INTERNATIONAL JOURNAL OF AGRICULTURE & BIOLOGY
ISSN Print: 1560–8530; ISSN Online: 1814–9596
22–0394/2023/29–2–159–163
DOI: 10.17957/IJAB/15.2015
http://www.fspublishers.org

Full Length Article

Molecular Characterization of New Begomovirus Complex Infecting *Gossypium hirsutum* in Sindh Province of Pakistan

Ifrah Anwar[1†], Kamran Rashid[1†], Imran Amin[2], Shahid Mansoor[2], Shabnum Shaheen[3], Ioly Kotta-Loizou[4], Mohsin Tariq[1], Amna Aslam[1] and Khadim Hussain[1*]

[1]*Department of Bioinformatics and Biotechnology, Government College University Faisalabad, Pakistan*
[2]*Argricultural Biotechnology Division, National Institute for Biotechnology and Genetic Engineering, Jhang Road, Faisalabad, Pakistan*
[3]*Department of Botany, Lahore College for Women University, Lahore, Pakistan*
[4]*Department of Life Sciences, Faculty of Natural Sciences, Imperial College London, UK*
[*]For correspondence: khadim787@gmail.com
[†]Contributed equally to this work and are co-first authors
Received 08 October 2022; Accepted 10 January 2023; Published 27 January 2023

Abstract

Begomoviruses are the causative agents of the cotton leaf curl disease, which is the main reason behind cotton (*Gossypium hirsutum*) losses in Pakistan as well as worldwide. In addition to annual CLCuD damages, two epidemic outbreaks have been reported in Pakistan during 1992–1993 and 2001–2002. A new recombinant strain, named cotton leaf curl Kokhran virus Burewala (CLCuKoV-Bur), is able to infect cotton varieties which were previously resistant to infection and have been reported as the leading begomovirus that infects cotton in all cotton-growing regions of Pakistan. Recombinant strains of satellite molecules; cotton leaf curl Multan alphasatellite (CLCuMA) and cotton leaf curl Multan betasatellite (CLCuMB) have been reported to accompany CLCuKoV-Bur. Samples of cotton plants showing the symptoms of leaf yellowing, curling, and reddening, were collected from different fields of Taluka Sakrand, Sindh, Pakistan in July 2018. Total genomic DNA extraction was performed by the CTAB method and diagnostic PCR amplifications were done using universal diagnostic primers. The full genome of clone IA6 (CLCuBuV PK: Sin: CCRI: Gossy: 18; MW166879) was amplified using begomovirus specific primers pair (uniNDVA3, uniNDVA4) and the alphasatellite specific primers DNA101 and DNA102 were used to amplify three alphasatellite genomes which were cloned (IA31, IA32 and IA33) and sequenced commercially. Full genome sequence analysis of virus clone showed the highest pairwise sequence identity (99.7%) with CLCuKoV-Bur isolates (accession numbers FR750321 and LN713477) from Punjab, Rahim Yar Khan and Sahiwal and alphasatellite clones IA31 (accession number MW166934) and IA32 (accession number MW166935) showed 96.7–95.6% homology with CLCuMA isolates while IA33 (accession number MW166936) has 92.7% identity with ToLCPKA (HE966421) reported from the district Jhang of Punjab. In the current study, we report for the first time the occurrence of a new alphasatellite named tomato leaf curl Pakistan alphasatellite (ToLCPKA) associated with CLCuKoV-Bur. The significance of this new begomovirus complex from the Sindh province is discussed. © 2023 Friends Science Publishers

