Pakistan J. Zool., vol. 39(3), pp. 165-170, 2007.

Coat Protein Gene Sequence Analysis of Three Begomovirus Isolates from Pakistan and their Affinities with other Begomoviruses

M. S. HAIDER, M. TAHIR, A. A. F. EVANS AND P. G. MARKHAM

School of Biological Sciences, University of the Punjab, Lahore, Pakistan (MSH, MT), Imperial College, London, Silwood Park Campus, Ascot, Berks, UK, SL5 7PY (AAFE) and John Innes Centre, Colney Lane, Norwich, UK, NR4 7UH (PGM)

Abstract.- Three virus infected host plant species (*Zinnia elegans, Solanum nigrum, and Ageratum conyzoides*) with suspected Geminivirus symptoms were screened for conserved regions of the coat protein genes using two degenerate primers from published sequences of the whitefly-transmissible geminiviruses from the Old World. The coat protein gene of three isolates designated Zinnia leaf curl virus (ZLCV), Solanum yellow leaf curl virus (SYLCV) and Ageratum yellow vein virus- Pakistan (AYVV-P) were PCR amplified, cloned and sequenced. Sequences were compared with previous known sequences of the begomoviruses by multiple sequence alignment and a phylogenetic tree was obtained. The viruses from Pakistan were found to be closely related to Indian subcontinent viruses *i.e.*, tomato leaf curl virus-India (TLCV-Ind), Indian cassava mosaic virus (ICMV) and cotton leaf curl virus (CLCuV). SYLCV was found to be a mechanically transmissible strain of TLCV-Ind and ZLCV was found to be the strain of AYVV-P.

Key words: Coat protein, PCR, Begomoviruses, Phylogeny.

INTRODUCTION

The Geminiviridae has three genera, which are divided on the basis of insect vector (Begomovirus if the vector is whitefly or Mastrevirus or Curtovirus if the vector is leafhopper) and genome arrangement. In addition to these three genera, a fourth genus has been proposed, to accommodate the treehopper-transmitted, tomato pseudu-curly top virus (Briddon and Markham, 1995).

The role of the geminivirus coat protein has been investigated by means of mutagenesis and agroinoculation. The isolation of a mutant of African cassava mosaic virus (ACMV) regenerated *in vivo* and defective in coat protein synthesis demonstrated that the coat protein of the whitefly-transmitted geminivirus is not essential for infection (Stanley and Townsend, 1986). A similar result has been reported by the investigation on an isolate of TLCV also known as Tomato leaf curl New Delhi virus (ToLCNDV) from India (Padidam *et al*., 1995a). In contrast to the above report, however, Gardiner *et al*. (1988) reported that the coding sequence of the coat protein of TGMV may be interrupted or substantially deleted without loss of infectivity, but the coat protein mutants showed reproducible delays in time of symptom appearance as well as reduced symptom development. The most attenuated symptoms appeared with a mutant in which the coat protein coding sequence was almost entirely deleted. The result indicated that the coat protein of TGMV was essential for infectivity. Similar studies with leafhopper transmitted geminiviruses BCTV and MSV showed that coat proteins of the monopartite geminiviruses are essential for spread of virus (Briddon *et al*., 1990; Boulton *et al*., 1989). More recent studies with TLCV-Ind showed that the coat protein was essential for the appearance of symptoms and spread of virus but did not affect the replication of the virus (Rigden *et al*., 1993).

The possible role of geminivirus coat protein in determining vector specificity in the transmission of geminiviruses was first demonstrated by Briddon *et al*. (1990). Using gene replacement to construct a chimeric clone in which the coat protein gene (AV1)

* Corresponding author: email:haider65us@yahoo.com

0030-9923/2007/0003-0165 $ 8.00/0

of the whitefly-transmitted ACMV was replaced by that of V2 of the leafhopper-transmitted BCTV, he then co-inoculated this chimeric DNA-A with DNA-B of ACMV onto *Nicotiana benthamiana*. Although plants were infected with the chimeric virus, the BCTV vector, *Circulifer tenellus*, could not transmit the virus by feeding on the infected plants. The leafhopper, however, did transmit the chimeric virus by membrane feeding on the sap extracted from the infected plants, and by injecting the chimeric virus purified from infected plants into the leafhopper's body. This result suggested that the coat protein might play a key role in determining the specificity of geminiviruses by their insect vector. The inability of the vector leafhopper to pick up the chimeric virus may be partly explained by the experimental host being a poor host plant of the leafhopper and partly by the possibility that the virus was not available in the tissue in which the leafhoppers fed. In addition, Mullineaux *et al*. (1984) demonstrated that leafhopper vectors could not transmit purified ss DNA or cloned ds DNA of leafhopper-transmitted MSV in the absence of any other proteins. These examples support the hypothesis that coat protein plays an important role in geminivirus transmission.

