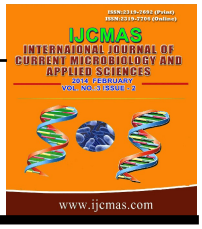
International Journal of Current Microbiology and Applied Sciences
ISSN: 2319-7706 **Volume 3 Number 2 (2014) pp. 489-497**
http://www.ijcmas.com



Original Research Article

In Silico Recombination Analysis of DNA-A sequence from Begomovirus reported in India: This identified recombinant is the evolution from other viruses prevailing at different geographical region of Pakistan and China.

Avinash Marwal, Rajneesh Prajapat and R K Gaur*

Department of Science, Faculty of Arts, Science and Commerce, Mody Institute of Technology and Science, Lakshmangarh, Sikar - 332311, Rajasthan, India

****Corresponding author***

A B S T R A C T

Keywords

Begomovirus; Marigold; Recombinant; Ornamental.

The Begomovirus DNA-A genome was isolated from an ornamental plant Marigold and is identified as a new recombinant species, sharing nucleotide identity with other isolates reported from China and Pakistan. The present study remarkably suggests that the exchange of DNA-A with other begomoviruses would create a new disease complex posing a serious threat to agriculture crops and horticulture ornamental plants production.

Introduction

Recombination has played, and continues to play, a pivotal role in geminiviral evolution and may be contributing to the emergence of new forms of geminiviruses because the high frequency of mixed infections of begomoviruses provides an opportunity for the emergence of new viruses arising from recombination among strains and / or species (Harrison and Robinson, 1999). In some cases, the recombinants exhibited a new pathogenic phenotype which is often more virulent than the parents (Zhou et al., 1997).

The specific recombination events including the recombination breaks and hot spots have not been reported so far in Marigold infecting begomovirus. It is also currently unknown as to whether the sequences in particular parts of the begomovirus genomes are exchangeable between different species and/or members of the same genus from different geographical locations. In addition, Northern India seems to be unusually rich in virus biodiversity; we investigated the extent of recombination events and examined their role in the evolution of virus in India and its neighboring countries.

The earlier reports shown the molecular characterization of complete genome of a begomovirus isolated from an ornamental plant Marigold (Marwal et al., 2013) i.e. *Ageratum enation virus* (AEV:

KC589699). Here we are presenting and highlighting the *In Silico* Recombination analysis approach for in depth study about the nature of the virus. Thus, in order to take a step forward to find a cure against such viruses that causes major crop loss worldwide.

Materials and Methods

Recombination between divergent genomes is believed to be a major mechanism by which diversity amongst viruses is generated (Robertson et al., 1995). To detect the possibility of recombination in begomovirus, Recombination Detection Program (RDP) was utilized, which is based on a pair wise scanning approach. It usually runs under Windows 95/98/NT/XP/VISTA/7 and couples a high degree of analysis automation with an interactive and detailed graphical user interface (Posada and Crandall, 2001). Using various recombination detection methods the conclusion of recombination studies were evaluated (Posada, 2002). The recombination breakpoint could be identified by using Recombination detection program [RDP], GENECONV, Maximum-Chi, BOOTSCAN, CHIMAERA, and 3SEQ methods. All these methods were implemented in RDP v.3.44 (Martin et al., 2005). Recombination positions were recognized by only RDP method and other methods such as GENECONV, Bootscan, Maxchi and Chimera were not found suitable for recombination analysis because of lowest recombination breakpoint detection accuracy.

The *Ageratum enation virus* (AEV: KC589699) DNA-A sequence were subjected to recombination analyses using RDP method used to drive automated recombination scan and the manual checking of automated analysis results. Analysis was allowed by employing Bonferroni correction with confidence greater than 95% (P value 0.05). In RDP analysis, the length of the window was set to 10 variable sites, and the step size was set to one nucleotide. P values were estimated by randomizing the alignment 1,000 times.

Results and Discussion

Recombination positions were observed in the *Ageratum enation virus* (AEV: KC589699) sequence (Figure 1). The schematic sequence display is where the results of automated recombination scans are presented and it is the part of the program that is used to drive the manual checking of automated analysis results. The colored rectangles correspond to sequence fragments, thus representing the recombinant, major and minor parents in a graphical representation of a sequence fragments that have been potentially derived through recombination from a sequence resembling the one named to the right of the rectangle.