Keywords: Begomoviruses; CLCuD; CLCuKoV-Bur; Alphasatellite; ToLCPKA

Introduction

Cotton (*Gossypium hirsutum*) is an economically significant crop in Pakistan and its protection is a major concern. The most important threat for cotton plants is cotton leaf curl disease (CLCuD). Over the last decade, it has been estimated that about 7.1 million bales of cotton were lost due to CLCuD (Mahmood-ur-Rahman *et al*. 2012; Zubair *et al*. 2017). The agents responsible for CLCuD are begomoviruses, which require single-stranded DNA betasatellite molecules that are pathogenicity determinants and help in development of typical disease symptoms in cotton (Amin *et al*. 2011). Some betasatellite-begomovirus complexes are also found to be associated with alphasatellites (Nawaz-ul-Rehman and Fauquet 2009; Anwar *et al*. 2020). Alphasatellites are also single-stranded DNA molecules associated with geminivirus and nanoviruses. Geminialphasatellites are 1.4 Kb in size and are associated with begomoviruses and mastreviruses from family *Geminiviridae* (Hamza *et al*. 2018). The begomovirus disease complexes are rapidly evolving by swapping different satellite components leading to recombination and

To cite this paper: Anwar I, K Rashid, I Amin, S Mansoor, S Shaheen, I Kotta-Loizou, M Tariq, A Aslam, K Hussain (2023). Molecular characterization of new begomovirus complex infecting *Gossypium hirsutum* in sindh province of Pakistan. *Intl J Agric Biol* 29:159–163

pseudorecombination which appears to be important for evading host resistance.

In the 1990s, CLCuD triggered a main loss of cotton in cotton growing districts of the Punjab province; however, there was a little effect in northern districts of the Sindh province while most cotton growing districts of southern Sindh remained unaffected. Although it has been reported that different begomoviruses affect different crops in Sindh, but there was no report in cotton till late nineties of the last century (Sanz *et al.* 2000). The begomoviruses and their associated whitefly vectors in Sindh were different from those which were reported in the Punjab province (Simón *et al*. 2003). During 1997–1998, CLCuD was reported to affect the cotton crop in the Sindh province as well, but the damage was not significant compared to that of Punjab (Mansoor *et al.* 1998). However, from 2003–2004, significant losses of cotton yield have been observed in Sindh, which has been proven to be associated with a new species of monopartite begomovirus which was reported by (Amrao *et al*. 2010) as cotton leaf curl Shahdadpur virus (CLCuShV).

A study on begomovirus diversity in cotton in two major cotton-growing provinces of Pakistan; Punjab and Sindh provinces reported as the dominant complexes the cotton leaf curl Kokhran virus Burewala strain (CLCuKoV-Bur) with the cotton leaf curl Multan betasatellite (CLCuMB) and a bipartite begomovirus tomato leaf curl New Dehli virus (Zaidi *et al*. 2016). Association of the cotton leaf curl Sindh alphasatellite and the cotton leaf curl Shahdadpur alphasatellite with the begomovirus-betasatellite complex have been reported (Amrao *et al*. 2010). Cotton leaf curl disease causing complex is changed after some time to evade the host resistance. Here we reported for the first time the association of tomato leaf curl Pakistan alphasatellite (ToLCPKA) and cotton leaf curl Multan alphasatellite (CLCuMA) with CLCuKoV-Bur in the Sindh province of Pakistan.

