

The nucleotide sequence of RNA-1 of Indian peanut clump furovirus*

J. S. Miller¹, S. V. Wesley², R. A. Naidu², D. V. R. Reddy², and M. A. Mayo¹

¹ Scottish Crop Research Institute, Invergowrie, Dundee, U.K.

² International Crop Research Institute for the Semi-Arid Tropics, Patancheru, Andhra Pradesh, India

Accepted July 17, 1996

Summary. The nucleotide sequence of RNA-1 of an isolate of the H serotype of Indian peanut clump virus (IPCV-H) was shown to comprise 5 841 nucleotides. The RNA contains three open reading frames (ORF) which are between nucleotides 133 and 3 522, nucleotides 3 526 and 5 103 (assuming expression by suppression of the ORF 1 termination codon) and nucleotides 5 168 and 5 539. The encoded polypeptides have M_r of 129 687 (p130), 60 188 (p60) and 14 281 (p14). ORF 2 is thought to be expressed by suppression of the termination codon of ORF 1 to produce a M_r 189 975 product (p190). p130 contains sequences characteristic of proteins with methyl transferase and NTP-binding properties and p190 contains these and sequences characteristic of RNA-dependent RNA polymerases. The nucleotide sequence of IPCV RNA-1 is similar to that of peanut clump virus (PCV) and corresponding encoded polypeptides are 88% (p130), 95% p60 and 75% (p14) identical. The sequences of the translation products are also similar to those of soil-borne wheat mosaic virus and barley stripe mosaic virus. Oligonucleotide primers, designed on the basis of the sequences of RNA-1 of IPCV and PCV, were effective in reverse transcription-PCR amplification of these RNAs and that of IPCV isolates of the serologically distinct L and T serotypes.

Introduction

Indian peanut clump virus (IPCV) is widespread in India where it induces clump disease in groundnut crops [15, 17]. The virus is thought to be transmitted by the soil-inhabiting fungus *Polymyxa graminis* which can persist in the soil for

*The nucleotide sequence data reported in this paper has been deposited at EMBL under accession number X99149.

several years [17]. Clump disease can be economically important in several regions [15].

IPCV has a bi-partite genome of positive-sense RNA molecules [16]. RNA-1 is about 6 kb and RNA-2 is about 4 kb, and each is encapsidated in a rod-shaped particle [16]. In vitro translation showed that the coat protein was encoded by RNA-2 [11] and nucleotide sequencing confirmed this [21]. The other principal translation products of IPCV RNA were those of RNA-1 and had M_r of about 143 000 and 195 000 [11].

IPCV resembles peanut clump virus (PCV) [15], which induces clump disease in groundnut crops in West Africa. Biological and physical properties of PCV has led to it being classified in the genus *Furovirus* [1].

Isolates of IPCV fall into one of three distinct serotypes, Hyderabad (H), Ludhiana (L) or Talod (T) [14]. The serotypes are found in geographically separated areas of India [14, 16], although isolates of the T serotype have been found at several locations [14]. The lack of serological cross-reaction between serotypes is a potential complication for serological diagnosis of infection and this, and the lack of genes for resistance to IPCV in groundnut germplasm [17], have prompted work on the molecular characterisation of the virus [13, 21].

In previous work, the coat protein gene of IPCV was identified as the 5'-proximal gene in RNA-2 [21]. In the work reported here, we have cloned and sequenced RNA-1 of IPCV-H and have located the three open reading frames it contains. The results show that RNA-1 of IPCV and of PCV are closely similar and also show, as demonstrated for PCV RNA-1 [7], that some proteins encoded by IPCV RNA-1 resemble those of both the type member of the genus *Furovirus*, soil-borne wheat mosaic virus (SBWMV), and the type member of the genus *Hordeivirus*, barely stripe mosaic virus (BSMV).

In other work [22], it was shown that hybridisation assays using a probe made from cDNA to the 3'-most 742 nucleotides of IPCV RNA-1 were able to detect RNA from isolates of any of the three serotypes of IPCV and PCV. From an alignment of the sequence of IPCV RNA-1 established in the work reported here with that of PCV RNA-1, pairs of oligonucleotide primers were selected for use in reverse transcription/polymerase chain reaction assays (RT-PCR) to detect isolates of IPCV-H and PCV. The primers were shown also to be applicable to the detection of isolates of IPCV serotypes L and T.

Materials and methods

Virus isolates and propagation

The Hyderabad isolate (H), Ludhiana isolate (L) and the D isolate of IPCV (D), which belong to the H, L and T serotypes [14], respectively, were propagated in *Phaseolus vulgaris* and purified as described by Reddy et al. [16].

cDNA synthesis and cloning

RNA was extracted from purified virus particles as described by Mayo and Reddy [11]. cDNA was synthesized as described by Gubler and Hoffman [6] using a commercial kit

(Boehringer) and cloned in *Sma* I-digested pUC19. Clones specific to IPCV RNA-1 were identified by Northern blotting.

