Accepted Manuscript

Isolation and sequence characterization of DNA-A genome of a new begomovirus strain associated with severe leaf curling symptoms of *Jatropha curcas* L.

Sushma Chauhan, Hifzur Rahman, Shaik G. Mastan, Pamidimarri D.V.N. Sudheer, Muppala P. Reddy

PII: S0378-1119(18)30436-0
DOI: doi:[10.1016/j.gene.2018.04.062](https://doi.org/10.1016/j.gene.2018.04.062)
Reference: GENE 42786

To appear in: *Gene*
Received date: 10 November 2017
Revised date: 15 April 2018
Accepted date: 20 April 2018

Please cite this article as: Sushma Chauhan, Hifzur Rahman, Shaik G. Mastan, Pamidimarri D.V.N. Sudheer, Muppala P. Reddy , Isolation and sequence characterization of DNA-A genome of a new begomovirus strain associated with severe leaf curling symptoms of *Jatropha curcas* L.. The address for the corresponding author was captured as affiliation for all authors. Please check if appropriate. *Gene*(2017), doi:[10.1016/j.gene.2018.04.062](https://doi.org/10.1016/j.gene.2018.04.062)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Title:

Isolation and sequence characterization of DNA-A genome of a new begomovirus strain associated with severe leaf curling symptoms of *Jatropha curcas* L

Sushma Chauhan^{1†}, Hifzur Rahman^{2†}, Shaik G. Mastan³, Pamidimarri D. V. N. Sudheer^{4*}, Muppala P. Reddy⁵

¹Department of Chemical and Biochemical Engineering, Dongguk University–Seoul, Seoul 100-715, Republic of Korea

²Department of Plant Biotechnology, Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India-641003.

³Department of Biotechnology, Aditya Degree and PG college, Adikavinannaya University, Laxminarayana Nagar, Kakinada, 533003, Andhra Pradesh, India.

⁴Department of Environmental Engineering, College of Engineering, Ajou University, Suwon, Gyeonggi-do, South Korea

⁵ Plant Stress Genomics and Technology Center, King Abdullah University of Science and Technology, Thuwal 23955-6900, Kingdom of Saudi Arabia

† authors contributed equally

***Corresponding Author:**

Pamidimarri D. V. N. Sudheer

Discipline of Wasteland Research, Central Salt and Marine Chemicals Research Institute (CSIR), G. B. Marg, Bhavnagar 364002, Gujarat, India

Present address:

Department of Environmental Engineering,

College of Engineering, Ajou University, Suwon, Gyeonggi-do, Republic of Korea.

Phone No: +8242-860-3096

E-mail: pdvnsudheer@ajou.ac.kr

Abstract

Begomoviruses belong to the family Geminiviridae are associated with several disease symptoms, such as mosaic and leaf curling in *Jatropha curcas*. The molecular characterization of these viral strains will help in developing management strategies to control the disease. In this study, *J. curcas* that was infected with begomovirus and showed acute leaf curling symptoms were identified. DNA-A segment from pathogenic viral strain was isolated and sequenced. The sequenced genome was assembled and characterized in detail. The full-length DNA-A sequence was covered by primer walking. The genome sequence showed the general organization of DNA-A from begomovirus by the distribution of ORFs in both viral and anti-viral strands. The genome size ranged from 2844 bp–2852 bp. Three strains with minor nucleotide variations were identified, and a phylogenetic analysis was performed by comparing the DNA-A segments from other reported begomovirus isolates. The maximum sequence similarity was observed with Euphorbia yellow mosaic virus (FN435995). In the phylogenetic tree, no clustering was observed with previously reported begomovirus strains isolated from *J. curcas* host. The strains isolated in this study belongs to new begomoviral strain that elicits symptoms of leaf curling in *J. curcas*. The results indicate that the probable origin of the strains is from Jatropha mosaic virus infecting *J. gossypifolia*. The strains isolated in this study are referred as Jatropha curcas leaf curl India virus (JCLCIV) based on the major symptoms exhibited by host *J. curcas*.

Keywords:

begomovirus, biodiesel, *Jatropha curcas* leaf curl India virus (JCLCV), *Jatropha curcas* L

Introduction

Jatropha curcas L, a member of Euphorbiaceae family, has gained significance as a potential biodiesel crop. Biodiesel obtained from *J. curcas* seed oil showed better performance compared to petrodiesel, making it a prime candidate among other commercially available biofuel sources (Sudheer Pamidiemarri et al., 2009; Pamidiemarri et al., 2010). *J. curcas* is promoted as an alternative non-conventional energy source because of the present energy crisis and current environmental issues. *J. curcas* is a promising crop that can grow in marginal areas, including drought-affected lands and semi-arid tropics. Several studies have been conducted on the short-term and long-term performance of *J. curcas* grown in marginal lands (Bressan et al., 2011a). As a member of the Euphorbiaceae family, *J. curcas* contains latex and a significant amount of phorbol esters, which makes the plant naturally resistant to the majority of disease pests. However, long-term field studies reported that *J. curcas* is susceptible to viral infections. Most *J. curcas* germplasms showed similar symptoms such as leaf curl and mosaic disease, which are caused by begomoviruses. The manifestation of leaf curl showed the upward curling of leaf, leaf vein clearing, reduced leaf size and necrosis in old leaves. Under severe infected conditions, plants showed stunted growth and resulted in no fruit formation or seed setting (Nigam et al., 2014).