Materials and Methods

In 2018, cotton plants showing the typical CLCuD symptoms (leaf yellowing, curling and reddening, vein-thickening and foliar discoloration) were collected from different fields of Taluka Sakrand (Sindh, Pakistan) using sterile disposable gloves for each sample to avoid the cross contaimination. The collected samples were packed in clean polyethene bags and labelled with location and date of collection. Those samples are transported in ice-filled cooler to the molecular virology lab. Genomic DNA was extracted from a leaf sample by using the modified CTAB method as described by Akram *et al*. 2017 (Doyle and Doyle 1990; Akram *et al.* 2017) and visualised with 1% (w / v) agarose gel. The DNA then served as template in polymerase chain reaction (PCR) amplifications with the begomovirus specific primers uniNDVA3 (5'-GAG CTC GTG CAG TTG TCC CCA TTG CCC GCG TCA C-3') and uniNDVA4 (5'-GAG CTC CAT AGG GGC TGT CGA AGT TG-3'), and the alphasatellite specific primers DNA101 (5'-CTG CAG ATA ATG TAG CTT ACC AG-3') and DNA102 (5'-CTG CAG ATC CTC CAC GTG TAT AG-3') (Bull *et al*. 2003). The amplicons were cloned in a TA plasmid vector and sequenced. Amplified PCR products were purified and sent for commercially available services of Sanger sequencing from Macrogen Inc. Soule, S. Korea. Received sequences of different amplicons of begomoviruses and alphasatellites were cleaned and assembled into a single consensus viral sequence using Lasergene (DNA-Star Inc., Madison, WI, USA) and submitted to the GenBank database. Potential genes in the begomovirus sequences and the associated betasatellite sequences were determined using the online NCBI tool ORFfinder (https://www.ncbi.nlm.nih.gov/orffinder/). Closely related sequences were retrieved from public databases using BLAST (Zhang *et al.* 2000) and multiple sequence alignments were performed using Clustal W, as implemented by MEGA 7.0 (Kumar *et al.* 2016). Pairwise sequence comparisons were done using MegAlign, as implemented by Lasergene and Sequence Demarcation Tool (SDT) (Muhire *et al.* 2014). Phylogenetic analysis was conducted by constructing a phylogenetic tree using the neighbor-joining algorithm as implemented by MEGA 7.0 with 1,000 bootstrap replicates. Abbreviations of begomovirus and their satellites names are shown as previously (Varsani *et al.* 2017).

Results

The complete nucleotide sequence of the begomovirus clone (IA6) (2759bp) (PK: Sin: CCRI: Gossy: 18 CLCuBuV_MW166879) confirmed its identity as a CLCuKoV-Bur strain. The IA6 clone had 99.7% sequence identity with CLCuKoV-Bur isolates (accession numbers FR750321 and LN713477) from Punjab Rahim Yar Khan and Sahiwal. The pairwise distance matrix (Fig. 1) of the IA6 clone with similar sequences from public databases demonstrated 99.7–99% identity with CLCuKoV-Bur strains. Phylogenetic analysis of the IA6 clone (PK: Sin: CCRI: Gossy: 18 CLCuBuV_MW166879) (Fig. 2) confirmed its homology with different CLCuKoV-Bur isolates from various regions of India and Pakistan.

From the same cotton sample, three alphasatellites were cloned (IA31, IA32 and IA33) consisting of two species. The sequence analysis of IA31 (1385 bp) (PK: Sind: Gossy: 18_CLCuMA_MW166934) and IA32 (1376 bp) (PK: Sind: Gossy: 18_CLCuMA_MW166935) revealed these as CLCuMA while IA33 was revealed as ToLCPKA. Specifically, the IA31 (accession number MW166934) and IA32 (accession number MW166935) clones (Fig. 3) showed 96.7–95.6% homology with CLCuMA sequences (accession numbers LT840038, LT840037, LT840024, LT840044), while the pairwise sequence alignment (Fig. 4) indicated that IA33 (1384 bp) (accession number MW166936) has 92.7% identity with ToLCPKA (HE966421) reported from the

