

Rolling circle amplification-based analysis of Sri Lankan cassava mosaic virus isolates from Tamil Nadu, India, suggests a low level of genetic variability

Akhilesh Kumar Kushawaha¹ · Ramalingam Rabindran² · Indranil Dasgupta¹ 

Received: 22 April 2017 / Accepted: 20 January 2018 / Published online: 1 February 2018
© Indian Virological Society 2018

Abstract Cassava mosaic disease is a widespread disease of cassava in south Asia and the African continent. In India, CMD is known to be caused by two single-stranded DNA viruses (geminiviruses), Indian cassava mosaic virus (ICMV) and Sri Lankan cassava mosaic virus (SLCMV). Previously, the diversity of ICMV and SLCMV in India has been studied using PCR, a sequence-dependent method. To have a more in-depth study of the variability of the above viruses and to detect any novel geminiviruses associated with CMD, sequence-independent amplification using rolling circle amplification (RCA)-based methods were used. CMD affected cassava plants were sampled across eighty locations in nine districts of the southern Indian state of Tamil Nadu. Twelve complete sequence of coat protein genes of the resident geminiviruses, comprising 256 amino acid residues were generated from the above samples, which indicated changes at only six positions. RCA followed by RFLP of the 80 samples indicated that most samples (47) contained only SLCMV, followed by 8, which were infected jointly with ICMV and SLCMV. In 11 samples, the pattern did not match the expected patterns from either of the two viruses and hence, were variants. Sequence analysis of an average of 700 nucleotides from

31 RCA-generated fragments of the variants indicated identities of 97–99% with the sequence of a previously reported infectious clone of SLCMV. The evidence suggests low levels of genetic variability in the begomoviruses infecting cassava, mainly in the form of scattered single nucleotide changes.

Keywords Cassava · Begomovirus · Rolling circle amplification · RCA-RFLP · Genetic diversity

Introduction

Cassava (*Manihot esculenta* Crantz, Family Euphorbiaceae) is widely grown in tropical regions of the world and its tuberous roots are used as a source of dietary carbohydrates in many countries. In India, cassava is grown in substantial amounts in the southern states of Andhra Pradesh, Kerala and Tamil Nadu (1, 2).

Geminiviruses (family *Geminiviridae*) cause significant yield losses in crop plants throughout the tropical and subtropical parts of the world [1–4]. Within the family *Geminiviridae*, members of the genus *Begomovirus* are transmitted by whiteflies of the species *Bemisia tabaci* [5] and consist of one (monopartite) or two (bipartite) circular single-stranded DNA components of about 2.7 kb in size. Cassava mosaic disease (CMD) is widespread in cassava in the African continent, India and Sri Lanka [1, 2, 6, 7]. Two begomoviruses are associated with CMD in the Indian subcontinent; Sri Lankan cassava mosaic virus (SLCMV) and Indian cassava mosaic virus (ICMV) [8–10], SLCMV being more common as compared to ICMV [11–13]. Following the initial report of an infectious SLCMV clone from Adivaram, Kerala, (SLCMV-Adi) on *Nicotiana benthamiana* [10] another cloned DNA of SLCMV, from

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13337-018-0432-x>) contains supplementary material, which is available to authorized users.

✉ Indranil Dasgupta
indranil58@yahoo.co.in

¹ Department of Plant Molecular Biology, University of Delhi, South Campus, Benito Juarez Road, New Delhi 110021, India

² Center for Plant Breeding and Genetics, Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu 641003, India

Attur, Tamil Nadu (SLCMV-Attur) has been recently shown to be infectious on *N. benthamiana* [14].

Apart from spontaneous variation arising due to the error-prone replication mechanism in RNA viruses [15], recombination has also been proposed to be a reason behind the frequent occurrence of variants [16–19], especially in geminiviruses, where, the intergenic region of the DNA has a role in the generation of variation through intermolecular recombination and pseudorecombination during mixed infection of viruses [20–24]. Earlier studies on the variability of cassava-infecting begomoviruses from India used the methods of PCR and PCR-RFLP [11, 13], both of which require prior information on nucleotide sequences of the viruses to be studied. To detect novel begomoviruses or recombinants within the known and novel begomoviruses, sequence-independent detection methods such as rolling circle amplification (RCA) are very useful [25, 26].