Begomoviruses (Family: Geminiviridae, genus: Begomovirus) is one among highly pathogenic virus group predominantly distributed in tropical and subtropical regions. The viruses in this family possess either monopartite or bipartite genome (Gutierrez, 1999). Bipartite genome containing viruses contains two circular DNA components, namely DNA-A and DNA-B, which share a common region (CR) of ~200 nucleotides. In begomoviral genome, the genes/ORFs are distributed on both genomes (in the case of bipartite genomes). The genes are aligned in both strands overlapping to one another. DNA-A harbours gene AV1 (*i.e.* coat protein) on the viral strand; three genes AC1 (replication initiation protein [Rep]), AC2 (transcription activator protein [TrAP]), and AC3 (replication enhancer protein) on complementary strand (Fauquet et al., 2008). DNA-B contains the gene BV1 (nuclear shuttle protein) in the viral strand and the gene BC1 (movement protein) in the complementary strand. Recent studies revealed that begomoviruses tend to co-evolved with host plants, resulting in genetic diversity *via* the recombination and exchange of genomes among themselves. This phenomenon has led to the emergence of potent pathogenic viral strains with increased virulence and extended host range (Wang et al., 2014)(Ramkat et al., 2011a; De Bruyn et al., 2012). Hence, a comprehensive study on the characterization of each viral strain in different crop plants is needed for the protection and improvement of crops. If neglected, begomoviruses could have a huge negative impact on field plantation and crop productivity (Ramkat et al., 2011b).

Since *J. curcas* being a potential biodiesel source, precast measures need to be developed for healthy agricultural practice. Begomovirus infections of *J. curcas* has been reported by several experimental field stations (Raj et al., 2008; Srivastava et al., 2015b; SK et al., 2016) and they were found to be a potential threat. Hence, the detailed characterization of the begomovirus genome is required to prevent crop damage. Srivastava *et al.* recently reported a begomovirus strain that showed leaf yellow mosaic disease in *J. curcas* (Srivastava et al., 2015b). However, in our field studies (CSIR-CSMCRI, Bhavnagar, Gujarat, India), the primary symptom of leaf curling was observed to have characteristic features of begomoviral infection. Hence, the present study was aimed to isolate viral DNA from infected

J. curcas and characterize it at the genomic level. Isolated sequences were used to study the phylogeny and evolution by comparing to sequence information available in the database. A comparative study was conducted on the full-length sequence of DNA-A and important structural genes (i.e. coatprotein [AC1] and replication associated proteins[AC2]). Comprehensive conclusions were derived from the results.

Materials and methods

Genomic DNA extraction

Genomic DNA was extracted using the CTAB protocol as described by Sudheer *et al.* (Sudheer Pamidimarri et al., 2009b) with the following minor modifications. The samples were collected from established *J. curcas* plants in the experimental field (21°75'N, 72°14'E) of Central Salt and Marine Chemicals Research Institute, Bhavnagar, Gujarat, India. Leaves with severe symptoms of leaf curling were carefully collected. About 0.1 g of leaf tissue was ground in liquid nitrogen and taken into 2 mL microcentrifuge tube. To the ground sample, 0.5 mL of extraction buffer (2% CTAB, 100 mM Tris-HCl, 3.5 M NaCl, 20 mM EDTA, 0.2 M β -Mercaptoethanol, 2% PVP, pH 8.0.) was added and mixture was incubated at 65°C for 90 min. The sample was then extracted with an equal volume of chloroform:isoamyl alcohol (24:1), and the supernatant was transferred to a new tube. The obtained supernatant was precipitated with 80% of ethanol for 10 min and centrifuged at 12,000 rpm for 20 min. The pellet was air dried and dissolved in 100 μ L of ultrapure water.

Amplification of viral DNA-A segment

The DNA-A segment was amplified by primers designed in a conserved region (5'-GGGGCCAGGATTGCAGAGGAAGATTGTCGGAATG-3') in the rep protein region (ACI). Among the seven samples collected, three samples showed the expected amplification size (~2.9 kb), which were purified using the QIAquick Gel Extraction Kit (Qiagen, Germany). The purified fragments were ligated into pTZ57R (Insta) T/A cloning vector (MBI Fermentas, USA) according to the manufacturer's protocol, and the recombinant plasmids were transformed into a competent *E. coli* (DH5 α) by the heat shock method (Green and Sambrook, 2012). The cells were then spread on Luria-Bertani (LB) agar selection plates containing appropriate concentrations of ampicillin, IPTG, and X-gal and then incubated at 37°C for 14 hr. The putative clone selection was carried out based on blue/white selection and checked for insert by colony PCR using an M13 universal primer set. Recombinant plasmids were extracted using the GenElute Plasmid Miniprep Kit (Sigma, USA). The sequencing analysis was performed using the ABI PRISM 3100 Genetic Analyzer (Applied Biosystems) using M13 universal primers (IDT, USA). The full-length sequence was obtained by primer walking *via* designing primers for every ~600 bp forward. All three sequences showed variations in nucleotide sequence, hence they were named as JCLCIV1, 2 and 3. The annotated sequences were submitted to the NCBI-GenBank database (NCBI accession numbers: JCLCIV1-GU451249, JCLCIV2-KU667096, and JCLCIV3-KU667097).