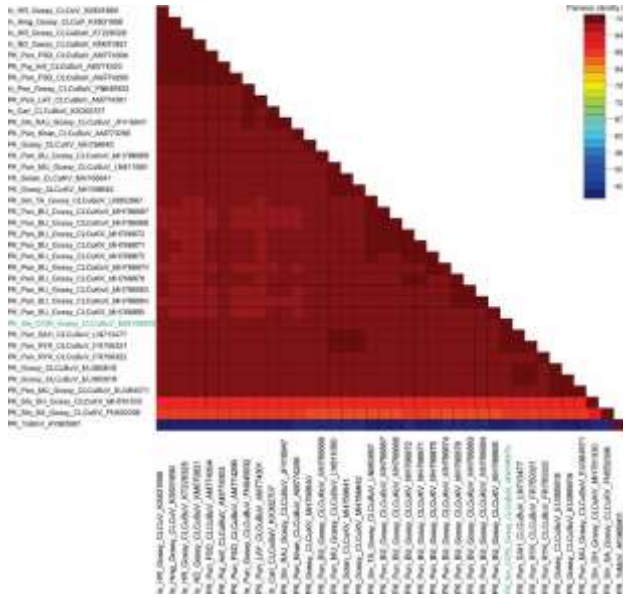

Fig. 1: Pairwise distance matrix of begomovirus sequences alinged by CLUSTAL W. The percent identity of the begomovirus IA6 clone (PK:Sin:CCRI:Gossy: CLCuBuV MW166879) with similar sequences from public databases is shown

Jhang district of Punjab. Phylogenetic analysis (Fig. 2) also showed that IA33 (PK: Sind: Gossy: 18_ToLCuPKA_MW166936) is grouping with different TLCuPKA isolates.

Discussion

Betasatellite associated monopartite begomovirus disease complex can significantly impact the cotton yield in Pakistan and Indian regions and was reported previously to occur most oftenly as multiple infection of distinct begomovirus species (Zaidi *et al*. 2016). This disease complex can readily advance by capturing additional new components, recombining and genetically modifying the existing components to overcome host resistance and to increase its host range (Sattar *et al*. 2013; Zubair *et al*. 2017). There are reports of more than a few monopartite begomoviruses involved in CLCuD in cotton in Pakistan. In the 1990s, minimum four different species of begomovirus, including CLCuKoV and CLCuMuV, CLCuAlV and PaLCuV were involved in CLCuD, either as single or mixed infections. In contrast, only one betasatellite molecule namely, CLCuMB, has been found to be part of the cotton leaf curl begomovirus complex (Briddon *et al*. 2004; Nawaz-ul-Rehman *et al*. 2017). Due to the emergence of the new recombinant CLCuKoV-Bur strain, the conventional resistant cotton varieties also become susceptible to CLCuD (Yasmeen *et al*. 2016).

Our current study focused on the diversity of begomoviruses and associated satellites in the Sindh province of Pakistan. Diversity studies in a cotton sample collected from Taluka Sakrand (Sindh province) revealed that it

