The Journal of Animal & Plant Sciences, 27(4): 2017, Page: 1245-1255
ISSN: 1018-7081

MOLECULAR CHARACTERIZATION OF A BEGOMOVIRUS AND ASSOCIATED SATELLITES FROM COTTON (*GOSSYPIUMHIRSUTUM* FROM DERA GHAZI KHAN DISTRICT OF PAKISTAN

A. Akram[1], K. Hussain[1*], N. Nahid[1], Mahmood-ur-Rahman[1], A. Nasim[1] and S. Shaheen[2]

[1]Department of Bioinformatics and Biotechnology, Government College University Faisalabad Pakistan
[2]Plant Taxonomy Lab, Department of Botany, Lahore College for Women University, Lahore, Pakistan
[*]Corresponding author e-mail: hussaink@gcuf.edu.pk

ABSTRACT

Cotton leaf curl disease is a major threat to cotton production in tropical and subtropical regions of the world. Begomoviruses are whiteflies-transmitted single-stranded DNA viruses belonging to the family *Geminiviridae* that cause leaf curl disease in cotton. In present study, a strain of tomato leaf curl betasatellite and *Xanthium strumarium* alphasatellite have been found associated with *Cotton leaf curl Kokhran virus* in *Gossypium hirsutum*. A begomovirus and associated molecules were amplified, cloned and sequenced. Partial sequence analysis for begomovirus showed 98.2% sequence identity with an isolate of *Cotton leaf curl Kokhran virus* (HQ257374). The partial sequence of betasatellite showed highest sequence identity of 96.4% with tomato leaf curl betasatellites (EF068245, FR819710). The complete sequence for alphasatellite showed 98.2% sequence identity with Xanthium strumarium alphasatellite (HF547408). This is the first report that tomato leaf curl betasatellite have been found in cotton and after second epidemic of cotton leaf curl disease this is the first report of parent strain of *Cotton leaf curl Kokhran virus* in the field.

Keywords: Cotton, Begomovirus, Betasatellite, Alphasatellite, Cotton leaf curl disease.

INTRODUCTION

Geminiviruses are plant viruses having circular single-stranded DNA (ss-DNA) genome of size 2.6-2.8Kb transmitted by insect vectors. Members of family *Geminiviridae* has been classified into seven genera (*Begomovirus*, *Mastrevirus*, *Curtovirus*, *Topocuvirus*, *Becurtovirus*, *Turncurtovirus* and *Eragrovirus*) based on their genome arrangement, insect vectors, sequence relatedness and host range (Brown, 2012). Among all seven genera only two of them (*Begomovirus* and *Mastrevirus*) have been reported to have associated satellite DNAs with them (Briddon *et al*., 2001; Kumar *et al*., 2014). Begomovirus is the largest genus having more than 280 species and have wide host range. They infect dicotyledonous plants and are transmitted by whiteflies (*Bemisia tabaci*). Begomovirus genome can be monopartite (single component genome) or bipartite (two component genome). In bipartite genome both components (DNA A and DNA B) are essential for *bonafied* infection. In the New World, begomoviruses are exclusively bipartite with one recently reported exception of a monopartite *Tomato leaf curl deformation virus* [ToLDeV; (Melgarejo *et al*., 2013)]. In Old World mostly monopartite begomoviruses are found but a few bipartite begomoviruses have also been reported (Ha *et al*., 2008).

Approximately 15 years ago, some ss-DNA satellite molecules were discovered associated with monopartite begomoviruses named as betasatellite (previously known as DNAβ) and alphasatellite (previously known as DNA-1) that are around half the size (approximately 1.3-1.4 Kb) of a begomovirus genome. Betasatellite-begomovirus complexes cause many economically important diseases in Africa and Asia, resulting in yield losses estimated at millions of US dollars (Mansoor *et al*., 1999, Mansoor *et al*., 2003). Betasatellites encodes only a single protein (betaC1) involved in the pathogenicity. They are dependent on the helper begomovirus for their replication and encapsidation. Begomovirus-associated alphasatellites are autonomously replicating satellite-like molecules. They encode one protein, the replication associated protein (Rep) that is replication initiator protein. They rely on the helper virus for movement, encapsidation and vector transmission. Therefore they are called as satellite-like since, by definition, satellites are not capable of independent replication (Briddon *et al*., 2004). The elective advantage of alphasatellites to begomovirus-betasatellite complexes remains uncertain. Satellites have shown to be associated with mostly Old World monopartite begomoviruses-betasatellite complexes but recently been identified with bipartite begomoviruses in the New World (Fiallo-Olive *et al*., 2012) as well as with a mastrevirus in India (Kumar *et al*., 2014). Begomoviruses have found associated with leaf curl and mosaic diseases in many crop plants as well as in weed plants such as croton, sweet potatoes and *Xanthium strumarium* (Hussain *et al*., 2011; Albuquerque *et al*., 2012; Mubin *et al*., 2012).

The economically most important disease caused by monopartite begomovirus-betasatellite complexes in Pakistan is cotton leaf curl disease (CLCuD) affecting *Gossypium hirsutum* L. CLCuD is characterized by typical symptoms of leaf curling, vein darkening, vein swelling and enations on undersides of leaves (Briddon and Markham, 2001). The disease was sporadic problem throughout southern Asia during early 1980s and it became epidemic in 1986 near Multan region and spread rapidly nearly all cotton growing regions of Punjab, Pakistan and moved eastwards to the east Punjab and Rajasthan provinces of India (Briddon *et al*., 2001). During late 1990s plant breeders developed cotton varieties showing resistance against *Cotton leaf curl Multan virus and Cotton leaf curl Khohkran virus*, the most prevailing begomovirus complex causing CLCuD (Rahman *et al*., 2005). But the resistant varieties succumbed to CLCuD in the area of Burewala near Multan Pakistan during 2001 and this for the beginning of a second epidemic (Mansoor *et al*., 2003). A new strain of begomovirus was associated with these con epidemic; Cotton leaf curl Kokhran virus strain Burewala (CLCuKoV-Bur; previously called Cotton leaf curl Burewala virus; Amrao *et al*., 2010 b; Rajagopalan *et al*., 2012). This new resistance breaking strain is dominant in all cotton growing regions of Punjab Pakistan as well as western India. Economic losses due this disease in cotton are significant and a challenge to modern agricultural biotechnology. The new resistance breaking virus complex associated with CLCuD is recombinant a of two species of begomovirus *Cotton leaf curl Multan virus* and *Cotton leaf curl Kokhran virus* (Amrao *et al*., 2010 b). This resistance-breaking recombinant strain was found associated with recombinant strain cotton leaf curl Multan betasatellite (Amrao *et al*., 2010 b) but no alphasatellite was found associated with this recombinant begomovirus complex (Amrao *et al*., 2010 b). In present study we survey cotton fields in Dera Ghazi Khan district of Punjab Pakistan and found begomovirus likely to be *Cotton leaf curl Kokhran virus* associated with a betasatellite likely to be tomoto leaf curl betasatellite and Xanthium strumarium alphasatellite (Hussain, 2013). Significance of this new complex in cotton has been discussed.