The RDP plot of *Ageratum enation virus* (AEV: KC589699) sequence conservation displayed a graphical overview of the sequence alignment that also indicates the portion of the alignment. Within the sequence part of the display, individual nucleotides are color coded according to their degree of conservation. When a recombination event selected, the ‘toggle sequence display’ button can be used to highlight nucleotide polymorphisms that contribute to the recombination signals depicted in the plot display. As per the schematic sequence display six recombination break point positions were identified in the begomovirus infecting Marigold by the RDP method.

The first evidence is given in Figure 2 where breakpoint begin from 2665th [position 2751 in alignments] position and ending breakpoint ends at 287th [position 309 in alignments] position. Approximate p-value for this region was 3.281 x 10^{-11}. The region probability (MC Uncorrected) was 1.125 E-14 and region probability (MC corrected) was 7.652 E-12. It suggested recombination in the common region (IR) and AV2 ORF fragment of the sequence. The major parent was identified as *Pedilanthus leaf curl virus* (JQ012790) identified in India and were found infecting *Cestrum nocturnum*. Whereas the minor parent was *Croton yellow vein mosaic virus* (FN678906) found infecting an ornamental plant *Alcea rosea* in Pakistan. This clearly indicates that this portion of recombinant fragment of common region (IR) and AV2 ORF is contributed from the two viruses prevailing at different geographical region, undoubtedly pointing towards the begomovirus evolution.

RDP looks for regions within a sequence alignment in which sequence pairs are sufficiently similar to suspect that they may have arisen through recombination. The second recombination was detected downstream of the first recombinant sequence where breakpoint begin from 388th [position 310 in alignments] position and ending breakpoint ends at 935th [position 962 in alignments] position (Figure 3). Approximate p-value for this region was 5.787 x 10^{-24}. Here the contribution of major parent in the RDP plot was by *Tomato leaf curl Karnataka virus* (FJ514798) infecting *Mentha viridis* in India. The minor parent was identified as *Croton yellow vein mosaic virus* (AJ507777) infecting *Croton bonplandianum* in India. The region probability (MC Uncorrected) was 3.694 E-26 and region probability (MC corrected) was 2.512 E-23. Hence recommended recombination was observed at the AV2 ORF, spanning completely AV1 ORF and the starting of AC3 ORF fragment of the sequence.

The third recombination event was discovered at the downstream of the second recombination event in the begomovirus DNA-A sequence, where breakpoint begin from 938th [position 965 in alignments] position and ending breakpoint ends at 1042th [position 1094 in alignments] position. Approximate p-value for this region was 1.335 x 10^{-12}. In this case the major parent was identified as *Ageratum enation virus* (JQ911765) causing disease in *Papaver somniferum* and was reported from India (Figure 4). The minor parent in the RDP plot was found to be *Tomato leaf curl Ranchi virus* (GQ994095) reported from India infecting Tomato sp. The region probability (MC Uncorrected) was 1.963 E-15 and region probability (MC corrected) was 1.335 E-12. The recombination was detected in the AC3 ORF fragment of the DNA-A sequence.

The fourth recombination event was discovered at the downstream of the third recombination event in the begomovirus DNA-A sequence, where breakpoint begin from 1042th [position 1095 in alignments] position and ending breakpoint ends at 1200th [position 1258 in alignments] position. Approximate p-value for this region was 1.660 x 10^{-08}. In this case the major parent was identified as *Ageratum enation virus* (JQ911765) causing disease in *Papaver somniferum* and was reported from India (Figure 5). The minor parent in the RDP plot was found to be *Tomato leaf curl Ranchi virus* (GQ994095) reported from India infecting Tomato sp. The region probability (MC Uncorrected) was

Figure.1 Diagram of the schematic sequence display representing the RDP recombination map of the recombinant fragments for the *Ageratum enation virus* (AEV: KC589699) sequence. Each color/pattern represents a sequence specific of a virus. The virus genome organization is represented under the diagram, positioning the different viral genes named according to the begomovirus convention.