The 5'- and 3'-extremities of IPCV RNA-1 were cloned using the method of RACE-PCR [5] utilising the 5'-Amplifinder RACE kit (Clontech) as described by the manufacturers. To obtain cDNA to the 3'-end of IPCV RNA-1, the 5'-phosphorylated and 3'-amino blocked anchor oligonucleotide (5'-CACGAATTCATCGATTCTGGAACCTTCAGAGG) was ligated directly to IPCV single-stranded RNA using T4 RNA ligase. First strand cDNA was made using as a primer 5'CCTCTGAAGGTTCCAGAATCGATAG (AP1) which was complementary to the anchor. A DNA fragment corresponding to the 3'-end of the RNA-1 genome was then amplified using AP1 and 5'-AGTGGTAGACTATTGACC (which corresponds to nucleotides 5 099 to 5 116 of IPCV RNA-1). To obtain cDNA to the 5'-end of the RNA-1, cDNA was synthesised using a primer complementary to nucleotides 434 to 448 (#137, 5'AAGTCTCCAACGTTC). The RACE procedure was then followed and the PCR was done using primer AP1 and a second, nested primer which was complementary to nucleotides 49 to 68 (5'-GTTGCAGGAATTGCGAACCT).

All PCR fragments were gel purified and cloned into the PCR cloning vector pGEM-T (Promega). Insert size was determined following digestion with *Pvu* II.

Nucleotide sequencing

Nucleotide sequences were determined by dideoxy chain termination [18]. Template DNA was either single-stranded M13 DNA or double-stranded DNA from cDNA clones [24]. All sequencing was done using the Sequenase enzyme (Amersham). The genome was sequenced in its entirety in both orientations, at least twice on independent clones.

Nucleotide sequence was assembled as described by Mayo et al. [12] and sequences were compared using programs in the GCG package [4].

RT-PCR

Primers were designed to be used in reverse transcription reactions and in subsequent PCR. Sequences were chosen which with minimum degeneracy matched the sequences of IPCV RNA-1 and PCV RNA-1 [7]. The primers were J1: 5'-GTATTCTGTGGTGTGGTTT-3' (nucleotide positions 1–20), J2: 5'-ATAAACCTCCCCAAAGGAGT-3' (nucleotide positions 1 933–1 952), J3: 5'-AT/CGACATGAAAGATGAAGTG-3' (nucleotide positions 1 769–1 788), J4: 5'-AGCCAAAATAGACTCCGCAA-3' (nucleotide positions 3 809–3 828), J5: 5'-GGTTTGGGTCAAAGTTTGAG-3' (nucleotide positions 3 527–3 546), J6: 5'-GTCAATAGTC/TTACCACTCAT-3' (nucleotide positions 5 096–5 115), J7: 5'-TGTT-TGCCAATGGAA/GAAAGC-3' (nucleotide positions 4 558–4 577), J8: 5'-TGGGACGG-ATATCGCTCCGT-3' (nucleotide positions 5 822–5 841).

To prime first strand cDNA synthesis, 100 ng of each downstream primer (J2, J4, J6 and J8) was added to 0.5 µg of RNA extracted from purified virus particles or to 5 µg of total RNA extracted from the leaf tissue of IPCV-infected *Nicotiana benthamiana*. The reaction components were 1 × PCR buffer, 4 mM MgCl₂, 1 mM dNTPs, 0.1 mg/ml BSA, 0.005 mM DTT, 20U RNaseIn (Boehringer) and 10 U/ml AMV reverse transcriptase (Boehringer). The mixture was incubated at 42 °C for 1 hour and then heated at 95 °C for 2 min to inactivate the reverse transcriptase. One to 5 µl of the reverse transcriptase products was then used in a PCR reaction, consisting of 1 × PCR buffer, 1.5 mM MgCl₂, 0.2 mM dNTPs, 1.5 µg/ml downstream primer, 3 µg/ml upstream primer and 1-2 U *Taq* polymerase (Amplitaq, Perkin Elmer). Cycling times and temperatures were 94 °C for 5 min (1 cycle, during which the polymerase was added), 94 °C for 1 min, 55 °C for 1 min, 72 °C for 2 min (30

