High genetic diversity in the coat protein and 3′ untranslated regions among geographical isolates of *Cardamom mosaic virus* from south India

T JACOB, T JEBASINGH, M N VENUGOPAL* and R USHA†

Department of Plant Biotechnology, School of Biotechnology, Madurai Kamaraj University, Madurai 625 021, India
**Indian Institute of Spices Research, Appangala 571 201, India*

†*Corresponding author (Fax, 91-452-2459105; Email, usha@mrna.tn.nic.in)*

A survey was conducted to study the biological and genetic diversity of *Cardamom mosaic virus* (CdMV) that causes the most widespread disease in the cardamom growing area in the Western Ghats of south India. Six distinct subgroups were derived based on their symptomatology and host range from the sixty isolates collected. The serological variability between the virus isolates was analysed by ELISA and Western blotting. The 3′ terminal region consisting of the coat protein (CP) coding sequence and 3′ untranslated region (3′UTR) was cloned and sequenced from seven isolates. Sequence comparisons revealed considerable genetic diversity among the isolates in their CP and 3′UTR, making CdMV one of the highly variable members of *Potyviridae*. The possible occurrence of recombination between the isolates and the movement of the virus in the cardamom tract of south India are discussed.

[Jacob T, Jebasingh T, Venugopal M N and Usha R 2003 High genetic diversity in the coat protein and 3′ untranslated regions among geographical isolates of *Cardamom mosaic virus* from south India; *J. Biosci*. **28** 589–595]

1. Introduction

Cardamom mosaic virus (CdMV) is the causative agent of the most widespread and devastating disease of small cardamom (*Eletteria cardamomum* Maton). The virus is transmitted by the banana aphid *Pentalonia nigronervosa* (Uppal *et al* 1945). Cardamom is cultivated in geographically distinct pockets in the high range regions of the Western Ghats over a stretch of 1000 km, covering Kerala, Tamil Nadu and Karnataka states of south India. Mayne (1951) reported the prevalence of the mosaic disease in the cardamom-growing hills of Tamil Nadu and Karnataka. The continued cultivation of cardamom for centuries has resulted in considerable variations in the mosaic disease of cardamom. Jacob and Usha (2001) have classified CdMV as a new member of the *Macluravirus* genus of *Potyviridae* based on the 3′ terminal sequence of the virus genome. The present study was undertaken to analyse the biological and molecular properties of the various isolates of the virus collected from all along the cardamom tract of south India.

2. Materials and methods

2.1 *Virus isolates*

A survey was conducted in the cardamom-growing area covering 402 plantations in 120 villages distributed in 20 districts spanning 6715 hectares in south India. In each plantation, 100 plants were selected for disease count. Sixty isolates depicting the mosaic symptoms, ranging from mild to severe mosaic, were collected and maintained in the glass house. All the isolates were passaged through

Keywords. *Cardamom mosaic virus*; coat protein; genetic diversity; geographical isolates

Abbreviations used: CdMV, *Cardamom mosaic virus*; CP, coat protein; PCR, polymerase chain reaction; 3′UTR, 3′ untranslated region; YMV, *Yam mosaic virus*.

aphid-transmission to seedlings of CV-37 cultivar of cardamom. Six symptomatologically distinct subgroups of CdMV were identified (table 1) based on the inoculation studies on the eight *Zingiberaceae* plants, namely *Eletteria cardamomum*, *Amomum muricatum*, *Amomum microstephanum*, *Amomum cannicarpum*, *Alpinia mutica*, *Hedychium flavescens*, *Zingiber cernuum* and *Aframomum melequeta*.

2.2 *ELISA and Western blotting*

Virus was extracted from ten isolates representing all the six subgroups, using the method described by Jacob and Usha (2001). The molecular weight of the coat protein of the virus isolates was determined by SDS PAGE. ELISA and Western blotting experiments were carried out on all the virus isolates using anti CdMV IgG produced against the Yeslur isolate of CdMV (Jacob and Usha 2002). Buffer was used as negative control for the antigen and normal rabbit (non-immunized) serum was used as the negative control for the antiserum.