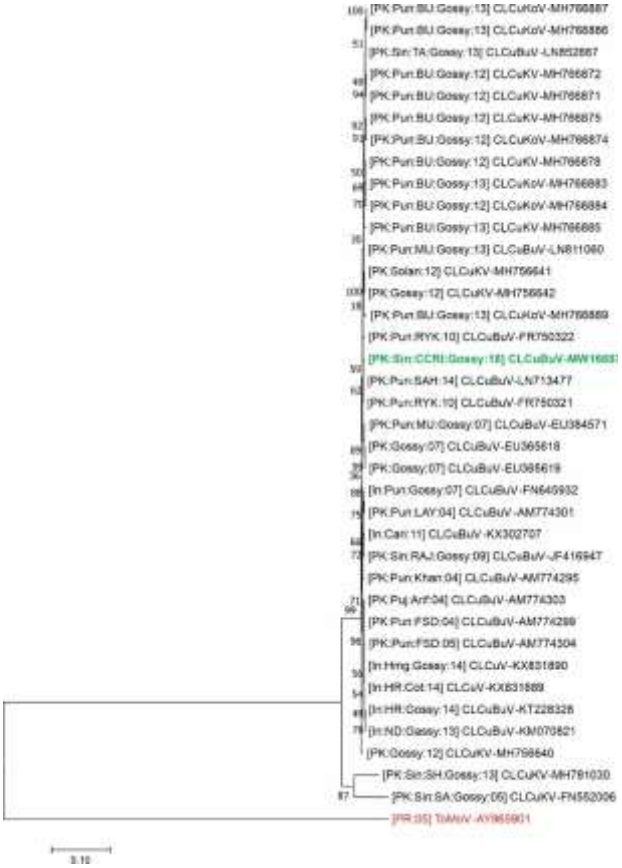

Fig. 2: Phylogenetic analysis of the begomoviruse clone IA6. (PK:Sin:CCRI:Gossy:18 CLCuBuV-MW166879). Based on the alignment of the CLCuKoV-Bur sequences, a neighbor-joining phylogenetic tree was constructed, with closely related begomovirus sequences. Isolates in green were obtained in the present study. Sequence accession number of each isolate is listed together with the strain descriptors (in square brackets), which include the country, location, host, and year of sampling. The tree was rooted using the sequence of tomato mottle virus (ToMoV) as an outgroup

contained CLCuKoV-Bur associated with two alphasatellites, which were previously reported in Punjab Pakistan (Akhtar *et al*. 2010). Although the role of alphasatellites in disease etiology is not clear, it has been reported that it is an important component of begomovirus disease complex and a target of the host defense (Nawaz-ul-Rehman *et al*. 2010). The begomovirus clone (IA6) from cotton had 99.7% identity with CLCuKoV-Bur isolated from Rahim Yar Khan and Sahiwal in the Punjab province. From the same cotton sample, two types of alphasatellites were isolated, namely CLCuMA and ToLCPKA, demonstrating the fact that multiple alphasatellites can simultaneously be associated with a begomovirus inside a single host. In 2016, a study also reported the presence of three alphasatellites with a begomovirus in a single cotton plant in Punjab Pakistan (Siddiqui *et al*. 2016). Another recent study carried out in north western states of India revealed that analyses of cotton

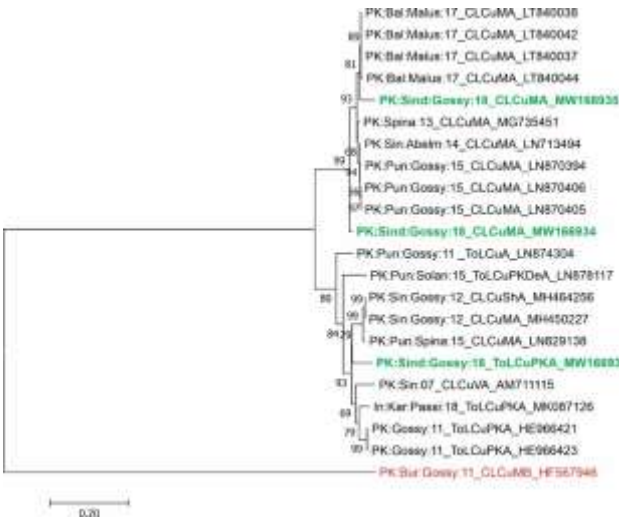

Fig. 3: Phylogenetic analysis of the alphasatellite clones IA31 (PK:Sind:Gossy:18_CLCuMA; MW166934), IA32 (PK:Sind:Gossy:18_CLCuMA; MW166935) and IA33 (PK:Sind:Gossy:18_ToLCuPKA; MW166936). A Neighbor-Joining phylogenetic tree was constructed based on alignments of the sequences of Cotton leaf curl Multan alphasatellite (CLCuMA), Tomato leaf curl alphasatellite (ToLCuA), Tomato leaf curl Pakistan defective alphasatellite (ToLCuPKDeA), Cotton leaf curl Shahdadpur alphasatellite (CLCuShA), Tomato leaf cur Pakistan alphasatellite (ToLCPKA) and Cotton leaf curl virus alphasatellite (CLCuVA). The tree was rooted using the sequence of cotton leaf curl Multan betasatellite (CLCUMB) as an outgroup