To generate additional information on begomoviruses associated with CMD in southern India, samples from cassava plants showing symptoms of CMD were collected from 80 locations in 9 districts of Tamil Nadu, India and RCA-RFLP was used to analyze the molecular diversity. As the first step, a conserved CP gene was amplified by PCR, using CP-gene specific primer (based on the Accession No AJ579307) followed by cloning and sequencing. Variability of amino acid sequences of CP was checked from 12 samples. In the second step, RCA-RFLP technique using the restriction enzymes *EcoRI*, was used to determine variability in the amplified begomoviral DNA patterns from all the samples. Some of the RCA-RFLP generated DNA fragments were cloned, partially sequenced and aligned with SLCMV-Attur genome (Acc. No KC424490) to analyse the mutational changes, protein variability or any hot spot of insertion or deletion. This results indicate that both ICMV and SLCMV display sporadic scattered single nucleotide changes in their DNA, which result in the accumulation of a low level substitutions in the viral proteins. This knowledge will help in developing effective anti-viral strategies to combat the disease with the confidence that it will be applicable uniformly across the regions surveyed, due to the genetic uniformity of the viruses.

Materials and methods

Collections of cassava stem cuttings

A total of more than 100 cassava fields (Supplementary Table 1, Fig. 1) were visited in end of September, 2011. Cassava plants aged three to 6 months of age were selected. The leaves of symptomatic and non-symptomatic

cassava and stem cuttings were collected and carried to the University of Delhi South Campus, New Delhi, India. Stem cuttings were planted in soil in pots in the glass house maintained at 30 °C with 16 h light and 8 h dark periods. Emerging fresh leaves, sprouted from the above stem cuttings were used as a source for the further experiments.

Isolation of genomic DNA from cassava leaves

Total genomic DNA was extracted from glasshouse-grown symptomatic cassava leaves using hot SDS method [27].

Cloning and sequencing of coat-protein genes

Genomic DNA extracted from symptomatic leaves were used to amplify the complete CP gene by PCR, using synthetic abutting oligonucleotide based on published nucleotide sequence of SLCMV-Adi (Acc. No AJ579307) [10]. The amplified DNA were cloned into pTZ57R/T vector (MBI Fermentas). DNA sequencing was performed through services provided at University of Delhi South Campus, New Delhi, India.

Rolling circle amplification (RCA)

RCA was carried out using the TempliPhi™ Amplification Kit (GE Healthcare, Amersham) as per manufacturer's recommendation for amplification and cloning purpose. Ten to twenty nanogram of total genomic DNA were dissolved in 5 µl of sample buffer, denatured for 3 min at 95 °C and cooled down to the room temperature. Then 5 µl of reaction buffer and 0.2 µl enzyme ø29 DNA polymerase were added and incubated at 30 °C for 18–20 h.

Cloning and sequencing of RCA fragments

Digestion with *EcoRI* restriction enzymes (Fermentas) was performed to linearize the concatameric RCA product and the resulting fragment was cloned with and without gel electrophoresis in pTZ57R vector (MBI Fermentas) and partially sequenced using a universal M13 primer [28, 29]. Randomly selected RCA-RFLP fragments, showing unusual banding patterns were sequenced.

Sequence analysis

Nucleotide sequences from the amplified fragments were searched by BLAST at NCBI server (www.ncbi.nlm.nih.gov/) and analysed by the software Gene Runner version 3.05. Nucleotide identities between cloned viral DNA molecule and other selected begomoviruses were analysed by the ClustalW method and NCBI BLAST server [30]. Multiple sequence alignment was performed by mega 6.0



Fig. 1 Map Showing the sites of collection (shown by dark dot) of cassava stem cuttings

software and ClustalW. Protein sequence variability was estimated by using the Protein Variability Server [30, 31].

Results

Collection of samples and observation of symptoms

Both mild and severe symptoms could be observed on naturally-infected cassava plants in the fields visited. However, most of the samples collected from the districts

of Dharmapuri, Erode, Salem and Villupuram showed mild mosaic symptoms. Plants showing symptoms of severe mosaic were few. Symptoms in the samples collected from Cuddalore district were very mild. The types of symptoms displayed are shown in Supplementary Table 1. In most cases, the symptoms displayed at the field were replicated in the leaves sprouted from the stem cuttings grown under glasshouse conditions.