2.3 *RNA extraction, RT-PCR and cloning of PCR products*

RNA was extracted from the purified virus preparations using 1% SDS and three volumes of acid phenol. The quantity and quality of the viral RNA was checked by MOPS formaldehyde agarose gel electrophoresis (Farrel 1993). First strand cDNA was synthesized from 1 µg of viral RNA using super script RNAse H^- reverse transcriptase (GIBCO, BRL) according to a modified protocol (Carninci

Table 1. Percentage identities between the different isolates of CdMV. Identities between the nucleotide sequence of CP alone in bold, nucleotide sequence of UTR alone is given in normal lettering, CP sequence identity is underlined and the identity among the 40 aa from the N-terminus of the CP is given in italics. The subgroup number is given in brackets in the table. EMBL accession numbers are given in brackets as follows. Yes, Yeslur (AF189125); Mad, Madikeri (AJ308474); Tha, Thalathamane (AJ308475); Apa, Appangala (AJ308472); Sir, Sirsi (AJ308473); Kur, Kursupara (AJ308476); Kat, Kattappana (AJ312774); Van, Vandiperiyar (AJ308477).

Isolates	Mad (II)	Tha (II)	Apa (II)	Sir (III)	Kat (IV)	Kur (V)	Van (IV)
Yes (I)	**82·7** 80·1 92·9 *62×5*	**82·7** 77·8 92·3 *60×0*	**82·7** 76·8 93·4 *62×5*	**77·2** 65·5 89·3 *47×5*	**78·2** 71·6 90·8 *47×5*	**76·8** 71·8 89·3 *42×5*	**76·4** 69·7 88·6 *51×3*
Mad (II)		**97·3** 97·3 97·4 *90×0*	**98·4** 99·1 98·2 *90×0*	**77·7** 71·2 87·5 *45×0*	**76·3** 67·8 89·3 *47×5*	**76·8** 70·0 88·6 *45×0*	**75·7** 70·8 87·5 *45×0*
Tha (II)			**98·1** 98·2 98·5 *95×0*	**77·5** 73·2 86·7 *42×5*	**76·6** 67·3 88·9 *47×5*	**77·1** 70·5 89·6 *47×5*	**75·4** 70·4 86·7 *42×5*
Apa (II)				**77·5** 70·6 87·8 *45×0*	**76·6** 67·8 89·6 *47×5*	**77·7** 70·5 89·3 *47×5*	**75·6** 70·8 87·5 *42×5*
Sir (III)					**76·5** 71·6 87·2 *42×5*	**77·4** 71·6 87·5 *47×5*	**76·4** 73·1 86·5 *48×7*
Kat (IV)						**84·4** 84·4 93·1 *67×5*	**89·9** 84·3 95·9 *87×5*
Kur (V)							**84·4** 88·8 91·57 *67×5*

et al 1998) from all the ten isolates. A specific forward primer CPn1 (5′GGCAACACAAACAGTGCCATC3′) and a general reverse primer (Poty1 + primer as described by Gibbs and Mackenzie 1997) complementary to the poly A tail of the RNA genome were used to amplify the coat protein (CP) coding sequences and 3′ untranslated region (3′UTR) from all the strains. The polymerase chain reaction (PCR) products were cloned into pGEM-T vector (Promega) and completely sequenced by automated sequencing and submitted in the EMBL database (Accession numbers are given in the legend to table 1).

2.4 *Sequence analysis*

The GCG software (Wisconsin package version 9.1) was used to find the open reading frame (ORF) and the proteolytic cleavage sites (PEPTIDESORT) in the CP sequences. The pair-wise alignment and the multiple alignments of the CP and 3′UTR of CdMV isolates were done by GAP (GCG software) and CLUSTALW programs respectively (Thompson *et al* 1994). A neighbour-joining phylogenetic tree was constructed using CLUSTALW, and the tree was drawn by DRAWGRAM or TREEVIEW of the PHYLIP package (Felsentstein 1989).