DNA-A sequence analysis and annotation

The sequences obtained by primer walking were assembled with the help of BioEdit DNA editing software (Version 7.2). The genome annotation was done by comparing with the begomoviruses reported in the NCBI database. Sequence mining was done from the database and sequences in FASTA format were downloaded for phylogenetic analysis. The sequences were aligned and edited with the help of BioEdit software. The genetic distances among the sequences were obtained with the help of MEGA6 software (Felsenstein, 1985), and the mean genetic distance was calculated by the sum of individual genetic distances divided by the number of samples. Similarities in the sequence of DNA and proteins were obtained by the MUSCLE analysis, which was available in the EMBL-EBI online service facility. Based on the maximum parsimony method including bootstrapping (1,000 replicates), the phylogenetic tree was obtained with the help of MEGA6 analysis software (Tamura et al., 2007). The phylogram was rooted using sweet potato leaf curl virus (KJ476509).

Detection of recombination events

Putative recombinant events were identified using the recombination detection program (RDP4, version 4.0). The strains listed in Table 1 were included in the analysis and analysis was conducted by taking 1,000 bootstrap replicates. The same analysis was conducted using the Croton yellow vein mosaic virus (EU727086), which was found to be strongly clustered in the phylogenetic analysis. For analysis, all other options were put at default settings specified in the software.

Results and discussion

J. curcas as a biodiesel source, has attracted the attention of several researchers. With the aim of establishing the species as a potential crop for oil production in semi-arid tropical land, CSMCRI India conducted major field studies in various conditions of dry and the semi-arid tropical land (Bressan et al., 2011b). The primary symptom of severe leaf curling (Fig. 1a), which was spread quickly by insects, was observed in our field plantation. It was suspected to be caused by begomoviral infection based on the symptoms. Begomoviral infection in *J. curcas* was first reported in 2004 in the southern part of India (Raj et al., 2008; Gao et al., 2010), and it was subsequently reported in the northern Indian provinces and other parts of the globe (Kashina et al., 2012; Snehi et al., 2012). Although previous studies were reported on begomoviral infection in *J. curcas*, the present study includes the characterization of the viral strain that caused severe leaf curl symptoms in our field plantation. This study provides preliminary information about the significant variability and pathogeny of these potent viral strains. In this study, leaf samples that showed leaf curling symptoms were collected and then subjected to viral DNA isolation, sequencing and characterization.

Viral DNA Isolation and sequencing

Leaf curling, the major symptom of begomovirus infection, was reported in a field survey of CSMCRI *J. curcas* plantations (Central Salt and Marine Chemicals Research Institute, Bhavnagar, Gujarat, India, experimental field [21_750N, 72_140E]) conducted in 2009. Seven samples with severe leaf curl symptoms were collected from infected plants in selected regions and then used for DNA extraction. The extracted DNA was subjected to PCR amplification with primers designed to amplify DNA-A. Among the seven samples, three samples displayed amplification with the

expected fragment size (Fig. 1b). The amplified fragments were cloned into the T-vector and sequenced. The sequence data obtained from three amplicons showed minor differences in size. The strains JCLCIV1 and 2 showed differences in the size of four nucleotides (GU451249 with 2844bp and KU667096 with 2848bp), which were confined to the non-coding region. However, sample JCLCIV3 (KU667097) showed a size of 2852bp, with either four or eight nucleotides size more than other isolates. Moreover, the differences were not confined to the non-coding regions. Because all three samples showed differences in their sequences, all the sequences were included in the further analysis. Further samples sequenced showed similarity with either GU451249 or KU667096.