MATERIALS AND METHODS

Sampling and DNA isolation: Infected cotton leaves showing typical disease symptoms of begomovirus infection were collected from different cotton fields in D G Khan district (Fig1). Leaves showing symptoms like leaf curling, vein swelling and yellowing and enations under side of the leaf were given particular importance for total plant genomic DNA extraction by using cetyl trimethyl ammonium bromide (CTAB) method (Doyle and Doyle, 1987).

Amplification and cloning of viral components: For amplification of begomovirus components rolling circle amplification (RCA) technique using phi29 DNA polymerase and random hexamer primers was used (TempliPhi TM, GE Healthcare). Amplified fragments were digested with restriction enzymes and ligated in to vector pTZ57R (Fermentas/Thermo Fisher Scientific, Massachusetts, USA) as per recommendations of the supplier.

Sequencing and Sequence analysis: Cloned molecules were sequenced completely using M13 based primers by Macrogen (Seoul, SouthKorea). Sequences were assembled and analyzed with the help of Lasergene package (DNASTAR, Madison, Wisconsin). Multiple sequence alignments were performed and phylogenetic trees were constructed using Clustal X. Trees were displayed and manipulated using Tree view software. Percentage identity was determined using MegAlign application. The reference sequences used in analyses were obtained from National Center for Biotechnology Information (NCBI) databases (https://www. ncbi.nlm. nih.gov/).

RESULTS AND DISCUSSION

Pakistan is a hub of begomoviruses infection since decades, which was a big reason of cotton crop destruction in terms of millions and billions dollars (Sanz *et al*., 2000). As Pakistan is an agriculture dependent country and its economy is based on such economical important crops so it is necessary at prior basis to save such crops (Wilkins *et al*., 2000). Cotton leaf curl disease is a major viral constraint faced by cotton growers in sub-continent. In 2002 on-wards disease reappeared in epidemic form in cotton varieties which were showing some tolerant behavior. During this second epidemic of single species of begomovirus *Cotton leaf curl Kokhran virus-Burewala strain* (CLCuKV-Bu) is dominating throughout Punjab Pakistan.

In this study the presence of parent strain of *Cotton leaf curl Kokhran virus* has been shown in infected cotton plants. The partial sequence of a begomovirus AA55 was done. The sequence of clone was submitted to database and accession number KX599186 assigned to it. The clone of begomovirus was partially sequenced for replication associated protein(Rep) gene and intergenic region. Rep is conserved gene in function and position among geminiviruses except in Mastrevirus which is indispensable for rolling circle replication of virus genome (Hanley-Bowdoin *et al*.,2004). The clone was determined to 786 nucleotides in length (coordinates: nucleotides 1967-2748 of *Cotton leaf curl Kokhran virus*; HQ257374). Comparison of the sequence of clone AA55 with sequences of same regions of begomoviruses available in data bases revealed the higher level of

nucleotide sequence identity to different isolates of *Cotton leaf curl Kokhran virus* (CLCuKV) and *Cotton leaf curl Rajasthan virus* ranging from 97.6-98.2% (Table 1). The highest identity value was shown with *Cotton leaf curl Kohkran virus*; HQ257374 isolated from cotton in India. To all other begomovirus sequences already available in the data bases this partial sequence showed less than 90% nucleotide sequence identity. We can say our clone AA55 is likely to be an isolate of *Cotton leaf curl Kokhran virus* present in cotton in D.G. Khan district of Punjab Pakistan. In phylogenetic dendogram our clone has been segregated along with different isolates of *Cotton leaf curl Kokhran virus* (CLCuKV) and *Cotton leaf curl Multan virus Rajasthan strain* (CLCuMV-Rajstrain; Fig1). This is a first report of CLCuKV and CLCuMV-Raj strain in cotton after second epidemic caused by CLCuKV-Bu. As CLCuKV-Bu is a recombinant strain of two parental viruses *Cotton leaf curl Multan virus* (CLCuMV) and *Cotton leaf curl Kokhran virus*. The partial sequences we are discussing here can be the counter part of CLCuKV-Bu. But it is not the case as the CLCuKV-Bu has higher sequence identity with CLCuKV in its virion-sense genes whereas complementary-sense genes have higher sequence identity with CLCuMV (Amrao *et al*., 2010 b). The partial sequence covering Rep gene sequence (a complementary-sense gene) are showing highest sequence identity with CLCuKV.

The partial sequences of a betasatellite molecule, which was cloned in vectorpTZ57R and designated as AA30 was obtained. The sequence is available in the sequence databases under accession number KX599188. The sequence of clone AA30 encompasses βC1 gene region and some downstream sequences towards the satellite conserved region (SCR) of betasatellite. The clone was determined to be 652 nucleotides in length. Alignment of the sequence of the clone AA30 with betasatellite sequences available in database (of same regions) revealed highest sequence identity 96.4% with two isolates of tomato leaf curl betasatellite (EF068245 and FR819710) and second highest nucleotide sequence identity is 95.2% with another isolate of tomato leaf curl betasatellite (HM989847; Table2). In the tropical and sub-tropical countries, chilli-infecting begomoviruses have emerged as a major threat by infecting economically important crops such as potato, bitter gourd, papaya, tomato and tobacco leading to complete crop loss (Mubin *et al*., 2009). The result shows betasatellite associated with Cotton leaf curl Kokhran virus may be tomato leaf curl betasatellite. In phylogenetic dendogram our clone AA30 has been grouped together with some isolates of tomato leaf curl betasatellite and tomato yellow leaf curl betasatellite in a separate clade (Fig 2).