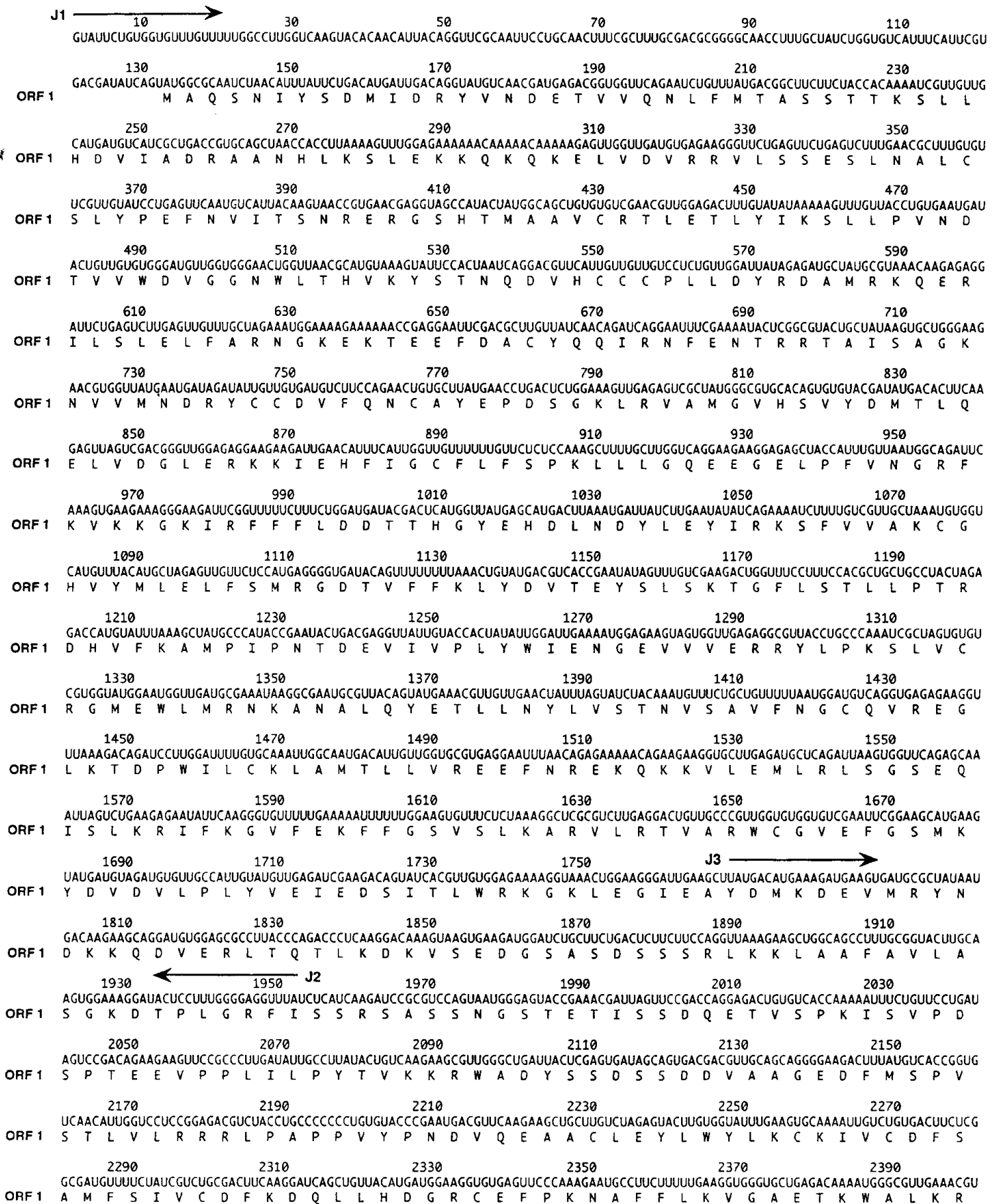


Fig. 1 (continued)