MATERIALS AND METHODS

Source of plant material

Three host plants (*Zinnia elegans, Solanum nigrum* and *Ageratum conyzoides*) showing begomovirus like symptoms were collected from the fields of Punjab (Multan and Bahawalpur zone), Pakistan. These infected samples were brought to the United Kingdom, and studies were conducted at John Innes Centre, Norwich. The viruses were maintained on the original/wild host plants or suitable alternate host plants (Table I).

Total DNA extraction from plant material

Total DNA extraction from healthy or infected plant tissue was carried out essentially as described by Parish and Kirby (1966), commonly known as the Kirby method. Tissues (0.5 mg) was ground in a pestle and mortar in liquid nitrogen, followed by adding 1 ml of Kirby buffer consisting of 1% TNE (Triisopropylnaphthalenesul-fonic acid sodium salt), 6% PMS (p-Minosalicylic acid (Sigma), 50 mM Tris-HCl (pH8.3), and 6% (v/v) buffer saturated phenol, and 1 ml of phenol/ chloroform was then added. The homogenate was transferred into an eppendorf tube and spun at 10 K for 10 minutes. The supernatant was transferred into a new tube, followed by two phenol/chloroform extractions. The DNA was ethanol precipitated, dried and resuspended in sterile distilled water.

PCR amplifications

Amplifications were performed in volumes of 100 µl containing 10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM MgCl, 0.01% (w/v) gelatin, 200 mM each of dATP, dTTP, dGTP, dCTP (Pharmacia), 0.05 mM of each primer, 100 ng of total nucleic acid, 2.5 U of *Taq* polymerase I (Amplitaq, Perkin Elmer Cetus) and were overlayed with 50 µl of mineral oil (Sigma) to reduce evaporation. Reactions were carried out in a Techne PHC-3 Thermal Cycler programmed for 35 cycles of 1 min at 94°C, 1 min at 55°C and 2 min at 72°C, using the fastest available transition between each temperature.

PCR amplification was attempted using two degenerate primers designed to conserved regions of the coat protein genes from published sequences of the whitefly-transmissible geminiviruses from the Old World (Haider, 1996).

The two degenerate oligonucleotide primers had the following sequences:

Primer A:5'-ATG(C/A/T) (G/C)(G/C/A) AAGCG (A/T) (C/A) C (A/C) G (G/C) (A/C) GATAT-3' (23 nucleotide, N-terminal)

Primer B: 5'TTAATT (T/G/C) (C/G/A) (A/T/C) (A/T/G) A (C/T) (A/T/C) (G/C) (C/A/T) (A/G) TCATA (G/A) AA (A/G) TA-3' (27 nucleotide, C-terminal). PCR product was detected by agarose gel electrophoresis using 1Kb ladder as marker.

Cloning and sequencing

PCR-amplified fragment of the coat protein gene of ZLCV, SYLCV and AYVV-P was ligated into pGEM-T cloning vector according to the

suppliers instructions. The clones in DH5α cells were screened for the correct insert through restriction digestion of the isolated plasmids.

Table I.- Virus codes, original host plants, symptoms and maintenance methods

Viruses	Host Plant	Symptoms	Maintenance method*
Zinnia leaf curl virus	*Zinnia elegans*	Leaf curl	(Gt)(It)
Solanum yellow leaf curl virus	*Solanum nigrum*	Yellow leaf curl	(Gt)(It)(Mi)
Ageratum yellow vein virus-Pakistan	*Ageratum conyzoides*	Yellow vein	(Vp)(It)

*Gt, Graft transmission; Vp, Vegetative propagation; It, Insect transmission; Mi, Mechanical inoculation.