The complete nucleotide sequence of clone AA33 of alphasatellite was found to be 1361 nt and showed the characteristics of a typical alphasatellite. It encodes a single 311 amino acids long protein on virion sense similar to nanovirus encoded rep (coordinates of gene76-1011 nt). Usually alphasatellite Rep protein is 315 amino acids long but in this case, there is deletion of four amino acids. Pairwise protein sequence alignment was performed with Rep protein of Gossypium darwinii symptomless alphasatellite (GoDSA) using Emboss Water on line tool (http://www.ebi.ac.uk/Tools/psa/emboss_water/). It showed that four amino acids (Lue, Gln, Gly, Try) present on position 50-53 of GoDSA Rep were deleted in Rep of AA33 alphasatellite. The complete sequence of the clone AA33 (accession number KX599187) was analyzed via phylogenetic dendogram and sequence identity with 29 different alphasatellites from a variety of hosts. It made a distinct clade with GoDSA, (FR877535), Xanthium symptomless alphasatellite (XSA, HF547408) and Cotton leaf curl Burewala alphasatellite (CLCuBuA, FN658730) as in Fig 3. Highest nucleotide identity was observed with Xanthium symptomless alphasatellite (XSA, HF547408) which is 98.2%, second highest identity is 97.8% with GoDSA and it has 96.6% identity with CLCuBuA. Other sequences showed less than 77.9% similarity (Table 3). The alphasatellites found in this complex is an isolate of Xanthium symptomless alphasatellite previously reported from *Xanthium strumarium* a weed growing in cotton fields (Mubin *et al.,* 2012).

Cotton is the backbone of economy of Pakistan and CLCuD is a major threat to cotton production in the country Monopartite begomovirus complex is responsible for epidemics of CLCuD. Genetic changes by recombination and pseudorecombination in the complex are driving force for the development of new complexes, which could be more virulent. In present study, we are reporting CLCuKoV, ToLCuB and XSA as complex for the first time. Further experimentation is needed for cloning of full genome of begomovirus and betasatellite which lead to infectivity efficiency of the complex.

Figure 1. Field sample of cotton plant showing typical disease symptoms of leaf curling, vein swelling and vein yellowing (Panel A). Healthy cotton plant is shown in Panel B.

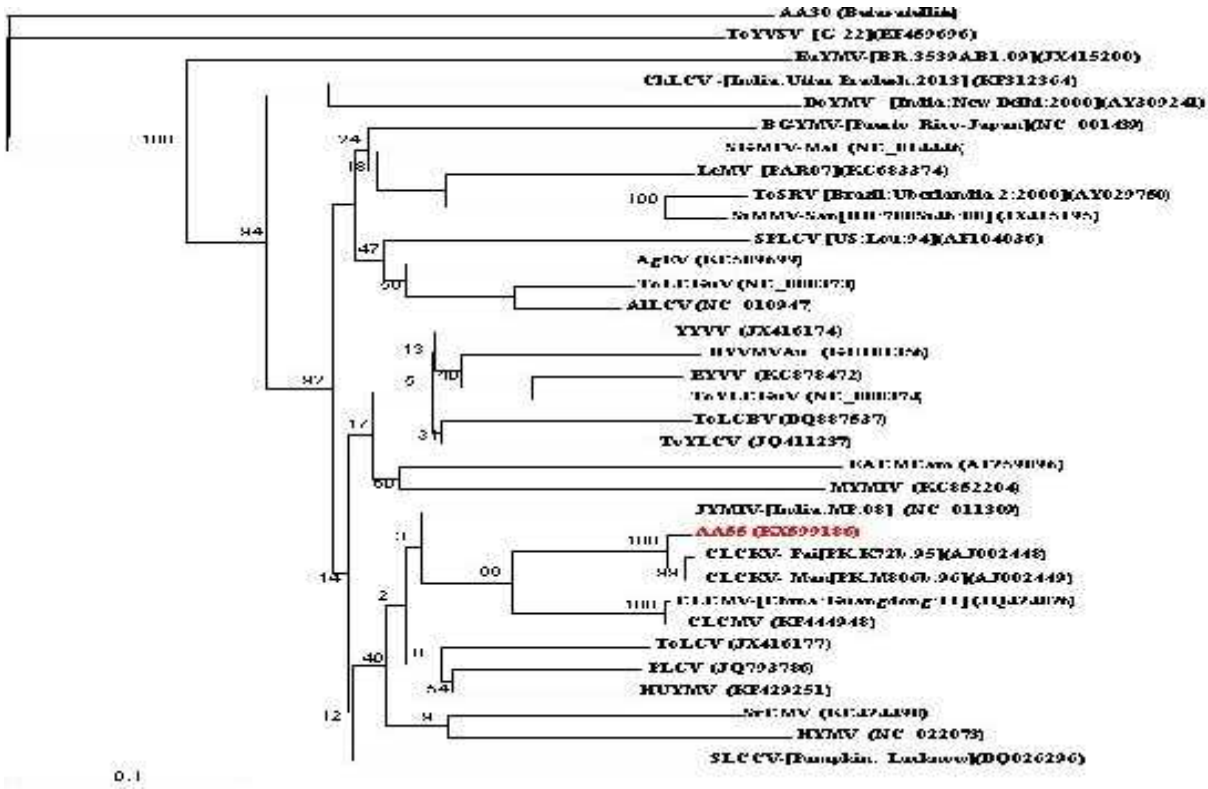

Figure 2. Phylogenetic dendograms based on the alignment of the selected begomovirus partial genome sequences. Vertical branches are arbitrary, horizontal branches are proportional to calculated mutation distances. Values at nodes indicate percentage bootstrap values (1000 replicates). Partial begomovirus sequences used for comparison were *Tomato yellow vein streak virus*–[G-22] (EF 459696), *Euphorbia yellow mosaic virus*-[BR:3539 AB1:09] (JX415200), *Chili leaf curl virus*-[India:UP:13](KF312364), *Dolichos yellow mosaic virus*--[India: New Delhi: 2000] (AY309241), *East African cassava mosaic Cameroon virus*- [Ivory Coast] (AF259896), *Mung bean yellow mosaic India virus* (KC852204), *Sweet potato leaf curl virus*-[US:Lou:94] (AF104036), *Ageratum enation virus* (KC 589699), *Tomato leaf curl Guangdong virus* (NC_008373), *Allamanda leaf curl virus* (NC_010947), *Bean golden yellow mosaic virus*-[Puerto Rico-Japan] (NC_001439), *Sida golden mosaic Florida virus*-Malvastrum (NC_014446), *Leonurus mosaic virus*-[PAR07] (KC683374), *Tomato severe rugose virus*-[Brazil: Uberlandia 2:2000] (AY029750), *Sida micrantha mosaic virus*-[BR:780Si4b:08] (JX415195),*Squash leaf curl China virus*-[Pumpkin: Lucknow] (DQ026296), *Okra yellow vein mosaic virus* Aurangabad (GU181356), *Jatropha yellow mosaic India virus*-[India:MP:08] (NC_011309), *Cotton leaf curl Kokhran virus*-Man [PK:M806b:96] (AJ002449), *Cotton leaf curl Kokhran virus*-Fai [PK:K72b:95] (AJ002448), *Cotton leaf curl Multan virus*–[China:Guangdong:11] (JQ424826), *Cotton leaf curl Multan virus* (KF444948), *Tomato leaf curl virus*-[Australia:Aurukun:2003] (JX416177), *Premna leaf curl virus* (JQ793786), *Hedyotis uncinella yellow mosaic virus* (KF429251), *Sri Lankan cassava mosaic virus*-India [Attur:2009] (KC424490), *Hemidesmus yellow mosaic virus* (NC_022073), *Tomato yellow leaf curl virus* (JQ411237), *Tomato leaf curl Bangalore virus*-[India:KeralaIV:2005] (DQ887537), *Honey suckle yellow vein virus*-[Australia:Ayr:1983] (JX416174), *Emilia yellow vein virus* (KC878472), *Tomato yellow leaf curl Guangdong virus*(NC_008374).

