A phylogenetic reexamination of *Cucumber mosaic virus* isolates based on 1a, 2a, 3a and 3b proteins

S. BALAJI[2], A.I. BHAT[1]* and S.J. EAPEN
Distributed Information Sub Centre for Bioinformatics, [1]Division of Crop Protection, Indian Institute of Spices Research, Calicut-673012

SUMMARY: The protein based phylogenetic trees based on 21 isolates of *Cucumber mosaic virus* estimated for 1a, 2a, 3a and 3b proteins, suggest that the 3a movement protein (MP) and 3b coat protein (CP) were congruent and indicated no significant difference (p-value= 0.699 > 0.05) in the branch lengths. The conservation might be due to the coordination of CP with MP. Hence it supports the earlier sub grouping of CMV based on 1a and 3a proteins in addition to coat protein. In the CP and MP phylogenies, on the basis of branch length the subgroup IA appears ancestral to the subgroup IB. Drastic variations of branch length was also observed in a few isolates. There are significant variations in the branch length between 1a and 2a proteins indicating that the RNAs had independent evolutionary histories.

Key words: *Cucumber mosaic virus*, 1a phylogeny, 2a phylogeny, 3a phylogeny, 3b phylogeny

INTRODUCTION

Cucumber mosaic virus (CMV) is the type species in the genus *Cucumovirus*, family *Bromoviridae* (Roossinck *et al.*, 1999a). CMV is one of the most economically important plant viruses and has a very wide host range including plants from approximately 365 genera and at least 85 families (Palukaitis *et al.*, 1992). CMV is a tripartite plus-sense RNA virus. Like most other plant viruses with divided genomes, the genomic RNAs are packaged in separate particles. This allows a larger genome to be packaged in a very simple virion but requires that multiple virus particles invade a single cell to initiate an infection (Roossinck, 2002). CMV possesses a genome consisting of three plus-sense, single-stranded RNA molecules, designated RNA1, RNA2 and RNA3 in decreasing order of molecular weight. The la protein and 2a protein encoded by RNA1 and RNA2, respectively, form the viral component of the replicase complex and are necessary for viral replication (Nitta *et al.*, 1988; Hayes and Buck, 1990; Palukaitis *et al.*, 1992). RNA2 also encodes a second protein, 2b, which functions in host-specific long-distance movement (Ding *et al.*, 1994, 1995). RNA3 encodes two proteins. The 3a protein is a cell-to-cell movement protein (MP), and the 3b protein is the capsid protein (CP), which is also involved in cell-to-cell movement and aphid-mediated CMV transmission from plant to plant (Canto *et al.*, 1997; Perry *et al.*, 1994; 1998).

A number of CMV strains reported from all over the world have been placed in to two subgroups I and II on the basis of serology (Wahyuni *et al.*, 1992; Hu *et al.*, 1995; Ilardi *et al.*, 1995), nucleic acid hybridization (Owens and Palukaitis, 1988), gene sequences (Owen *et al.*, 1990; Szilassy *et al.*, 1999), restriction fragment length polymorphism (RFLP) (Rizos *et al.*, 1992; Sialer *et al.*, 1999) and also by peptide mapping of the coat protein (CP) and nucleotide sequence identity (Palukaitis *et al.*, 1992). More recently a system for microarrays was developed to detect and differentiate CMV serogroups and subgroups (Deyong *et al.*, 2005). Phylogenetic analysis of CMV strains subdivided subgroup I into IA and IB on the basis of gene sequences (Palukaitis and Zaitlin, 1997; Roossinck *et al.*, 1999b; Roossinck, 2002). Further, Asian strains of CMV have been placed in subgroup IB (Roossinck *et al.*, 1999). Recently, complete phylogenetic analysis of 15 strains of CMV showed that the trees estimated from the open reading frames (ORFs) located on the different RNAs were not congruent and did not completely support the sub grouping indicated by the CP ORF, indicating that different RNAs had independent evolutionary histories (Roossinck, 2002). In order to confirm this, in the present study, protein based phylogeny (1a, 2a, 3a and 3b proteins) of 21 strains of CMV and its branch length significance of different subgroups was carried out to get more reliable phylogeny of the expressed genes.

MATERIALS AND METHODS

Dataset

Twenty one sequences of CMV 1a protein, 2a protein, 3a proteins and 3b proteins were retrieved from GenBank [http:// www.ncbi.nlm.nih.gov/Genbank/index.html]. The sequence data were belonging to different CMV infecting hosts like *Cucurbita pepo, Nicotiana glutinosa, Raphanus sativus, Arachis hypogaea, Lactuca saligna, Lycopersicon*

* *Corresponding author's E-mail: aib65@yahoo.co.in*
[2] *Present Address: Department of Biotechnology, Manipal Institute of Technology, Manipal University, Manipal.*

esculentum, Cucumis sativus, Pinellia ternate, Spinacia oleracea etc., and geographical locations such as China, Hungary, Italy, Japan, South Korea, Spain, Taiwan and USA were used for the analysis (Table 1).

Sequence alignment

Multiple sequence alignment was performed using the software program CLUSTALW (1.83) (Thompson *et al.*, 1994) with default parameters such as 30% delay divergent sequence and 0.50 as DNA transitions weight. Finally they were visually verified to check the gap usage.

Phylogenetic analysis

Peanut stunt virus (PSV) sequence was used as an out-group and added to the alignment profile. Alignments were analysed and phylogenetic relationships among sequences were established using different procedures: Neighbour-Joining (NJ) (Saitou and Nei, 1987), Fast Minimum Evolution (FastME) (Desper and Gascuel, 2002), Unweighted Pair Group Method with Arithmetic Mean (UPGMA) (Sokal and Michener, 1958) and Fitch-Margoliash (FM) (Fitch and Margoliash, 1967). Trees and genetic distances were based on 10,000 replicates in order to assess the degree of confidence for each branch on the trees. Heuristic searches were completed with maximum parsimony. Absolute distances and pairwise distances were calculated for all pairwise combinations of operational taxonomic units (OTUs). In order to find out the significant difference in the branch length (stated by Roossinck, 2002) of the phylogenetic tree derived from different proteins of different subgroups I A, IB and II, the paired student's *t*-test was performed.