Genetic diversity of cloned CP genes of SLCMV in southern India

Genetic diversity of twelve complete CP genes were randomly cloned from field grown sprouted leaves from stem cuttings collected from the districts of Cuddalore (2), Dharmapuri (2), Salem (4), Theni (1) and Villupuram (3) (Supplementary Table 1, shown in bold). Although attempts were made to clone CP genes from isolates showing severe symptoms, only the above 12 could be cloned, which showed mild symptoms. Upon multiple alignment, out of the 256 amino acid residues, changes detected at only six positions (Supplementary Fig. 1). Out of a total of 12, only five CP sequences showed substitution mutations. One CP sequence (CP7) showed substitutions at three positions. CP10, CP41, CP46 and CP79 showed substitutions in single positions only as compared to SLCMV-attur. In CP7, at three substitution involved a change of a polar amino acid threonine was changed to basic amino acid arginine, non-polar leucine to non-polar isoleucine and N to polar glycine. In CP10, CP41 and CP46 non-polar amino acid alanine were changed to polar serine amino acid, non-polar valine and polar serine respectively. In CP79, charged aspartic amino acid was changed to polar tyrosine (Supplementary Figure 1).

Analysis of cassava infecting begomoviruses by RCA-RFLP

RCA-RFLP was used to investigate the presence of ICMV and/or SLCMV in sprouted leaves of cassava stem cuttings collected from the fields. The patterns obtained after RCA-RFLP were analyzed and compared with that expected from in silico analysis of the nucleotide sequences of ICMV and SLCMV, using the available full length sequences from publicly available nucleotide databases. The expected band patterns based on the sequence of SLCMV-Attur (KC424490) were 1844, 794, 70 and 50 bp for SLCMV DNA-A and 2722 and 16 bp for SLCMV DNA B Acc. No (AJ579308). Three expected bands were 1623, 923 and 120 for ICMV DNA-A (AY730035) and 2406, 221 and 16 bp for ICMV DNA-B (AY730036). Hence, two bands above 2.0 kb indicate the presence of both ICMV and SLCMV. Such patterns were found in eight samples (sample 2, 20, 27, 30, 46, 56, 60 and 78, Fig. 2). Eleven samples produced band patterns, which were not expected (samples 31, 41, 42, 47, 52, 54, 58, 60, 69, 72 and 78), based on the known sequence information. Fourteen samples showed no amplification and the remaining 47 samples indicated the presence of only SLCMV (Fig. 2). There was no sample which indicated the presence of only ICMV.

Thirty one RCA-RFLP generated fragments were randomly cloned and sequenced to an average length of 700 bp. When the sequences were searched for maximum identities with database entries using publicly accessible databases, they showed between 97 and 99% identities with entries described as either as ICMV or SLCMV DNA-A or DNA-B (Supplementary Table 2), isolated from the southern Indian states of Kerala or Tamil Nadu. When the sequenced portions were aligned with the known sequences of SLCMV/ICMV, 15 fragments covered the AC4 gene and 3 covered the AC2 gene completely. This enabled the comparison of the derived amino acid sequences of AC2 and AC4 proteins encoded by the cloned DNA fragments. When these amino acid sequences were studied for variability, 5 amino acid residues were found to be variable in AC2 and 11 in AC4 (Fig. 3a, b), compared to the amino acid sequences of the corresponding viral proteins derived from the infectious clone of SLCMV (SLCMV-Attur, KC424490), although there were an average 2–3 single base changes for each 100 bases analyzed in the nucleotide sequences.