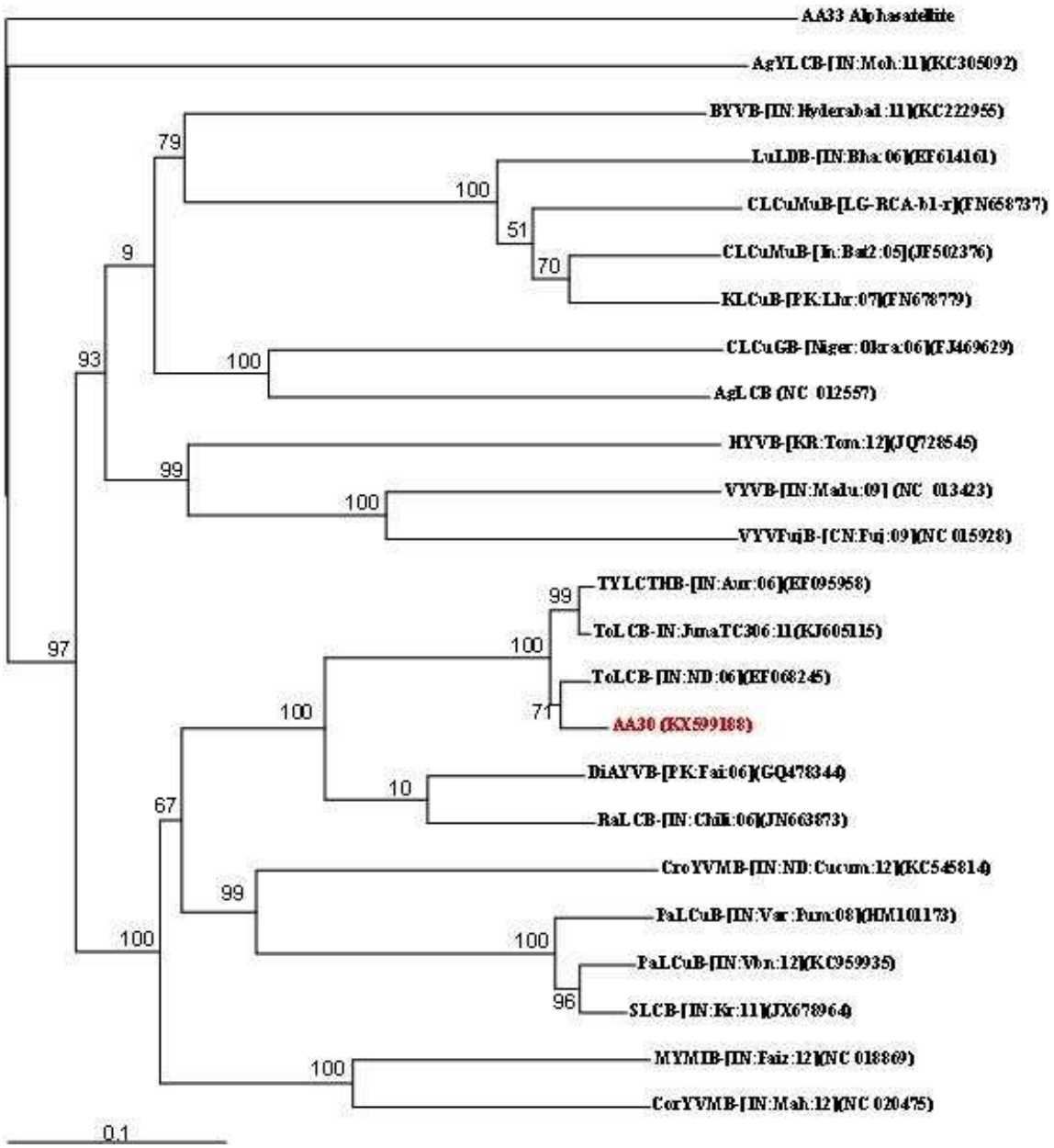

Figure3. Phylogenetic dendograms based on the alignment of the selected betasatellite partial genome sequences. Vertical branches are arbitrary, horizontal branches are proportional to calculated mutation distances. Values at nodes indicate percentage bootstrap values (1000 replicates). Partial betasatellite sequences used for comparison were Ageratum yellow leaf curl betasatellite (KC305092), Cotton leaf curl Gezira betasatellite-[okra:Niger] (FJ469629), Ageratum leaf Cameroon betasatellite (NC_012557), Bhendi yellow vein betasatellite (KC222955), Cotton leaf curl Burewala betasatellite (FN658737), Ludwigia leaf distortion betasatellite [India:Bhangha:Hibiscus] (EF614161), Cotton leaf curl Multan betasatellite (JF502376), Kenaf leaf curl betasatellite (FN678779), Honey suckle yellow vein mosaic betasatellite (JQ728545), Vernonia yellow vein betasatellite (NC_013423), Vernonia yellow vein Fujian virus betasatellite (NC_015928), Mung bean yellow mosaic India virus associated betasatellite (NC_018869), Corchorus yellow vein mosaic betasatellite (NC_020475), Croton yellow vein mosaic betasatellite (KC545814), Papaya leaf curl virus betasatellite (HM101173), Papaya leaf curl virus betasatellite (KC959935), Sunflower leaf curl virus betasatellite (JX678964), Rose leaf curl betasatellite (GQ478344), Radish leaf curl betasatellite (JN663873), Tomato leaf curl virus-associated DNA beta(EF095958),Tomato leaf curl virus-associated DNA beta (EF068245).