RESULTS

Four different CMV proteins such as 1a protein, 2a protein, 3a protein (movement protein (MP)) and 3b protein (coat/capsid protein (CP)) in 21 strains were used in this study at the initial stage, later justified the consistency of branching pattern in CP and MP with 48 strains. The divergence of subgroups I and II was clearly seen in the 1a phylogeny estimation and the further divergence of the subgroup I into IA and IB was obvious and indicated that it can be used to confirm the subgroups (Fig. 1A). The 2a phylogeny estimation showed difference in the

Table 1. Coat Protein, 1a protein, 2a protein and 3a proteins of 21 strains of CMV from GenBank. '*' indicates that already the classification was confirmed and available in data base

No	Isolates	Country	Specific Host	Subgroup	A (1a protein)	B (2a protein)	C (3a protein)	D (3b/CP)
1	**FNY**	USA	*Cucurbita pepo*	**IA**	P17769	NP_049324	Q00271	AAA86502
2	**Ns**	Hungary	*Nicotiana glutinosa*		CAE45702	CAD54665	CAD54667	CAD54668
3	**Rs**	Hungary	*Raphanus sativus*		CAD54664	CAD57160	CAD57162	CAD57163
4	**CS_ch**	China	*Arachis hypogaea*	IA	AAR89474	AAR89475	AAR89477	AAR89478
5	**Ca**	China	*Arachis hypogaea*		AAR89474	AAR89471	AAR89469	AAM81361
6	**Ri-8**	Spain: Barcelona	*Lycopersicon esculentum*		CAJ65580	CAJ65581	CAJ65583	CAJ65584
7	**Pl-1**	Spain: Barcelona	*Lycopersicon esculentum*		CAJ65575	CAJ65576	CAJ65578	CAJ65579
8	**Y**	Japan	Not mentioned	**IA**	BAA02105	BAA02106	P18028	AAA46420
9	**Phy**	China	Not mentioned		ABD61578	ABD73003	ABD73005	ABD73006
10	**LS**	USA: New York	*Lactuca saligna*	**II**	AAL60175	AAL60176	AAD45246	AAD45247
11	**Ix**	USA	*Cucumis sativus, Solanum lycopersicum, Spinacia oleracea*	**IB**	AAC54620	AAC54616	AAC54618	Q66120
12	**O**	Japan	*Cucumis sativus*	**IA**	P20122	BAA01061	P16491	BAA00297
13	**Tfn**	Italy	*Lycopersicon esculentum*		CAB75952	CAB75953	CAB75955	CAB75956
14	**TN**	Japan	*Lycopersicon esculentum*		BAD15372	BAD15370	BAD15368	BAD15369
15	**Bx**	China	*Pinellia ternate*	**I***	ABD64217	ABD64218	ABD64221	ABD64220
16	**leg**	Japan	Iso. Iizuka, Str. Legume	**IA**	BAA03887	BAA03890	BAA03888	BAA03889
17	**Nt-9**	Taiwan	*Cucumis sativus, Solanum lycopersicum, Spinacia oleracea*	IB	O40976	BAA21694	O40979	BAA21697
18	**SD**	China	Not mentioned	**IB**	AAC24006	BAA13070	BAA23800	BAA23801
19	**Mf**	South Korea	Not mentioned	**IA***	CAB77386	CAB77387	CAB77389	CAB77390
20	**Pepo**	Japan	Not mentioned	**IA**	BAD00040	BAD00041	Q89785	AAD17925
21	**Cb7(Ib)**	China	*Lycopersicon esculentum*	**IB***	ABN12317	ABG79771	ABN12318	ABN12319

