Arch Virol (1994) 134: 271–278

Archives of Virology

Printed in Austria

The coat protein of Indian peanut clump virus: relationships with other furoviruses and with barley stripe mosaic virus

S. V. Wesley[1,*], M. A. Mayo[1], C. A. Jolly[1], R. A. Naidu[2], D. V. R. Reddy[2], M. K. Jana[3], and V. K. Parnaik[4]

[1] Scottish Crop Research Institute, Invergowrie, Dundee, U.K.
[2] ICRISAT, Patancheru, Andhra Pradesh,
[3] Indian Institute of Technology, Kharagpur, West Bengal, and
[4] Center for Cellular and Molecular Biology, Hyderabad, Andhra Pradesh, India

Accepted September 28, 1993

Summary. The 5'-most open reading frame of the c.4kb RNA-2 of Indian peanut clump furovirus (IPCV) encodes a protein of 208 amino acids. This protein is thought to be the coat protein of IPCV because its amino acid composition and M_r closely resemble those reported for IPCV coat protein and because its amino acid sequence is 61% identical to that of the coat protein of peanut clump virus (PCV) from West Africa. The extent of the sequence identity between IPCV and PCV coat proteins confirms previous conclusions that the viruses are distinct rather than strains of one virus. The sequences of the coat proteins of IPCV and PCV were between 18% and 26% identical to those of other furoviruses and those of unrelated tobamoviruses and tobraviruses. In contrast, the coat protein sequences were 37% (IPCV) and 36% (PCV) identical to that of the coat protein of barley stripe mosaic hordeivirus (BSMV). This similarity between the coat proteins of viruses from different groups (= genera) is unusual but is consistent with previous reports of sequence relatedness in various genes between certain furoviruses and BSMV.

Introduction

Peanut clump is a soil-borne disease of groundnut crops [26]. In West Africa, the disease is caused by infection with peanut clump virus (PCV) which is transmitted by the plasmodiophoromycete fungus *Polymyxa graminis* [27]. PCV particles are rod-shaped and of two lengths and PCV has been classified with other fungus-transmitted bi-partite genome viruses with rod-shaped

* Present address: ICRISAT, Patancheru, Andhra Pradesh, India.

particles in the furovirus group [3]. In India, a similar but serologically distinct virus, named Indian peanut clump virus (IPCV), induced clump disease [20]. No serological relation between PCV and IPCV could be detected by immunosorbent electron microscopy [18, 20, 21], by immunoblotting [15] or by ELISA [18]. The results of nucleic acid hybridization tests suggested between 23% and 41% sequence homology between any of three strains of IPCV and PCV [21]. Some authors have considered IPCV and PCV to be strains of one virus [4] but others [21, 22] have referred to them as separate viruses.

IPCV is thought to have a bi-partite genome consisting of a c. 5.4kb and a c. 4kb species [21]. Smaller RNA species were detected in preparations from purified virus particles [16] but their origins were not established. Translation in vitro of IPCV RNA-2 yielded coat protein and little else [16]. This suggested that the coat protein gene was near the 5'-end of the RNA, which is the position in the genome of the coat protein genes of other furoviruses [4]. In this paper we report the nucleotide sequence of the part of RNA-2 which encodes the coat protein of IPCV and comparisons among the amino acid sequences of coat proteins of other viruses with rod-shaped particles including the recently determined sequence of PCV coat protein [14].

Materials and methods

Propagation and purification of IPCV

The Hyderabad isolate of IPCV [18] was propagated in *Nicotiana clevelandii* and purified as described by Reddy et al. [21].

Preparation and cloning of IPCV RNA

RNA was extracted from purified virus particles as described by Mayo and Reddy [16]. It was annealed with random composition hexa-deoxynucleotides and reverse transcribed [11], and the resulting cDNA was cloned in *Sma*I-cut pUC19 [17]. Cloned cDNA was used as a probe in Northern blotting [23] of RNA extracted from IPCV particles in order to assign the cDNA clones to RNA-1 or RNA-2.

Nucleotide sequence determination

Sequence was determined by using dideoxy chain termination as described by Mayo et al. [17]. Sequence at the 5'-end of the RNA was obtained by primer extension as described by Fichot and Gerard [7]. The oligonucleotide used was complementary to the sequence between nucleotides 38 and 53 (Fig. 1). The sequence was assembled as described by Mayo et al. [17].