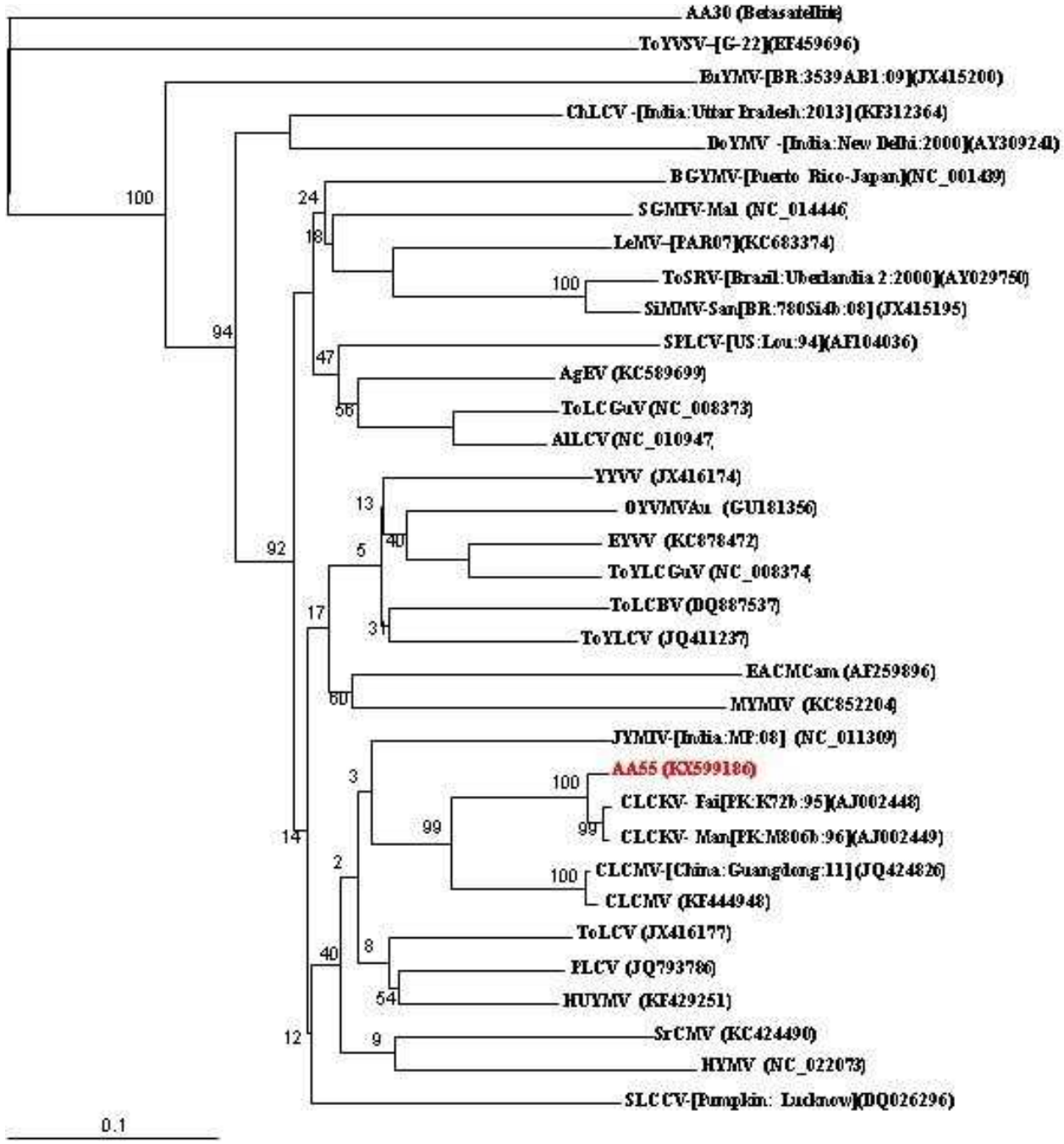

Figure4. Phylogenetic dendograms based on the alignment of the selected alphasatellite genome sequences. Vertical branches are arbitrary, horizontal branches are proportional to calculated mutation distances. Values at nodes indicate percentage bootstrap values (1000 replicates). Alphasatellite complete genome sequences used for comparison were Ageratum yellow vein Singapur alphasatellite, Tomato leaf curl New Dehli alphasatellite, Gossypium darwinii symptomless alphasatellite, Cyamopsistetragonoloba leaf curl alphasatellite, Ageratum conyzoides associated symptomless alphasatellite, Bendhi yellow vein mosaic alphasatellite,, Cassava mosaic Madagascar alphasatellite, Cotton leaf curl Gezira alphasatellite, Gossypium mustilinum symptomless alphasatellite, Croton yellow vein mosaic alphasatellite, Vernonia yellow vein Fujian virus alphasatellite, Holly hock yellow vein symptomless alphasatellite.

Table 1. Nucleotide sequence identity percentage of begomovirus clone AA55 with begomoviruses from databases