Genome structure and organization of DNA-A segment

The sequence of DNA segment from isolated viral strains showed DNA sizes ranging from 2844–2852. In a BLAST search (Blastn, NCBI database), the full-length sequence of the isolated virus genome showed a similarity of 95-98% with virus strains isolated from the *J. gossypifolia* host. The database search also confirmed that the amplified segment was DNA-A. The maximum similarity of 83% was observed in the Croton yellow vein mosaic virus (EU727086), whose host also belongs to the Euphorbiaceae family. The full-length sequence of JCLCIV viral strains showed 65% similarity with *Jatropha* leaf yellow mosaic Katarniaghat virus (JN135236 and JN698954) (Srivastava et al., 2015b) and 72.24% identical to the *Jatropha* yellow mosaic India virus (FJ177030) (Snehi et al., 2012). In the present study, the isolated strain sequences did not show significant homology with the sequences reported to cause leaf mosaic symptoms in *J. curcas*. Hence, the strains from this study were referred as *Jatropha curcas* leaf curl India virus (JCLCIV). JCLCIV showed higher homology with other begomovirus sequences isolated from various species of *Jatropha*. The present strain showed a significant similarity to the DNA sequences and structural organizations of other characterized begomoviruses with bipartite genomes. Hence, the JCLCIV strains isolated in the present study were considered bipartite. The gene sequences in the virus genome segment were distributed in both directions on viral and anti-viral strands. In the viral (sense) strand contained two ORFs (AV1 and AV2), and the anti-viral (antisense) strand had four ORFs (AC1, AC2, AC3 and AC5). All genes are overlapping, which is the characteristic feature of begomovirus. The genes of the sense and antisense strands were segmented into separate regions. One predicted ORF AC5, which is designated as C5 protein, whose function has not been determined till date, and it was found to be integrated within AV1 and AV2 in the antisense direction.

Among the three isolates, two strains (GU451249 and KU667096) showed differences in the sizes of four nucleotides, which were confined to the non-coding region. The differences were detected at the 16th position ('T' insertion) and the 65th position ('GAA' insertion). However, both were inserted as repeat sequences, and they showed no significant effects in the structure and the functional aspects of the virus. No differences between the two sequences were observed in the nucleotide sequence in the coding regions. When KU667097 was compared with the other two strains, GU451249 & KU667096, we found changes in both the coding and non-coding regions, leading to changes in the amino acid sequence. Among these changes, the genes in the sense strand genes did not show any variation, whereas in the antisense strand AC2 and AC1 genes showed differences. In the case of AC1, there was a minor difference at the C-terminal amino acid sequence. However, in the case of AC2, there was a major difference at the beginning of the 30 amino acid sequence at the N-terminal end. The protein BLAST search (Blastp suite, NCBI database) using the N-

terminal amino acid sequences of AC2 did not show similarities to the other begomovirus AC2 protein sequences. Sequences GU451249 and KU667096 showed similarities to the AC2 protein sequences of other reported isolates of *Jatropha* leaf curl virus and *Jatropha* mosaic virus. The overall genome structure and gene distribution are shown in Figure1c. The study was extended to examine the diversity of the coat protein and the replication associated protein (rep), which were then compared with the diversity of full-length sequence. The mean genetic distance in whole genome DNA-A segment was found to be 0.804; whereas, the coat protein and rep protein showed genetic distances of 0.345 and 0.327, respectively. Similar results were observed in other ORFs, indicating that the coding sequences were less diverse than the non-coding sequences (García-Andrés et al., 2007).

Genetic similarity analysis of JCLCIV with the reported begomoviruses

The JCLCIV DNA-A segment sequence showed the maximum similarity with Croton yellow mosaic India virus (FN645901). The genetic similarity with *Jatropha* leaf curl viruses in the NCBI nucleotide BLAST showed the maximum identity of 98% with the *Jatropha* mosaic Lucknow virus (KF652018) and minimum identity of 94% and 96% with the *Jatropha* leaf curl virus DNA (LC080676). However, no previous studies reported on these isolates. Two strains from the host *J. curcas* (Srivastava et al., 2015b) were reported recently; however, they were found to be very different from the strain sequences isolated in this study. The analysis of genetic distance and the generation of the pairwise genetic distance matrix was performed based on the reported begomovirus sequences (Table 1). The highest genetic similarity (0.833) was seen in the Euphorbia yellow mosaic virus, which was also clustered in the phylogenetic tree (Fig. 2). Strains that were isolated from *J. curcas* (Srivastava et al., 2015b) showed significantly less genetic similarity (0.682) as well as less nucleotide identity (65%) (Table 1), which aligns with the results of the phylogenetic analysis conducted in the present study. The phylogenetic analysis showed that JCLCIV belongs to an old world begomovirus group (Fauquet et al., 2008).

Genetic diversity within begomoviral strains infecting *Jatropha* species

Among the viral strains discussed in this section, the genome DNA-A segment showed significant variations in the size ranging from a minimum of 2743 (FJ177030) to a maximum of 2859 (LC080676). These variations were not observed in the coding region. The overall genetic distance among the viral strains infecting *Jatropha* species was found to be 0.338, whereas the mean pairwise genetic distance was found to be 0.174. The strains isolated in this study showed significant similarity with *Jatropha* mosaic Lucknow virus, which has been well characterized and isolated from ornamental species of *Jatropha* (Ala-Poikela et al., 2005; Polston et al., 2014; Srivastava et al., 2015a). The major nucleotide variations observed were transitions/transversions rather than deletions/insertions. The estimated transition/transversion bias (R) was found to be 0.93, which indicated little preference for either transition or transversion. No correlation was observed between genetic distance and host species divergence. Overall, genetic diversity among the viruses infecting different species of *Jatropha* showed narrow genetic divergence. However, the strains isolated from the host with yellow mosaic symptoms (JN135236 and JN698954) showed major diversion and did not cluster with any strains isolated in this study (Srivastava et al., 2015b).

