TOMATO LEAF CURL PATNA VIRUS CAUSING TOMATO LEAF CURL DISEASE IN BANGLADESH

MOHAMMAD NURUL ISLAM*, ANEESA ANSARI1 AND RH SARKER

Department of Botany, University of Dhaka, Dhaka-1000, Bangladesh

Keywords: ToLCV-JB, Monopartite, Tomato, Geminivirus, ToLCV-Patna

Abstract

Tomato leaf curl virus (ToLCV) has appeared as a potential threat to the tomato production in the world. ToLCV, a member of the family Geminiviridae may contain either bipartite or monopartite genome. The genetic nature of a monopartite ToLCV isolate characterized from the tomato leaf curl diseased samples of Jamalpur district, Bangladesh (ToLCV-JB) has been reported. The products of rolling circle amplification (RCA) were digested, cloned and sequenced. Sequence analysis revealed the features of begomovirus genome organization in the ToLCV-JB isolate, containing six open reading frames. BLAST analysis showed 100% sequence similarity with tomato leaf curl Patna virus (EU862323.1) and more than 80% similarity with other reported monopartite begomoviruses. Hence, the virus isolate was registered as Tomato leaf Curl Patna virus-[Bangladesh:Jamalpur:2014] isolate ToLCV-JB (Genebank Accession: KU933675.1) according to the suggestion of NCBI. Recombination analysis also did not show any genetic exchange between ToLCV-JB and ToLCV-Patna virus. Moreover, they belong to the same cluster as observed in phylogenetic analysis. The present work suggests the possibility of cross-border spread of ToLCV-Patna viruses without mutation and this could pose a threat to tomato production in Bangladesh as well as in the Asian continent.

Introduction

Geminiviruses are usually transmitted by insects and cause significant economic and yield losses to many crops (Moffat 1999, Varma and Malathi 2003). Members of the family geminiviridae are divided into nine genera, i.e. Becurtovirus, Begomovirus, Capulavirus, Curtovirus, Eragrovirus, Grablovirus, Mastrevirus, Topocuvirus and Turncurtovirus based on viral vector, genome organization and host range (Zerbini et al. 2017). These viruses possess either monopartite (DNA-A) or bipartite (DNA-A and DNA-B) genome, which is single stranded of approximately 2.7 kb size and contains ~220 bp common region (CR) (Ahmed 1991, Navot 1991). DNA-A is required for viral replication, encapsidation, transmission and spread in its host (Ahmed 1991). Studies have confirmed that when monopartite genome only was inoculated into a susceptible host, it gave rise to disease symptoms like those of the field (Dry et al. 1993).

Tomato leaf curl virus (ToLCV) belongs to the genus Begomovirus, transmitted by whitefly (Bemisia tabaci) and is commonly found in tropical and subtropical climates (Chakraborthy et al. 2003, Fazeli et al. 2009). ToLCV is a major causative agent for the serious reduction of tomato production if transmitted at a very early stage of plant development. The infected plants become severely stunted, leaflets curls upward, become distorted and have yellow margins. The discovery of ToLCV in Bangladesh together with its spread has provoked the need to manage ToLCV situations (Kamal et al. 2015). Both monopartite (Green et al. 2001) and bipartite (Maruthi et al. 2005, 2006) ToLCV isolates have been reported earlier in Bangladesh with limited information. Therefore, the study of the identity, distribution, molecular variability, and the threat that these arising geminiviruses pose to tomato production in Bangladesh has become very important. In this report, molecular identification of a novel monopartite virus causing tomato leaf curl disease in Bangladesh showing 100% sequence similarity with tomato leaf curl Patna virus was described.

*Author for correspondence: mnurul@du.ac.bd. 1Department of Genetic Engineering and Biotechnology, Jessore University of Science and Technology, Jessore-7408, Bangladesh.
Materials and Methods

Leaf samples were collected from the tomato growing areas of Jamalpur, Bangladesh. More than 90% of the plants in the field showed severe symptoms of tomato leaf curl diseases. Begomovirus infection was suspected due to the presence of whiteflies. The tomato plants exhibiting severe leaf curling, stunted growth, distortion of leaf lamina and reduced leaf size symptoms associated with ToLCV infection were selected for sampling. The genomic DNA was extracted from the collected leaves using modified CTAB method of Islam and Mukherjee (2011).