Amino acid sequence analysis

Sequences were compared by using programs in the GCG package [5] and CLUSTALV [13]. The significance of the clustering detected by CLUSTALV was assessed by bootstrapping 1000 samples. Amino acid sequences for comparisons were of the coat proteins of beet necrotic yellow vein furovirus (BNYVV) [1], barley stripe mosaic hordeivirus (BSMV) [12], Nicotiana velutina mosaic furovirus (NVMV) [19], peanut clump furovirus (PCV) [14], soil-borne wheat mosaic furovirus (SBWMV) [24] and tobacco mosaic virus (TMV) [28].

Results

Figure 1 shows the nucleotide sequence of the 5'-terminal 1128 residues of RNA-2. The sequence has been submitted to the EMBL database under Accession Number 76658. The non-coding leader is 501 nucleotides long and is followed by an open reading frame (ORF) which is the only potential ORF in the sequence. The encoded protein was identified as the coat protein of IPCV because the M_r of 23017 is close to the value of 24500 estimated by gel electrophoresis [16] and because the amino acid composition was a significant match (Chi squared = 18.2, with 15 degrees of freedom) to that reported for an Indian isolate of PCV (of unknown serotype) [6]. The only pronounced mis-match was between the 23 serine residues deduced from the sequence (Fig. 1) and the 12 found experimentally [6]. The putative sequence of IPCV coat protein also had strong homologies with the sequence of the coat protein of PCV [14].

Table 1 shows the result of comparisons made by using GAP among the sequences of the coat proteins of IPCV, PCV, SBWMV, BNYVV, NVMV, BSMV and TMV. The only marked similarities were 61% identity between IPCV and PCV coat proteins, and 37% and 36% identity between BSMV coat protein and those of IPCV and PCV, respectively. Otherwise there were no greater similarities among furovirus coat proteins than any had with TMV coat protein. Similarities between furovirus coat proteins and those of other

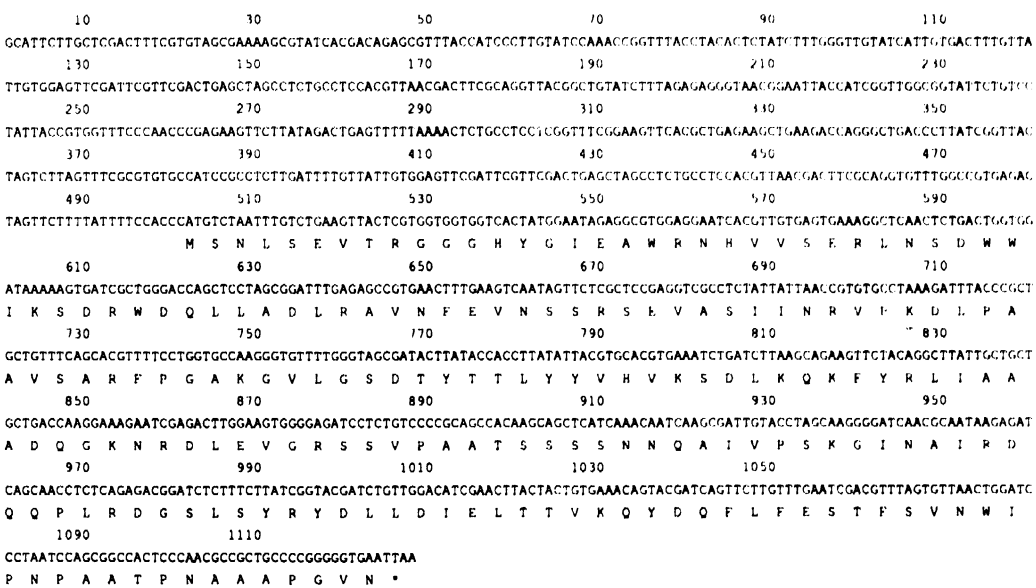

Fig. 1. Nucleotide sequence of the 5' 1128 nucleotides of IPCV RNA-2. The sequence is shown with that of the encoded coat protein. * signifies the termination codon