The phylogenetic analysis showed all strains grouped together except *Jatropha* yellow mosaic India virus (isolated from the host *J. gossypifolia*), and the *Jatropha* leaf yellow mosaic virus isolates (isolated from the host *J. curcas*) (Fig. 3). The same observation was made in the previous phylogram (Fig. 2). This result indicated that the origin of *Jatropha* leaf yellow mosaic virus (FJ11030, JN135236 and JN698954) was distinct from the others and might have evolved from a common ancestral viral strain. To examine this a recombinant event test was conducted. The results showed a recombinant event found with strain AJ97126, the Tomato Yellow leaf curl virus reported in China. In this analysis, the strains isolated in the present study also showed a recombinant event with AJ97126. Surprisingly, the divergent strain FJ11030, the *Jatropha* yellow mosaic India virus also showed a recombinant event with AJ97126. The results indicated that the same recombinant event occurred in Y15934, GU19958 with EU38457, which was expected to be a minor parent. This result demonstrated that these two diverse groups infecting the *Jatropha* species might have originated from a common ancestral strain in a single recombinant event and then evolved separately.

Genetic similarity and phylogenetics of JCLCIV

The analysis of genetic similarity showed that, the present isolated strains had the highest percentage of similarity to Croton yellow vein mosaic virus (EU727086). The strains that were recently isolated and characterized from the *J. curcas*, *Jatropha* leaf yellow mosaic katarniaghat virus (JN698954 & JN135236) and *Jatropha* yellow mosaic India virus (FJ177030) did not show any significant homology with JCLCIV. As in other begomoviruses where the nucleotide difference reflected amino acid variation, JCLCIV strains also showed a similar tendency (Table 1).

A phylogenetic tree was constructed with 29 begomovirus isolates, including three JCLCIV isolates from the present study. Sweet potato leaf curl virus was used to root the phylogram because it was reported to be a distinct group to the Tomato Yellow leaf curl virus (AJ971265 & AJ971524), which showed the highest similarity with the present isolates (Pardina et al., 2012). The phylogram showed no significant major grouping, but it formed multiple minor clades with three to four members in each clade. The Euphorbia yellow mosaic virus (FN435995) was clustered with the present JCLCIV isolates in the phylogram. The clerodendron yellow mosaic virus (EF408037), which showed significant similarity to JCLCIV, was joined as a separate branch with JCLCIV strains and Euphorbia yellow mosaic virus (FN435995). The recently reported virus isolates from *J. curcas*, *Jatropha* leaf yellow mosaic virus strains (JN135236 and JN698954) showed considerable genetic distance and were distinctly separated from the present isolates. In contrast, the members of *Jatropha* mosaic Lucknow virus (KC513823, KF652078 and KF652077) and other uncharacterized members (KM023145, KF652077, LC080676, EU798996, FJ177030, JN698954 and JN135236) were clustered with the JCLCIV isolates (Figure S1, supporting material). These results suggested that the leaf curl virus isolated in our study and the isolated putative uncharacterized isolates in the gene bank might have the same ancestral origin but evolved separately.

To determine the ancestral origin and evolution of JCLCIV and the *Jatropha* yellow mosaic virus, a further recombinant event detection analysis was conducted. The probable recombinant parent obtained with JCLCIV and JYMV in the previous analysis was taken, and the recombinant event test was conducted with the other begomovirus genome DNA-A segment (Martin and Rybicki, 2000). The analysis clearly showed a major recombination event found with the

Tomato Yellow curl virus (AJ971265) (Figure 3). To verify this, the same analysis was conducted using other isolates clustered with the *Jatropha* leaf curl viruses in the phylogenetic tree (Figure 3). The results revealed that the same recombination event occurred in the similar region (Figure 4 b,c). The analysis of the Croton yellow vein mosaic virus (EU727086) showed that the recombinant event occurred with the same major parent Tomato Yellow curl virus (AJ971265) and the JCLCIV. The recombinant region was also found to be similar. However, the recombinant region was found to be much narrow between JCLCIV and the Tomato Yellow curl virus (AJ971265) (Figure2).

The recombinant event test was also performed on JYMV (JN698954 and JN135236) reported by Srivastava *et al.*, (2015). The results showed a major recombinant event with Chilli leaf curl virus (HM992939) (Figure S2, supporting material). Surprisingly, the Tomato Yellow curl virus (AJ971265) and Chilli leaf curl virus (HM992939) were the major parents found in JCLCIV and JYMV. This result suggests that these viral strains are phylogenetically close. These findings were also correlated with their close genetic distance (0.02) and phylogenetic grouping. However, the recombination events detected were in the different regions, proximal in case of JCLCIV and distal in the case of JYMV (positions are only based on the liner sequence used in the analysis; note that the begomovirus genome is circular). The results imply that they could have originated from distinct parents with common ancestral origin. When the analysis was conducted with uncharacterized isolated strains from host *Jatropha* *sp.*, the results showed cross-strain recombination in multiple positions among the isolates (data not shown). The reason might be that JCLCIV evolved from the *Jatropha* leaf curl virus, whose host is *J. gossypifolia*. The JLCV infecting *J. gossypifolia* could have the ability to infect *J. curcas*. Since these species are phylogenetically very close, which was revealed in our previous studies (Sudheer Pamidimarri *et al.*, 2009; Sudheer Pamidimarri *et al.*, 2009a); JLCV might have evolved to JCLCIV during the adaptation with host *J. curcas* or vice versa. Since, *J. gossypifolia* is a native species in India, the prior condition is more probable.