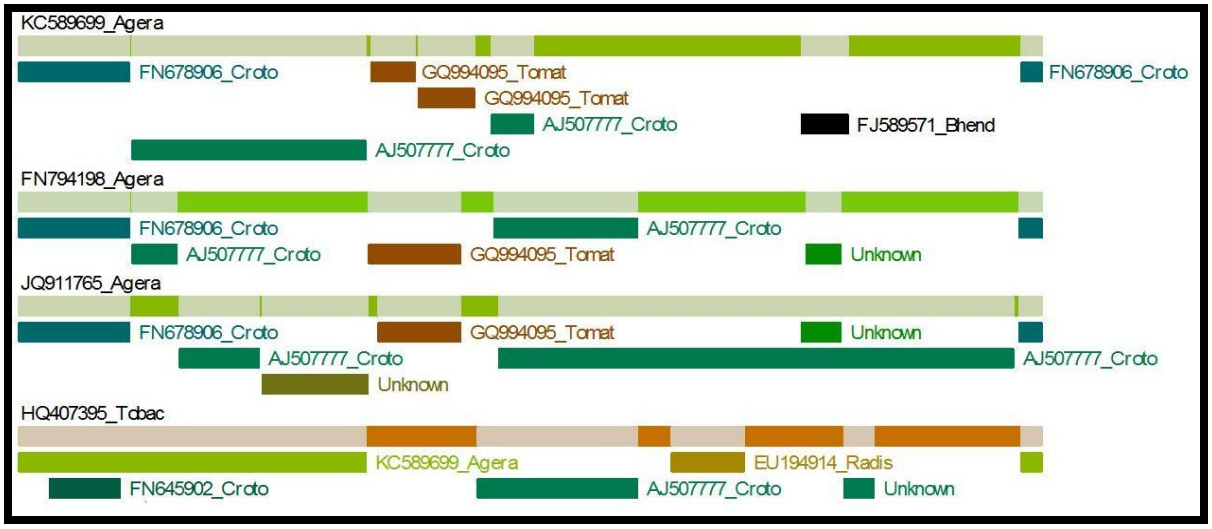


Figure.2 An RDP pairwise identity plot for the piece of sequence from the major parent (JQ012790_Pedil) Uppermost bares indicating positions of informative sites; pink region indicates breakpoint positions suggested by the RDP software method. The pairwise identity plot have major parent: minor parent plot (JQ012790_Pedil: FN678906_Croto; yellow), major parent: recombinant plot (JQ012790_ Pedil: KC589699_ Agera; dark blue) and minor parent: recombinant plot (FN678906_ Croto: KC589699_ Agera; purple).