Table 1. Results of pairwise GAP comparisons between coat proteins

	IPCV	PCV	SBWMV	BNYVV	NVMV	BSMV	TMV
				% identity			
IPCV		61	18	18	23	37	22
PCV	80		18	18	26	36	23
SBWMV	37	37		21	19	20	24
BNYVV	44	38	46		22	15	17
NVMV	41	40	40	42		22	16
BSMV	56	59	37	38	44		20
TMV	39	46	44	38	34	41	
				% similarity			

Values shown are the percent of identical residues (% identity) or the percent of matches of chemically similar amino acids (% similarity) obtained when comparing coat proteins of the viruses indicated. Default values of gap weight = 3 and length weight = 0.1 were used to obtain the alignments

tobamoviruses or tobacco rattle and pepper ringspot tobraviruses were essentially the same as those obtained with TMV coat protein (data not shown).

The strong similarities among coat proteins of IPCV, PCV and BSMV (Table 1) were also obvious when the sequences were compared by using the NJTREE option in CLUSTALV. Figure 2 shows the 'unrooted tree' obtained by comparing coat proteins of IPCV, PCV, BSMV, SBWMV, NVMV and

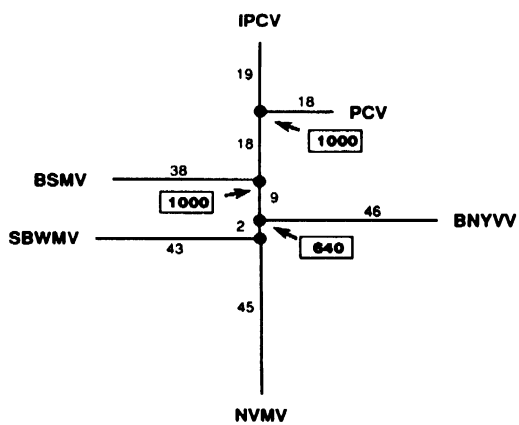

Fig. 2. An unrooted tree generated by CLUSTALV illustrating the similarity distances among the coat proteins of *IPCV* (Indian peanut clump furovirus), *PCV* (peanut clump virus), *BSMV* (barley stripe mosaic hordeivirus), *SBWMV* (soil-borne wheat mosaic furovirus), *NVMV* (Nicotiana velutina mosaic furovirus), and *BNYVV* (beet necrotic yellow vein furovirus). Unboxed numbers indicate the distances between the nodes; the boxed numbers are the numbers out of a bootstrapped sample of 1000 which gave the same branching pattern as that shown

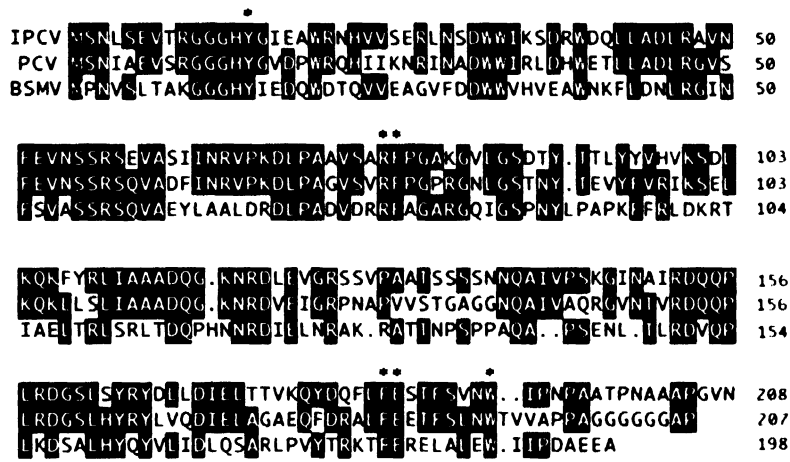

Fig. 3. Alignment of the coat proteins sequences of *IPCV* (Indian peanut clump furovirus), *PCV* (peanut clump virus), and *BSMV* (barley stripe mosaic hordeivirus). Amino acid matches are shown in reverse contrast. Numbering excludes gaps inserted by CLUSTAL V. * indicates the residues common to coat proteins from viruses with rod-shaped particles and mentioned in the text

BNYVV. The only clustering detected was of IPCV with PCV and to a lesser extent of IPCV and PCV with BSMV; these clusters were judged significant because they occurred in all the bootstrapped samples (Fig. 2). 'Trees' which included tobamovirus and tobravirus coat proteins showed essentially the same relationship among furovirus sequences as that shown in Fig. 2. Except for IPCV and PCV, the different furoviruses were as dissimilar from each other as they were from tobamoviruses or tobraviruses.