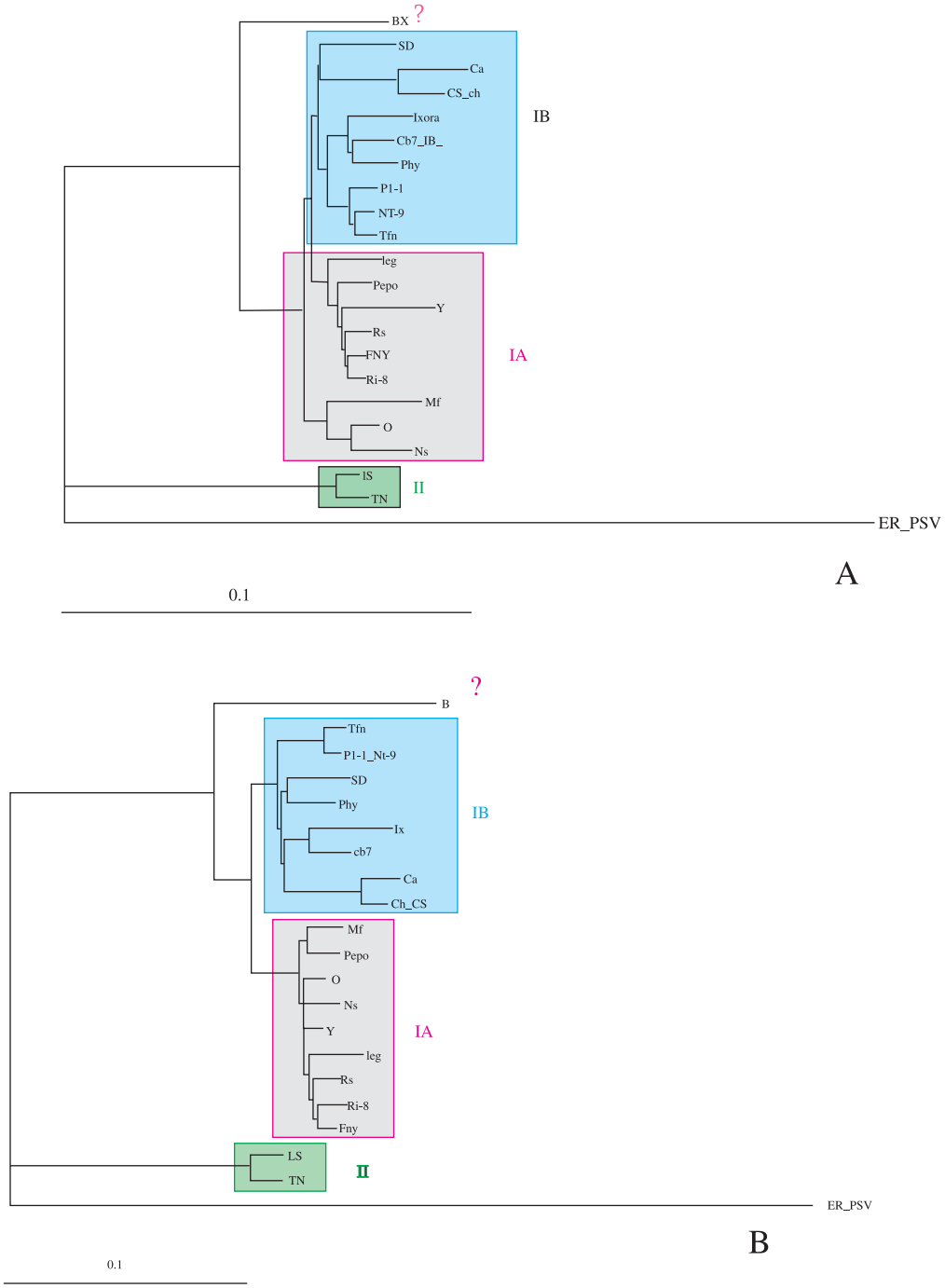


Fig. 1: Phylogenetic trees constructed using Neighbour-Joining (NJ), Fast Minimum Evolution (FastME), Unweighted Pair Group Method with Arithmetic Mean (UPGMA) and Fitch-Margoliash (FM) showed similar pattern. Phylogenetic tree constructed using FM with global optimization is shown. Peanut stunt virus *(PSV) sequence was used as an out-group. (A) The divergence of subgroups I and II is obvious in the 1a phylogeny; (B) 2a phylogeny shows the difference in the branching pattern compared with 1a tree.*

branching pattern compared with the 1a phylogeny. In 2a protein phylogeny the branching patterns of all the Ia strains were precise, in contrast to 1a protein phylogeny (Fig. 1B). But the branching pattern of IB strains in both 1a and 2a protein phylogeny showed similar trend. The CP phylogeny (Fig. 1D) was compact than the 3a (Fig. 1C) which yielded more branching within the groups. Visual comparison of CP tree with other trees such as 1a, 2a and 3a proteins showed that CP was compact and less radially evolved, the CP of Bx strain, isolated from *Pinellia ternate* in China was compactly grouped into IB (Fig. 1D), the similar grouping was not observed in 1a,