2410 2430 2450 2470 2490 2510
 ORF 1 CCCACUUCACAGCAGGUCGGACCAUAUCUGUGAAAAUUUCUGAGAUGAUGACAUUGGAGUUGACCCCGUUGUUGGAAAAAGCAGAACGAUGAAGUGCGUAGUUAUUUCCCU
 P T S Q Q V G H Q Y C V K F S E N D D H M E L T P V S W K K Q N D E V R S I F P
 2530 2550 2570 2590 2610 2630
 ORF 1 CAAGGAUUGAGCGAUGGGUGGUACAUUUUCUGACUUGACUUUUUGAUGAUGAGUGGUUAUUUUUACAAGUUGGUUGCUAUGUACCUUACUUCUAAAAAACAACUAAAAAGUU
 Q G L S D G W Y M F S D L T F L M N E W L I F N K L V A M Y P T L Q K N Q L K V
 2650 2670 2690 2710 2730 2750
 ORF 1 AGGUUGAUGAUGGAGUACCGUGGUUGGAAAAUCUACCGUUAUUUGAUAUUUGAUGAUGAUGAAGCAGGUUGUGCUGCGUAGGAAGAAGCGACCGAUGAUGAAGAGAGA
 R L I D G V P G C G K S T W I L N N C D L D K Q V V L A E G R E A T D D L R K R
 2770 2790 2810 2830 2850 2870
 ORF 1 UUCACUGAAAAGGGUUUCUAGAAAGCGUUGUGAGGAGGGUGAGAACUGGACUCCUUUUGCUUAAAACCUUGACGCGUGGUUUCAAUUUUUACUUUGAUGAAGCUCUACG
 F T E K G F P R K R C E E R V R T V H S F M L K P L T R G F N S F H F D E A L M
 2890 2910 2930 2950 2970 2990
 ORF 1 GCCAUGCUGUUAUGAUUUUAUUUGCGGGCGUAUGUUGAGAGCUAGGGAAGUGAUCUGCAGGGUGAUUCUAAAACAGAUCCUUUUUAUCAUACCGGUUGAACAGAUUACCUUAAGGUAU
 A H A G M I Y I C G R M L R A R E V I C Q G D S K Q I P F I N R V E Q I T L R Y
 3010 3030 3050 3070 3090 3110
 ORF 1 GCUUCGUUUAAUGAUGAAGGAGUAUGUAGAAAGACUACAGGUGCCAUUGGAUGUCAUCUUAUUAUAAAUAAGAGAUUAUCAUACAGGUGAUGAUGAUGGUGGUUCUCA
 A S F N V V E R E Y V R K T Y R C P L D V I Y Y L N K K R Y Y Q G D D I V G F S
 3130 3150 3170 3190 3210 3230
 ORF 1 AAGACAACAUUCUGUUAUCAAAAGUCAAAAACUUCUGGAUUCACCUUUUGGUGAAGCUUCCGAAAGAACUGUGCAUUUUUGACAUUCCUGCAAGCUGAGAAAGAGGAAGUUGCU
 K T T H S V D T K S K T S G F T S L V K L P K E P V H Y L T F L Q A E K E E V A
 3250 3270 3290 3310 3330 3350
 ORF 1 AAACAUUUGCGGGAGUGAAAGGUGCUACGGUGUCGACUUAUCUAGAGGCGCAGGGGAAGACGUUUGAUGCGUAAAUUCUGGUCAGGUGAAAAUGACAGACAAGUAGUUUAUUCUGGA
 K H L A G V K G A T V S T I H E A Q G K T F E C V N L V R L K M T D N E L Y P G
 3370 3390 3410 3430 3450 3470
 ORF 1 GGGCAAAGGCAGGCCUACACAAUUGUUGUUGACAAGGCAUACCAAGUAGUUCUUAUUCUUGUUGAAGACAGCGUUUAUGAGGUAUUCUCUGCGUUGAAGGUGAUGUG
 G A K A E P Y T I V G L T R H T R S L V Y Y S V V E D R L Y E D I S A L K D V M
 3490 3510 3570 3590
 ORF 1/2 GAGGACCAGUUGCUGAAGUGCUCUCACUCUGAGCAGACCAAAGUAGCGGUUUUGGUCAAAGUUUGAGUCCAUUCGUGGUAGCAGAUAGAAAGGUCAUGGCCCUUGAUGUUGGGGAUCUUGG
 E D Q L L K C S H S E Q T K * R F G S K F E S I V V A D R K V M A P D V G D L V
 3610 3630 3650 3670 3690 3710
 ORF 2 ACCAUUCAAGACUUGUACGAUCGGACUUUCCUGGUAAUUCACUUGGAUUAACGUUUUGACGGGUAUACUGUUGCUUUAUUAACCUAGAGUUAAGAGAUUUUAACUGUAGAAUUGCA
 T I Q D L Y D R T F P G N S T L D S S F D G Y T V A S S N L E L E I S N C K I A
 3730 3750 3770 3790 3850 3870 3890 3910 3930 3950
 ORF 2 CCUAACAAGUCUAUUAGAGGGUUCCAAGAAAGAUUCGUUUUCUCCGAGAUUGAGGACCGCUAUGCCUGAGAAGCGUCAAGGUUCCUUUGCGGAGUCUAUUUUGGCUCUGCGUAAGAGA
 P N K S I R G F Q E K I C F S P R L R T A M P E K R Q G S F A E S I L A L R K R
 3970 3990 4010 4030 4050 4070
 ORF 2 ACAGAUGAGCGUCUCUUAAGGUGGUGGAAAAACAUAACCAACCGGCAAAAAUAUAAUGCUCGUGAUUUGCGUUCUUGGAUCAAAUUGAUGUGGCACGUAACAUCUUAUGAUCAG
 T D E A A L R W W E K Q S T T A K N Q M L A D W R S L D Q I D V C T Y N F M I K
 4090 4110 4130 4150 4170 4190
 ORF 2 AAUGACGUUAAACCGAAGUUGGAUUUGACACCCCAAAGUGAAUUAUGCAGCUCUCAAACUGUUGCUAUAUCCUGAAAAUAAGUAAACCGUUGUUGGGCCGAUUAUAGGAAAUCAAU
 N D V K P K L D L T P Q S E Y A A L Q T V V Y P E K I V N G L F G P I I K E I N
 4210 4230 4250 4270 4290 4310
 ORF 2 GAACGGGUGUUGCAGCUCUUAAGCCAAAUGUGUUUGUUAACACCGCGGAUGACUUCUGAAGAGUUGAGUCGACAGCCGAGUAUCUUUUUCCAGGUGACGAUUCGAGGUUGUUGAAUU
 E R V L S A L R P N V F V N T R M T S E E L S R T A E Y L F P G D E F E V V E I
 4330 4350 4370 4390 4410 4430
 ORF 2 GAUUUUCCAAGUAUGACAAGUCAAAAACUUAUUGCACAUCAGAAGUUUAUAGACUUUACGAGCAGUUUGGUUUAUUGUUAUUAUGAAGUAUCUUAUGGAAAAUACAAAACUCAG
 D F S K Y D K S K T S L H I R M V I R L Y E Q F G L N G Y M K Y L W E K S Q T Q
 4450 4470 4490 4510 4530 4550 J7 -----
 ORF 2 ACAGUCGUCAAAGAUAGAAACUUAUGCGGUUGAAGCUUACUUCUUGUACCAAGCAGAAAAUCAGGCAACUGUGACACAUAGGUUUAACACAUUUUUUUAUGUUUGCAUUGCUGUAGCUGU
 T V V K D R N Y G V E A Y I L Y Q Q K S G N C D T Y G S N T Y S S M F A L L D C
 4590 4610 4630 4650 4670
 ORF 2 UUGCCAAUGGAAAAAGCUGUAUCUCUUAUUGGUGGUGAUGACUUAUUUAUUUUUCAAAGGGGACUGUUGUUAUAGCCUUGUGGUCGUUUAAGCUUCUCUGUGGAAUUUUGAC
 L P M E K A V Y S I F G G D D S L I L F P K G T V V N D P C G R L A S L W N F D
 4690 4710 4730 4750 4770 4790
 ORF 2 UGUAAUCAAUGAAUUUAAAGUUCGGCUUUUUGGUAAGUUAUGAUAUUCAGUUGCGCGGAGUACCGUUUUAAGCCGACCUUUGAAGUUAUAAAGCUGGUAACAAAGACA
 C K S M K F K V P A F C G K F M I P V A G R Y R F E P D P L K L I T K L G N K T