The plant genomic DNA was subjected to rolling circle amplification (RCA) analysis using the TempliPhi 100 amplification kit (GE Healthcare) for amplifying the viral genome. The begomovirus specific, BGCPF (5’-TGTGARGGYCCWTGYAARGTYCA-3’) and BGCPR (5’-TASARGCATGWTACANGCCATATAC-3’) primers (Pandey et al. 2010) were used to amplify the coat protein gene from the RCA products in a standard PCR reaction for initial confirmation. The RCA products were digested with BamHI restriction enzyme to get a band of ~2.7 kb size for cloning into suitable vector. The digested products were checked on 1% agarose gel run at 80V. The digested RCA product was finally cloned into the pGreen cloning vector using BamHI restriction site.

Using the vector backbone primers, [pGFwd (5’- CTCTTCGCTATTACGCCAG -3’) and pGRev(5’-GTGAGCGGATAACAATTT -3’)], the viral genome (~2.7 kb) was amplified. The thermocycling conditions included 1 cycle at 95ºC for 5 mins., 30 cycles at 94ºC for 1 min., 58ºC for 30 secs., 72ºC for 3 min. and 1 cycle of 72ºC for 10 mins. The amplified product was purified using the PCR purifying kit (Favorgen) and the nucleotide sequence was determined by primer walking on both strands.

The sequence data of the virus were compared with the other reported isolates of the ToLCV from the Indian subcontinent and from other part of the world. The open reading frames (ORFs) of the nucleotide sequences were analyzed using NCBI-ORFfinder tool (Wheeler et al. 2003) and Bioedit (version 7.2) software (Hall 1999). All the ORFs that could express proteins were used in protein-protein BLAST searches (BLASTp) to identify potential homologs of these in the NCBI non-redundant protein sequences database and a nested search strategy to restrict the search to geminivirus protein sequences (Altschul 1990).

Nucleotide/amino acid sequence similarities were checked by aligning it with other reported ToLCV isolates using Clustal W (Version 1.82) (Thompson et al. 1994) and MEGA 6.0 softwares (Tamura et al. 2013). Using multiple sequence alignment, the phylogenetic trees were constructed by the neighbor joining majority rule consensus. The degree of similarity shared by the full length genome of various viruses was graphically represented using a neighbor joining tree (1000 bootstrap replicates).

For the identification of potential recombination event, analysis was performed on DNA-A sequences using RDP4 software (Darren et al. 2015). The sequences were also used to study the pairwise comparison between different isolates and an average profile for the selected cluster of viruses that was used to calculate for these two categories taking 100 nt along the genome sequence. The analyses were performed with default settings for the detection methods, a Bonferroni-corrected p-value cutoff of 0.05. The approximate recombinant sequence(s) and breakpoint positions inferred for every detected potential recombination event were manually checked and adjusted where necessary using the extensive phylogenetic and recombination signal analysis features available in RDP4. Once a set of unique recombination events was identified, a breakpoint map containing the positions of all clearly identifiable breakpoints was compiled. The sequences used for the analysis include KU955675, EU862323, U38239.1 and EU910140.1.
Results and Discussion

Tomato leaf curl disease (ToLCD) is a whitefly-transmissible disease which has seriously affected the production of tomato in Bangladesh (Alam et al. 2016). Tomato plants were observed and collected from Jamalpur district that showed severe dwarfing, accompanied with curling, deformation, mottling and puckering of the leaves, typical for ToLCD (Fig. 1 a,b).

To identify viral agents potentially associated with ToLCD, the plant genomic DNA was isolated from the symptomatic leaves. Using the genomic DNA, the Φ29 polymers based rolling circle amplification (RCA) was performed to amplify the ToLCV genome. Restriction digestion of RCA product with BamHI showed amplification of ~2.7 kb (Fig. 2a). The presence of begomovirus in the infected plant samples was confirmed by the amplification of DNA in a earlier report (Pandey et al. 2010). For further confirmation, the RCA products were subjected to PCR using begamovirus coat protein specific forward and reverse primers. The PCR product showed the amplification at expected size of the amplicon (~500 bp) (Fig. 2b), again indicating the presence of begamovirus infection of tomato plants which were also confirmed by sequence comparison in BLAST search analysis.