3. Results and discussion

The subgroups I (Yeslur), II (Appangala, Madikeri and Thalathmanae from Coorg district) and III (Sirsi) originate from Karnataka state and subgroup IV (Kattappana and Vandiperiyar from Idukki district), V (Kursupara) and VI (Palakkad) originate from Kerala state (figure 1). The subgroups I, II, V (figure 2) and VI (not shown) produced mild or severe mosaic symptoms or yellowing with the first four host plants tested, namely *E. cardamomum*, *A. muricatum*, *A. microstephanum* and *A. cannicarpum*. Subgroup III (figure 2) produced either mosaic or necrotic or chlorotic flecks with all the five susceptible plants. Subgroup IV produced veinal chlorosis with mosaic symptoms. SDS PAGE of selected isolates showed a minor difference in the mobility of their CPs. The smaller size of the CP of the Vandiperiyar and Valparai isolates as compared to the other isolates could be attributed to the proteolytic degradation of these two proteins. Although the number of amino acid residues in the Vandiperiyar and Kursupara isolates is identical, the former has a larger number of proteolytic cleavage sites, as predicted by the PEPTIDESORT program of the GCG package (data not shown). Such an analysis of the Valparai isolate could not be done, as the CP gene of this isolate has not been sequenced. No protein band could be observed from the Sirsi isolate as it had aggregated irreversibly. The Wayanadu isolate showed multiple bands (figure 3a).

ELISA with anti CdMV IgG that was raised against the Yeslur isolate of CdMV (Jacob and Usha 2002) showed considerable differences in the reaction (table 2). Yeslur isolate showed the maximum reaction since the antiserum was derived from the same strain. All other isolates showed significantly lower values, approximately 1/10th of that with the Yeslur isolate. The negative controls gave average OD values of 0·008 to 0·009. All the isolates gave ELISA readings three times greater than the negative controls. Among these, the isolates from Appangala and Madikeri (Coorg district) showed much higher values indicating serological relatedness with the Yeslur isolate

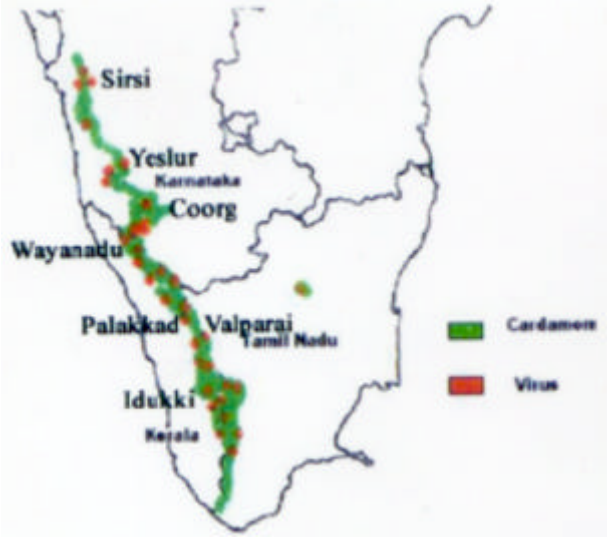

Figure 1. Map showing the cardamom tract and CdMV infected regions.

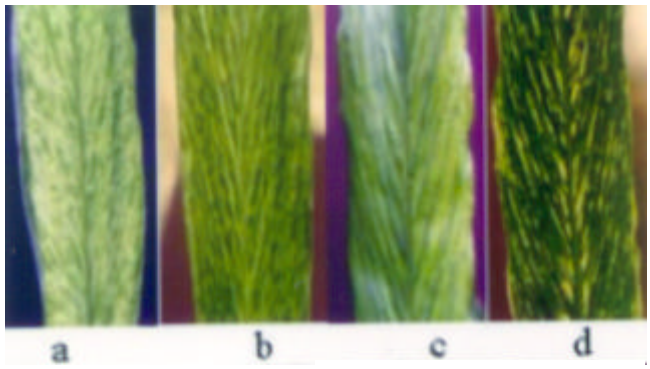

Figure 2. Symptoms produced by some of the CdMV isolates on cardamom leaves. (**a**) Sirsi (subgroup III): whitish green with discontinuous dark green islands. (**b**) Yeslur (subgroup I): light green with prominent cholrotic streaks. (**c**) Appangala (subgroup II): continuous dark green stripes along the veins. (**d**) Kursupara (subgroup V): prominent yellow mosaic patches along the veins.

of CdMV compared to the geographically distinct isolates (table 2).