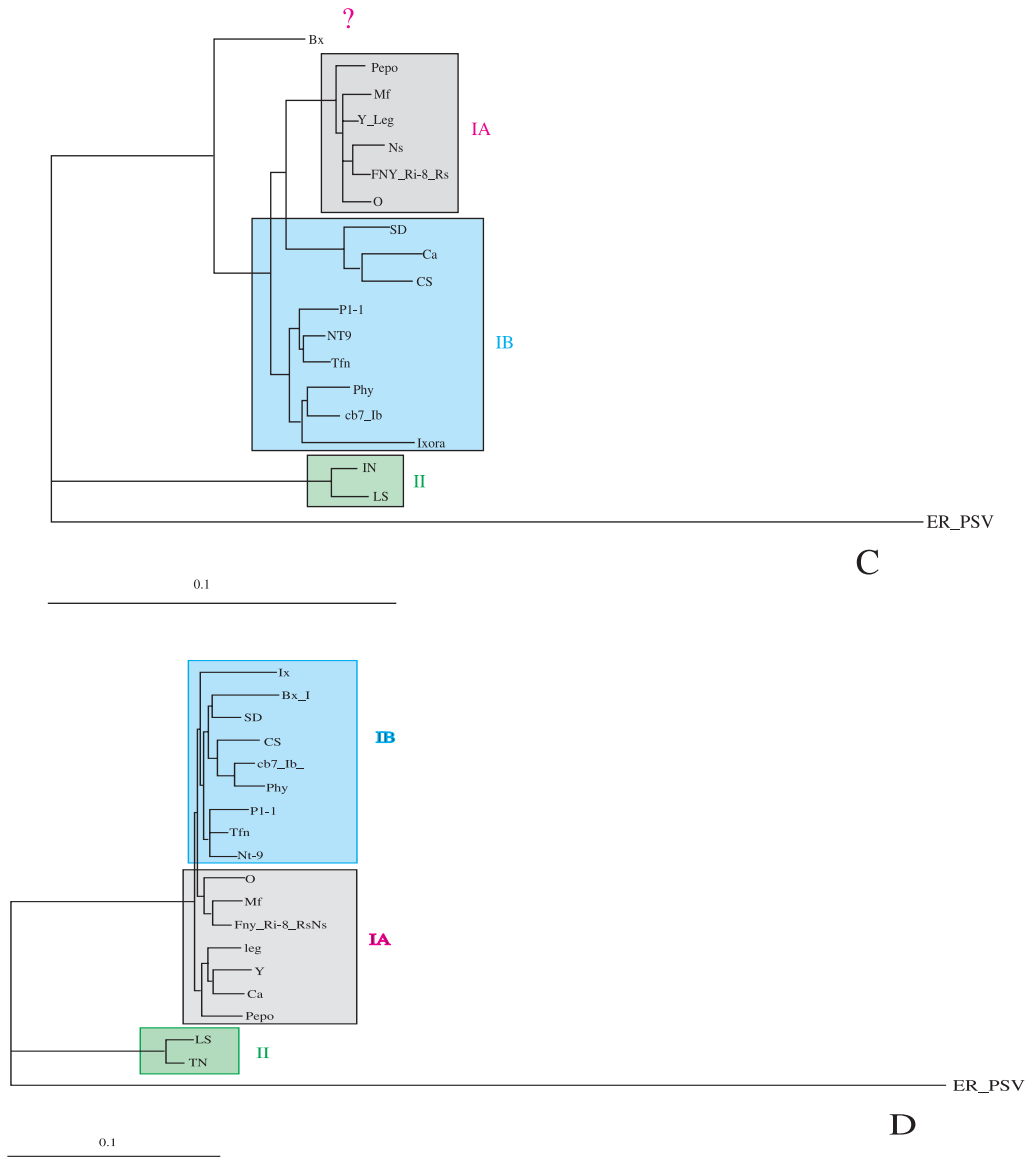


Fig. 1: Phylogenetic trees constructed using Neighbour-Joining (NJ), Fast Minimum Evolution (FastME), Unweighted Pair Group Method with Arithmetic Mean (UPGMA) and Fitch-Margoliash (FM) showed similar pattern. Phylogenetic tree constructed using FM with global optimization is shown. Peanut stunt virus *(PSV) sequence was used as an out-group. (C) 3a phylogeny shows more branching pattern compared with 3b tree; (D) 3b phylogeny shows the clades are compact compared with 3a tree.*

2a and 3a trees and it was far deviated from the IA and IB subgroups (Fig. 1A-C). The Bx strain was not related to any of the strains used and stands separately in the phylogeny (except the CP tree), showed a different phylogenetic pattern. It neither falls in group IA nor in IB and it was also different from group II, indicating the geographical variation in Asia. Similarly the Ix strain showed divergence with other strains.

Since we concentrated on the protein based phylogeny, we eliminated 3'NTR and 5'UTR to resolve the phylogeny. To verify the differences in the branch lengths, we have performed a paired student's *t*-test by assuming the null hypothesis (H_0) that there is no significant difference in the branch lengths of CP and 3a trees. Hence, the null hypothesis is H_0: $\mu 1 = \mu 2$. Our statistical findings on the branch lengths of 21 strains (in common to all coded proteins) confirmed that there is no significant difference between the branch lengths of CP and MP trees (Table 2 a, b); this guided us to focus on more strains of CMV. Hence we have collected 48 strains of CMV in common to CP and MP. The calculated *p*-value= 0.699 > 0.05 (95% confidence level) concluded to reject the alternate hypothesis (H_a: $\mu 1 \neq \mu 2$) and strongly supports the null hypothesis, hence we propose that there is no significant difference between the branch lengths of CP and MP. Moreover in the CP (Fig. 2A) and MP (Fig. 2B)

Table 2(a). Estimation of phylogenetic tree length variations among 21 common isolates of CMV from four genes (A: 1a protein, B: 2a protein, C: 3a protein, D: 3b protein/CP).

	Isolates	A	B	C	D
1	FNY	0.00208	0.00679	0.00358	0
2	Ns	0.03017	0.00743	0.00709	0
3	Rs	0.00311	0.00814	0.00358	0
4	CS_ch	0.03445	0.0733	0.0476	0.02753
5	Ca	0.04027	0.07778	0.04809	0.02163
6	Ri-8	0.00208	0.01004	0.00358	0
7	Pl-1	0.01833	0.05408	0.03564	0.01887
8	Y	0.01858	0.00074	0	0.02565
9	Phy	0.0231	0.05146	0.03769	0.03
10	LS	0.13544	0.23121	0.16626	0.17852
11	Ix	0.02685	0.07392	0.05389	0.03844
12	O	0.02252	0.00373	0.00358	0.01381
13	Tfn	0.01807	0.05635	0.0291	0.00969
14	TN	0.13802	0.23169	0.16287	0.17439
15	Bx	0.05641	0.12342	0.05616	0.03677
16	leg	0.01094	0.01795	0	0.01938
17	Nt-9	0.01847	0.05408	0.02926	0.01421
18	SD	0.02235	0.05739	0.04105	0.01835
19	Mf	0.03325	0.01418	0.00364	0.00458
20	Pepo	0.00453	0.01222	0.0037	0.01898
21	Cb7(Ib)	0.02235	0.05896	0.03453	0.02541
Average		0.037483	0.056352	0.036709	0.0322
Variance		0.009009	0.013562	0.009978	0.010509

Table 2(b). Phylogenetic branch length comparisons among four proteins (A: 1a protein, B: 2a protein, C: 3a protein (MP), D: CP in 21 common isolates of CMV. 'NS'- Not significant, 'S' - Significant.

Proteins	P value	Result
AB	0.002	<0.05 'S'
AD	0.95	>0.05 'NS'
CD	0.187	>0.05 'NS'
BC	0	<0.05 'S'
AC	0.258	>0.05 'NS'
BD	0	<0.05 'S'