Full-length sequencing of a single clone of approximately 3.0 kb amplicon was obtained using the vector backbone specific primers, i.e. pGFwd and pGRev. BLASTN analysis using the reference sequence as a query revealed that it showed 99% similarity with tomato leaf curl Patna virus isolate ToCPV-IN, complete genome and 97% similarity with tomato leaf curl Patna virus segment DNA-A, complete sequence. Therefore, the sequenced genome of this study was considered to be a monopartite and is referred to as DNA-A since it shows the maximum homology to the DNA-A of previously characterized begamovirus. The ToLCV-JB is named as tomato leaf curl patna virus-[Bangladesh, Jamalpur, 2014] and is registered to NCBI database under the accession number KU955675.

At this stage, the presence of additional genomic components of ToLCV, as well as that of DNA satellites in begomoviruses with monopartite genomes cannot be excluded. To gain the insights into this, RCA product was digested with BamHI, SacI and HindIII restriction enzymes and predicted to cut at a single site of the ToLCV-JB circular DNA. The digested products generated a single fragment of about 2.7 kb showing the same electrophoretic mobility as the PCR product of the full-length ToLCV-JB DNA molecule used as a loading control. Additional bands...
were never detected in the agarose gel stained with ethidium bromide and no fragments of 3.0 kb or smaller sizes were observed in the negative controls. This supports that ToLCV-JB is a newly identified geminivirus with a monopartite genome.

![Fig. 2. Molecular confirmation of the presence of ToLCV-JB as causal agent of the ToLCD through (a) Rolling Circle Amplification (RCA) of total DNA from three samples digested with BamHI (lanes 1, 2 and 3), (b) PCR of the begomovirus coat protein specific primers showing amplification at 500 bp, in both cases M denotes 1 kb plus size marker lane, (c) genome map of tomato leaf curl virus showing the sense (AV1 and AV2) and complementary (AC1, AC2, AC3, AC4 and AC5) strand ORFs.]

The nucleotide sequence was translated into protein using BioEdit software 7.0 and also by using the program ORF Finder (https://www.ncbi.nlm.nih.gov/orffinder/). Seven ORFs were identified on both the strands of ToLCV-JB circular DNA (Fig. 2c, Table 1). Half of the complementary sense strand codes for two putative proteins while the antisense strand codes for the remaining four proteins by an intergenic region (IR) containing transcription regulatory sequences including TATA boxes at positions 176 - 179 and 2827 - 2830 for transcription of negative strand and positive strand, respectively) and the nine-nucleotide sequence (TAATATTAC, at positions 2601-2609), conserved in almost all geminiviruses and positioned within a short palindromic sequence forming a stem-loop structure.

Based on BLASTP analysis and pairwise comparisons with related geminiviruses, the predicted V1 protein of ToLCV-JB had the highest amino acid sequence identity with the coat protein (CP) of tomato leaf curl Patna virus (E-value: 1.48e-57). The conserved domain helps in the formation of the protein coat and in assembling the viral particle, was identified at the position 156 - 929aa of ToLCV-JB V1 protein. Also, ToLCV-JB AV2 shared the highest sequence identity with the cognate V2 movement protein of tomato leaf curl Patna virus with a conserved domains at 208 - 909 (pfam 01524 and E-value of 1.93e-37). In the sense strand, AC5 was also found to be overlapping AV1 and AV2 regions in an anti-sense orientation and contains the conserved region of 289 - 414 (pfam 08464 and E-value of 7.25e-12) similar to the replication initiator (Rep) associated proteins in the geminivirus (Table 1). Although, its function is still unknown.

The ORF AC1 of ToLCV-JB (1384 - 2460 bp) codes for the replication initiator protein (Rep) of geminiviruses and it is essential for the preparation of a replisome-specific initiator enzyme and synthesis of the replisome. The AC4 region codes for putative proteins comprising of 309 aa for which no homologue protein in the database was found. AC2 and AC3 help for transcription activator protein and replication enhancer protein, respectively and both of the region overlaps that of AC1 protein.
Table 1. ORFs and encoded proteins in the ToLCV-JB genome.

