Virus Genes (2013) 46:581–584
DOI 10.1007/s11262-013-0888-0

# Complete nucleotide sequence of a monopartite *Begomovirus* and associated satellites infecting *Carica papaya* in Nepal

**M. S. Shahid · S. Yoshida · G. B. Khatri-Chhetri · R. W. Briddon · K. T. Natsuaki**

Received: 28 December 2012 / Accepted: 28 January 2013 / Published online: 9 February 2013


**Abstract** *Carica papaya* (papaya) is a fruit crop that is cultivated mostly in kitchen gardens throughout Nepal. Leaf samples of *C. papaya* plants with leaf curling, vein darkening, vein thickening, and a reduction in leaf size were collected from a garden in Darai village, Rampur, Nepal in 2010. Full-length clones of a monopartite *Begomovirus*, a betasatellite and an alphasatellite were isolated. The complete nucleotide sequence of the *Begomovirus* showed the arrangement of genes typical of Old World begomoviruses with the highest nucleotide sequence identity (>99 %) to an isolate of *Ageratum yellow vein virus* (AYVV), confirming it as an isolate of AYVV. The complete nucleotide sequence of betasatellite showed greater than 89 % nucleotide sequence identity to an isolate of Tomato leaf curl Java betasatellite originating from Indonesian. The sequence of the alphasatellite displayed 92 % nucleotide sequence identity to Sida yellow vein China alphasatellite. This is the first identification of these components in Nepal and the first time they have been identified in papaya.

**Keywords** *Geminivirus* · *Begomovirus* · Betasatellite · Alphasatellite · Papaya

**Electronic supplementary material** The online version of this article (doi:10.1007/s11262-013-0888-0) contains supplementary material, which is available to authorized users.

M. S. Shahid · S. Yoshida · K. T. Natsuaki
Department of International Agricultural Development, Tokyo University of Agriculture, Tokyo 156-8502, Japan

M. S. Shahid (✉)
Department of Biosciences, COMSATS Institute of Information Technology, Sahiwal 57000, Pakistan
e-mail: shafiqinayat@gmail.com

G. B. Khatri-Chhetri
Department of Plant Pathology, Institute of Agriculture and Animal Science Tribhuvan University, Chitwan 44209, Nepal

R. W. Briddon
Agricultural Biotechnology Division, National Institute for Biotechnology and Genetic Engineering, Faisalabad, Pakistan

Plant viruses in the family Geminiviridae are characterized by small, twinned isometric particles containing genomes consisting of single-stranded DNA (ssDNA). The family comprises four genera (*Mastrevirus*, *Curtovirus*, *Topocuvirus*, and *Begomovirus*), distinguished by insect vector, host range and genomic characteristics [1]. The majority of *Geminiviruses* fall into the genus *Begomovirus*. Members of this genus are transmitted by the whitefly *Bemisia tabaci*, infect only dicotyledonous plants and have either monopartite or bipartite genomes. In the Old World (OW), the majority of begomoviruses have monopartite genomes and most associate with ssDNA satellites known as betasatellites (earlier known as DNA β [2, 3]) and satellite-like components known as alphasatellites (previously called DNA 1 [3]).

Leaves of four papaya plants exhibiting severe leaf curling, vein darkening, vein thickening, and a reduction in leaf size, typical of *Begomovirus* infection, were collected from three gardens in Darai village (Rampur, Nepal) in 2010 (Fig. 1). Leaf tissues were desiccated with silica gel for 2 weeks and stored at 4 °C until use. Total DNA was extracted as described previously [4]. The extracted DNA was used as a template for rolling-circle amplification (RCA) [5], yielding a high molecular weight product (results not shown). Digestion of the RCA product with *Bam*HI resulted in an ~2.7 kb fragment which was cloned

**Fig. 1** Papaya plants originating from Nepal showing leaf curl and *yellow* mosaic symptoms typical of *Begomovirus* infection

in pUC118 (Takara Bio. Inc., Japan). Betasatellite and alphasatellite components were PCR amplified using β01/β02 and DNA101/DNA102 specific primers, respectively [6, 7]. Multiple clones were obtained in each case, and one clone in each case (clones R3, R7, and R8 for the *Begomovirus*, betasatellite, and alphasatellite, respectively) was selected and sequenced.