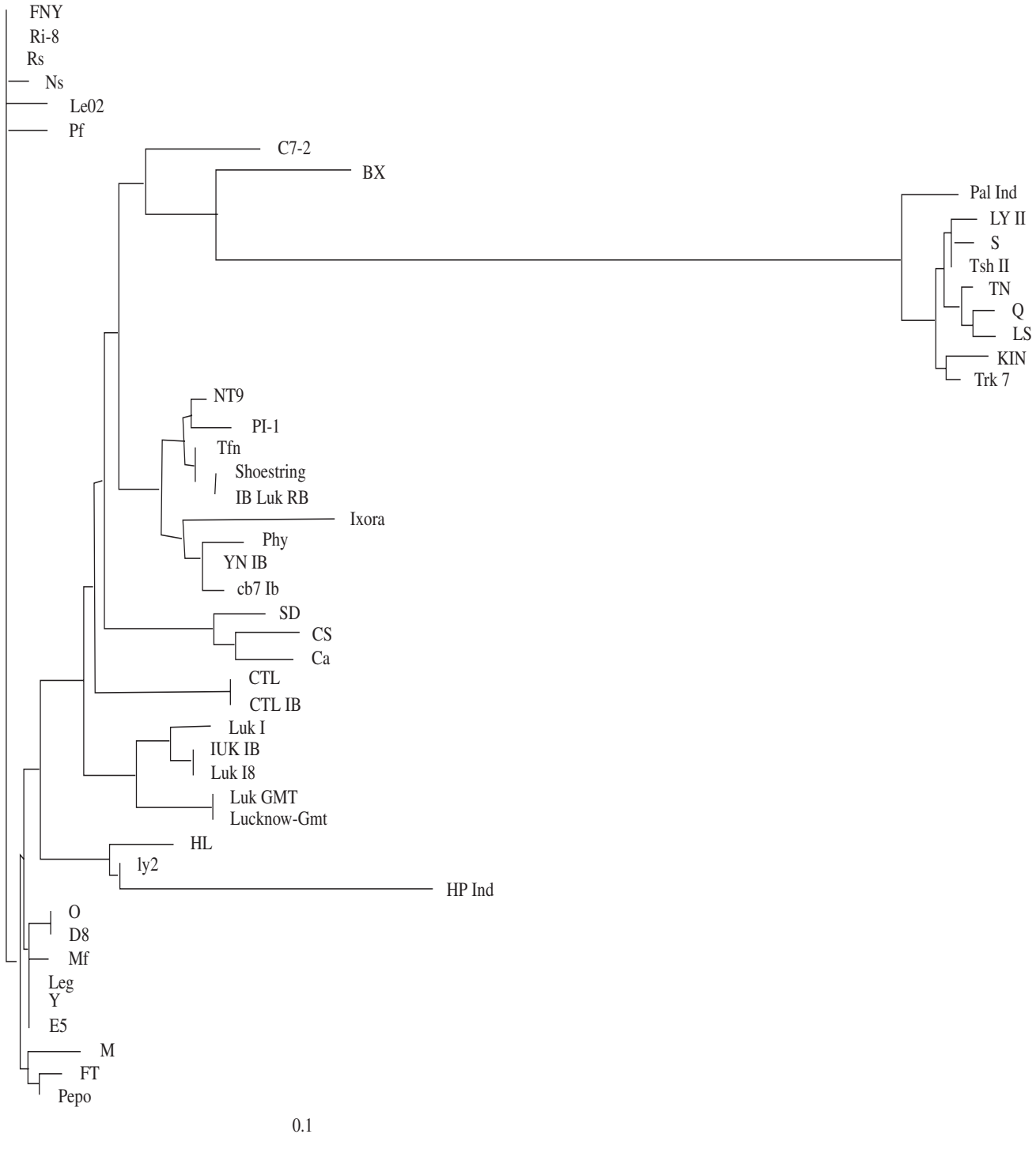
phylogenies, on the basis of branch length, the subgroup IA appears ancestral to the IB subgroup. In addition to the 21 strains (in common to all proteins), a few more isolates were also included in the analysis. They included 9 strains (AD combination), 27 strains (CD combination), 1 strain (BC combination), 10 strains (AC combination) and 3 strains (BD combination). All these combination also yielded the similar phylogeny derived from 21 strains (in common to all pairs, data not shown).

The Indian isolate HP showed major deviation (when compared to Bx and Ix) in the MP tree of the subgroup IB, whereas in CP it was compactly packed. One other Indian isolate (Palampur) which is also from Himachal Pradesh, inclined to subgroup II in both CP and MP trees. The phylogenetic grouping of 48 strains (which includes 7 strains of India) based on the branch length is shown in the Fig. 2. Similarly within Lucknow strains there were deviations. When all the proteins were considered, there are significant variations in the branch length between 1a and 2a proteins, 2a and 3a proteins as well as CP and 2a proteins. There were no significant differences in branch length comparisons between CP and 1a proteins, CP and 3a proteins as well as 1a and 3a proteins.