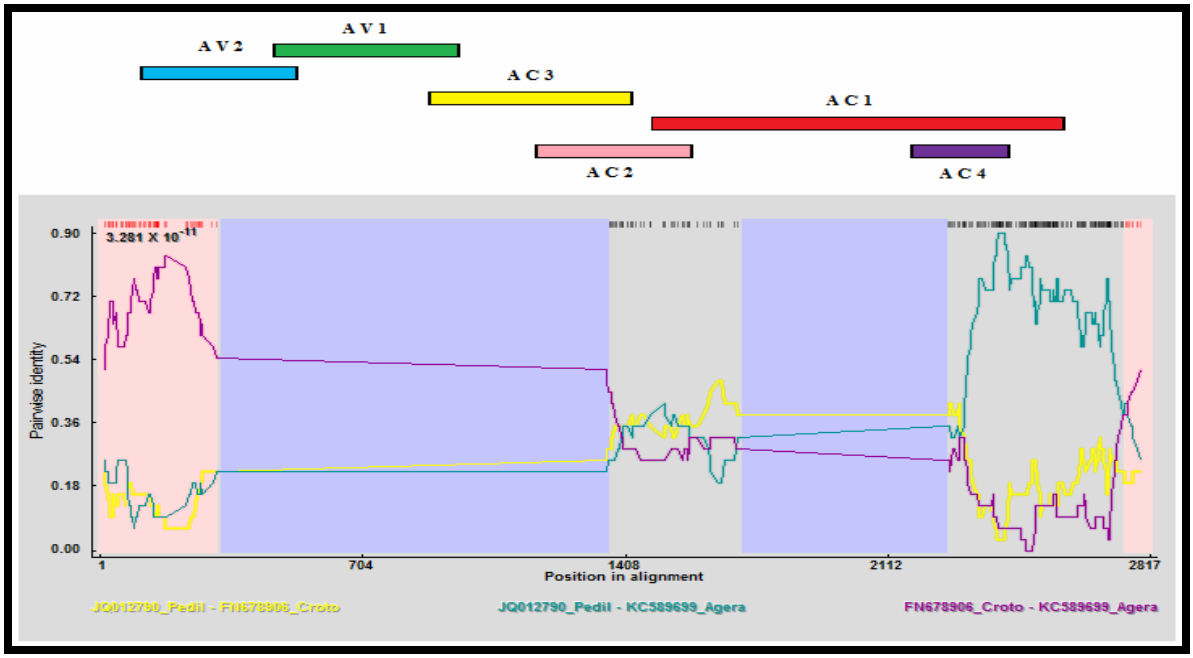


Figure.3 An RDP pairwise identity plot for the piece of sequence from the major parent (FJ514798_Tomat) Uppermost bares indicating positions of informative sites; pink region indicates breakpoint positions suggested by the RDP software method. The pairwise identity plot have major parent: minor parent plot (FJ514798_Tomat: AJ507777_Croto; yellow), major parent: recombinant plot (FJ514798_ Tomat: KC589699_ Agera; dark blue) and minor parent: recombinant plot (AJ507777_ Croto: KC589699_ Agera; purple).

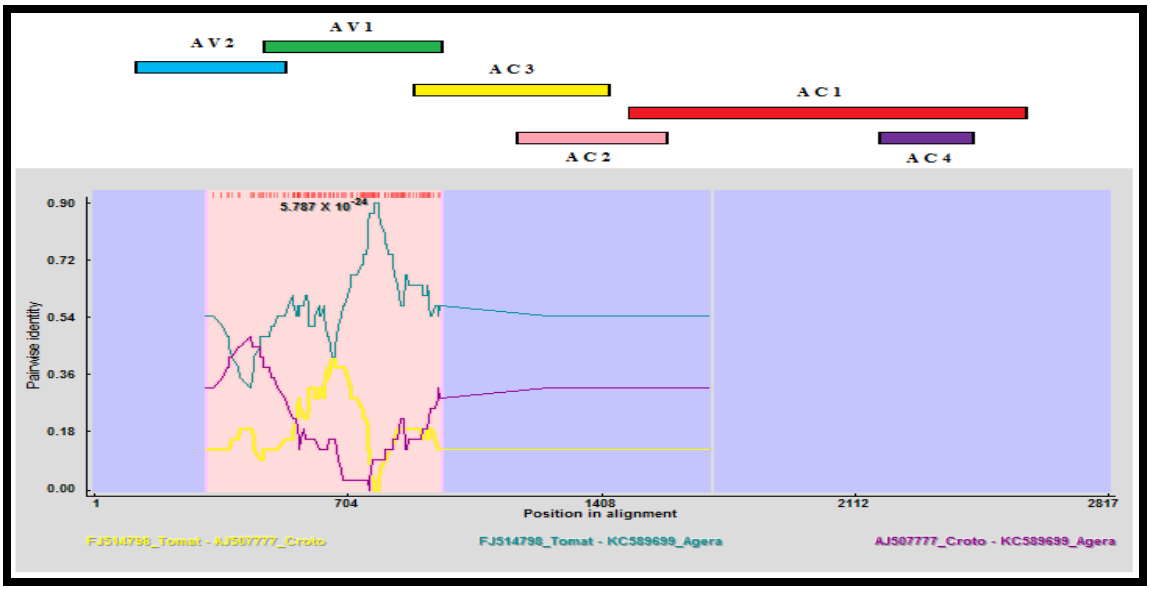


Figure.4 An RDP pairwise identity plot for the piece of sequence from the major parent (JQ911765_Agera) Uppermost bares indicating positions of informative sites; pink region indicates breakpoint positions suggested by the RDP software method. The pairwise identity plot have major parent: minor parent plot (JQ911765_Agera: GQ994095_Tomat; yellow), major parent: recombinant plot (JQ911765_Agera: KC589699_ Agera; dark blue) and minor parent: recombinant plot (GQ994095_Tomat: KC589699_ Agera; purple).