The complete nucleotide sequence of the *Begomovirus* isolated from papaya was found to be 2753 bp (accession no. KC282641). Analysis of the sequence showed the genome organization of the virus to be typical of OW begomoviruses [1] with four open reading frames (ORFs) in the complementary sense (C1, C2, C3, and C4) and two in the virion-sense (V2 and V1; Table 1). A predicted stem-loop structure, containing the sequence TAATATTAC within the loop, was identified which is conserved between all *Geminiviruses*. This marks the origin of virion-strand DNA replication for *Geminiviruses* [8].

Comparisons to sequences available in the databases showed the sequence of R3 to have the highest levels of nucleotide sequence identity (99.5 %) to an isolate of *Ageratum yellow vein virus* (AYVV) originating from Japan (AYVV-[JR:Ishi05:05]; AB306314) isolated from tomato [9] but less than 87 % identity to the sequences of other *Begomovirus* species (results not shown). A more detailed analysis showed the predicted amino acid sequences of all ORFs to have greater than 98 % amino acid sequence identity to the corresponding ORFs of AYVV-[JR:Ishi05:05] with the exception of ORF V2 that showed the highest level of identity (98 %) to the V2 of an AYVV isolate from China (AYVV-[CN:F11]; Table 1). A phylogenetic analysis shows R3 to cluster with AYVV isolates, being most closely related to an isolate originating from Japan (AYVV-[JR:Ishigaki05:05], AB306314; Fig. 2a). For begomoviruses, the threshold cut-off value for distinguishing species from strains currently rests at 89 % [10]. This indicates that the *Begomovirus* identified in papaya is an isolate of AYVV, for which we propose the isolate descriptor AYVV-[Nepal:Rampur3:Papaya:2010].

The complete nucleotide sequence of R7 was determined to be 1356 nucleotides (acc. no. KC282642). The sequence shows all the features typical of betasatellites [11], with a single open reading frame in the complementary sense (known as βC1; coordinates 190–546), a region of sequence rich in adenine (A-rich; coordinates 705–1013, with 60 % adenine content), and a sequence motif highly conserved between all betasatellites, known as the satellite conserved region (SCR; coordinates 1240–14). The SCR contains a predicted hairpin structure with similarity to the origin of virion-strand DNA replication of *Geminiviruses* [11]. The βC1 gene is predicted to encode a 118 amino acid protein which shows greater than 97 % amino acid sequence identity to the βC1 product of isolates of tomato leaf curl Java betasatellite (ToLCJaB) isolated from tomato and *Ageratum*

**Table 1** Positions and coding capacity of predicted genes for the AYVV isolated from papaya

| ORF | Start codon (nucleotide coordinates) | Stop codon (nucleotide coordinates) | Predicted size (no. of amino acids) | Predicted molecular weigh (kDa) | Predicted highest amino acid sequence identity (%) |
|---|---|---|---|---|---|
| V2 | 133 | 483 | 83 | 9.81 | 98 (AYVV-[CN:F11]) |
| CP | 293 | 1,066 | 257 | 29.81 | 99 (AYVV-[JR:Ishi:05]) |
| Ren | 1,467 | 1,063 | 129 | 15.13 | 100 (AYVV-[JR:Ishi:05]) |
| TrAP | 1,615 | 1,208 | 128 | 14.96 | 99 (AYVV-[JR:Ishi:05]) |
| Rep | 2,603 | 1,521 | 343 | 39.50 | 98 (AYVV-[JR:Ishi:05]) |
| C4 | 2,446 | 2,156 | 91 | 10.50 | 100 (AYVV-[JR:Ishi:05]) |