Nucleotide sequences of the coat protein genes of ZLCV, SYLCV and AYVV-P were determined by dideoxynucleotide chain termination sequencing (Sanger *et al*., 1980) using Sequence Version II (USB) and [-^{35}P]dATP (New England Nuclear). Sequencing products were resolved by electrophoresis using Acrylamide solutions from Severn Biotech.

1xTBE buffer was used as the running buffer. The gel was fixed in 2L of 5% (v/v) methanol and 5% (v/v) acetic acid, transferred to filter paper (Whatman 3 MM) and dried in a gel dryer (Slab Gel Drier Model 483, Bio-Rad) at 80°C for 90 minutes before being exposed to x-ray film (Fuji) at room temperature overnight in an X-ray cassette and developed using a developing machine.

Computer analyses

Sequence data analysis, such as multiple sequence alignment, sequence comparisons, constructing phylogenetic trees *etc*., was carried out by the programs from the sequence analysis package distributed by the University of Wisconsin Genetics Computer Group (UWGCG) (Devereux *et al*., 1984). The GCG package is loaded and run in VAX/ VMS computer at the Computer Department of the John Innes Centre. Sequences of the coat protein gene were analyzed in Localpileup (GCG-package), and multiple sequence alignment file was further proceeded for Localpileup Figure.

RESULTS

The coat protein gene was successfully amplified for ZLCV, SYLCV and AYVV-P (Fig. 1). As a positive control, amplification was also obtained from CLCuV, all the amplified fragments were of the same size approximately 0.75–0.80 kb. and amplification was not obtained from healthy plant DNA Fragments were successfully cloned into pGEM-T (Fig. 2) and sequenced.

The nucleotide sequence of the coat protein gene for ZLCV (Accession No. *AM040438*), SYLCV (Accession No. *AM040437*) and AYVV-P (Accession No. *AJ810825*) were analyzed in Localpileup (GCG package), together with selected sequences of whitefly-transmitted geminiviruses. Phylogenetic analysis of the nucleotide sequences is

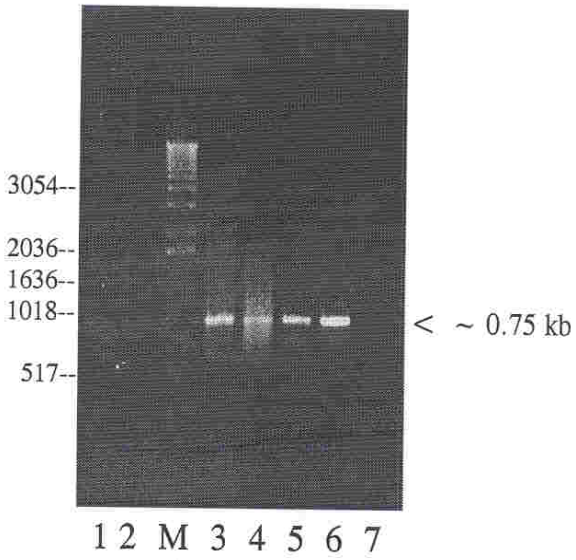

Fig. 1. Ethidium bromide stained agarose gel, Samples were the amplification product resulting from PCR reactions (coat protein) containing nucleic acids extracted from EPYVV and PLCV infected *N. benthamiana* (lane1 & 2), respectively ZLCV infected tomato(lane 3), AgYVV infected *Ageratum conyzoides* (lane 4), SYLCV infected *N. benthamiana* (lane 5), CLCuV infected *N. tabaccum* (lane 6), healthy

N. benthamiana (Lane 7). The sizes (bp) of co-electrophoresed markers are given. Arrow marks the position of amplified band, approximately 750bp.