In summary, the DNA-A segments of JCLCIV from host *J. curcas* showing severe leaf curl symptoms were isolated and characterized in this study. The sequencing results showed that the strains were phylogenetically similar to Euphorbia yellow mosaic virus (FN435995). The Tomato yellow curl virus (AJ971265) and the Chilli leaf curl virus (HM992939) were the major ancestral parents for JCLCIV and JYMS respectively revealed by recombinant event analysis. These results were also correlated with phylogenetic analysis. The recombinant events detected in this study clearly indicated that these two strains could have originated from distinct parents of probable common ancestral origin. When the same analysis was applied to the uncharacterized strains isolated from the host *Jatropha*, the results showed that the strains had cross-strain recombination in multiple positions among the isolates. The results showed that JCLCIV could have originated from JLCV upon host adaptation. The present study provides information about the origin and the evolution of JCLCIV, which will help in designing preventive measures and to generate viral resistant cultivars of *J. curcas* plant using genetic modification and other biotechnological approaches.

Acknowledgements

We are thankful to the Council for Scientific and Industrial Research, New Delhi, India for its financial support for the SRF and research associate fellowships.

References

- Ala-Poikela, M., Svensson, E., Rojas, A., Horko, T., Paulin, L., Valkonen, J.P.T. and Kvarnheden, A., 2005. Genetic diversity and mixed infections of begomoviruses infecting tomato, pepper and cucurbit crops in Nicaragua. *Plant Pathology* 54, 448-459.
- Bressan, R., Reddy, M., Chung, S., Yun, D., Hardin, L. and Bohnert, H., 2011a. Stress-adapted extremophiles provide energy without interference with food production. *Food Security* 3, 93-105.
- Bressan, R.A., Reddy, M.P., Chung, S.H., Yun, D.J., Hardin, L.S. and Bohnert, H.J., 2011b. Stress-adapted extremophiles provide energy without interference with food production. *Food Security* 3, 93-105.
- De Bruyn, A., Villemot, J., Lefeuvre, P., Villar, E., Hoareau, M., Harimalala, M., Abdoul-Karime, A.L., Abdou-Chakour, C., Reynaud, B., Harkins, G.W., Varsani, A., Martin, D.P. and Lett, J.-M., 2012. East African cassava mosaic-like viruses from Africa to Indian ocean islands: molecular diversity, evolutionary history and geographical dissemination of a bipartite begomovirus. *BMC Evolutionary Biology* 12, 228.
- Fauquet, C.M., Bridson, R.W., Brown, J.K., Moriones, E., Stanley, J., Zerbini, M. and Zhou, X., 2008. Geminivirus strain demarcation and nomenclature. *Archives of Virology* 153, 783-821.
- Felsenstein, J., 1985. Confidence Limits on Phylogenies: An Approach Using the Bootstrap. *Evolution* 39, 783-791.
- Gao, S., Qu, J., Chua, N.-H. and Ye, J., 2010. A new strain of Indian cassava mosaic virus causes a mosaic disease in the biodiesel crop *Jatropha curcas*. *Archives of Virology* 155, 607-612.
- García-Andrés, S., Accotto, G.P., Navas-Castillo, J. and Moriones, E., 2007. Founder effect, plant host, and recombination shape the emergent population of begomoviruses that cause the tomato yellow leaf curl disease in the Mediterranean basin. *Virology* 359, 302-312.
- Green, M.R. and Sambrook, J., 2012. *Molecular Cloning: A Laboratory Manual*, 4th Edition. CSH press Cold Spring Harbor, New York.
- Gutierrez, C., 1999. Geminivirus DNA replication. *Cellular and Molecular Life Sciences CMLS* 56, 313-329.
- Kashina, B.D., Alegbejo, M.D., Banwo, O.O., Nielsen, S.L. and Nicolaisen, M., 2012. Molecular identification of a new begomovirus associated with mosaic disease of *Jatropha curcas* L. in Nigeria. *Archives of Virology* 158, 511-514.
- Martin, D. and Rybicki, E., 2000. RDP: detection of recombination amongst aligned sequences. *Bioinformatics* 16, 562-563.
- Nigam, K., Srivastava, S., Kumar, D., Verma, Y. and Gupta, S., 2014. Detection and analysis of leaf curl virus from *Jatropha*. *International Journal of Scientific and Research Publications* 4, 1-6.
- Pamidimarri, D.V.N.S., Mastan, S., Rahman, H. and Reddy, M., 2010. Molecular characterization and genetic diversity analysis of *Jatropha curcas* L. in India using RAPD and AFLP analysis. *Molecular Biology Reports* 37, 2249-2257.
- Pardina, P.R., Luque, A., Nome, C., Colomba, E.L., Delgado, S.F. and Di Feo, L., 2012. First report of Sweet potato leaf curl virus infecting sweet potato in Argentina. *Australasian Plant Disease Notes* 7, 157-160.
- Polston, J.E., Londoño, M.A. and Capobianco, H., 2014. The complete genome sequence of New World *jatropha* mosaic virus. *Archives of Virology* 159, 3131-3136.
- Raj, S.K., Snehi, S.K., Kumar, S., Khan, M.S. and Pathre, U., 2008. First molecular identification of a begomovirus in India that is closely related to Cassava mosaic virus and causes mosaic and stunting of *Jatropha curcas* L. *Australasian Plant Disease Notes* 3, 69-72.
- Ramkat, R., Calari, A., Maghuly, F. and Laimer, M., 2011a. Occurrence of African cassava mosaic virus (ACMV) and East African cassava mosaic virus – Uganda (EACMV-UG) in *Jatropha curcas*. *BMC Proceedings* 5, P93-P93.