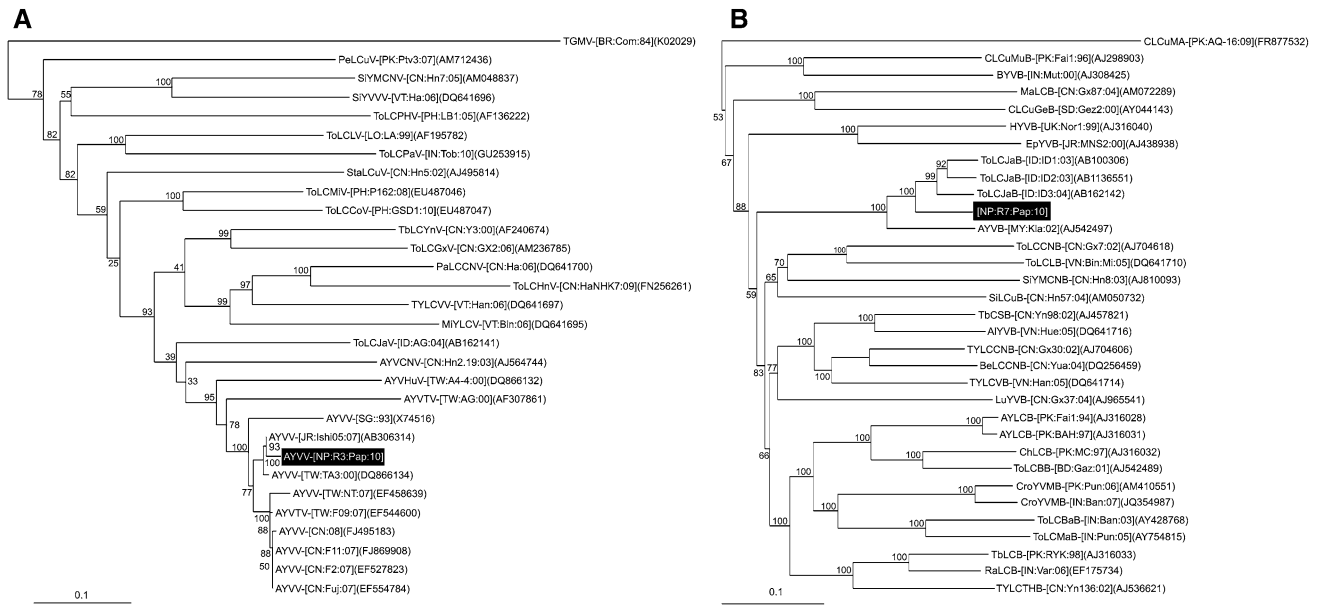


**Fig. 2** Phylogenetic dendrograms based upon alignments of selected *Begomovirus* genome (or DNA A component **a** and betasatellite sequences **b**. Vertical branches are arbitrary; horizontal branches are proportional to calculated mutation distance. *Values at nodes* indicate percentage bootstrap values (1,000 replicates). *Begomovirus* acronyms used are listed in Supplementary Table 1. This tree was arbitrarily rooted on the sequence of *Tomato golden mosaic virus* (TGMV) DNA A, a distantly related *Begomovirus*, as outgroup. The isolate descriptors used are as given in Fauquet et al. [10]. The betasatellite acronyms used are listed in Supplementary Table 1. The betasatellite tree was arbitrarily rooted on the sequence of Cotton leaf curl Multan alphasatellite (CLCuMA), an unrelated sequence of similar size, as outgroup. The betasatellite isolate descriptors used are as given in Briddon et al. [2]

*conyzoides* from Indonesia. A phylogenetic analysis based upon an alignment of the sequence of R7 with selected betasatellite sequences from the databases shows R7 to segregate with the isolates of ToLCJaB (Fig. 2b). Overall it shows the highest nucleotide sequence identity (85.5 %) to ToLCJaB-[ID:ID2:03](AB113651). The descriptor ToLCJaB-[Nepal:Rampur7:Papaya:2010] is proposed for this isolate.

The complete nucleotide sequence of R8 was determined to be 1359 bp (acc. no. KC282643) and was most closely related to Sida yellow vein China alphasatellite (SiYVCNA)-[CN:Y340:10] (FN806782), with 92 % nucleotide sequence identity. This molecule has an arrangement typical of previously characterized alphasatellites [7], containing a single ORF in the virion-sense which is predicted to encode a 315aa product with similarity to the replication-associated protein (Rep; a rolling-circle replication initiator protein; coordinates 79–1026) of nanoviruses, an A-rich sequence (coordinates 982–1223 with 49 % adenine content) and a predicted hairpin structure with the loop sequence TAGTATTAC typical of alphasatellites [7]. The predicted sequence of the Rep shows the highest levels of amino acid sequence identity (96 %) to the Rep of SiYVCNA. In a phylogenetic analysis, the alphasatellite isolated from papaya segregated with SiYVCNA (Supplementary Fig. 1). These findings indicate that the alphasatellite characterized here is an isolate of SiYVCNA for which the isolate descriptor SiYVCNA-[Nepal:Rampur8:Papya:2010] is proposed. Only a single isolate of SiYVCNA has been identified previously, in a *Sida acuta* plant with yellow vein symptoms infected with two begomoviruses (*Kenaf leaf curl virus* [FN806777] and *Malvastrum yellow vein Baoshan virus* [FN806779]) and *Malvastrum yellow vein* Yunnan betasatellite [FN806780]. The identification here of SiYVCNA with AYVV and ToLCJaB is consistent with the suggestion that alphasatellites show little, if any, host or helper virus specificity [7].