<table>
<thead>
<tr>
<th>ORF</th>
<th>Nucleotide position</th>
<th>Orientation</th>
<th>No. of amino acids</th>
<th>BLASTP (E-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AV1</td>
<td>156-929</td>
<td>Sense</td>
<td>257</td>
<td>Geminivirus coat protein/nuclear export factor BR1 family, (1.48e-57)</td>
</tr>
<tr>
<td>AV2</td>
<td>50-346</td>
<td>Sense</td>
<td>98</td>
<td>Geminivirus V1 protein; Disruption of the V1 gene in Tomato yellow leaf curl virus (TYLCV), (1.93e-37)</td>
</tr>
<tr>
<td>AC1</td>
<td>1384-2460</td>
<td>Complementary</td>
<td>358</td>
<td>Geminivirus Rep catalytic domain (3.46e-48)</td>
</tr>
<tr>
<td>AC2</td>
<td>1071-1478</td>
<td>Complementary</td>
<td>135</td>
<td>The AC2 gene product transactivates expression of TGMV coat protein gene, and BR1 movement protein. (5.73e-15)</td>
</tr>
<tr>
<td>AC3</td>
<td>926-1330</td>
<td>Complementary</td>
<td>134</td>
<td>Replication enhancer protein (3.91e-28)</td>
</tr>
<tr>
<td>AC4</td>
<td>2007-2315</td>
<td>Complementary</td>
<td>102</td>
<td>C4 protein in Tomato yellow leaf curl geminivirus (4.49e-25)</td>
</tr>
<tr>
<td>AC5</td>
<td>138-413</td>
<td>Complementary</td>
<td>91</td>
<td>Geminivirus AC4/5 conserved region (7.25e-12)</td>
</tr>
</tbody>
</table>

Domains and E-value obtained from Pfam analysis.

All the ORFs of ToLCV-JB were used to compare with the ORF of the published data (Table 2). The amino acid sequences of ToLCV-JB virus was almost identical (98%) to that of Tomato leaf curl Patna virus while the full genome showed 98% similarity (Saeed and Samad 2016) which is a monopartite virus. Moreover, the full length genome of the characterized isolate shows 76 - 80% homology of ToLCV-Ban4 (AF165098) another well characterized monopartite ToLCV (Muniyappa et al. 2000). The coat proteins of the ToLCV-JB isolate displayed 77 - 98% homology with other ToLCV isolates and contains two basic domains of KVRRR and KR at the N-terminus. The pre-coat protein (AV2) region has the closest homology with ToLCV-Patna (EU862323.1) of 98% as shown in Table 2. The coat protein gene of whitefly transmitted geminiviruses typically ends with a double stop codon (TAATAA), the first and second nucleotide, respectively of the stop codon of the complementary AC3 gene (Zhou et al. 1998, Pandey et al. 2010) as observed in the case of ToLCV-JB, a previously characterized monopartite ToLCV (data not shown).

The AC2 protein region of ToLCV-JB has the closest homology again with ToLCV-Patna of 98% as shown in Table 2. The alignment of the AC2 protein with other viruses revealed that like Tomato leaf curl yellow AC2 protein (Wezel et al. 2002), a putative zincfinger motif C36-X1-C38-X7-C46-X6-H53-X4-H58C59 and four potential phosphorylation sites (T52, S61, Y68, and S74) are present. The identity of AC4 protein of ToLCV-JB has 100% similarity with ToLCV-Patna again proving the presence that it is a monopartite (Saeed and Samad 2016).