Western blotting (figure 3b) with the antiserum mentioned above showed a strongly positive reaction with the Yeslur isolate and to a lesser extent with the Coorg strains (Appangala and Thalathamane). There was only a negligible reaction with the other virus isolates (figure 3b).

PCR with CPn2 and Poty1 + resulted in an amplification of a nearly 1 kb DNA fragment from Appangala, Thalathamane, Madikeri, Sirsi, Kursupara, Kattappana and Vandiperyar isolates. But Wayanadu, Palakkad and Valparai isolates failed to give amplification with the same set of primers.

In Southern hybridization, the probe (the cloned CP coding sequence from the Yeslur isolate) hybridized with the PCR products from all the isolates tested (Appangala, Sirsi, Kursupara and Vandiperiyar) (not shown).

Sequences of all the isolates showed a single ORF of 813–819 and a 3′UTR of 200–221 bases (figure 4). Translation of the coding sequences showed CP of 271–273 amino acid residues.

Multiple alignments of amino acid sequences (figure 4a) of CP showed a high level of homology in the core and C terminal region of the CP. On the other hand, the first 40 amino acids from the N-termini showed considerable variations with a minimum identity of 42·5% between Yeslur and Kursupara and a maximum identity of 95% between Appangala and Thalathamane.

The extent of identity between pairs of distinct geographical isolates is not uniform in the two regions namely CP and 3′UTR. For e.g. CP amino acid sequence identity between Vandiperiyar and Thalathamane is 86·7% and that between Kursupara and Kattappana is 93·1%. But the identity at the N terminal portion between the above pairs of isolates is 42·5% and 67·5% respectively and the identity between the 3′UTRs is 70·4% and 84·4% respectively. This confirms the presence of a significant amount of variability within the CP, N terminus of the CP and UTR independent of each other. Existence of such a non-linear variation points towards a divergent evolution among the CdMV strains.

The CP and 3′UTR sequences are commonly used as markers of genetic relatedness of potyviruses (Frenkel *et al* 1989, 1992; Shukla *et al* 1994). The identity between the nucleotide sequences (CP coding region) of the isolates of CdMV ranges from 98·4% (Madikeri vs Appangala) to 75·4% (Thalathmane vs Vandiperiyar). The identity in the CP region reported among the strains of a number of potyviruses is generally higher [90 to 99% with an average value of 95% (Shukla and Ward 1988)]. The thirteen isolates of *Papaya ring spot virus* from more than five countries showed 88% or more identity in their CP (Bateson *et al* 1994, 2002). *Yam mosaic virus* (YMV) is one of the potyviruses known to have a high level of diversity among its strains with only 70% identity in the CP coding sequences (Aleman-Verdaguer *et al* 1997). A recent report (Bousalem *et al* 2000) on the genetic diversity among potyviruses shows YMV to be the most variable potyvirus with respect to the CP sequence. The minimum identity (47%) observed was in the N-terminal region of YMV strains. This was greater than the identity observed in the N-terminal region among the CdMV strains where it is 42·5%. *Zucchini yellow mosaic virus*, one of the most

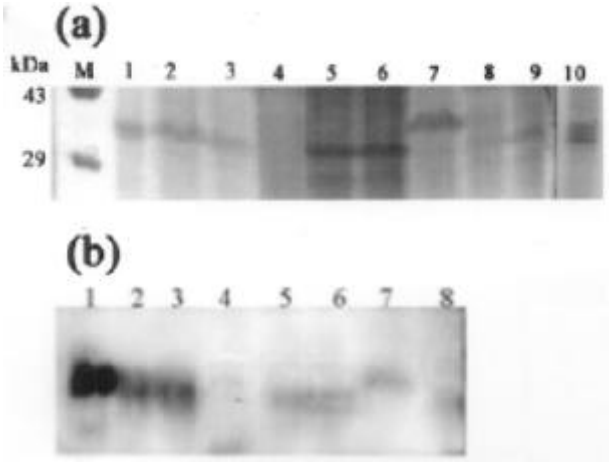

Figure 3. (**a**) SDS PAGE of some CdMV isolates: 1. Yeslur, 2. Madikeri, 3. Appangala, 4. Sirsi, 5. Valparai, 6. Vandiperiyar, 7. Kurusupara, 8. Myladumpara, 9. Kattappana, 10. Wayanadu. (**b**) Western blot of some isolates of CdMV: 1. Yeslur, 2. Madikeri, 3. Appangala, 4. Sirsi, 5. Valparai, 6. Vandiperiyar, 7. Kurusupara, 8. Myladumpara.