Fig. 1 (continued)

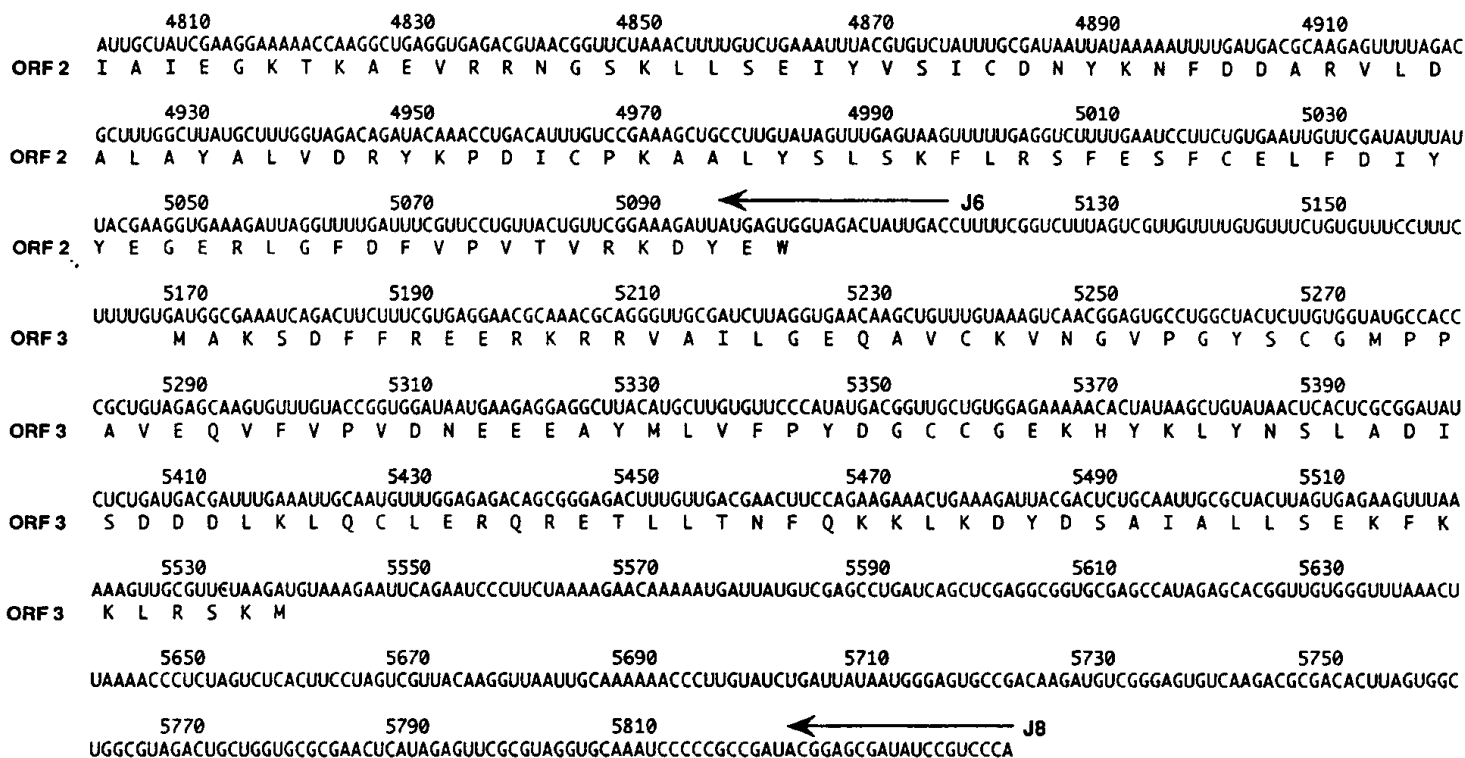


Fig. 1. Nucleotide sequence of IPCV RNA-1. The positions of the primers J1 to J8 are indicated by arrows which show the direction of the primed synthesis. Open reading frames are indicated as ORF to the left. * indicates termination codons

cycles) and 72 °C for 5 min (1 cycle). Products were analysed by electrophoresis in a 10% acrylamide gel.

Results

Nucleotide sequence and genome organization

The nucleotide sequence (EMBL database accession number X99149) was determined from cloned cDNA except for 15 and 802 nucleotides at the 5'- and 3'-ends respectively. These terminal sequences were obtained using 5'- and 3'-RACE procedures. Figure 1 shows the sequence of the 5841 nucleotides (nts) together with the location of the three open reading frames detected. ORF 1 is from nt 133 to UGA at nts 3523 to 3525. The next ORF, if starting with an AUG codon, would be between nts 3575 and 5103. However, in vitro translation of IPCV RNA-1 yielded major products with M_r of about 143 000 (the major product) and about 195 000 [11], which would suggest that the 143 000 product is the ORF 1 product and that the 193 000 product arises by suppression of the ORF 1 termination codon. Similar UGA codon suppression is known to occur during translation of tobacco rattle virus (TRV) RNA-1 [23] and, as for RNA-1 of PCV [7] which is similar to IPCV RNA-1 (see below), the codon context of the UGA terminating ORF 1 is very similar to that around UGA codons in TRV RNA-1, pea early browning virus RNA-1 [8] and SBWMV RNA-1 [19] which are also thought to be suppressed during translation. Therefore, ORF 2 is taken