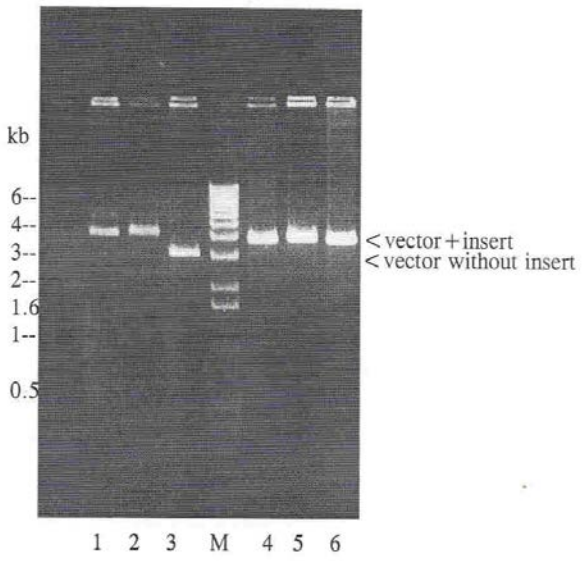

Fig. 2. Ethidium bromide stained agarose gel, samples were the products obtained by digestion with *Pst1*, linearised cloned fragments (in pGEM-T vector) of coat proteins of ZLCV (Lane 1), SYLCV (Lane 2), vector without insert (Lane 3), AgYVV-P individual colony of the same insert, (Lane 4,5 and 6). The sizes (bp) of co-electophoresed markers are given.

shown in Figure 3. The phylogenetic tree clearly separated Old World and New World viruses. Old World WTG's (spell out) showed further grouping, all the Indian-Pakistan viruses fall in the same group and showed segregation from the other groups in the old world.

The highest sequence homology (97%) was observed between SYLCV and TLCV-In, suggesting that SYLCV could be a mechanically transmissible isolate of TLCV-In.

Nucleotide sequences from the N-terminal (nt 22-201, 60 amino acids) region of the coat protein genes were analyzed more critically, by comparing the pair wise sequences of a number of WTGs from different parts of the World (Table II). Comparison of pairs of sequences show that those of AYVV-Pakistan and AYVV-Singapore are 81% identical, suggesting that these are two different viruses infecting *Ageratum conyzoides* under natural conditions.

However, AYVV-P show highest sequence homology (93%) with ZLCV from Pakistan, which suggests ZLCV be the strains of the AYVV-P. New World viruses showed 50-59% homology with AYVV-P and 46-58% with ZLCV.

SYLCV showed similarity with AYVV-P and ZLCV. CLCuV from Pakistan also showed 82% homology with ZLCV and 81% with AYVV-P which indicates that CLCuV is a different virus from either of these two isolates. SYLCV showed even less similarity (78%) with CLCuV, which is good evidence for these viruses being different from one another. New World viruses showed a similarity of 45-54% with SYLCV.

DISCUSSION

The results confirmed that the viruses (strains) belong to the family Geminiviridae and genus *Begomovirus* (Briddon and Markham, 1995). Previously, we reported the transmission of these viruses by the whitefly (Haider *et al*., 2003a) and DNA hybridization studies (Haider *et al*., 2003b) that support our present findings. Analyses of the sequence homologies revealed that SYLCV is very closely related to TLCV-In. However, despite the high level of sequence homology with TLCV-In, SYLCV showed significant biological differences with TLCV-In on symptoms, transmission and host range (Haider *et al*., 2003a). TLCV-In produced leaf curl symptoms with no yellowing whereas SYLCV produce irregular leaf curling and yellowing. Also the former was not mechanically transmitted (Padidam *et al*., 1995a) whereas the latter was mechanically transmitted (Haider, 1996).

Sequence homology between ZLCV and AYVV-P underlined a very close relationship and the host range for both was very similar, except that ZLCV did not infect *Ageratum conyzoides.* AYVV-P, on the other hand, could infect *Zinnia elegans* that caused different symptoms. They could be considered host adapted strains of the same virus. AYVV infecting *Ageratum* in Singapore was found to be significantly different virus when compared with AYVV-P. Nucleotide sequence similarity of

the complete (*i.e.*, 771 nucleotides for AYVV-P) coat protein gene the most conserved gene in WTG's was shown to be only 77% homologus and