Table 2. ELISA reading (at 410 nm) of various isolates of CdMV.

	Dilution of antiserum	
Name of the isolate	1 : 16000	1 : 32000
Sirsi	0·053	0·030
Mudigere	0·14	0·045
Sakalespur	0·580	0·245
Yeslur	1·25	0·620
Appangala	0·400	0·150
Thalathamane	0·150	0·060
Madikeri	0·210	0·103
Valaparai	0·060	0·024
Wayanadu	0·065	0·025
Nelliyampathi	0·080	0·030
Kursupara	0·070	0·026
Kattappana	0·065	0·028
Vandiperiyar	0·077	0·030
Myladumpara	0·075	0·031

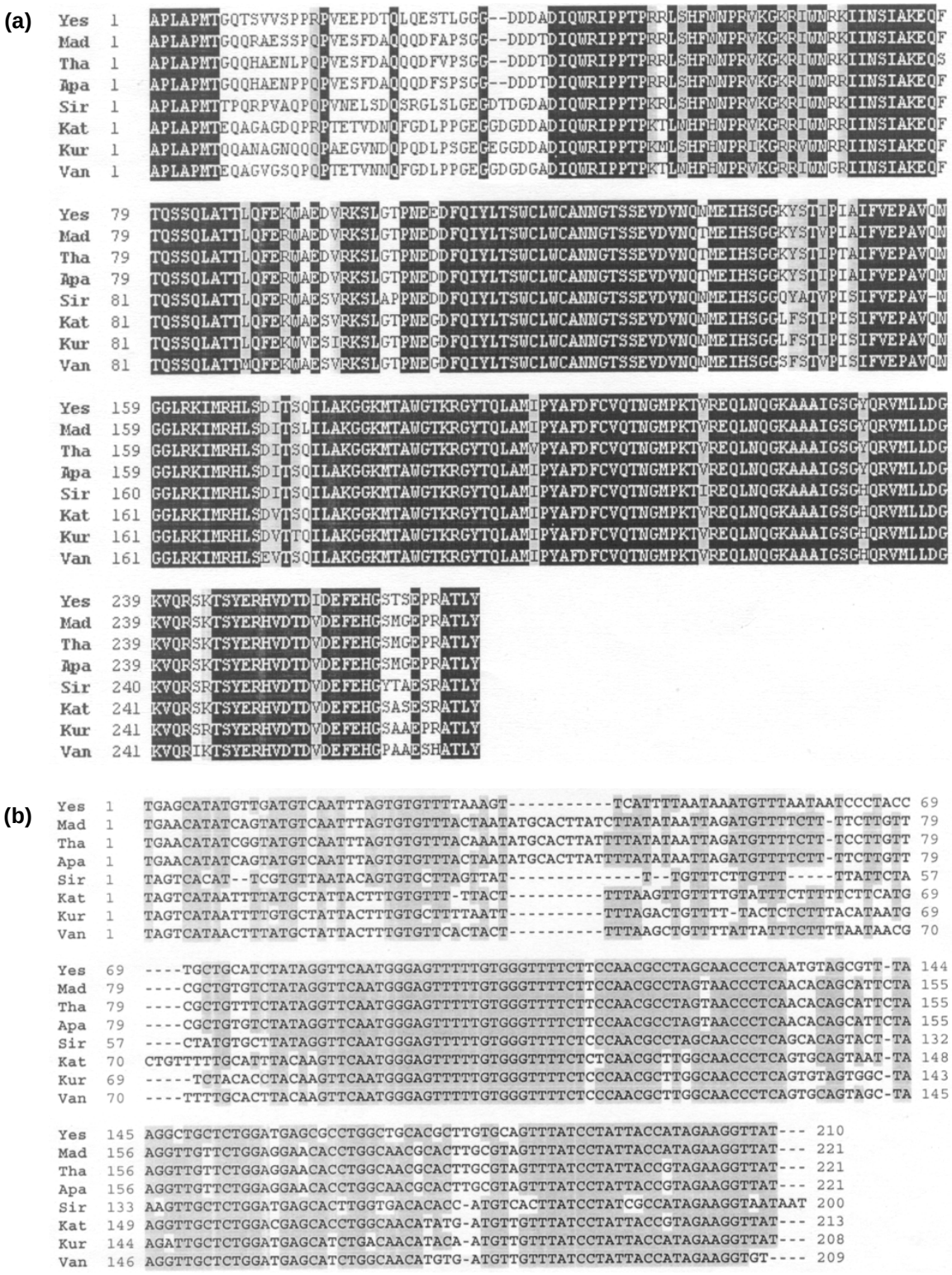

Figure 4. (**a**) Multiple alignment of the amino acid sequence of CP from eight isolates of CdMV. (**b**) Multiple alignment of the nucleotide sequences of the 3′UTR from eight isolates of CdMV. Expansions of abbreviations are given in the legend to table 1. Numbering is with respect to the Yeslur isolate.

studied potyviruses for its strainal variation, showed a divergence of 26% or less in their N-terminal part of their CP among 110 isolates of different geographical origins (Desbiez *et al* 2002).

With respect to the 3′UTR sequence, *Potato virus Y* (PVY) showed the largest diversity, followed by *Bean yellow mosaic virus* (BYMV) and YMV with a mean identity of 87·4%, 91·4% and 92·3% respectively among their strains (Bousalem *et al* 2000). In the case of CdMV the identity among the distant strains in the 3′UTR ranges from 65·5% to 67·3% (table 1). Therefore, based on the current data on the variability among the strains, CdMV can be considered as one of the most variable viruses among the members of *Potyviridae* with respect to the CP and 3′UTR sequences.

The N-terminal portion of the CP (the first forty amino acids) shows the least identity among the isolates, the lowest value being 42·5% between Yeslur and Kursupara isolates. The extent of identity between pairs of distinct geographical isolates is not uniform throughout the sequence under consideration, namely