DISCUSSION

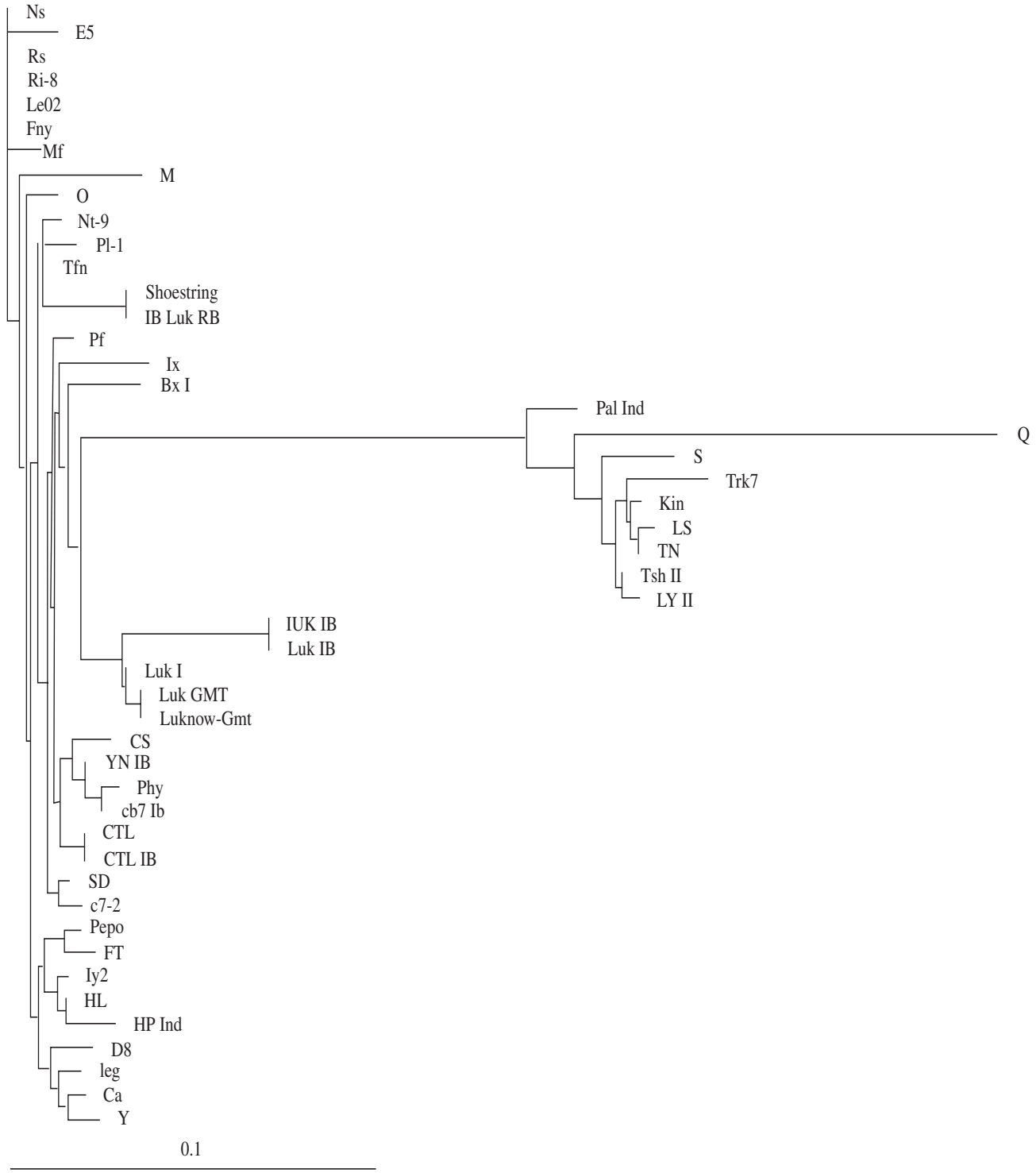
Present findings on the branch lengths based on 48 strains of CMV coding for CP and MP showed that there is no significant difference between the branch lengths of the CP and MP trees. This is in contrast to Roossinck (2002) who reported using 15 strains of CMV that there is a significant difference in the degree of branching and overall branch lengths between CP and MP (also indicated that there was not enough data to resolve the tree), reflecting different constraints on the evolution of each ORF (Roossinck, 2002). As far as is known, the CP interacts predominantly with itself or with the viral RNA and has little interaction with the host. CP interactions with the aphid vector are also probably minimal and mostly nonspecific, since more than 85 species of aphids can transmit CMV (Edwardson and Christie, 1991). The similar clustering pattern of all strains in the CP and MP could also be correlated with the recent finding that the MP of CMV forms ribonucleoprotein (RNP) complexes with viral RNA, capable of trafficking from cell-to-cell throughout the infected plant only in the presence of the CMV CP (Andreev *et al.,* 2004). The amino-terminal proximal sequences of the CMV CP interact with viral RNA and are required for the formation of stable virions and it is necessary for cell-to-cell movement (Kaplan *et al.,* 1998). Hence, some of the evolutionary constraints on the MP are also imposed by the coordination of CP.

Upon comparison of the 2a tree with the 1a tree, our statistical analysis of the branch length of 1a and 2a proteins is congruent with Roossinck (2002) indicating that the difference in the 1a and 2a trees are due to the host constraints. This may indicate that the evolution of the 2a protein is not constrained by virus-host interactions and thus is not the replicase component that interacts with the host factors but rather that it interacts with the 1a and/or itself, as well as with the viral RNAs (Roossinck, 2002). Recently, the results of Kim *et al.* (2006) suggest that the Tobacco CMV 1a interacting protein 2 (Tcoi2) as a novel host factor that is capable of interacting and phosphorylating methyltransferase domain of CMV 1a protein. Therefore the difference in the 1a and 2a trees may be due to the host constraints. With respect to 2a protein, subgroups IA and IB are found on separate clades, while 1a protein tree showed that few strains of the subgroup IA were related with the subgroup IB (based on the distance).



A

Fig. 2A: The grouping patterns of 3b protein (coat protein). The phylogenetic grouping of 48 strains is shown based on the branch length. Peanut stunt virus *(PSV) sequence was used as an out-group and it is not shown here due to disruption of the tree. Apart from the 21 strains listed in Table 1, additional 27 isolates of CMV belonging to both the subgroups were used. The accession numbers for the these strains are listed in bracket (Coat protein). HL (BAB69049), C7-2 (BAA07675), E5 (BAA07677, BAA07676), Pf (AAK52977, CAD58625), Q (AAA46414, P03604), Trk7 (AAA208884, Q83252), FT (BAA05847, Q66139), Kin (CAA78279), M (BAA01399), S (AAC16859), CTL (ABN05379), Tsh (ABN03949), LY (II) (AAM10540), Ly2 (CAC13146), D8 (BAA31136), CTL_IB (ABN05379), LUK_IB (ABM46606), YN_IB (ABN12316), Shoestring (ABM46608), Le02 (CAJ38270), Luk_GMT (ABM55263), Lucknow-GMT (EF178298), Luk_IB (ABM46608), LUK_IB (ABM46606), Luk_I (ABI95868), HP_Ind (CAI84629), Pal_Ind (CAL48248).*



B

Fig 2B: The grouping patterns of 3a protein (movement protein). The phylogenetic grouping of 48 strains is shown based on the branch length. Peanut stunt virus *(PSV) sequence was used as an out-group and it is not shown here due to disruption of the tree. Apart from the 21 strains listed in Table 1, additional 27 isolates of CMV belonging to both the subgroups were used. The accession numbers for these strains are listed in bracket (Movement protein). HL (BAB69048), C7-2 (BAA07674), E5 (BAA07676), Pf (CAD58625), Q (P03604), Trk7 (Q83252), FT (Q66139), Kin (Q06938), M (Q00272), S (Q66134), CTL (ABN05378), Tsh (ABN03948), LY (II) (AAM10539), Ly2 (CAC13145), D8 (BAA31135), CTL_IB (ABN05378), LUK_IB (ABM46605), YN_IB (ABN12315), Shoestring (ABM46607), Le02 (CAJ38269), Luk_GMT (ABM55262), Lucknow-GMT (EF178298), Luk_IB (ABM46607), LUK_IB (ABM46605), Luk_I (ABI95868), HP_Ind (CAH41945), Pal_Ind (CAL48247).*