Table 1. Percentage identity between nucleotide sequence of parts of RNA-1 of IPCV and PCV and between the IPCV RNA-encoded proteins and corresponding proteins of PCV, SBWMV or BSMV

Region ^a	Length (nts)	%Identity with ^b			
		RNA-1 of PCV	corresponding polypeptides of		
			PCV	SBWMV ^c	BSMV ^c
5'NCR	132	86	–	–	–
ORF 1	3 393	78	88	38	37
ORF 2	1 581	79	95	56	56
IG	61	62	–	–	–
ORF 3	375	75	75	27	30
3'NCR	299	86	–	–	–

^a NCR: non-coding region, IG: intergenic region between ORF 2 and ORF 3

^b Measured by using GAP

^c Corresponding region from SBWMV RNA-1 (ORF 1, 2) and RNA2 (ORF 3) and RNA α (ORF 1) and RNA γ of BSMV (ORF 2, 3)

to start immediately downstream of the ORF1 UGA (nt 3 526) and terminates at nt 5 103. A third ORF (nts 5 168 to 5 539) follows a non-coding region of 61 nts.

Similarity between IPCV RNA-1 and PCV RNA-1

Comparisons with the nucleotide sequences of other virus RNA showed that IPCV RNA-1 strongly resembles PCV RNA-1. By using GAP, the sequences were aligned with only two insertion/deletions of > 3 nucleotides. These were “deletions” in IPCV RNA-1 of 48 nucleotides starting at nt 1 980 and 16 nucleotides starting at nt 5 155.

The 5'-terminal seven nucleotides of IPCV RNA-1 are the same as those in RNA-2 [21], but the 5'-non-coding region (NCR) is 133 nucleotides long in contrast to the 501 nucleotides of the 5'-NCR of RNA-2 [21]. The six 5'-terminal nucleotides are the same as the six nucleotides shared between RNA-1 and RNA-2 of PCV [7]. The RNA-1 sequences of the two viruses are identical for the 5'-terminal 27 nucleotides and the 3'-terminal 20 nucleotides, presumably reflecting the strong similarity between RNA-1 of IPCV and PCV.

Table 1 lists the percent identities between IPCV and PCV in each region of RNA-1. Resemblance is closest (86%) in the terminal NCR; in the 3'-terminal 276 nucleotides there are only 30 mismatches between RNA-1 molecules of IPCV-H and PCV and 15 between RNA-1 and RNA-2 of PCV. The resemblance of IPCV to PCV is least in the intergenic region, although this region is perhaps too short to make the estimate meaningful. The similarities between the RNA-1 of IPCV

Table 2. Percentage identities among fragments of the RNA-1 nucleotide sequences of serotypes of IPCV and PCV

Region ^a	Viruses compared				
	H/T	H/L	H/PCV	T/PCV	L/PCV
63–561	–	92	86	–	84
2 956–3 360	89	–	76	75	–
4 734–5 425	88	–	76	76	–

^a Position of the nucleotide sequence fragments compared in the complete sequence of RNA-1 of IPCV-H

and PCV are more striking than those between RNA-2 molecules. The 5'-NCR and the coat protein gene in IPCV RNA-2 [21] are 55% and 62% identical respectively to those of PCV [9].

During other sequencing work with RNA-2 of isolates of IPCV belonging to serotypes L and T, fragments were obtained which matched the sequence of RNA-1 of H. Table 2 shows the similarities between these fragments and the corresponding sequences of IPCV-H RNA-1 or PCV RNA-1 as well as the similarities between H and PCV RNAs in the same regions. The percentage identity between sequences of isolates from different serotypes of IPCV were greater (88%, 89% and 92%) than the percentage identity between any of the IPCV isolates and PCV (75% to 86%).

Polypeptides encoded by IPCV RNA-1

The translation product of ORF 1 has a calculated M_r of 129 687 (p130) and that of ORF 1 and ORF 2 together has a calculated M_r of 189 875 (p190). Like the p131 and p191 proteins of PCV [7], which they resemble closely (Table 1), p130 and p190 contain sequence motifs which are thought to be characteristic of proteins with methyl transferase and helicase activity and, only in p190, sequence motifs thought to be characteristic of RNA-dependent RNA polymerases. The translation product of ORF3 (M_r 14 281; p14) does not contain motifs strongly suggestive of function. The calculated translation products of ORF 1 and ORF 1 + ORF 2 thus are close to the sizes estimated for the products of in vitro translation of RNA-1 of IPCV (L serotype) [11]. All three translation products are similar in amino acid sequence to those of PCV RNA-1 (Table 1). p14 of IPCV-H and PCV are 75% identical and ORF 3 nucleotide sequences are also 75% identical. In contrast, the nucleotide sequences of the ORF-1, and in particular the ORF 2, are less identical than are the proteins encoded by the ORFs (Table 1).