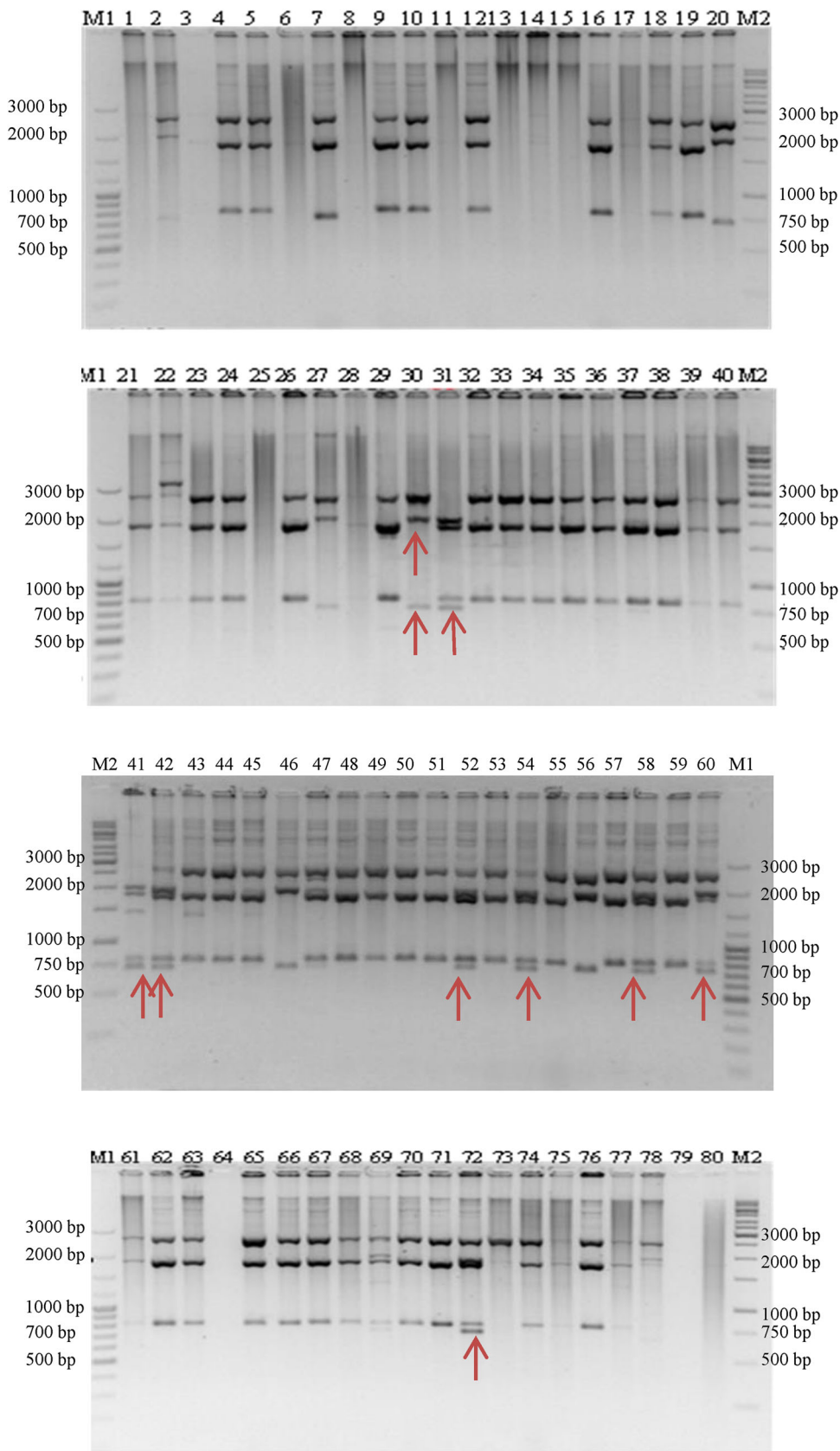
Overall, it can be said that although some samples showing severe symptoms were analysed along with a large number of samples showing mild symptoms, no specific genetic relationship could be established between the symptoms and the genetic nature of the viruses.

Discussion

This work was undertaken to try and generate additional information on the nature and variability of cassava infecting begomoviruses in southern India. An attempt was also made to detect any other begomoviral DNA or satellites DNAs, which may be associated with CMD-affected cassava in samples, other than what has been reported earlier [11, 13]. RCA is able to generate sequence non-specific amplification products derived from begomoviruses in plants [25] and hence, was considered appropriate for this investigation.

With that in mind, symptomatic cassava samples were analyzed for the presence of new begomoviral DNAs by using RCA-RFLP. The choice of *EcoRI* as a selecting enzymes was advantageous because it was capable of producing discrete sets of DNA fragments representing both DNA components of ICMV and SLCMV. The results indicated that a vast majority of the samples analyzed (47 out of 67 which produced patterns) contained only the two SLCMV DNA components and a minor fraction contained the components of both ICMV as well as SLCMV (8). This agreed with the earlier reports of surveys performed using PCR and PCR-RFLP [11, 13] and reinforced the view that SLCMV is more widespread than ICMV in India.