The Bx, Ix and HP strains were compactly grouped into IB in CP tree. Similar grouping was not observed in 1a, 2a and 3a trees and it was far deviated from the IA and IB subgroups. The Palampur strain (from India) was grouped in to the subgroup II and showed a similar pattern of branching in both CP and MP trees. We suppose the strains Bx, Ix and HP which tend towards subgroup II in all the trees except in CP may have resistance-breaking capacity like systemic chlorosis observed in RNA 2 of subgroup II CMV strains (Sleat and Palukaitis, 1990; Zaitlin *et al.*, 1994). Therefore the divergence of Bx, Ix and HP strains may claim a new classification (Group IC) if sequences of more variant strains are available.

In contrast to Roossinck's statement i.e., based on the RNA 3 ORFs where the IA strains fall out as a clade within the IB strains, the present protein based phylogeny on 3a protein (Fig 2B) showed the subgroup IA as an ancestor to subgroup IB (based on the distance) and the IB clade is clearly separated from IA. Our analysis of CMV strains in the 3a tree (the cluster composed of strains S, LY, Tsh, Q, LS, TN, KIN, TrK7) is well in agreement with Salanki *et al.* (1994) who proved the earlier biological and serological classification of the RNA 3 molecule of Trk7-CMV in subgroup II. A high degree of homology was found in the strains Q and Kin of CMV (Salanki *et al.*, 1994). The strain Q is closely related to LS and TN in the 3a tree, whereas in the CP tree it was much deviated from the other strains of sub-group II.

Takeshita *et al.* (2004) suggested that the dynamics of competitive interactions between the two CMV subgroups (I and II) could be characterized by exclusive infection and multiplication of the individual viruses in cowpea plants. Phylogenetic analysis by Lin *et al.* (2004) indicated a naturally occurring reassortant between subgroup IA and IB isolates and potential reassortants between subgroup IA isolates, suggesting that genetic exchange by reassortment contributed to the evolution of the California CMV population. Analysis of various population genetics parameters and distribution of synonymous and non synonymous mutations revealed that different coding regions and even different parts of coding regions were under different evolutionary constraints, including a short region of the 2b gene for which evidence suggests possible positive selection. Hence it is clear that some RNA appears to have a different evolutionary history. The increased genetic diversity afforded by reassortment may allow for host range expansion and may allow for recovery of strains bearing deleterious mutations. However, if the reassortant event occurs between two strains that are already highly divergent, it may cause problems in protein-protein interactions, such as those required for a fully functional replicase, or in recognition of promoter sequences, RNA packaging signals, and other intra-RNA functions. In these cases the reassortment event may be common but the new viruses may be eliminated by strong negative selection (Roossinck, 2002). The deviant strains such as Bx, Ix and HP are suspected to have strong negative selection observed in the branching pattern.

Clearly the phylogenetic analysis of worldwide distributed strains based on the protein sequence confirm that the reassortment events have played an important role in CMV. It is clear that apart from CP, 1a and 3a proteins can be used for the grouping of CMV. In CMV, the 1a protein is expressed at very low levels compared to the 3a protein and the CP (Gallie and Kobayashi, 1994) and the evolutionary trees of the 3a and 3b proteins were more compact because of non significant variations in the branch lengths. Owen *et al.*, (1990) stated that the limited variation shown by RNA 3 of strains of CMV with different passage histories, isolated in different countries over a 50 year period, suggests that the maintenance of the highly conserved nucleotide sequence may be important for other viral RNA functions or interactions (Andreev *et al.*, 2004). Hence we suggest the usage of CP and 3a proteins for CMV sub grouping.

REFERENCES

Andreev, I.A., Hyon Kim, S., Kalinina, N.O., Rakitina, D.V., Fitzgerald, A.G., Palukaitis, P. and Taliansky, M.E. 2004. Molecular interactions between a plant virus movement protein and RNA: force spectroscopy investigation. *J. Mol. Biol*. **339:** 1041-1047.

Canto, T., Prior, D.A.M., Hellwald, K.H., Oparka, K.J. and Palukaitis, P. 1997. Characterization of cucumber mosaic virus. IV. Movement protein and coat protein are both essential for cell-to-cell movement of cucumber mosaic virus. *Virology* **237:** 237-248.

Desper, R. and Gascuel, O. 2002. Fast and accurate phylogeny reconstruction algorithms based on the minimum-evolution principle. *J. Comp. Biol*., **9:** 687-705.

Deyong, Z., Willingmann, P., Heinze, C., Adam, G., Pfunder, M., Frey, B. and Frey, J.E. 2005. Differentiation of *Cucumber mosaic virus* isolates by byhridization to oligonucleotides in a microarray format. *J. Virol. Methods.* **123:** 101-108.

Ding, S.W., Anderson, B.J., Haase, H.R. and Symons, R.H. 1994. New overlapping gene encoded by the cucumber mosaic virus genome. *Virology* **198:** 593-601.

Ding, S.W., Li, W.X. and Symons, R.H. 1995. A novel naturally occurring hybrid gene encoded by a plant RNA virus facilitates long distance virus movement. *EMBO J.* **14:** 5762-5772.

Edwardson, J.R. and Christie, R.G. 1991. Cucumoviruses CRC Handbook of viruses infecting legumes. CRC Press, Boca Raton, Fla, p. 293-319.

Fitch, W.M. and Margoliash, E. 1967. *Construction of phylogenetic trees*. *Science* **155:** 279-284.

Gallie, D.R. and Kobayashi, M. 1994. The role of the 3-untranslated region of non-polyadenylated plant viral mRNAs in regulating translational efficiency. Gene **142:**159-165.

Ilardi, V., Mazzei, M., Loreti, S., Tomassoli, L. and Barba, M. 1995. Biomolecular and serological methods to identify strains of *Cucumber mosaic cucumovirus* on tomato. *EPPO Bull.* **25:** 321-327.