Figure 3 illustrates the alignment obtained between the coat proteins of IPCV, PCV and BSMV and the extent and location of the amino acids identical in two of the three proteins. Some of the residues (asterisked in Fig. 3) which were the same in all three sequences (Y 14, R 78, F 79, F 186, E 187, W 194) correspond to those identified [10] as being present in analogous positions in the coat proteins of both tobraviruses and tobamoviruses (Fig. 3). Most of these residues are also present in the coat protein sequences of other furoviruses.

Although the sequence alignments do not suggest any feature distinguishing furovirus coat proteins from others, Fauquet et al. [6] were able, by a multi-dimensional classification of the amino acid compositions of plant virus coat proteins, to detect a clustering of furoviruses, and a similarity of the whole cluster with BSMV.

Discussion

Some authors have regarded IPCV and PCV as being strains of one virus [4] whereas others have considered the viruses to be separate on the grounds of a lack of serological relationship and only limited nucleotide sequence homo-

logy [20, 21]. An additional complication is that some isolates of IPCV from different regions in India are relatively distantly related serologically [18, 21] and have only 50% to 60% apparent nucleotide sequence homology [21]. The value of 61% sequence identity between the coat proteins of IPCV and PCV supports the separation of these viruses into different species. A similar value of sequence identity between potyvirus coat proteins was proposed by Shukla and Ward [25] to indicate that two viruses be considered as distinct rather than strains of one virus; values of >90% were considered to indicate that two isolates were strains of the same virus. Similarly, the coat proteins of different tobamoviruses are between 28% and 83% identical in pairwise comparisons [9].

Although the furoviruses compared in our tests have been assigned to the furovirus group because of biological and physical properties [3], they differ in their coat protein sequences as much from each other as they do from viruses in other taxonomic groups. In contrast, IPCV and PCV are relatively closely related. However, a surprising result of our sequence comparisons was that both IPCV and PCV coat proteins were shown to have striking homologies with BSMV coat protein. There have been reports of sequence similarities between non-structural proteins of BSMV and those of the furoviruses BNYVV [2], SBWMV [24] and PCV [14] but this is the first report of a sequence relationship between the coat proteins of a furovirus and BSMV. It has been suggested that the ancestors of furoviruses and hordeiviruses may have undergone recombination to exchange genes or parts of genes [24]. Our results add further emphasis to this suggestion and point to the difficulties of basing a classification solely on the sequences of the genes of viruses.

By analogy with the sequence homologies among tobamovirus coat proteins, the value of 37% homology between IPCV and BSMV would suggest that the two viruses belong to the same genus. However, there are many distinguishing features that justify classification of the viruses into different genera [8]. Nevertheless, the results do suggest that any taxonomic clustering of genera of rod-shaped viruses should place furoviruses and hordeiviruses in close proximity.

Acknowledgements

We thank Drs. G. Jonard and K. Richards and their colleagues for helpful discussions and for sharing their results prior to publication. This work was funded in large part by the Overseas Development Administration, Plant Sciences Programme; submitted as Journal Article No. JA 1515 by ICRISAT.

References

1. Bouzoubaa S, Ziegler V, Beck D, Guilley H, Richards K, Jonard G (1986) Nucleotide sequence of beet necrotic yellow vein virus RNA-2. J Gen Virol 67: 1689–1700
2. Bouzoubaa S, Quillet L, Guilley H, Jonard G, Richards KE (1987) Nucleotide sequence of beet necrotic yellow vein virus RNA-1. J Gen Virol 68: 615–626
3. Brunt AA (1991) Furovirus group. In: Francki RIB, Fauquet CM, Knudson DL, Brown F (eds) Classification and Nomenclature of Viruses. Vth Report of the International