Leaf curl disease of papaya is a widespread problem across Asia and has been shown to be associated with a number of distinct begomoviruses and betasatellites [12–14]. This is the first identification of AYVV and ToLCJB in papaya and also the first time these two components have been identified in Nepal. AYVV is widespread across Asia whereas ToLCJaB has so far only been identified in Indonesia [15, 16]. The geographic distribution of SiYVCNA remains unclear; only a single isolate so far having been characterized originating from China. Nevertheless, the results obtained here may suggest that Nepal is a cross-road where begomoviruses from northern and southern Asia meet, leading to novel associations. However, further studies to determine the diversity of begomoviruses and associated satellites from this country are needed since so far only three begomoviruses (*Mungbean yellow mosaic India virus*, *Ageratum enation virus*,

and *Tomato leaf curl Gujarat virus* [1, 17, 18]), a single betasatellite (Tomato yellow leaf curl Thailand betasatellite [11]) and no alphasatellites have been identified there.

**Acknowledgments** This research was partly supported by an Advanced Research Project of the Nodai Research Institute, Tokyo University of Agriculture, Japan. RWB is supported by the Higher Education Commission (Govt. of Pakistan) under the "Foreign Faculty Hiring Program." The authors are also grateful to Mr. B. J. Pudashini to help in collecting samples.

## References

1. J.K. Brown, C.M. Fauquet, R.W. Briddon, M. Zerbini, E. Moriones, J. Navas-Castillo, in *Geminiviridae*, ed. by A.M.Q. King, M.J. Adams, E.B. Carstens, E.J. Lefkowitz (Elsevier Inc, London, 2011), pp. 351–373
2. R.W. Briddon, J.K. Brown, E. Moriones, J. Stanley, M. Zerbini, X. Zhou, C.M. Fauquet, Arch. Virol. **153**, 763–781 (2008)
3. R.W. Briddon, J. Stanley, Virology **344**, 198–210 (2006)
4. J.J. Doyle, J.L. Doyle, Phytochemistry **19**, 11–15 (1987)
5. D. Haible, S. Kober, H. Jeske, J. Virol. Methods **135**, 9–16 (2006)
6. R.W. Briddon, S.E. Bull, S. Mansoor, I. Amin, P.G. Markham, Mol. Biotechnol. **20**, 315–318 (2002)
7. R.W. Briddon, S.E. Bull, I. Amin, S. Mansoor, I.D. Bedford, N. Rishi, S.S. Siwatch, M.Y. Zafar, A.M. Abdel-Salam, P.G. Markham, Virology **324**, 462–474 (2004)
8. B.M. Orozco, L. Hanley-Bowdoin, J. Virol. **7**, 148–158 (1996)
9. T. Andou, A. Yamaguchi, S. Kawano, K. Kawabe, S. Ueda, M. Onuki, J. Gen. Plant Pathol. **76**, 287–291 (2010)
10. C.M. Fauquet, R.W. Briddon, J.K. Brown, E. Moriones, J. Stanley, M. Zerbini, X. Zhou, Arch. Virol. **153**, 783–821 (2008)
11. R.W. Briddon, S.E. Bull, I. Amin, A.M. Idris, S. Mansoor, I.D. Bedford, P. Dhawan, N. Rishi, S.S. Siwatch, A.M. Abdel-Salam, J.K. Brown, Y. Zafar, P.G. Markham, Virology **312**, 106–121 (2003)
12. L.S. Chang, Y.S. Lee, H.J. Su, T.H. Hung, Plant Dis. **87**, 204 (2003)
13. P. Singh-Pant, P. Pant, S. Mukherjee, S. Mazumdar-Leighton, Arch. Virol. **157**, 1217–1232 (2012)
14. X. Wang, Y. Xie, X. Zhou, Virus Genes **29**, 303–309 (2004)
15. T. Kon, S.H. Hidayat, S. Hase, H. Takahashi, M. Ikegami, Phytopathology **96**, 517–525 (2006)
16. T. Kon, K. Kuwabara, S.H. Hidayat, M. Ikegami, Arch. Virol. **152**, 1147–1157 (2007)
17. M.S. Shahid, P.B. Pudashini, G.B. Khatri-Chhetri, M. Ikegami, K.T. Natsuaki, New Dis. Rep. **25**, 30 (2012)
18. M.S. Shahid, M. Ikegami, K.T. Natsuaki, Australas. Plant Dis. **7**, 85–89 (2012)