Fig. 2 RCA/*EcoRI* RFLP patterns of begomoviral DNAs associated with cassava mosaic disease. Sample collection sites for 1–80 are shown in Supplementary Table 1. Red arrow indicated the ICMV. M1 = GeneRuler™ 100 bp DNA Ladder and M2 = GeneRuler™ 1 Kb Ladder are size markers (color figure online)



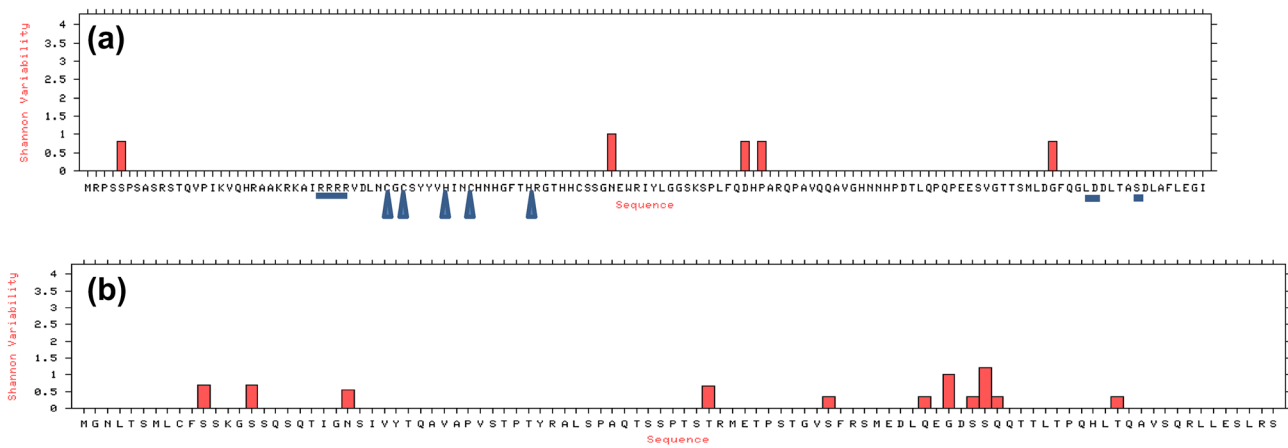


Fig. 3 Sequence variability plot of the AC2 and AC4 amino acid sequences of cassava infecting geminiviruses represented according to Shannon's method [31]. A score higher than 0 indicate variable amino acid positions. Sequence variability ranges from 0 (only one amino acid type is present at that position) to 4.322 (all 20 amino acids are equally represented in that position). **a** The bar and

arrowhead (blue triangle) indicate the locations of the predicted motifs in AC2 (29 RRRR 32 nuclear localization signal), C 37 , C 39 , H 44 , H 54 , and C 47 (promoter activation). Conserved amino residues of the C-terminal activation domains of AC2 in SLCMV-Attur, 121 LDD 123 and 127 S are indicated by bar. **b** Myristoylation motif in AC4 is shown by bar (color figure online)

Interestingly, a study has recently shown evidence of a possible suppression of ICMV accumulation by SLCMV, in case of joint infection of naturally infected cassava plants collected from the state adjoining Tamil Nadu, Kerala; the ICMV emerging only after "recovery" of the plant from SLCMV infection [32]. Our present observation of low incidence of joint infection with both ICMV and SLCMV and previous reports [11, 13] indicated that ICMV can co-exist with SLCMV in many plants, indicating the existence of possibly more interacting relationships between the two viruses, other than that reported [32].

Eleven samples produced RCA-RFLP patterns not expected from the known sequences of ICMV or SLCMV. To determine whether they represent begomoviral sequences hitherto not reported from cassava, nucleotide sequences were generated from 31 such fragments. Which also included some samples showing the above unexpected patterns. These represented various portions of the begomoviruses and showed high sequence identities (97–99%) with known ICMV and SLCMV sequences, indicating them to be variants of known cassava-infecting begomoviruses, not representing any novel begomoviruses. The most likely explanation for the presence of a novel pattern could be random point mutations in the viral DNA bringing about changes in the recognition sites for *EcoRI*.