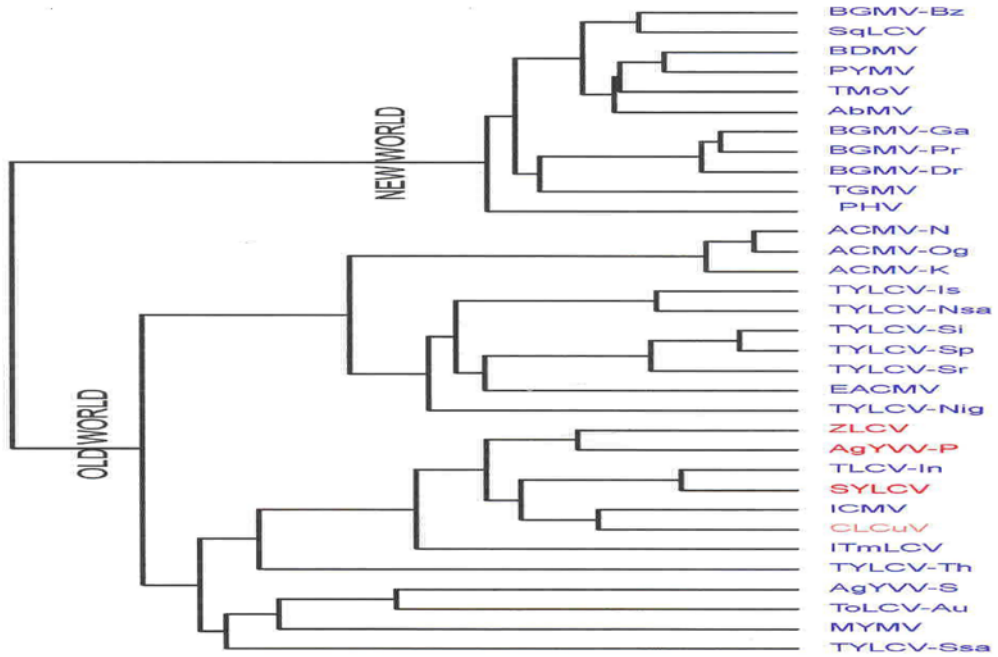

Fig. 3. Phylogenetic tree based on coat protein amino acid sequences of old and new world WTG's. Cyan and magenta colour indicating the viruses from Pakistan.

Table II. Sequence similarity between published data and the Pakistan viruses based upon nucleotide sequences from N-terminal region of the coat protein genes.

	AYVV-P			**Old**	**World**	**Viruses**	
AYVV-P	100%	ACMV-K					
ACMV-K	80	100%	TLCV-In				
TLCV-In	83	76	100%	**SYLCV**			
SYLCV	82	76	97	100%	ICMV		
ICMV	76	79	73	73	100%	**CLCuV**	
CLCuV	81	76	77	78	80	100%	**ZLCV**
ZLCV	93	81	82	82	79	82	100%
AYVV-S	81	75	75	74	76	77	80
TYLCV-T	78	77	73	75	79	78	80
TYLCV-Si	77	77	70	68	75	75	78
TYLCV-Sr	77	77	70	68	75	75	78
TYLCV-Sp	76	75	68	67	73	73	77
TYLCV-Is	81	77	75	74	77	77	83
EACMV	76	79	73	72	76	73	77
ToLCV-Au	74	77	67	67	71	73	73
ITmLCV	77	72	71	71	73	74	78
			New	**World**	**Viruses**		
TMoV	52	48	51	49	43	54	51
BDMV	56	47	51	54	42	50	46
BGMV-Bz	51	46	52	51	50	53	51
SqLCV	53	47	50	50	48	53	55
TGMV	59	53	52	45	55	60	58
PHV	52	50	47	50	50	51	50
PYMV	50	47	50	49	46	52	49

Bold letters showing the viruses from Pakistan.

the translated peptide sequence reached only 88% homology. Both the viruses infect *Ageratum conyzoides* under natural conditions and produce similar symptoms on this wild host plant. Similar conclusions were reached by Swanson *et al*. (1992).

The similarities and differences among Pakistani whitefly-transmitted geminiviruses (WTGs) are put in perspective by including in the analysis the CPs of other WTGs from a variety of hosts sequenced. It can then be seen that sequence divergence is independent of the natural host of each virus (Fig. 3). Thus on the basis of their CPs, WTGs from Pakistan in the Old World resemble WTGs from other or the same hosts in Old World countries more closely than they resemble New World WTGs. Comparable similarities in epitope profiles of the particles of WTGs from the Americas, and differences between those and Old World WTGs were found by Swanson *et al*. (1992).