CP and 3′UTR. For instance, the identity of the N-terminal portions between the Vandiperiyar and Thalathmane isolates (pair A) is 42·5% and that between Kursupara and Kattappana (pair B) is 67·5%. However the identity in the entire CP among the members of pair A is 86·7% and that between members of pair B is 93·1%. The identity in the 3′UTR for pair A is 70·4% and for pair B is 84·4%. Such a non-uniform distribution of identities might reflect the possible recombination between the various isolates of the virus.

The isolates from Kerala and Karnataka show conserved sequences within themselves. In the CP, positions 35, 62, 143 and 171 and in the 3′UTR, positions 24, 25, 81, 83, 120, 122 and 174 are conserved within each group (figure 4). However, the variable amino acid residues in the CP of the Kerala and Karnataka isolates are chemically similar in nature, pointing towards the conservation of the structural and functional similarities of the CP of the virus during the course of evolution.

One of the Karnataka isolates (from Sirsi) shows several unique features in common with the Kerala isolates in their CP and UTR sequences but different from the rest of the Karnataka isolates (Yelsur and Coorg). The Kerala and the Sirsi isolates have an amber stop codon (UAG) but the rest of the Karnataka isolates have an ochre (UGA) stop codon (data not shown). The former group has two additional amino acid residues after the 34th position compared to the latter. Also the former group has Gln, Lys, Ser and His at positions 33, 49, 149 and 230 respectively with respect to the numbering of the Yeslur isolate but for the latter group they are Gly, Arg, Ala and Tyr respectively (figure 4a). UTR also shows similar conservations within the above groups at nucleotide positions 835, 836, 862, 946, 952, and 992 with respect to the above numbering.