Hayes, R.J. and Buck, K.W. 1990. Complete replication of a eukaryotic virus RNA *in vitro* by a purified RNA-dependent RNA polymerase. *Cell* **63:** 363-368.

Hu, J.S, Li, H.P, Barry, K., Wang, M. and Jordan, R. 1995. Comparison of dot blot ELISA and RT-PCR assays for detection of two *Cucumber mosaic virus* isolates infecting banana in Hawaii. *Plant Dis.* **79:** 902-906.

Kaplan, I.B, Zhang, L. and Palukaitis, P. 1998. Characterization of cucumber mosaic virus. V. Cell-to-cell movement requires capsid protein but not virions. *Virology.* **246:** 221-231.

Kim, M.J., Ham, B.K. and Paek, K.H. 2006. Novel protein kinase interacts with the Cucumber mosaic virus 1a methyltransferase domain. *Biochem Biophys Res Commun.* **340:** 228-235.

Lin, H.X., Rubio, L., Smythe, A.B. and Falk, B.W. 2004. Molecular population genetics of *Cucumber mosaic virus* in California: evidence for founder effects and reassortment. *J Virol.* **78:** 6666-6675.

Nitta, N., Takanami, Y., Kuwata. S. and Kubo. S. 1988. Inoculation with RNAs I and 2 of cucumber mosaic virus induces viral RNA replicase activity in tobacco mesophyll protoplasts. *J. Gen. Virol.* **69:** 2695-2700.

Owen, J., Shintaku, M., Aeschleman, P., Tahar, S.F. and Palukaitis, P. 1990. Nucleotide sequence and evolutionary relationships of *Cucumber mosaic virus* (CMV) strains, CMV RNA 3. *J. Gen. Virol.* **71:** 2243-2249.

Owens, J. and Palukaitis, P. 1988. Characterization of *Cucumber mosaic virus*. I. Molecular heterogeneity mapping of RNA 3 in eight CMV strains. *Virology* **69:** 496-502.

Palukaitis, P. and Zaitlin, M. 1997. Replicase-mediated resistance to plant virus diseases. *Adv. Virus Res.* **48:** 349-377.

Palukaitis, P., Roossinck, M.J., Dietzgen, R.G. and Francki, R.I.B. 1992. Cucumber mosaic virus. *Adv Virus Res.* **41:** 281-348.

Perry, K.L., Zhang, L. and Palukaitis, P. 1998. Amino acid changes in the coat protein of cucumber mosac virus differentially affect transmission by the aphids *Myzus persicae* and *Aphis gossypii*. *Virology* **242:** 204-210.

Perry, K.L., Zhang, L., Shintaku, M.H. and Palukaitis, P. 1994. Mapping determinants in cucumber mosaic virus for transmission by *Aphis gossypii*. *Virology* **205:** 591-595.

Rizos, H., Gunn, L.V., Pares, R.D. and Gillings, M.R. 1992. Differentiation of *Cucumber mosaic virus* isolates using the polymerase chain reaction, *J. Gen. Virol.* **73** 2099-2103.

Roossinck, M.J. 2002. Evolutionary history of *Cucumber mosaic virus* deduced by phylogenetic analyses; *J. Virol.* **76:** 3382-3387.

Roossinck, M.J., Bujarski, J., Ding, S.W., Hajimorad, R., Hanada, K., Scott, S. and Tousignant, M. 1999a. Family *Bromoviridae*. In: M. H. V. vanRegenmortel, C. M. Fauquet, and D. H. L. Bishop (Ed.), *Virus Taxonomy–Seventh Report of the International Committee on Taxonomy of Viruses*. Academic Press, San Diego, Calif. pp. 923-935.

Roossinck, M.J., Zhang, L. and Hellward, K. 1999b. Rearrangements in the 5' nontranslated region and phylogenetic analyses of *Cucumber mosaic virus* RNA 3 indicate radial evolution of three subgroups. *J. Virol.* **76:** 6752-6758.

Saitou, N. and Nei, M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4:** 406-425.

Salanki, K., Thole, V., Balazs, E. and Burgyan, J. 1994. Complete nucleotide sequence of the RNA 3 from subgroup II of cucumber mosaic virus (CMV) strain: Trk7. *Virus Res.* **3:** 379-384.

Sialer, M.M., Cillo, F., Barbarossa, L. and Gallitelli, D. 1999. Differentiation of *Cucumber mosaic virus* subgroups by RTPCR RFLP. *J. Plant Pathol.* **81:** 145-148.

Sleat, D.E. and Palukaitis, P. 1990. Induction of tobacco chlorosis by certain cucumber mosaic virus satellite RNAs is specific to subgroup II helper strains. *Virology* **176:** 292-295.

Sokal, R.R. and Michener, C.D. 1958. A statistical method for evaluating systematic relationships. *University of Kansas Scientific Bulletin* **38:** 1409-1438.

Szilassy, D., Sala´nki, K. and Bala´zs, E. 1999. Molecular Evidence for the existence of two distinct subgroups in *Cucumber mosaic cucumovirus*. *Virus Genes* **18:** 221-227.

Takeshita, M., Shigemune, N., Kikuhara, K., Furuya, N. and Takanami, Y. 2004. Spatial analysis for exclusive interactions between subgroups I and II of Cucumber mosaic virus in cowpea. *Virology*. **328:** 45-51.

Thompson, J.D, Higgins, D.G. and Gibson, T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22:** 4673-4680.

Wahyuni W.S., Dietzen R.G., Handa, K. and Francki, R.I.B. 1992. Serological and biological variation between and within subgroup I and II strains of *Cucumber mosaic virus*. *Plant Pathol.* **41:** 282-297.

Zaitlin, M., Anderson, J.M., Perry, K.L., Zhang, L. and Palukaitis, P. 1994. Specificity of replicase-mediated resistance to cucumber mosaic virus. *Virology* **201:** 200-205.