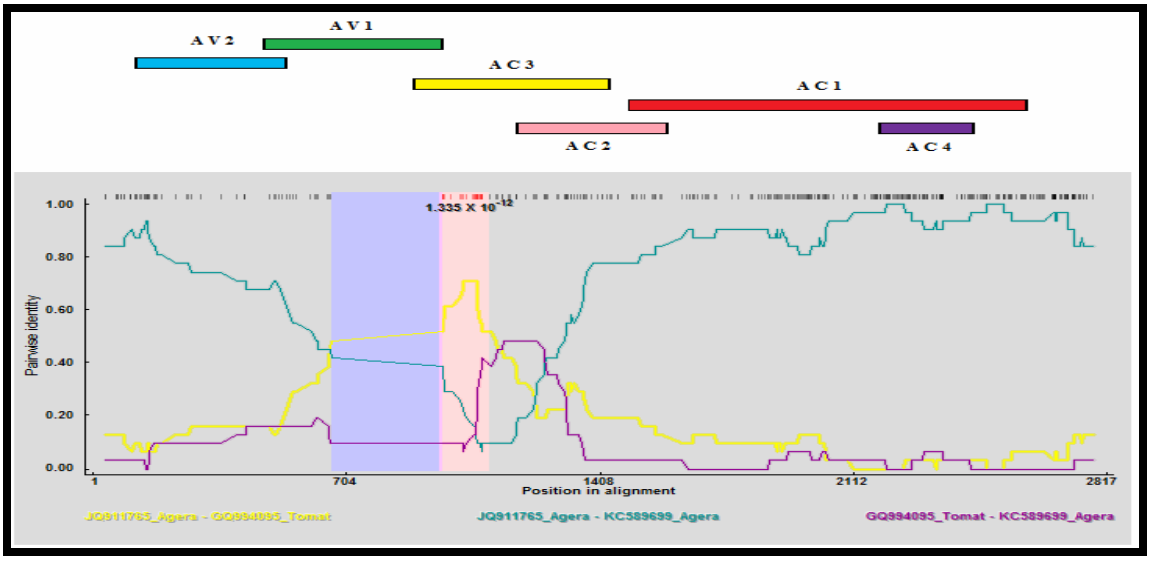


Figure.5 An RDP pairwise identity plot for the piece of sequence from the major parent (JQ911765_Agera) Uppermost bares indicating positions of informative sites; pink region indicates breakpoint positions suggested by the RDP software method. The pairwise identity plot have major parent: minor parent plot (JQ911765_Agera: GQ994095_Tomat; yellow), major parent: recombinant plot (JQ911765_Agera: KC589699_ Agera; dark blue) and minor parent: recombinant plot (GQ994095_Tomat: KC589699_ Agera; purple).