The phylogenetic tree (figure 5) constructed with the CP and 3′UTR sequences using CLUSTALW (Thompson *et al* 1994) also shows that the Sirsi isolate clusters closer to the Kerala isolates rather than with the rest of the Karnataka isolates. However, geographically Sirsi is the farthest place in Karnataka from Kerala. Cardamom mosaic disease spreads by the propagation of infected rhizomes also (Varma 1962). Therefore, the sequence comparisons might give a clue to the movement of the virus-infected propagules along the cardamom tract between geographically distant plantations. The phylogenetic tree constructed based on the entire CP and that based on the forty amino acid residues from the N terminus of the CP shows a branching pattern (figure not shown) similar to the tree in figure 5. For the tree constructed from the nucleotide sequence alignment of 3′UTR alone, the Sirsi isolate clusters among the Karnataka isolates rather than with the Kerala isolates. This might point to a possible recombination event between the Sirsi and the Kerala isolates. Virus-infected rhizomes carried from one plantation to the other could have resulted in the spread and thereby mixed infection of a plant by two different virus isolates. This could have resulted in the emergence of recombinants as has been postulated for YMV (Bousalem

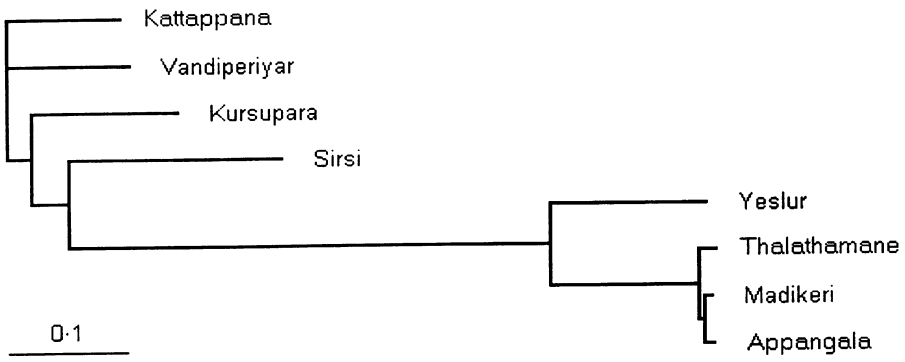

Figure 5. Phylogenetic tree constructed from the nucleotide sequence of the CP coding region and 3′UTR from eight isolates of CdMV.

et al 2000) and *Papaya ringspot virus* (Bateson *et al* 2002). Single and double recombinants have also been detected when squash plants were co-bombarded with mixtures of constructs of *Zucchini yellow mosaic virus* (Gal-On *et al* 1998). A natural recombinant has also been detected by analysis of nucleotide sequences in the potyvirus *Plum pox virus* (Cervera *et al* 1993). Based on the genetic diversity data, the CP(s) of CdMV can be chosen for the development of transgenic cardamom for virus resistance. The CP gene of the Sirsi isolate will be an ideal choice as it lies between both the Karnataka and Kerala isolates and is therefore expected to provide resistance to isolates from both the clusters.

5. Conclusion

The sequence analysis of the CP and 3′UTR of the various geographically distinct isolates of CdMV from south India reveals that CdMV has the maximum genetic diversity within its isolates among the members of *Potyviridae*. The perennial nature of cardamom, its continued cultivation over several centuries and the exchange of infected rhizomes between plantations would have provided ample opportunities for this RNA virus to mutate and recombine.

Acknowledgements

This work was done as part of the Centre for Plant Molecular Biology project funded by the Department of Biotechnology, New Delhi. The Council of Scientific and Industrial Research, New Delhi, is thanked for research fellowships to Jacob and Jebasingh and Ms Archana Somanath is thanked for help in virus extractions. We acknowledge Dr R Naidu and Dr P Rajan for their support for the survey of cardamom plantations. Facilities availed from Genetic Engineering Research Unit and Bioinformatics Center of the School of Biotechnology are gratefully acknowledged. We thank Prof. G Kulandaivelu for providing the swing-out rotor for ultracentrifugation.