From the recombination analysis, it is clear that there is no recombination between ToLCV-Patna and ToLCV-JB (Fig. 3). Although, ToLCV-Patna viruses was used as a major parent backbone for ToLCV-Kerala and minor parent for ToLCV-Banglore. Based on BLASTP searches and pairwise comparisons of amino acid sequences of several proteins and nucleotide sequences of full-length genomes, ToLCV-JB was identified as the geminivirus most closely related to ToLCV-Patna (Table 2 and Fig. 3). In fact, when phylogenetic trees were generated by the maximum-likelihood method using the full genome encoded by ToLCV-JB and representative members of all genera in the family Geminiviridae, ToLCV- JB and ToLCV-Patna clustered together in the same branch, clearly separated from all the other geminiviruses (Fig. 4).
<table>
<thead>
<tr>
<th>Accession</th>
<th>Virus species/virus isolates</th>
<th>Identity species (%)</th>
<th>AV1-aa (%)</th>
<th>AV2-aa (%)</th>
<th>AC1-aa (%)</th>
<th>AC2-aa (%)</th>
<th>AC3-aa (%)</th>
<th>AC4-aa (%)</th>
<th>AC5-aa (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EU862323.1</td>
<td>Tomato leaf curl Patna virus segment DNA-A</td>
<td>97</td>
<td>98</td>
<td>98</td>
<td>98</td>
<td>95</td>
<td>95</td>
<td>100</td>
<td>97</td>
</tr>
<tr>
<td>AJ810157.1</td>
<td>Stachytopheta leaf curl virus - [Hn34]</td>
<td>84</td>
<td>87</td>
<td>89</td>
<td>85</td>
<td>81</td>
<td>85</td>
<td>91</td>
<td>89</td>
</tr>
<tr>
<td>EU910140.1</td>
<td>Tomato leaf curl Kerala virus isolate ToLCV-K5</td>
<td>83</td>
<td>81</td>
<td>85</td>
<td>88</td>
<td>-</td>
<td>-</td>
<td>94</td>
<td>85</td>
</tr>
<tr>
<td>KU955998.1</td>
<td>Tomato leaf curl Bangladesh virus isolate Dhk</td>
<td>82</td>
<td>-</td>
<td>-</td>
<td>88</td>
<td>-</td>
<td>-</td>
<td>97</td>
<td>-</td>
</tr>
<tr>
<td>FM164938.1</td>
<td>Tomato leaf curl Pakistan virus,</td>
<td>82</td>
<td>-</td>
<td>-</td>
<td>88</td>
<td>77</td>
<td>79</td>
<td>94</td>
<td>-</td>
</tr>
<tr>
<td>DQ116884.1</td>
<td>Tomato leaf curl Pakistan virus isolate Rahim Yar Khan 1 clone PT7</td>
<td>82</td>
<td>80</td>
<td>-</td>
<td>88</td>
<td>78</td>
<td>80</td>
<td>94</td>
<td>-</td>
</tr>
<tr>
<td>EU910141.1</td>
<td>Tomato leaf curl Kerala virus isolate ToLCV-K3</td>
<td>82</td>
<td>80</td>
<td>85</td>
<td>87</td>
<td>-</td>
<td>-</td>
<td>94</td>
<td>85</td>
</tr>
<tr>
<td>EF175733.1</td>
<td>Radish leaf curl virus</td>
<td>82</td>
<td>-</td>
<td>-</td>
<td>86</td>
<td>78</td>
<td>79</td>
<td>94</td>
<td>-</td>
</tr>
<tr>
<td>U38239.1</td>
<td>Tomato leaf curl virus - Bangalore II</td>
<td>80</td>
<td>81</td>
<td>-</td>
<td>81</td>
<td>79</td>
<td>80</td>
<td>80</td>
<td>-</td>
</tr>
<tr>
<td>FJ487911.1</td>
<td>Euphorbia leaf curl virus isolate Fujian</td>
<td>79</td>
<td>80</td>
<td>-</td>
<td>81</td>
<td>77</td>
<td>78</td>
<td>84</td>
<td>-</td>
</tr>
<tr>
<td>DQ989326.1</td>
<td>Chilli leaf curl virus-India [India: Papaya: 2015]</td>
<td>79</td>
<td>-</td>
<td>-</td>
<td>80</td>
<td>78</td>
<td>79</td>
<td>82</td>
<td>-</td>
</tr>
<tr>
<td>FJ859571.1</td>
<td>Bhendi yellow vein Bhubaneswar virus isolate OYBH1U</td>
<td>78</td>
<td>-</td>
<td>-</td>
<td>82</td>
<td>-</td>
<td>-</td>
<td>86</td>
<td>-</td>
</tr>
<tr>
<td>DQ629101.2</td>
<td>Tomato leaf curl New Delhi virus isolate ToLCND-CTS</td>
<td>78</td>
<td>-</td>
<td>-</td>
<td>81</td>
<td>77</td>
<td>77</td>
<td>79</td>
<td>-</td>
</tr>
<tr>
<td>AY297924.1</td>
<td>Tomato leaf curl Iran virus from Iran</td>
<td>78</td>
<td>78</td>
<td>-</td>
<td>79</td>
<td>79</td>
<td>79</td>
<td>80</td>
<td>85</td>
</tr>
<tr>
<td>DQ116881.1</td>
<td>Pepper leaf curl Bangladesh virus</td>
<td>78</td>
<td>-</td>
<td>-</td>
<td>78</td>
<td>78</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>AJ507777.1</td>
<td>Croton yellow vein mosaic virus</td>
<td>78</td>
<td>-</td>
<td>-</td>
<td>81</td>
<td>77</td>
<td>78</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DQ852623.2</td>
<td>Tomato leaf curl Banglore virus - [India: Kerala II: 2005]</td>
<td>76</td>
<td>77</td>
<td>-</td>
<td>77</td>
<td>-</td>
<td>-</td>
<td>79</td>
<td>-</td>
</tr>
</tbody>
</table>
In the last few years, identification of divergent geminiviruses in tomato plants is a common issue, but the complete identification of ToLCV in Bangladesh has not been reported. However, the presence of ToLCV in Bangladesh has been reported by Kamal et al. (2015). In this study, we report the identification and characterization of novel monopartite geminivirus which infects the tomato plant of Jamalpur district of Bangladesh. A circular DNA was amplified from the plant genomic DNA using the rolling amplification mechanism. It is a technique which can detect the presence of DNA A/B or beta-satellite without any prior knowledge of the presence of the virus. The amplified viral DNA was digested with BamHI to form a single stranded DNA, which can be visualized at ~2.7 kb. The coat protein is the most conserved region of the ToLCV genome (Fauquet et al. 1995, Wyatt and Brown 1996), therefore the BGCP forward and reverse degenerate primer pairs were used to amplify the conserved region of ToLCV coat protein genes, as it was designed to amplify universal begamovirus coat protein gene (Pandey et al. 2010). The digested RCA product was cloned into a vector and sequenced. The homology based searches in databases for similar sequences and comparisons with other viruses (Table 2) were consistent with the possibility that such circular DNA was the genomic component of a geminivirus related to ToLCV-Patna, a divergent geminivirus recently identified in India (Saeed and Samad 2016).