AA30 Betasatellite	***																															
AF104036 SPLCV	3.7	***																														
AF259896 EACMCamV	6.4	63	***																													
AJ002448 CLCKV-Fai	4	69.1	65.5	***																												
AJ002449 CLCKV-Man	4	69.1	65.7	99.3	***																											
AY029750 ToSRV	4.7	66.4	63.7	64.7	64.5	***																										
AY309241 DoYMV	3.5	55.5	47.9	48.7	48.9	56.5	***																									
DQ026296 SLCCV	5.5	65.2	65.2	73.4	73.8	64.2	52.8	***																								
DQ887537 ToLCBV	7.7	69.4	65.9	70.9	71.3	66.7	49.4	73.3	***																							
EF459696 ToYVSV	8.9	7.9	4.5	4.2	4.2	2.7	4.5	10.1	6.1	***																						
GU181356 OYVMV	4.5	67.1	65.4	76.8	76.8	65.9	36.1	68.4	75.3	3	***																					
JQ411237 ToYLCV	5.5	68.7	68.7	74.1	73.9	66.6	55.6	72.9	80.5	4	77.5	***																				
JQ424826 CLCuMV	3.9	64.2	64.7	83.5	83.4	68.2	48.1	69.4	70.6	2.9	76.8	72.3	***																			
JQ793786 PLCV	6.6	69.1	68.6	78.7	78.8	59.3	50.6	74.5	73.9	5.2	73.1	78.5	79.2	***																		
JX415195 SiMMV	4	63.9	62.4	63.4	63	89.4	48.2	62.4	64.2	3.7	63.7	66.4	67.7	60	***																	
JX415200 EuYMV	3.5	19	35	43	43.2	31.3	21.7	19.3	34.5	3	45	35.5	44.2	18.7	33.9	***																
JX416174 YYVV	6.4	68.7	66.1	72.3	72.6	60	49.9	70.6	80	4.2	75.8	79.7	72.4	78.7	65.9	33.3	***															
JX416177 ToLCV	6.4	68.9	66.9	77	77.5	60	49.6	72.8	74.1	2.9	72.4	78	77	83.4	61	33.4	78	***														
KC424490 SrCMV	6.4	66.1	55.1	73.8	73.3	61	49.4	69.2	68.7	6.4	67.1	71.9	70.4	77.5	59.8	26.4	70.8	76.8	***													
KC589699 AgEV	6.1	69.2	61.7	70.9	71.4	72.6	49.2	69.2	71.9	6.9	73.3	73.6	74.8	74.1	67.7	46.2	71.8	74.8	66.2	***												
KC683374 LeMV	9.7	59	57.1	63.2	63.2	69.7	49.1	58.2	61.3	6.6	63.4	61.5	66.2	64.2	75.3	51.4	61.5	65.2	59.8	67.1	***											
KC852204 MYMIV	3.4	62	64.4	66.1	66.1	59.3	47.2	66.2	70.1	7.2	66.9	71.3	61.5	66.2	58.3	42.5	71.6	68.4	64.2	59.8	57.8	***										
KC878472 EYVV	5.7	63	66.9	69.6	69.9	66.9	53.8	70.8	77.3	3.7	80.7	79.2	72.8	74.3	66.9	35.6	80.2	72.9	65.5	73.8	59.8	68.1	***									
KF312364 ChLCV	3.4	69.4	63.4	65.5	65.9	62	55.8	65.5	67.2	5.5	65.5	66.6	62.9	68.7	59.2	44.2	66.9	68.4	67.4	68.7	58.5	61.5	65.2	***								
KF429251 HUYMV	2.9	71.3	65.2	78.5	78.5	63.9	48.6	74.8	73.3	3.5	75.1	77	78.7	84.2	63.7	42.9	76	82.7	75.6	76.3	65.2	68.2	73.4	70.4	***							
KF444948 CLCuMV	3.9	67.1	64.4	82.7	82.5	67.9	48.2	69.2	69.9	2.9	76.3	71.6	99.2	78.5	66.7	44.2	72.1	77	70.3	74.1	66.1	60.8	71.8	63.4	78	***						
KX599186 AA55	4	69.2	65.7	97.8	98	64	48.7	74.3	71.1	4	77.3	73.8	84.4	78.5	62.9	42.9	72.3	77.1	73.1	71.3	64	66.1	68.1	66.6	78.7	84	***					
NC_001439 BGYMV	2.4	67.6	63.2	66.6	66.6	68.1	58.7	68.6	66.1	2.5	63	68.9	68.1	67.9	62.9	27.4	66.9	67.9	63.2	66.9	63.7	63.7	55.1	62.7	68.9	67.7	66.2	***				
NC_008373 ToLCGuV	6.4	75.6	67.4	72.4	72.9	71.6	48.9	71.8	76.1	5.7	73.4	77.1	71.8	76.6	68.1	29.2	77.5	75.6	70.8	77.3	64	65.9	75.6	70.9	77.3	71.4	72.4	69.7	***			
NC_008374 ToYLCGuV	3.4	67.6	68.2	73.1	73.6	56.5	53.4	73.1	81.7	3.7	79.7	80.2	70.3	75.8	62.9	32.8	81.8	75.1	70.1	72.6	57.5	70.9	84.5	68.2	75.6	69.6	72.9	64	78	***		
NC_010947 AlLCV	2.9	75.3	68.1	73.6	73.8	67.9	53.6	72.8	75.1	4	72.4	77.3	73.4	78	67.2	42.9	77.5	75.3	69.9	78	65.2	65.9	76	71.9	79	72.9	73.4	70.1	90.6	77.3	***		
NC_011309 JYMV	6.1	69.9	62.5	79.2	79.7	67.1	57	72.3	75	8.1	77	77	74.8	77.5	63.7	43	76.1	80.2	73.4	73.6	61.5	66.4	70.8	69.9	77.6	74.1	79.3	67.4	74.1	75.5	74.8	***		
NC_014446 SGMFV	4	52.4	52.4	52.1	52.3	60.3	46.1	52.4	51.3	3.9	55.8	53.3	57.1	55	64.4	51.4	53.8	52.6	50.1	60	64.7	50.4	54.8	48.6	55.8	56.5	52.3	52.9	58.5	52.6	58.5	52.8	***	
NC_022073 HYMV	9.2	67.4	62.4	69.1	68.7	53.9	50.9	65	67.7	3.7	65.5	70.6	67.2	73.1	60.3	28.2	68.4	72.6	73.9	64.5	55	61.7	62.9	66.9	73.3	66.9	68.9	64.7	69.1	68.2	71.1	71.1	52.6	***
	AA30 Beta	AF104036 SPLCV	AF259896 EACMCamV	AJ002448 CLCKV-Fai	AJ002449 CLCKV-Man	AY029750 ToSRV	AY309241 DoYMV	DQ026296 SLCCV	DQ887537 ToLCBV	EF459696ToYVSV	GU181356 OYVMV	JQ411237 ToYLCV	JQ424826 CLCuMV	JQ793786 PLCV	JX415195 SiMMV	JX415200EuYMV	JX416174 YYVV	JX416177 ToLCV	KC424490 SrCMV	KC589699 AgEV	KC683374 LeMV	KC852204 MYMIV	KC878472 EYVV	KF312364 ChLCV	KF429251 HUYMV	KF444948 CLCuMV	KX599186 AA55	NC_001439 BGYMV	NC_008373 ToLCGuV	NC_008374 ToYLCGuV	NC_010947 AlLCV	NC_011309 JYMV	NC_014446 SGMFV	NC_022073 HYMV

Table 2. Nucleotide sequence identity percentage of betasatellite clone AA30 with betasatellites from databases