The low variability of the SLCMV and ICMV DNA mainly manifested as scattered single nucleotide changes. This was evident both in the RCA-derived as well as PCR-derived sequences analyzed. The analysed sequences indicated no incidence of recombination or hot-spots of mutation in the RCA-derived fragments. The sequence analysis also showed the presence of scattered point mutations evenly distributed in the sequenced fragments

(data not shown). To determine whether the above scattered point mutations can result in any major changes at the protein level, the derived amino acid sequences of AC2 and AC4 were aligned with those of SLCMV-Attur. Of the 135 amino acid sequence of AC2, amino acid changes were seen only at 5 positions and of the 100 amino acid AC4, changes were found at 11 positions. This reinforces the conclusion that cassava begomoviruses display low levels of scattered point mutations, unlikely to affect any viral protein. In addition, none of the sequences of the cloned RCA fragments showed any homology to any satellite sequences known to be associated with begomoviruses, making their existence very unlikely in CMD-affected cassava.

The lack of variability in cassava begomoviruses in India is in stark contrast to the evidence presented by other workers [16, 33] showing high levels of genetic diversity in begomoviruses isolated from tomato [16] and various satellites [33]. One possible reason behind the low variability of cassava begomoviruses could be linked to the vegetative propagation of cassava practiced by farmers, resulting in accumulation of begomoviruses representing those which are carried forward along with the infected propagules (stem cuttings) in the locations surveyed. The begomoviruses displaying high variability infect annual plants and are transmitted by whiteflies in the field. This raises interesting questions about the role of insect vectors in the generation of variability in plant viruses, an aspect which needs more in-depth study.

Acknowledgements AKK gratefully acknowledges fellowships from University of Delhi and Department of Biotechnology, Government of India. This work was funded by the Grant Number BT/PR/13094/

PBD/16/866/2005 by Department of Biotechnology, Government of India to ID.