- Ramkat, R.C., Calari, A., Maghuly, F. and Laimer, M., 2011b. Biotechnological approaches to determine the impact of viruses in the energy crop plant *Jatropha curcas*. *Virology Journal* 8, 386-386.
- SK, S., SS, P., G, G., V, S., SK, R. and V, P., 2016. The Current Status of New Emerging Begomovirus Diseases on *Jatropha* Species from India. *Journal of Plant Pathology & Microbiology* 7, 2157-7471.
- Snehi, S.K., Srivastava, A. and Raj, S.K., 2012. Biological Characterization and Complete Genome Sequence of a Possible Strain of Indian cassava mosaic virus from *Jatropha curcas* in India. *Journal of Phytopathology* 160, 547-553.
- Srivastava, A., Kumar, S., Jaidi, M. and Raj, S.K., 2015a. Characterization of a novel begomovirus associated with yellow mosaic disease of three ornamental species of *Jatropha* grown in India. *Virus Research* 201, 41-49.
- Srivastava, A., Kumar, S., Jaidi, M. and Raj, S.K., 2015b. Molecular characterization of a new begomovirus associated with leaf yellow mosaic disease of *Jatropha curcas* in India. *Archives of Virology* 160, 1359-1362.
- Sudheer Pamidimarri, D.V.N., Pandya, N., Reddy, M. and Radhakrishnan, T., 2009. Comparative study of interspecific genetic divergence and phylogenetic analysis of genus *Jatropha* by RAPD and AFLP. *Molecular Biology Reports* 36, 901-907.
- Sudheer Pamidimarri, D.V.N., Chattopadhyay, B. and Reddy, M., 2009a. Genetic divergence and phylogenetic analysis of genus *Jatropha* based on nuclear ribosomal DNA ITS sequence. *Molecular Biology Reports* 36, 1929-1935.
- Sudheer Pamidimarri, D.V.N., Singh, S., Mastan, S., Patel, J. and Reddy, M., 2009b. Molecular characterization and identification of markers for toxic and non-toxic varieties of *Jatropha curcas* L. using RAPD, AFLP and SSR markers. *Molecular Biology Reports* 36, 1357-1364.
- Tamura, K., Dudley, J., Nei, M. and Kumar, S., 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. *Molecular Biology and Evolution* 24, 1596-1599.
- Wang, G., Sun, Y., Xu, R., Qu, J., Tee, C., Jiang, X. and Ye, J., 2014. DNA-A of a highly pathogenic Indian cassava mosaic virus isolated from *Jatropha curcas* causes symptoms in *Nicotiana benthamiana*. *Virus Genes* 48, 402-405.

Table1.Sequence similarity at nucleotide (nt) and amino acid (aa) level showed by viral DNA-A segment JCLCIV isolated from *J. curcas* compared with the selected begomovirusmembers.