The genomic organization of ToLCV-JB is shared by all the other monopartite geminiviruses. Moreover, the identity level between the full length genomic sequences and amino acid composition of the putative coat protein and replication proteins support a closer relationship
between ToLCV-JB and ToLCV-Patna than with other viruses (Table 2). Moreover, ORFs of all the other proteins, AV2, AC2, AC3 and AC5 showed above 95% similarity with ToLCV-Patna. The silencing gene AC4 showed 100% similarity with AC4 of ToLCV-Patna, suggesting that the virus somehow got transferred from India to Bangladesh.

Overall, the most solid evidence of a close relationship between ToLCV-JB and ToLCV-Patna relied on phylogenetic analyses. The phylogenetic tree was constructed on the basis of complete genome sequence. Trees were prepared using Treeview programs and are based on 100 bootstrapped data sets. All the sequences used in this analysis were collected from GenBank. The database accession numbers for each isolate are mentioned in the text (Table 2). The new isolate mentioned in this study is placed in a box (Fig. 4). In all phylogenetic trees, regardless of whether they were generated by maximum-likelihood or neighbour-joining methods, and whether the full-length genome of representative geminiviruses were considered, ToLCV-JB and ToLCV-Patna always clustered in the same clade, clearly separated from all the other geminiviruses with high bootstrap values (Fig. 4). Therefore, the close similarities in genome composition, organization and expression of ToLCV-JB and ToLCV-Patna, and their clear phylogenetic relationships support that ToLCV-JB is the second member of a Patna virus that has migrated from India to Bangladesh.

Acknowledgements
Authors are thankful to The World Academy of Sciences (TWAS), Triesty, Italy for providing financial support to carry out this research work under the grants number (15-034 RG/BIO/AS_I – FR3240287052).

References


(Manuscript received on 26 June, 2018; revised on 14 August, 2018)