Committee on Taxonomy of Viruses. Springer, Wien New York, pp 377–379 (Arch Virol [Suppl]2)
4. Brunt AA, Richards KE (1989) Biology and molecular biology of furoviruses. Adv Virus Res 36: 1–32
5. Devereux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 12: 387–395
6. Fauquet C, Desbois D, Fargette D, Vidal G (1988) Classification of furoviruses based on the amino acid composition of their coat proteins. In: Cooper JI, Asher MJC (eds) Viruses with fungal vectors. Association of Applied Biologists, Wellesbourne, pp 19–37
7. Fichot O, Gerard M (1990) An improved method for sequencing of RNA templates. Nucleic Acids Res 18: 6162
8. Francki RIB, Fauquet CM, Knudson DL, Brown F (1991) Classification and Nomenclature of Viruses. Vth Report of the International Committee on Taxonomy of Viruses. Springer, Wien New York (Arch Virol [Suppl 2])
9. Gibbs AJ (1986) Tobamovirus classification. In: Van Regenmortel MHV, Fraenkel-Conrat H (eds) The plant viruses, vol 2: The rod-shaped viruses. Academic Press, New York, pp 167–180
10. Goulden MG, Davies JW, Wood KR, Lomonossoff GP (1992) Structure of tobraviral particles: a model suggested from sequence conservation in tobraviral and tobamoviral coat proteins. J Mol Biol 227: 1–8
11. Gubler U, Hoffman BJ (1983) A simple and very efficient method for generating cDNA libraries. Gene 25: 263–269
12. Gustafson G, Armour SL (1986) The complete nucleotide sequence of RNA ß from the type strain of barley stripe mosaic virus. Nucleic Acids Res 14: 3895–3909
13. Higgins DG, Bleasby AJ, Fuchs R (1992) CLUSTALV: improved software for multiple sequence alignment. CABIOS 8: 189–191
14. Manohar SK, Guilley H, Dollet M, Richards K, Jonard G (1993) Nucleotide sequence and genetic organization of peanut clump virus RNA 2 and partial characterization of deleted forms. Virology 195: 33–41
15. Mayo MA, Rajeshwari R (1987) Detection of relatedness among strains of peanut clump virus (PCV) by immunoblotting. Annual Report of Scottish Crop Research Institute for 1986, Invergowrie, pp 184–185
16. Mayo MA, Reddy DVR (1985) Translation products of RNA from Indian peanut clump virus. J Gen Virol 66: 1347–1351
17. Mayo MA, Robinson DJ, Jolly CA, Hyman L (1989) Nucleotide sequence of potato leafroll luteovirus RNA. J Gen Virol 70: 1037–1051
18. Nolt BL, Rajeshwari R, Reddy DVR, Bharathan N, Manohar SK (1988) Indian peanut clump virus isolates: Host range, symptomatology, serological relationships, and some physical properties. Phytopathology 78: 310–313
19. Randles JW, Rohde W (1990) Nicotiana velutina mosaic virus: evidence for a bipartite genome comprising 3-kb and 8-kb RNAs. J Gen Virol 71: 1019–1027
20. Reddy DVR, Rajeshwari R, Iizuka W, Lesemann DE, Nolt BL, Goto T (1983) The occurrence of Indian peanut clump, a soil-borne virus disease of groundnuts (*Arachis hypogea*) in India. Ann Appl Biol 102: 305–310
21. Reddy DVR, Robinson DJ, Roberts IM, Harrison BD (1985) Genome properties and relationships of Indian peanut clump virus. J Gen Virol 66: 2011–2016
22. Reddy DVR, Nolt BL, Hobbs HA, Reddy AS, Rajeshwari R, Rao AS, Reddy DDR, McDonald D (1988) Clump disease in India: isolates, host range, transmission and management. In: Cooper JI, Asher MJC (eds) Viruses with fungal vectors. Association of Applied Biologists, Wellesbourne, pp 239–246

23. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
24. Shirako Y, Wilson TMA (1993) Complete nucleotide sequence and organization of the bipartite RNA genome of soil-borne wheat mosaic virus. Virology 195: 16–32
25. Shukla DD, Ward CW (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
26. Thouvenal JC, Fauquet C (1981a) Further properties of peanut clump virus and studies on its natural transmission. Ann Appl Biol 97: 99–107
27. Thouvenal JC, Fauquet C (1981b) Peanut clump virus. CMI/AAB Descriptions of Plant Viruses, no. 235
28. Zaitlin M, Israel HW (1975) Tobacco mosaic virus. CMI/AAB Descriptions of Plant Viruses, no. 151

Authors' address: Dr. M. A. Mayo, Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA, U.K.

Received June 18, 1993