Table 1 also shows amino acid sequence comparisons between the IPCV translation products and the corresponding proteins of the type member of the

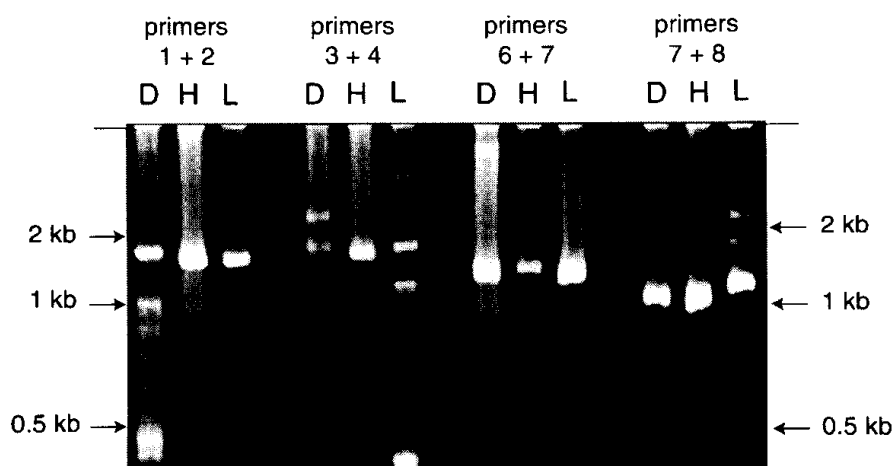


Fig. 2. Reverse transcription-PCR detection of IPCV RNA using different primer pairs. DNA products were analysed by electrophoresis in a 5% polyacrylamide gel, stained with ethidium bromide and photographed in UV light. The RNA samples were of isolates D, H and L (serotypes T, H and L respectively). The positions of marker DNA of known size are shown on either side. The bar indicates the top of the gel

genus *Furovirus*, soil-borne wheat mosaic virus and barely stripe mosaic virus. The percentage identity values are very similar to those reported between these proteins and the translation products of PCV RNA-1 [7].

RT-PCR detection assay

Four pairs of primers (J1 to J8) were designed for use in RT-PCR to amplify overlapping parts of RNA-1 of both IPCV-H and PCV. The positions of the primers are shown in Fig. 1. With each of the four primer pairs, RT-PCR yielded a prominent product between 1 kb and 2 kb (Fig. 2). These products were approximately the size of those predicted from the known sequence (Fig. 1) although small differences were detectable between the PCR products from isolates in the different serotypes suggesting that the RNA-1 molecules of different IPCV isolates differ in size. Other bands were presumably of DNA resulting from primer mismatching. Amplification done using primers J1 and J2 with RNA extracted from *N. benthamiana* leaves infected with IPCV also produced a fragment of the expected size (not shown).

Discussion

The results show that the RNA-1 molecules of IPCV-H and PCV are more similar than are the 5' regions of the RNA-2 molecules of the two viruses. The NCR of the RNA-1s are 86% identical whereas the 5' NCR of the RNA-2s are 55% identical. The most similar regions of the polypeptides encoded by the RNA-1s are the ORF 2 products, the putative polymerase region, which are 95% identical. Other RNA-1 encoded proteins are 75% identical and the RNA-2-

encoded coat proteins are 62% identical [21]. IPCV and PCV induce similar diseases in groundnut crops and the viruses have been considered by different authors as strains of one virus [2, 16] or distinct viruses [17, 21]. PCV occurs in a range of serological variants [10] and IPCV is known to occur in three serotypes between which there is little or no cross-reaction [14, 16]. Thus, although 95% sequence identity between polymerase amino acid sequences would suggest that the viruses being compared are variants of one virus species, the range of variation among isolates of IPCV, and perhaps PCV, together with the larger differences between other genes argue that a taxonomic distinction between IPCV and PCV is both justified and of practical value. The limited sequence comparisons between fragments of RNA-1 molecules of IPCV isolates of serotype L and T (Table 2) are consistent with this view in that the IPCV isolates were more alike than any was like PCV. Nevertheless, further sequence information, particularly from serologically distinct isolates of PCV, is needed for sound conclusions to be reached.

Because there are 61 possible codons in a RNA sequence but these encode one of only 20 amino acids, some nucleotide changes do not alter the encoded amino acid sequence. Thus in a gene derived from a common ancestor, nucleotide sequence divergence would be expected to be greater than amino acid sequence divergence; constraints as to permitted mutations would determine the actual ratio. In comparisons between IPCV and PCV (Table 1) the 15K genes were 75% identical in both nucleotide sequence and amino acid sequence whereas ORF 1 and its encoded protein were 78% and 88% identical and ORF 2 and its encoded protein were 79% and 95% identical. This suggests that the nucleotide sequence of the 15K gene has been constrained more than the sequences of the other ORF and thus that it contains functions other than that of being a mRNA.

As has been discussed by others [7, 9, 21], there are similarities among the gene products of PCV and IPCV and those of the type member of the genus *Furovirus*, SBWMV. However, there are also similarities between proteins of BSMV (genus *Hordeivirus*) and structural proteins of IPCV and PCV [21] or non-structural proteins of IPCV (p130, p190 and p14, Table 1) and PCV [7, 9]. The accommodation of these similarities in the taxonomy of furoviruses and hordeiviruses is a problem for future consideration.

IPCV and PCV infect a wide range of species including graminaceous crop species [17, 20], and Delfosse et al. [3] have reported that IPCV can cause a disease in wheat crops in India. Such graminaceous plants are good hosts for the virus vector *Polymyxa graminis*. These observations raise the possibility that graminaceous hosts could be infected by viruses which are related to IPCV or PCV but which do not infect groundnuts. If this were so, acquisition of the genetic determinant for groundnut infection, perhaps by recombination, could represent a source of novel pathogens. A broad specificity cDNA probe was devised on the basis of RNA-1 sequences and this has been shown to be capable of detecting infection by any of the 3 serotypes of IPCV as well as one isolate of PCV [22]. This probe would probably be capable of detecting such a novel virus.