In WTGs, the sequence homology between strains is usually very high (~90-98%), but shows different biological properties such as variation in symptom severity and patterns as well as host range. So far the molecular basis of these variations is still unclear although the difference of sequences between these strains can be established. If more sequences of different strains of a given geminivirus can be obtained, it may be possible to identify the substitution which modify gene functions and give rise to the variations of biological properties. This information will be particularly helpful for understanding the functions of geminivirus genes and hopefully will lead to the development of new methods for control of geminivirus associated diseases.

REFERENCES

BOULTON, M.I., STEINKELLNER, H., DONSON, J., MARKHAM, P.G., KING, D.I. AND DAVIES, J.W., 1989. Mutational analysis of the viron sense genes of maize streak virus. *J. Gen. Virol.,* **70**: 2309-2323.

BRIDDON, R.W. AND MARKHAM, P.G., 1995. Geminiviridae. In: *Virus taxonomy classification and nomenclature of viruses* (eds. F.A. Murphy, C.M. Fauquet, S.A. Bishop, Ghabrial, A.W. Jarvis, G.P. Martelli M.A. Mayo and M.D. Summers), pp. 158-165. Springer-Verlag, Wien, New York.

BRIDDON, R.W., PINNER, M.S., STANLEY, J. AND MARKHAM, P.G., 1990. Geminivirus coat protein replacement alters insect specificity. *Virology,* **177**:85-94.

DEVEREUX, J., HAEBERLI, P. AND SMITHIES, O., 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucl. Acids Res.,*. **12**: 387-395.

GARDINER, W.E., SUNTER, G., BRAND, I., ELMER, J.S., ROGERS, S.G. AND BISARO, D.M., 1988. Genetic analysis of tomato golden mosaic virus: the coat protein is not required for systemic spread or symptom development. *EMBO J.*, **7**:899-904.

HAIDER, M.S., BEDFORD, I.D., EVANS, A.A.F. AND MARKHAM, P.G., 2003A. Geminivirus transmission by different biotypes of the whitefly *Bemisia tabaci* (Gennadius). *Pakistan J. Zool.,* **35:** 343-351.

HAIDER, M.S., MUNEER, B., EVANS, A.A.F. AND MARKHAM, P.G., 2003b Dot blot hybridisation and PCR based detection of begomoviruses from the cotton growing regions of Punjab, Pakistan. *Mycopath,* **1**: 195-203

HAIDER, M.S., 1996. *Characterization of whitefly-transmitted geminiviruses from Pakistan*. Ph.D. thesis, Library University of London.

MULLINEAUXM, P.M., DONSON, J., MORRIS-KRSINICH, B.A.M., BOULTON, M.I. AND DAVIES, J.W., 1984. The nucleotide sequence of maize streak virus DNA. *EMBO J.*, **3**: 3063-3068.

PADIDAM, M., BEACHY, R.N. AND FAUQUET, C.M., 1995a. Tomato leaf curl geminivirus from India has a bipartite genome and coat protein is not essential for infectivity. *J. Gen. Virol.*, **76**:25-35.

PARISH, J.H. AND KIRBY, K.S., 1966. Reagents which reduce interactions between ribosomal RNA and rapidly labelled RNA from rat liver. *Biochim. Biphys. Acta,* **129**: 554-562.

RIGDEN, J.E., DRY, I.B., MULLINEAUX, P.M. AND REZAIAN, M.A., 1993. Mutagenesis of the virion-sense open reading frames of tomato leaf curl geminivirus. *Virology,* **193**: 1001-1005.

SANGER, F., COULSON, A,R,, BARRELL, G,G., SMITH, A.J.H. AND ROE, B.A., 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *J. mol. Biol.,* **143**:161-178.

STANLEY, J. AND TOWNSEND, R., 1985. Characterization of DNA forms associated with cassava latent virus infection. *Nucl. Acids Res.,* **13**:2189-2205.

SWANSON, M.M., VARMA, A., MUNIYAPPA, V. AND HARRISON, B.D., 1992. Comparative epitope profiles of the particle proteins of whitefly-transmitted geminiviruses from nine crop legumes in India. *Annls appl. Biol.,* **120:** 425-433.

THURING, W.R., SANDERS, J.P.M. AND BORST, P., 1975. A freeze-squeeze method of recovering long DNA from

agarose gels. *Anal. Biochem.*, **66**: 213-220.

(Received 5 February 2007, revised 16 May 2007)