV-BU2
YMV-TOG	–	97.8	77.2	83.5	73.5	73.2
YMV-COT	97.4	–	76.8	83.7	74.1	74.3
YMV-12	77.3	76.3	–	76.3	70.0	70.0
YMV-CAR	82.2	82.2	76.3	–	73.2	73.7
YMV-BU1	75.7	75.0	70.4	73.0	–	97.1
YMV-BU2	76.3	75.7	71.1	71.7	95.7	–

Above the diagonal: nucleotide sequence identity. Below the diagonal: amino acid sequence identity

The deduced amino acid sequences obtained for the six YMV isolates were aligned with that of the N-terminal region of the CP of PVY-D [16] (Fig. 3). This alignment revealed in all the YMV isolates the presence of consensus motifs characteristic of the potyvirus genus. Among them is the “DAG box” (position 13 to 15, Fig. 3) probably involved in the aphid transmissibility of the virus [1] and the putative cleavage site (Q/A, Fig. 3) between the NIB and the CP cistron. Other consensus motifs were also present in the C-terminal part of the NIB cistron (data not shown).

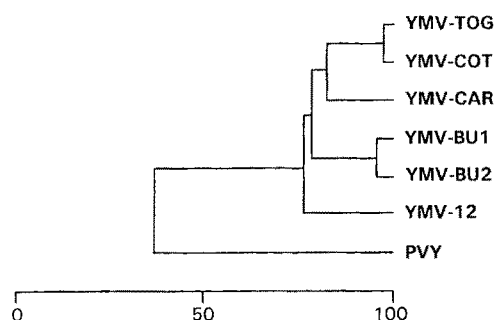


Fig. 4. Dendrogram built from the amino acid sequence alignment presented in Fig. 3. Vertical distances are arbitrary, horizontal distances are proportional to percent amino acid sequence divergence, shown in the scale below the figure

The sequence comparisons of the N-terminal parts of the CP of the six YMV isolates revealed identities ranging from 70.0% to 97.8% and from 70.4% to 97.4% at the nucleotide and at the amino acid level respectively (Table 1). At the amino acid level, high sequence identity were observed between YMV-BU1 and YMV-BU2 (95.7%) and between YMV-TOG and YMV-COT (97.4%). The other sequence identities observed between the six YMV isolates ranged from 70.4% to 82.2% (Table 1).

The comparison of the six YMV sequences revealed that the N-terminal part of the CP encompass a 152 aa fragment for the YMV-TOG, YMV-COT, YMV-12 and YMV-CAR isolates. However, this fragment is only 140 aa long for the YMV-BU1 and YMV-BU2 isolates. Indeed, the YMV-BU1 and YMV-BU2 isolates lack a stretch of 12 aa from position 34 to 45 in comparison to the four other isolates (Fig. 3).

In order to show the relationships between the six YMV isolates and PVY, a dendrogram was built from the multiple amino acid sequence alignment (Fig. 4). This dendrogram showed that YMV-TOG and YMV-COT and that YMV-BU1 and YMV-BU2 were respectively closely related. YMV-CAR shared an average sequence identity of 82.2% in comparison to YMV-TOG and YMV-COT while YMV-BU1 and YMV-BU2 presented 78.8% sequence identity with YMV-CAR, YMV-COT and YMV-TOG. The average sequence identity between YMV-12 and the five other YMV isolates was 75.9%. PVY, the type member of the *Potyvirus* genus, shared an average sequence identity of 35% with the six YMV isolates.

Amino acid sequence comparison between the six YMV isolates and 28 potyviruses previously sequenced in the considered region (listed in [20]) exhibit a broader divergence (identity level below 50%) (data not shown).

Discussion

The aim of our study on the N-terminal part of the CP cistron was to analyse the taxonomical relationship between YMV and members of the *Potyvirus* genus

and to evaluate the variability between YMV isolates of different origins. The comparison of the entire CP cistron has been widely used to establish taxonomic relationships among potyvirus members [14, 17–19, 26]. The sequence data show a bimodal distribution of CP sequence identities [17–19]. Known distinct potyviruses vary in CP amino acid sequence identity between 38% to 71%, while known strains of viruses range from 90% to 99%.