The probe is currently being used in field diagnostic work. The current work has yielded primers for use in RT-PCR assays for IPCV, and by virtue of sequence match, the isolate of PCV described by Herzog et al. [7]. Although cumbersome for diagnostic work, except where extreme sensitivity is needed, this method is of potential value in conjunction with survey work as it allows the subsequent identification of the RNA-1 sequence amplified.

Acknowledgements

This work was funded in large part by the Overseas Development Administration Plant Sciences and Crop Protection Programmes and we acknowledge support from the Scottish Office Agriculture, Environment and Fisheries Department. This paper is submitted as Journal Article number JA 1900 by ICRISAT.

References

1. Brunt AA (1995) Genus Furovirus. In: Murphy FA, Fauquet CM, Bishop DHL, Ghabrial SA, Jarvis AW, Martelli GP, Mayo MA, Summers MD (eds) *Virus Taxonomy. Classification and Nomenclature of Viruses: Sixth Report of the International Committee on Taxonomy of Viruses*. Springer, Wien New York, pp 445–449 (Arch Virol [Suppl] 10)
2. Brunt AA, Richards KE (1989) *Biology and molecular biology of furoviruses*. Adv Virus Res 36: 1–32
3. Delfosse P, Reddy AS, Devi P, Murthy AK, Wesley SV, Naidu RA, Reddy DVR (1995) A disease of wheat caused by Indian peanut clump virus (IPCV). Plant Dis 79: 1074
4. Devereux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 12: 387–395
5. Frohmann MA, Dush MK, Martin GR (1988). Rapid production of full-length cDNA from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc Natl Acad Sci USA 85: 8998–9002
6. Gubler U, Hoffman BJ (1983) A simple and very efficient method for generating cDNA libraries. J Mol Biol 227: 1–8
7. Herzog E, Guilley H, Manohar SK, Dollet M, Richards K, Fritsch C, Jonard G (1994) Complete nucleotide sequence of peanut clump virus RNA 1 and relationships with other fungus-transmitted rod-shaped viruses. J Gen Virol 75: 3147–3155
8. Macfarlane SA, Taylor SC, King DI, Hughes G, Davies JW (1989) Pea early browning virus RNA 1 encodes four polypeptides including a putative zinc-finger protein. Nucleic Acids Res 17: 2245–2260
9. Manohar SK, Guilley H, Dollet M, Richards K, Jonard G (1993) Nucleotide sequence and genetic organization of peanut clump virus RNA 2 partial characterization of deleted forms. Virology 195: 33–41
10. Manohar SK, Dollet M, Dubern J, Gargani D (1995) Studies on variability of peanut clump virus: symptomatology and serology. J Phytopathol 143: 233–238
11. Mayo MA, Reddy DVR (1985) Translation products of RNA from Indian peanut clump virus. J Gen Virol 66: 1347–1351
12. Mayo MA, Robinson DJ, Jolly CA, Hyman L (1989) Nucleotide sequence of potato leafroll luteovirus RNA. J Gen Virol 70: 1037–1051
13. Mayo MA, Miller JS, Wesley SV, Naidu RA, Reddy DVR (1993) Current research on Indian peanut clump virus at SCRI. In: Reddy DVR, McDonald D, Moss JP (eds) *Working together on groundnut virus disease*, ICRISAT, Patancheru, p 46

14. Nolt BL, Rajeshwari R, Reddy DVR, Bharatan N, Manohar SK (1988) Indian peanut clump virus isolates: host range, symptomatology, serological relationships, and some physical properties. *Phytopathology* 78: 310–313
15. Reddy DVR, Rajeshwari R, Iizuka N, Lesemann DE, Nolt BL, Goto T (1983) The occurrence of Indian peanut clump, a soil-borne virus disease of groundnuts (*Arachis hypogaea*) in India. *Ann Appl Biol* 102: 305–310
16. Reddy DVR, Robinson DJ, Roberts IM, Harrison BD (1985) Genome properties and relationships of Indian peanut clump virus. *J Gen Virol* 66: 2011–2016
17. 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
18. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 74: 5463–5467
19. Shirako Y, Wilson TMA (1993) Complete nucleotide sequence and organization of the bi-partite RNA genome of soil-borne wheat mosaic virus. *Virology* 195: 16–32
20. Thouvenal JC, Fauquet C (1981) Further properties of peanut clump virus and studies on its natural transmission. *Ann Appl Biol* 97: 99–107
21. Wesley SV, Mayo MA, Jolly CA, Naidu RA, Reddy DVR, Jana MK, Parnaik VK (1994) The coat protein of Indian peanut clump virus: relationships with other furoviruses and with barley stripe mosaic virus. *Arch Virol* 134: 271–278
22. Wesley SV, Miller JS, Devi PS, Delfosse P, Naidu RA, Mayo MA, Reddy DVR, Jana MK (1996) Sensitive broad spectrum detection of Indian peanut clump virus by non-radioactive methods. *Phytopathology* (in press)
23. Zerfass K, Beier H (1992) The leaky UGA termination codon of tobacco rattle virus RNA is suppressed by tobacco chloroplast and cytoplasmic tRNAs^{Trp} with CmCA anticodon. *EMBO J* 11: 4167–4173
24. Zhand H, Scholl R, Browse J, Somerville C (1988) Double stranded DNA sequencing as a choice for DNA sequencing. *Nucleic Acids Res* 16: 1220

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

Received April 25, 1996