KX599188 AA30	***																							
AA33 Alphasatellite	4	***																						
EF068245 ToLCB	96.2	2.6	***																					
EF095958 TYLCTHB	93.4	3.1	95.4	***																				
EF614161 LuLDB	18.4	2.8	48.3	20.4	***																			
FJ469629 CLCuGB	40.5	2.6	41.7	52.1	38.3	***																		
FN658737 CLCuMuB	35	4.3	36.2	17.5	63.2	19.6	***																	
FN678779 KLCuB	56	4.3	52	56	79.6	46.6	68.6	***																
GQ478344 DiAYVB	71.6	1.8	72.5	72.7	36.5	55.8	45.6	54.4	***															
HM101173 PaLCuB	54.9	2.3	55.5	54.9	23.2	13.8	23.6	20.4	58.1	***														
JF502376 CLCuMuB	32.5	4.3	44.5	45.6	78.7	37.4	73.2	87	48.6	23.9	***													
JN663873 RaLCB	73.2	2	74.2	73.6	35.1	43.3	34.7	56.1	84.8	56.1	33	***												
JQ728545 HYVB	42.2	4.6	38.7	41	17.8	27.1	24.8	21.9	33.7	19.2	34.4	36.3	***											
JX678964 SLCB	57.8	2.3	58.4	58.1	19.5	54	43.6	43.3	66.6	91.9	19.6	61.3	14.7	***										
KC222955 BYVB	56.7	4.6	49.2	50.5	45.7	43.6	31.9	45.4	39.9	11.7	42.3	58.7	17.3	22.7	***									
KC305092 AgYLCB	12.1	2.9	11.7	11.5	3.2	6.3	7.1	10	10.9	4.3	9.8	11.5	5.4	4.8	4	***								
KC545814 CroYVMB	56.6	3.1	56.4	57.7	5.2	15.2	16.7	4.9	53.1	56.3	19.9	54	26.1	62.6	21	10.4	***							
KC959935 PaLCuB	57.1	2.5	56.9	57.4	19.9	47.7	20.7	19.5	61	91.9	20.7	59.4	12.7	94.9	21.9	10.3	63.3	***						
KJ605115 ToLCB	94	3.1	96.3	98.8	17.9	36.7	35.4	56.9	72.1	54.8	44.6	73.6	42	57.2	49.4	12.1	57.2	56.4	***					
NC_012557 AgLCB	53.4	2.5	51.5	57.8	21.9	56.7	19.6	19.2	52.6	13.2	19.5	32.5	28.1	39.1	41.1	8	11.2	54.9	50.5	***				
NC_013423 VYVB	36	2.3	36.8	36.5	27.1	16	16.6	35	35.7	11.8	15.2	40.3	26.7	28.7	35.6	4.3	15.2	26.4	36.5	35.4	***			
NC_015928 VYVFujB	29.3	3.2	26.5	26.8	20.2	8.7	3.4	21.8	28.5	19.5	21.2	30.2	31	27.1	37.3	3.8	24.8	17.3	26.8	38.5	63.2	***		
NC_018869 MYMIB	53.4	2.9	53.8	39.9	19.2	15.6	5.5	26.1	57.7	54.4	23.2	33.4	29.4	53.8	20.2	4.6	52.8	36.5	54.6	13.8	25.9	32.2	***	
NC_020475 CorYVMB	54.8	2.5	54.8	55.7	27.1	26.7	17.2	49.7	54	54.1	26.5	53.4	32.1	55.5	6	5.1	53.2	54.9	55.8	13	26.7	27.1	72.1	***
	KX599188 AA30	AA33 Alphasat	EF068245 ToLCB	EF095958 TYLCTHB	EF614161 LuLDB	FJ469629 CLCuGB	FN658737 CLCuMuB	FN678779 KLCuB	GQ478344 DiAYVB	HM101173 PaLCuB	JF502376 CLCuMuB	JN663873 RaLCB	JQ728545 HYVB	JX678964 SLCB	KC222955 BYVB	KC305092 AgYLCB	KC545814 CroYVMB	KC959935 PaLCuB	KJ605115 ToLCB	NC_012557 AgLCB	NC_013423 VYVB	NC_015928 VYVFujB	NC_018869 MYMIB	NC_020475 CorYVMB

Table 3. Nucleotide sequence identity percentage of alphasatellite clone AA33 with alphasatellites from databases

KX599188 AA33	***																		
FJ956707 AYVSGA	23.1	***																	
FN554583 CLCuGA	45.5	26.2	***																
FN658716 BYVMA	69.6	27.4	42.3	***															
FN658717 BYVMA	69.6	27.4	42.3	100	***														
FR772085 AgCSA	79.7	21.9	40.3	79.3	79.3	***													
FR772086 HYVSA	43.3	15.1	73.5	44.1	44.1	44.8	***												
FR772087 GoMSA	52.5	25.4	54.4	47	47	53	72.9	***											
FR873573 CLCuMA	62	14.6	49	82.7	82.7	64.6	52.1	53.4	***										
FR877532 CLCuMA	76.8	16.7	48.9	83.2	83.2	65.2	51.7	55	88.9	***									
FR877534 CrYVMA	49	25.8	56.3	48.5	48.5	48.6	65.2	82.4	54.9	55.2	***								
FR877535 GoDSA	96.1	21.5	47	69.9	69.9	77.9	44.9	52.7	61.9	63.6	48.5	***							
GU385877 CyTLCuA	75.3	21.6	44.7	71	71	85.2	45.9	51.1	70.1	78	49.5	75.6	***						
HE984149 CMA	61.6	14.2	39.4	66.5	66.5	57.1	37.9	40.7	64.5	65.9	42.3	62.1	63.1	***					
HQ343234 CLCuA	60.8	15.6	40.8	70.3	70.3	72.4	53	54.8	73.3	71.6	54.7	60	68.3	64.3	***				
JQ041697 ToLCuNDA	16.8	48.6	13.1	28.5	28.5	30.2	25.3	6.4	16.1	17	23.6	23.3	29.1	14.5	17.5	***			
KC305095 GLCuA	75.1	15.9	47.4	69.2	69.2	83.5	47.3	52.6	71.9	81.5	51.6	74.1	93.8	62.8	73.6	17.7	***		
KC959932 VYVFujA	42.8	26.2	57.1	50.6	50.6	44.6	65.1	75.9	50.3	53.3	75.3	45.1	46.4	40.5	53.2	12.6	54.3	***	
NC_018628 CMMA	63.8	25.6	47.9	68.9	68.9	67.5	46.4	47.6	66.5	67.6	49.4	63.5	64.1	68.4	68.2	15.9	63.8	50.9	***
AJ298903 CLCuMB	9	8.5	2.2	7.2	7.2	8.5	2.1	3.6	5.8	4.7	1.8	8.7	8.5	5.7	3.8	8.2	4.2	2.7	1.7	***
	KX599188 AA33	FJ956707 AYVSGA	FN554583 CLCuGA	FN658716 BYVMA	FN658717 BYVMA	FR772085 AgCSA	FR772086 HYVSA	FR772087 GoMSA	FR873573 CLCuMA	FR877532 CLCuMA	FR877534 CrYVMA	FR877535 GoDSA	GU385877 CyTLCuA	HE984149 CMA	HQ343234 CLCuA	JQ041697 ToLCuNDA	KC305095 GLCuA	KC959932 VYVFujA	NC_018628 CMMA	AJ298903 CLCuMB

Acknowledgments: This research was supported through a start-up research grant no. PM-IPFP/HRD/HEC/2012/ 3543 funded by Higher Education Commission (HEC) Pakistan awarded to Khadim Hussain.