Name	Accession number	DNA-A	AV1(nt)	AV1(aa)	AC1(nt)	AC1(aa)
Jatropha curcas leaf curl India virus	GU451249	-	-	-	-	-
Jatropha curcas leaf curl India virus	KU667096	99.91	100	100	99.82	99.82
Jatropha curcas leaf curl India virus	KU667097	99.88	100	100	99.32	96.23
Jatropha mosaic Lucknow Virus	KC513823	94.73	98.23	98.41	94.23	97.36
Jatropha mosaic Lucknow Virus	KF652077	95.21	98.26	97.16	93.24	97.14
Jatropha mosaic Lucknow Virus	KF652078	93.41	97.35	97.84	92.40	96.25
Gossypium punctatum mild leaf curl virus	EU384575	70.93	66.39	63.09	76.44	79.94
Cotton leaf curl Alabad virus	AJ002452	70.16	65.85	62.66	75.16	80.5
Gossypium darwinii symptomless virus	NC011804	66.62	65.3	61.37	67	67.04
Mesta yellow vein mosaic virus	KJ462083	67.37	66.8	58.8	67.92	69.27
Mesta yellow vein mosaic virus	KJ462074	67.27	66.94	58.8	67.46	69.92
Cotton leaf curl Kokhran virus	AJ002449	65.71	67.21	61.37	66.76	69.55
Euphorbia leaf curl virus	KT259282	67.66	65.57	60.52	69.81	70.87
Euphorbia leaf curl virus	NC005319	67.74	65.98	60.52	69.72	71.43
Pedilathus leaf curl virus	AM712436	67.53	66.76	61.8	69.8	69.36
Tobacco curly shoot virus	GU199584	66.84	66.94	60.94	67.86	70.19
Papaya leaf curl virus	Y15934	70.82	67.21	61.37	76.7	78.27
Papaya leaf curl virus	FM955601	71.82	68.95	61.37	77.16	80.78
Croton yellow vein mosaic virus	AJ507777	71.14	68.17	58.8	76.15	80.5
Croton yellow vein mosaic virus	EU727086	74.07	67.35	57.94	80.85	81.62
Ageratum yellow vein Sri Lanka virus	AF314144	67.22	66.39	60.52	68.97	71.59
Chilli leaf curl virus	KP235539	67.3	66.26	62.66	68.78	69.92
Jatropha yellow mosaic India virus	FJ177030	72.24	64.89	60.52	73.2	69.64
Chilli leaf curl virus	HM992939	69.48	65.16	60.94	77.21	45.35
Tomato yellow leaf curl China virus	AJ971265	69.98	66.12	59.66	76.7	78.55
Tomato yellow leaf curl China virus	AJ971524	69.87	66.12	59.66	76.43	78.55
Clerodendron yellow mosaic virus	EF408037	70.44	67.21	60.94	75.23	78.83

Asystasia begomovirus	NC016575	63.25	65.84	57.33	76.44	61
Euphorbia Yellow mosaic virus	FN435995	61.17	67.75	61.4	59.81	56.41
Sweet potato leaf curl Henan4 virus	KJ476509	57.86	58.7	42.67	64	60.96
Jatropha leaf yellow mosaic Katarniaghat virus	JN698954	65.14	66.12	63.09	67.62	66.86
Jatropha leaf yellow mosaic Katarniaghat virus	JN135236	65.4	66.12	63.09	67.62	66.86

Figure 1

Figure 2

Figure3

Figure 4

Figure legends:

Figure 1. (a) Infected leaf of *J. curcas* showing symptoms of leaf curling of variable intensity by the three strains of JCLCIV included in the present study, (i) JCLCIV1, (ii) JCLCIV2, (iii) JCLCIV3; (b) Genome of JCLCIV amplified from infected leaves; (c) Genomic organization of Jatropa leaf curl India virus showing predicted ORFs along with their position. Direction of arrows shows their orientation in sense (clockwise direction) and antisense strand (anti-clockwise direction). IR, intergenic region AV1, coat protein; AV2, pre-protein; AC1, replication associated protein; AC2, transactivator protein; AC3, replication enhancer protein; AC5, C5 protein (predicted to be replication enhancer protein), size scale in the base pair.

Figure 2. Phylogenetic tree showing the genetic relationships of Jatropa curcas leaf curl India virus (JCLCV) with other begomoviral isolates from various plant hosts. The distance tree is generated by neighbour joining method with 1000 bootstrap replicates and values were given in the nodes of the tree.

Figure 3. Genetic distance tree of begomovirus isolated from the host species *Jatropha*. The distance tree is generated by the neighbour joining method with 1,000 bootstrap replicates; the values are given in nodes of the tree.

Figure 4. Possible recombinant regions detected in JCLCIV by RDP3 considering viral sequences included in Table 1. (a) The line diagram shows the structure of the genome and genes arranged in the genome of the virus; (b) Recombination event shown by JCLCIV (GU451249) with Tomato Yellow curl virus (AJ971265); (c) Recombinant event detected by Croton yellow mosaic virus (EU727086) with Tomato Yellow curl virus (AJ971265).

Abbreviations	Expansion
JCLCV	Jatropha curcas leaf curl India virus
CR	Common region
Rep	replication initiation protein
TrAP	transcription activator protein
CTAB	Cetyl trimethylammonium bromide
EDTA	Ethylenediaminetetraacetic acid
RDP	recombination detection program
ORF	open reading frame

Highlights:

- Begomovirus eliciting acute leaf curling symptoms were identified from India. The strains isolated in this study are referred as *Jatropha curcas* leaf curl India virus (JCLCIV) based on the major symptoms exhibited by host *J. curcas*
- The isolated strains showed bipartite genome with genome size of 2844 bp–2852 bp.
- Isolated strains showed maximum similarity with Euphorbia Yellow mosaic virus (FN435995).
- The results indicate that the probable origin of the strains is from *Jatropha* mosaic virus infecting *J. gossypifolia*.
- recombinant event analysis revealed that, these two strains could have originated from distinct parents of probable common ancestral origin
- The present study provides information about the origin and the evolution of JCLCIV reported from India.