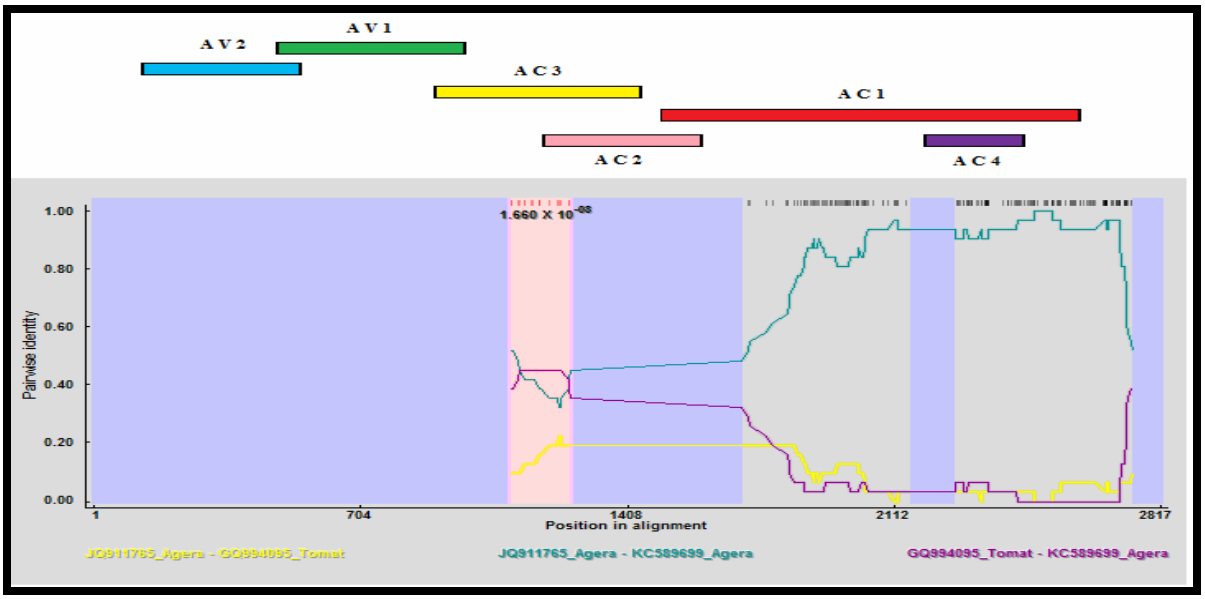


Figure.6 An RDP pairwise identity plot for the piece of sequence from the major parent (GU199583_Tobac) Uppermost bares indicating positions of informative sites; pink region indicates breakpoint positions suggested by the RDP software method. The pairwise identity plot have major parent: minor parent plot (GU199583_Tobac: AJ507777_Croto; yellow), major parent: recombinant plot (GU199583_ Tobac: KC589699_ Agera; dark blue) and minor parent: recombinant plot (AJ507777_ Croto: KC589699_ Agera; purple).

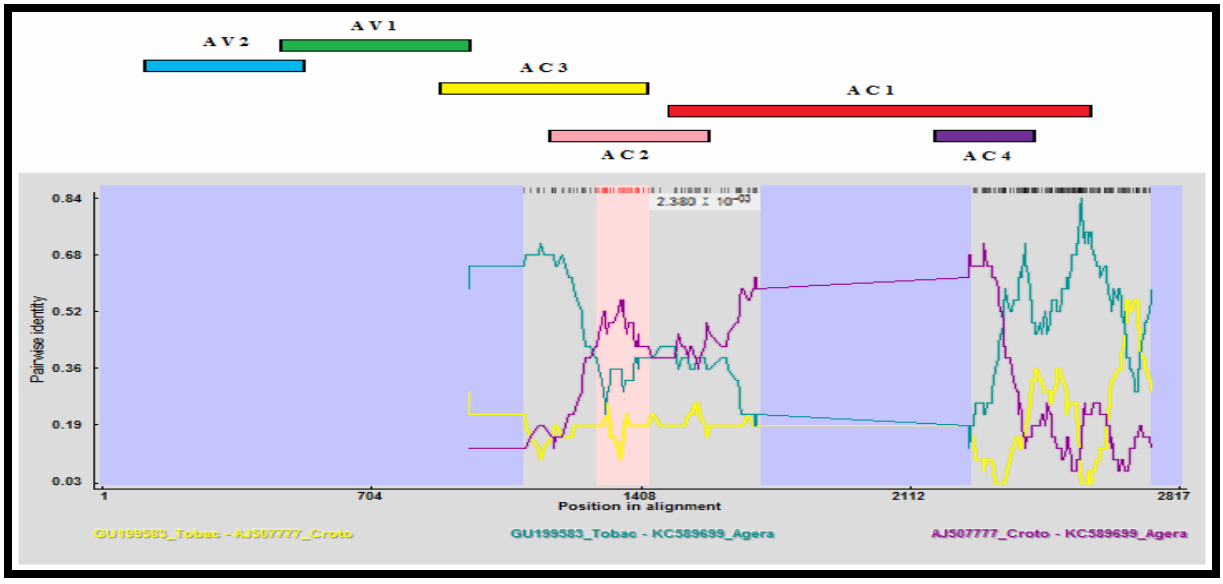
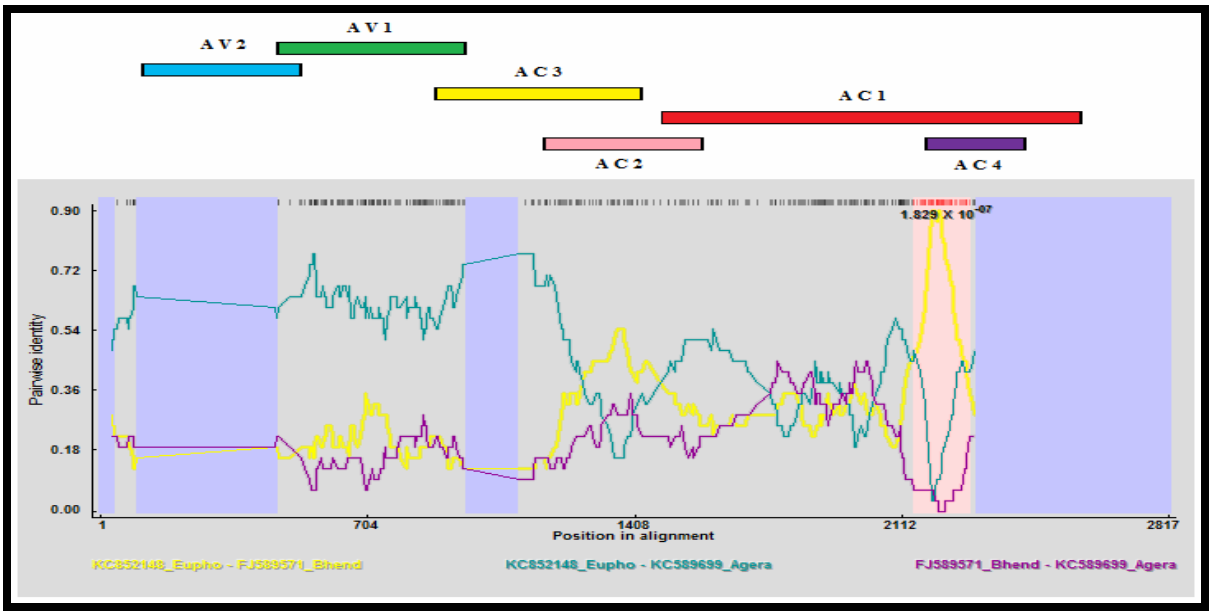