References

1. Thresh JM, Otim-Nape GW, Legg JP, Fargette D. African cassava mosaic disease: the magnitude of problem? *Afr J Root Tuber Crops*. 1997;2:13–9.
2. Patil BL, Fauquet CM. Cassava mosaic geminiviruses: actual knowledge and perspectives. *Mol Plant Pathol*. 2009;10:685–701.
3. Navas-Castillo J, Fiallo-Olive E, Sanchez-Campos S. Emerging virus diseases transmitted by whiteflies. *Annu Rev Phytopathol*. 2011;49:219–48.
4. Lozano-Duran R. Geminiviruses for biotechnology: the art of parasite taming. *New Phytol*. 2016;210:58–64.
5. Sanderfoot AA, Lazarowitz SG. Getting it together in plant virus movement: cooperative interactions between bipartite geminivirus movement proteins. *Trends Cell Biol*. 1996;6:353–8.
6. Abraham A. Farm Bulletin No 17. New Delhi: Indian Council of Agricultural Research; 1956.
7. Alagianagalingam MN, Ramakrishnan K. Cassava mosaic in India. *South Indian Hortic*. 1966;14:71–2.
8. Hong YG, Robinson DJ, Harrison BD. Nucleotide sequence evidence for the occurrence of three distinct whitefly-transmitted geminiviruses in cassava. *J Gen Virol*. 1993;74(Pt 11):2437–43.
9. Saunders K, Salim N, Mali VR, Malathi VG, Briddon R, Markham PG, Stanley J. Characterisation of Sri Lankan cassava mosaic virus and Indian cassava mosaic virus: evidence for acquisition of a DNA B component by a monopartite begomovirus. *Virology*. 2002;293:63–74.
10. Dutt N, Briddon RW, Dasgupta I. Identification of a second begomovirus, Sri Lankan cassava mosaic virus, causing cassava mosaic disease in India. *Arch Virol*. 2005;150:2101–8.
11. Patil BL, Rajasubramaniam S, Bagchi C, Dasgupta I. Both Indian cassava mosaic virus and Sri Lankan cassava mosaic virus are found in India and exhibit high variability as assessed by PCR-RFLP. *Arch Virol*. 2005;150:389–97.
12. Rothenstein D, Briddon RW, Haible D, Stanley J, Frischmuth T, Jeske H. Biolistic infection of cassava using cloned components of Indian cassava mosaic virus. *Arch Virol*. 2005;150:1669–75.
13. Borah BK, Dasgupta I. PCR-RFLP analysis indicates that recombination might be a common occurrence among the cassava infecting begomoviruses in India. *Virus Genes*. 2012;45:327–32.
14. Kushawaha AK, Rabindran R, Dasgupta I. Phylogenetic analysis and biolistic infectivity of a cloned Sri Lankan cassava mosaic virus DNA-A from Tamil Nadu, India on *Nicotiana benthamiana*. *Acta Virol*. 2015;59:57–63.
15. Elena SF, Sanjuan R. Adaptive value of high mutation rates of RNA viruses: separating causes from consequences. *J Virol*. 2005;79:11555–8.
16. Kirthi N, Maiya SP, Murthy MR, Savithri HS. Evidence for recombination among the tomato leaf curl virus strains/species from Bangalore, India. *Arch Virol*. 2002;147:255–72.
17. Ndunguru J, Legg JP, Aveling TA, Thompson G, Fauquet CM. Molecular biodiversity of cassava begomoviruses in Tanzania: evolution of cassava geminiviruses in Africa and evidence for East Africa being a center of diversity of cassava geminiviruses. *Virol J*. 2005;2:21.
18. Sharma S, Rabindran R, Robin S, Dasgupta I. Analysis of the complete DNA sequence of rice tungro bacilliform virus from southern India indicates it to be a product of recombination. *Arch Virol*. 2011;156:2257–62.
19. Rishishwar R, Mazumdar B, Dasgupta I. Diverse and recombinant begomoviruses and various satellites are associated with Bhendi yellow vein mosaic disease of okra in India. *J Plant Biochem Biotechnol*. 2015;24(4):470–5.
20. Hou YM, Gilbertson RL. Increased pathogenicity in a pseudorecombinant bipartite geminivirus correlates with intermolecular recombination. *J Virol*. 1996;70:5430–6.
21. Pita JS, Fondong VN, Sangare A, Otim-Nape GW, Ogwal S, Fauquet CM. Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. *J Gen Virol*. 2001;82:655–65.
22. Martin DP, van der Walt E, Posada D, Rybicki EP. The evolutionary value of recombination is constrained by genome modularity. *PLoS Genet*. 2005;1:e51.
23. Lima AT, Sobrinho RR, Gonzalez-Aguilera J, Rocha CS, Silva SJ, Xavier CA, Silva FN, Duffy S, Zerbini FM. Synonymous site variation due to recombination explains higher genetic variability in begomovirus populations infecting non-cultivated hosts. *J Gen Virol*. 2013;94:418–31.
24. Sobrinho RR, Xavier CA, Pereira HM, Lima GS, Assuncao IP, Mizubuti ES, Duffy S, Zerbini FM. Contrasting genetic structure between two begomoviruses infecting the same leguminous hosts. *J Gen Virol*. 2014;95:2540–52.
25. Haible D, Kober S, Jeske H. Rolling circle amplification revolutionizes diagnosis and genomics of geminiviruses. *J Virol Methods*. 2006;135:9–16.
26. Varsani A, Roumagnac P, Fuchs M, Navas-Castillo J, Moriones E, Idris A, Briddon RW, Rivera-Bustamante R, Murilo Zerbini F, Martin DP. Capulavirus and Grablovirus: two new genera in the family Geminiviridae. *Arch Virol*. 2017;162:1819–31.
27. Dellaporta S, Wood J, Hicks JB. A plant DNA minipreparation: version II. *Plant Mol Biol Report*. 1983;1:19–21.
28. Homs M, Kober S, Kepp G, Jeske H. Mitochondrial plasmids of sugar beet amplified via rolling circle method detected during curtovirus screening. *Virus Res*. 2008;136:124–9.
29. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. 1994;22:4673–80.
30. Wieczorek P, Obrepalska-Steplowska A. The N-terminal fragment of the tomato torrado virus RNA1-encoded polyprotein induces a hypersensitive response (HR)-like reaction in *Nicotiana benthamiana*. *Arch Virol*. 2016;161:1849–58.
31. Garcia-Boronat M, Diez-Rivero CM, Reinherz EL, Reche PA. PVS: a web server for protein sequence variability analysis tuned to facilitate conserved epitope discovery. *Nucleic Acids Res*. 2008;36:W35–41.
32. Karthikeyan C, Patil BL, Borah BK, Resmi TR, Turco S, Pooggin MM, Hohn T, Veluthambi K. Emergence of a latent Indian cassava mosaic virus from cassava which recovered from infection by a non-persistent Sri Lankan cassava mosaic virus. *Viruses*. 2016;8:264.
33. Kumar V, Singh D, Singh AK, Chakraborty S. Molecular diversity, recombination and population structure of alphasatellites associated with begomovirus disease complexes. *Infect Genet Evol*. 2017;49:39–47.