The PCR procedure utilizing degenerate primers designed to amplify the region coding for the C-terminal half of the Nib polymerase and the N-terminal half of the CP is a rapid and sensitive technique for detecting and characterizing members of the *Potyvirus* genus [3]. The cloning and sequencing of the six YMV amplified fragments obtained with the degenerate primers POT1 and POT2 revealed potyvirus specific consensus motifs, thus confirming the classification of YMV as a potyvirus member on the base of its biological, serological and microscopical properties [20]. Among these motifs, the “DAG box” [1] is in agreement with the aphid transmissibility of the YMV [24].

YMV-BU1 and YMV-BU2 lacked a stretch of 12 aa from position 34 to 45 within the CP cistron in comparison to the four other isolates. Similar deletions have been observed within some strains of PPV [7] and for SMV-N and WMV-2 [28]. Examination of the nucleotide and amino acid sequences of the six YMV isolates revealed no evidence for partial CP duplication [4, 9, 10, 27].

Since the sequence identities between the six YMV isolates and all the other potyviruses were below 50%, we can conclude that YMV is a distinct potyvirus. The high sequence identity observed between YMV-BU1 and YMV-BU2 and between YMV-TOG and YMV-COT (95.5% and 97.4% respectively) suggests that those YMV isolates are closely related strains of YMV. The other sequence identities between the six YMV isolates ranged from 70.4% to 82.2%, within the range of variation seen for this region in other potyvirus strains [20]. These identities were calculated for the N-terminal half of the CP cistron which includes the most variable region [17, 18, 20, 26]. As shown in Fig. 3, most of the sequence differences are located in the first 67 aa domain of the protein which is known to differ markedly between individual viruses. The core region of the CP (residues 67–152), shows much fewer amino acid substitutions in the YMV isolates compared to those seen in comparison with PVY (Fig. 3). These observations lead us to suggest that the six YMV isolates should be considered as different strains of a single virus.

References

1. Atreya P, Atreya C, Pirone T (1991) Amino acid substitutions in the CP result in the loss of insect transmissibility of a plant virus. *Proc Natl Acad Sci USA* 88: 7887–7891
2. Chirgwin J, Przybyla A, MacDonald R, Rutter W (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* 18: 5294–5299
3. Colinet D, Kummert J (1993) Identification of a sweetpotato feathery mottle virus isolate from China (SPFMV-CH) by the polymerase chain reaction with degenerate primers. *J Virol Methods* 45: 149–159