Figure.7 An RDP pairwise identity plot for the piece of sequence from the major parent (KC852148_Eupho) Uppermost bares indicating positions of informative sites; pink region indicates breakpoint positions suggested by the RDP software method. The pairwise identity plot have major parent: minor parent plot (KC852148_ Eupho: FJ589571_Bhend; yellow), major parent: recombinant plot (KC852148_ Eupho: KC589699_ Agera; dark blue) and minor parent: recombinant plot (FJ589571_ Bhend: KC589699_ Agera; purple).



2.442 E-11 and region probability (MC corrected) was 1.660 E-08. The recombination was detected in the AC2 and AC3 ORF fragment of the DNA-A sequence. In recombination event third and fourth, the minor parent and minor parent were found to be same.

The fifth recombination was again detected in the AC2 and AC3 ORF fragment of the DNA-A sequence. Beginning breakpoint position was 1237th [position 1295 in the alignment] and ending breakpoint position was 1366th [position 1424 in the alignment]. The region probability (MC Uncorrected) was 2.492 E-02 and region probability (MC corrected) was 16.9. Approximate p-value for this region was 2.380 x 10^{-03}. In this case the major parent was identified as *Tobacco curly shoot virus* (GU199583) causing disease in *Alternanthera philoxeroides* and was reported from China (Figure 6). The minor parent in the RDP plot was found to be *Croton yellow vein mosaic virus* (AJ507777) infecting *Croton bonplandianum* in India.

The final and the sixth recombination were detected in the AC4 ORF region as well the overlapping region of AC1 ORF. In this case the breakpoint begins from 2074th [position 2146 in alignments] position and ending breakpoint ends at 2210th [position 2286 in alignments] position. Approximate p-value for this region was 1.829 x 10^{-07}. The major parent was encountered to be *Euphorbia leaf curl virus* (KC852148) causing leaf curl disease in *Euphorbia pulcherrima* reported from China and the minor parent was discovered as *Bhendi yellow vein Bhubhaneswar virus*

(FJ589571) responsible for begomovirus accounted in India (Figure 7). The region probability (MC Uncorrected) was 2.690 E-10 and region probability (MC corrected) was 1.829 E-07.

In begomoviruses the recombination hot-spots map to complementary-sense gene transcription initiation and termination sites and virion strand origins of replication. The reason complementary gene transcription initiation and termination sites may be more predisposed to recombination than other sites is possibly that these are the regions where the most frequent clashes between transcription and replication complexes occur (Lefeuvre et al., 2007).