References

Aleman-Verdaguer M E, Goudou-Urbino C, Dubern J, Beachy R N and Fauquet C 1997 Analysis of sequence diversity of P1, HC, P3, NIb and CP genomic regions of several yam mosaic potyvirus isolates: implications for intraspecies molecular diversity of potyviruses; *J. Gen. Virol.* **78** 1253–1264

Bateson M F, Henderson J, Chaleeprom W, Gibbs A J and Dale J L 1994 Papaya ringspot potyvirus, isolate variability and the origin of PRSV type P (Australia); *J. Gen. Virol.* **75** 3547–3553

Bateson M F, Lines, R E, Revill P, Chaleeprom W, Ha C V, Gibbs A J and Dale J L 2002 On the evolution and molecular epidemiology of the potyvirus *Papaya ringspot virus*; *J. Gen. virol.* **83** 2575–2585

Bousalem M, Douzery E J P and Fargette D 2000 High genetic diversity, distant phylogenetic relationships and intraspecies recombination events among natural populations of Yam mosaic virus, a contribution to understanding potyvirus evolution; *J. Gen. Virol.* **81** 243–255

Carninci P, Nishiyama Y, Westover A, Itoh M, Nagaoka S, Sasaki N, Okazaki Y, Muramatsu M and Hayashizaki Y 1998 Thermo-stabilization and thermo-activation of thermo-labile enzymes by trehalose and its application for synthesis of full length cDNA; *Proc. Natl. Acad. Sci. USA* **95** 520–524

Cervera M T, Riechmann J L, Martin M T and Garcia J A 1993 3′-Terminal sequence of the plum pox virus PS and o6 isolates: evidence for RNA recombination within the potyvirus group; *J. Gen. Virol.* **74** 329–334

Desbiez C, Wipf-Scheibel and Lecoq H 2002 Biological and serological variability, evolution and molecular variability of Zucchini yellow mosaic virus (ZYMV, potyvirus) with special reference to Caribbean islands; *Virus Res*. (in press)

Farrell R E Jr 1993 RNA *Methodologies*: *A laboratory guide for isolation and characterization* (San Diego, California: Academic Press)

Felsenstein J 1989 PHYLIP-Phylogeny Inference Package (version 3.2); *Cladistics* **5** 164 (http://evolution.genetics.washington.edu/phylip.html)

Frenkel M J, Ward C W and Shukla D D 1989 The use of 3′non coding nucleotide sequences in the taxonomy of potyviruses: application to watermelon mosaic virus 2 and Soybean mosaic virus-N; *J. Gen. Virol.* **70** 2775–2783

Frenkel M J, Jilka J M, Shukla D D and Ward C W 1992 Differentiation of potyviruses and their strains by hybridization with the 3′ non-coding region of the viral genome; *J. Virol. Methods* **36** 51–62

Gal-On A, Meiri E, Raccah B and Gaba V 1998 Recombination of engineered defective RNA species produces infective potyvirus *in planta*; *J. Virol*. **72** 5268–5270

Jacob T and Usha R 2001 3′ terminal sequence analysis of the RNA genome of the Indian isolate of *Cardamom mosaic virus*: A new member of the genus *Macluravirus* of *Potyviridae*; *Virus Genes* **23** 81–88

Jacob T and Usha R 2002 Expression of *Cardamom mosaic virus* coat protein in *Escherichia coli* and its Assembly into filamentous aggregates; *Virus Res*. **86** 133–141

Gibbs A and Mackenzie A 1997 A primer pair for amplifying part of the genome of all potyvirids by RT-PCR; *J. Virol. Methods* **63** 9–16

Mayne M W 1951 A report on cardamom cultivation in South India; *Bulletin* (ICAR, New Delhi) **50** p 62

Shukla D D, Ward C W and Brunt A A 1994 *The Potyviridae* (Wallingford: CAB International)

Shukla D D and Ward C W 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

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

Uppal B N, Varma P M and Capoor S P 1945 Mosaic disease of cardamom; *Curr. Sci*. **14** 208–209

Varma P M 1962 The banana aphid (*Pentalonia nigronervosa* Coq.) and the transmission of ‘katte’ disease of cardamom; *Indian Phytopathol*. **15** 1–10

MS received 18 March 2003; accepted 4 July 2003

Corresponding editor: SHAHID JAMEEL