4. Frenkel M, Jilka J, McKern N, Strike P, Clark Jr J, Shukla D, Ward C (1991) Unexpected sequence diversity in the amino-terminal ends of the coat protein of strains of sugarcane mosaic virus. *J Gen Virol* 72: 237–242
5. Higgins D, Sharp P (1989) Fast and sensitive multiple sequence alignments on a micro-computer. *CABIOS* 5: 151–153
6. Higgins D, Bleasly A, Fuchs R (1991) Clustal V: improved software for multiple sequence alignment. *CABIOS* 8: 189–191
7. Maiss E, Timpe U, Jelkmann W, Casper R (1989) Comparison of two different plum pox virus isolates on nucleic acid level. *Acta Horticult* 235: 313–317
8. Marchoux G (1980) Pathologie des ignames en Guadeloupe; maladies virales. In: Les colloques de l'I.N.R.A., l'Igname. *Seminaire International de Pointe-à-Pitre*, 28 July–02 August 1980, 93–105
9. Pappu S, Pappu H, Rybicki E, Niblett C (1994) Unusual amino-terminal sequence repeat characterizes the capsid protein of dasheen mosaic potyvirus. *J Gen Virol* 72: 239–242
10. Pappu S, Pappu H, Lastra R, Niblett C (1994) Variability in the length of the amino terminal sequence contributes to the capsid protein diversity among dashen mosaic potyvirus isolates. *Arch Virol* 136: 407–413
11. Pearson W, Lipman D (1988) Improved tools for biological sequence comparison. *Proc Natl Acad Sci USA* 85: 2444–2448
12. Reckhaus P, Nienhaus F (1981) Etiology of a virus disease of white yam *D. rotundata* in Togo. *J Plant Dis Prot* 88: 492–509
13. Robaglia C, Durand-Tardif M, Tronchet M, Boudazin G, Astier-Manificier S, Casse-Delbart F (1989) Nucleotide sequence of potato virus Y (N strain) genomic RNA. *J Gen Virol* 70: 935–947
14. Rybicki E, Shukla D (1992) Coat protein phylogeny and systematics of potyviruses. In: Barnett OW (ed) *Potyvirus taxonomy*. Springer, Wien New York, pp 283–297 (*Arch Virol* [Suppl] 5)
15. Sambrook J, Fritsch E, Maniatis T (1989) *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Press, New York
16. Shukla D, Thomas J, McKern N, Tracy S, Ward C (1988) Coat protein of potyviruses. 4. Comparison of biological properties, serological relationships, and coat protein amino acid sequences of four strains of potato virus Y. *Arch Virol* 102: 207–219
17. Shukla D, Ward C (1988) Amino acid sequence homology of coat proteins as a basis for identification and classification of the potyvirus group. *J Gen Virol* 69: 2703–2710
18. Shukla D, Ward C (1989) Structure of potyvirus coat proteins and its application in the taxonomy of the potyvirus group. *Adv Virus Res* 36: 273–314
19. Shukla D, Ward C (1989) Identification and classification of potyviruses on the basis of coat protein sequence data and serology. *Arch Virol* 106: 171–200
20. Shukla D, Ward C, Brunt A (1994) *The potyviridae*. CAB International, Wallingford
21. Terry E (1976) Incidence, symptomatology and transmission of a yam virus in Nigeria. *Proceedings of the Fourth Symposium of the International Society of Tropical Roots Crops*. Cali, Columbia, pp 170–173
22. Thouvenel J-C, Fauquet C (1977) Une mosaïque de l'igname (*Dioscorea cayenensis*) causée par un virus filamenteux en Côte d'Ivoire. *Comptes Rendus hebdomadaires des Seances de l'Académie des Sciences* 284: 1947–1949
23. Thouvenel J-C, Fauquet C (1979) Yam mosaic, a new potyvirus infecting *Dioscorea cayenensis* in the Ivory Coast. *Ann Appl Biol* 93: 279–283
24. Thouvenel J-C, Fauquet C (1986) Yam mosaic virus. *AAB Description of Plant Viruses*, No. 314

25. Ward C, Shukla D (1991) Taxonomy of potyviruses: current problems and some solutions. *Intervirology* 32: 269–296
26. Ward C, McKern N, Frenkel M, Shukla D (1992) Sequence data as the major criterion for potyvirus classification. *Arch Virol [Suppl]* 5: 283–297
27. Xiao X, Frenkel M, Teakle D, Ward C, Shukla D (1993) Sequence diversity in the surface-exposed amino-terminal region of the coat protein of seven strains of sugarcane mosaic virus correlates with their host range. *Arch Virol* 132: 399–408
28. Yu M, Frenkel M, McKern N, Shukla D, Strike P, Ward C (1989) Coat protein of potyviruses. 6. Amino acid sequences suggest watermelon mosaic virus 2 and soybean mosaic virus-N are strains of the same potyvirus. *Arch Virol* 105: 55–64

Authors' address: Dr. O. Duterme, Faculté Universitaire des Sciences Agronomiques, Laboratoire de Pathologie Végétale, Passage des, Déportés 2, B-5030 Gembloux, Belgium.

Received August 8, 1995