Recombination sites have been reported in both the DNA and RNA viruses (Prasanna and Rai, 2007). Presumably, the different pathotypes could simultaneously infect a host cell and exchange genetic materials through recombination. The recombination observed between geographically separated isolates probably represents older events, which may have occurred before their present separation (Gagarinova et al., 2008). Movement of vectors and/or infected plant materials could be another factor for the gene flow between the widely separated locations (Rojas et al., 2005).

A recombination may result in significant changes in the biological properties of virus isolates with the ability to adopt and sustain in different environmental conditions. The interesting findings were the evolution of *Ageratum enation virus* (AEV: KC589699) DNA-A sequence from begomovirus prevailing in India, as well as from other isolates existing in Pakistan and China. It is also possible that exchange of begomovirus DNA-A genome symptoms in *Abelmoschus esculentus* could extend the virus host range thereby emergence of new diseases in cultivated crop plants and other ornamental plants. One factor favoring the spread of begomoviruses among these plants is that many dicotyledonous species in India are hosts of whiteflies of the *B. tabaci* complex, which are the known or likely vectors of all the viruses. Perhaps this is the first report of recombination in begomovirus infecting Marigold, which would provide significant information for understanding the diversity and evolution of begomoviruses in India.

Acknowledgement

The authors would like to acknowledge a vote of thanks to Department of Biotechnology (DBT project No. BT/PR13129/GBD/27/197/2009) and Department of Science and Technology (DST project no SR/FT/LS-042/2009), India for their financial support.

References

Gagarinova, A. G., M. Babu, M. V. Strömvik and Wang A. 2008. Recombination analysis of *Soybean mosaic virus* full-length genome sequences reveals evidence of RNA recombination between distinct pathotypes. Virol. J. 4: 111-120.

Harrison, B. D., and Robinson D. J. 1999. Natural genomic and antigenic variation in whitefly-transmitted geminiviruses (begomoviruses). Annual Review of Phytopathology. 37: 369-398.

Lefeuvre, P., D. P. Martin, M. Hoareau, F. Naze, H. Delatte, M. Thierry, A. Varsani, N. Becker, B. Reynaud and Lett J. M, 2007. Begomovirus 'melting pot' in the south-west Indian Ocean

islands: molecular diversity and evolution through recombination. J Gen Virol. 88: 3458–3468.

Martin, D. P., D. Posada, K. A. Crandall and Williamson C. 2005. A modified BOOTSCAN algorithm for automated identification of recombinant sequences and recombination breakpoints. AIDS Research and Human Retroviruses. 21: 98-102.

Marwal, A., A. K. Sahu, D. K. Choudhary and Gaur R. K. 2013. Complete nucleotide sequence of a begomovirus associated with satellites molecules infecting a new host *Tagetes patula* in India. Virus Genes. 47(1): 194 – 198.

Posada, D., 2002. Evaluation of methods for detecting recombination from DNA sequences: empirical data. Molecular Biology and Evolution. 19: 708-717.

Posada, D., and Crandall, K. A. 2001. Evaluation of methods for detecting recombination from DNA sequences: Computer simulations. Proceedings of the National Academy of Science. USA. 98: 13757-13762.

Prasanna, H. C., and Rai, M. 2007. Detection and frequency of recombination in tomato-infecting begomoviruses of South and Southeast Asia. Virol. J. 4: 111-120.

Robertson, D. L, B. H. Haha and Sharp, P. H. 1995. Recombination in AIDS virus. J. Mol. Evol. 40: 249-259.

Rojas, A., A. Kvarnheden, D. Marcenaro and Valkonen J. P. T. 2005. Sequence characterization of *Tomato leaf curl Sinaloa virus* and *Tomato severe leaf curl virus*: phylogeny of New World begomoviruses and detection of recombination. Archives of Virology. 150: 1281-1299.

Zhou, X., Y. Liu, L. Calvert, C. Munoz, G. W. Otim-Nape et al., 1997. Evidence that DNA-A of a geminivirus associated with severe cassava mosaic disease in Uganda has arisen by interspecific recombination. Journal of General Virology. 78: 2101-2111.