REFERENCES

Albuquerque L.C., A.K. Inoue-Nagata, B. Pinheiro, R.O. Resende, E. Moriones and J. Navas-Castillo (2012). Genetic diversity and recombination analysis of sweepoviruses from Brazil. Virol. J. 9(1): 241.

Amrao L., S. Akhter, M.N. Tahir, I. Amin, R.W. Briddon and S. Mansoor (2010 a). Cotton leaf curl disease in Sindh province of Pakistanis associated with recombinant. Virus. Res. 153:161-165.

Amrao L., I. Amin, M.S. Shahid, R.W. Briddon and S. Mansoor (2010 b). Cotton leaf curl disease in resistant cotton is associated with a single begomovirus that lacks an intact transcriptional activator protein. Virus. Res. 152:153-163.

Briddon R.W., S.E. Bull, I. Amin., S. Mansoor, I.D. Bedford, N. Rishi, S.S. Siwatch, Y. Zafar, A.M. Abdel-Salam and P.G. Markham (2004). Diversity of DNA1; a satellite like molecule associated with monopartite begomovirus-DNAb complexes. Virology. 324:462-474.

Briddon R.W., S. Mansoor, I.D. Bedford, M.S. Pinner, K. Saunders, J. Stanley, Y. Zafar, K.A. Malik and P.G. Markham (2001). Identification of DNA components required for induction of cotton leaf curl disease. Virology. 285:234-243.

Briddon R.W. and P.G. Markham (2001). Complementation of bipartite begomovirus movement functions by topocuviruses and curtoviruses. Arch. Virol. 146:1811-1819.

Brown J.K., C.M. Fauquet, R.W. Briddon, M. Zerbini, E. Moriones and Navas-Castillo (2012). Virus Taxonomy. In: King A MQ, Lefkowitz E, Adams MJ, Carstens EB (ed), Ninth Report of the International Committee on Taxonomy of Viruses. London, UK, Associated Press, pp.351-373.

Doyle J.J. and J.L. Doyle (1987). A rapid isolation procedure for Phytochem Bull. 19:11-15.

Fiallo-Olive E., Y. Martinez-Zubiaur, E. Moriones and J. Navas-Castillo (2012). A novel class of DNA satellites associated with New World begomoviruses. Virology. 427:151-157.

Ha C., S. Coombs, P. Revill, R. Harding, M. Vu and J. Dale (2008). Molecular characterization of begomoviruses and DNA satellites from Vietnam: additional evidence that the New World geminiviruses were present in the Old World prior to continental separation. J. Gen. Virol. 89:312-326.

Hanley-Bowdoin L., S.B. Settlage and D. Robertson (2004). Reprogramming plant gene expression: a pre requisite to geminivirus DNA replication. Mol. Plant. Pathol. 5:149-156.

Hussain K., M. Hussain, S. Mansoor and R.W. Briddon (2011). Complete nucleotide sequence of a begomovirus and associated betasatellite infecting croton (Crotonbon plandianus) in Pakistan. Arch. Virol. 156:1101-1105.

Hussain K., M.A. Mehmood, N. Nahid, A.Q. Khan, A. Akram, Mahmood-ur-Rahman., F. Azeem and S. Shaheen (2013). Molecular characterization of begomovirus associated alphasatellite from an asymptomatic weedplant; Xanthium strumarium L. Pakistan J. Life. and Social Sci 11: 233-237.

Kumar J., S.P. Singh and R. Tuli (2014). Association of satellites with a mastrevirus in natural infection: complexity of Wheat dwarf India virus disease. J. Virol. 88:7093-104.

Mansoor S., I. Amin, S. Iram, M. Hussain, Y. Zafar, K.A. Malik and R.W. Briddon (2003). Break down of resistance in cotton to cotton leaf curl disease in Pakistan. Plant. Pathol. 52:784.

Mansoor S., S. H. Khan, A. Bashir, M. Saeed, Y. Zafar, K.A. Malik, R.W. Briddon, J. Stanley and P.G. Markham (1999). Identification of a novel circular single-stranded DNA associated with cotton leaf curl disease in Pakistan. Virology. 259:190-199.

Melgarejo T.A., T. Kon, M.R. Rojas, L. Paz-Carrasco, F.M. Zerbini and R.L. Gilbertson (2013). Characterization of a New World monopartite begomovirus causing leaf curl disease of tomato in Ecuador and Peru reveals a new direction in geminivirus evolution. J. Virol. 87:5397-5413.

Mubin M., R.W. Briddon and S. Mansoor (2009). Complete nucleotide sequence of chili leaf curl virus and its associated satellites naturally infecting potato in Pakistan. Arch Virol. 154:365-368.

Mubin M., S. Akhtar, I. Amin, R.W. Briddon and S. Mansoor (2012). *Xanthium strumarium*: a weed host of components of begomovirus-betasatellite complexes affecting crops. Virus Genes. 44(1):112-9.

Rahman M., D. Hussain, T.A. Malik and Y. Zafar (2005). Genetics of resistance against cotton leaf curl disease in *Gossypium hirsutum*. Plant. Pathol. 54:764-772.

Rajagopalan P.A., A. Naik, P. Katturi, M. Kurulekar, R.S. Kankanallu and R. Anandalakshmi (2012). Dominance of resistance-breaking cotton leaf curl Burewala virus (CLCuBuV) in north western India. Arch. Virol. 157:855-68.

Sanz A.I., A. Fraile, F. García-Arenal, X. Zhou, D.J. Robinson, S. Khalid, T. Butt and B.D. Harrison (2000). Multiple infection, recombination and genome relationships among begomovirus isolates found in cotton and other plants in Pakistan. J. Gen. Virol. 81:1839-1849.

Wilkins T.A., K. Rajasekaran and D.M. Anderson (2000). Cotton biotechnology. Crit. Rev. Plant. Sci. 19:511-550.