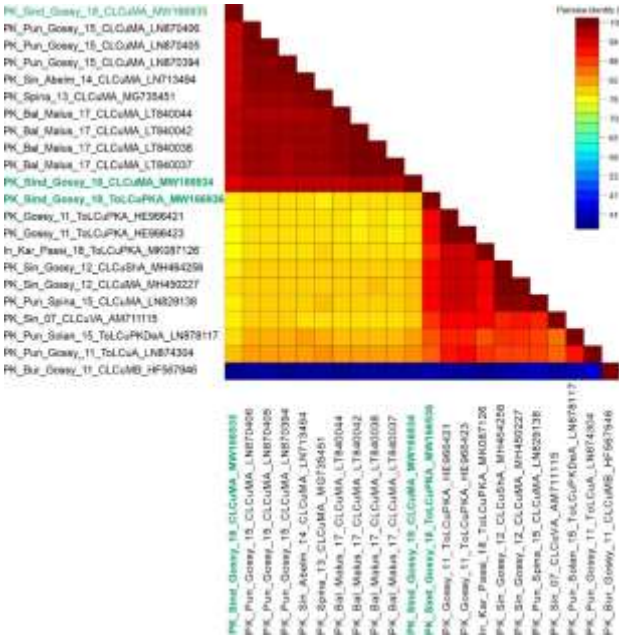

Fig. 4: Pairwise distance matrix of alphasatellite sequences aligned by CLUSTAL W using sequence demarcation tool. The % identity of alphasatellite clones IA31(PK:Sind:Gossy:18 CLCuMA MW166934), IA32 (PK:Sind:Gossy:18 CLCuMA MW166935) and IA33 (PK:Sind:Gossy:18 ToLCuPKA MW166936) with similar sequences from public databases is shown

field symptomatic samples for three consecutive years, two different begomovirus species Cotton leaf curl Multan virus (with three different strains named as Cotton leaf curl Multan virus-Rajasthan, Cotton leaf curl Multan virus-Pakistan, Cotton leaf curl Multan virus-Faisalabad) and Cotton leaf Kokhran virus-Burewala strain were identified. CLCuMV-Raj strain; the most abundantly found begomovirus strain is this study was associated with Cotton leaf curl Multan betasatellite (CLCuMB) and Gossypium darwinii symptomless alphasatellite (GDarSLA), and Croton yellow vein mosaic alphasatellite (CrYVMoA). CLCuKoV-Bur was found second most abundant in cotton samples and associated satellites were CLCuMB and GDarSLA (Biswas *et al.* 2020). These studies support our findings that there is no species barrier in the association of satellites with helper begomoviruses as alphasatellites; tomato leaf curl Pakistan alphasatellite (ToLCPKA) was identified in cotton. The CLCuMA isolates from cotton have the highest similarity with those isolated from *Malus pumila* originating from Pakistan.

Conclusion

This study focused on characterizing the begomoviruses and associated satellites present in the Sindh province, Pakistan. Pakistan's economy heavily depends on cotton and the major cause that decreased cotton yields is CLCuD. The causative agent of CLCuD was recognized as a begomovirus. Interestingly, begomoviruses possess the strategy of recombination and pseudorecombination, allowing them to overcome the plant defense mechanisms and control their machinery for propagating themselves. This is a serious problem as the virus can modify itself by recombination and evolve into a more virulent complex by capturing diverse components. In current study it is concluded that CLCuKoV-Bur associates with different alphasatellites to cope with host defense for bona fide infection. To the best of our knowledge this is the first report of ToLCuPKA associated with CLCuD in the Sindh region of Pakistan.

Acknowledgments

All authors are thankful to Higher Education Commission (HEC) Pakistan, for providing funds through project No. 5584/Punjab/NRPU/R&D/HEC/2016, which was awarded to Khadim Hussain.

Author Contributions

KH designed and directed the project. IA, IA and KH were involved in sample collection and performing the experimental work. IA, KR, SM and KH analyzed the data. KR, AA, KH, MT, SS and IKL wrote the manuscript. IKL and MT were involved in the revision of the final file of the manuscript.

Conflicts of Interest

Authors declare no conflict of interests and all authors read and approved the manuscript and agreed to submit it in IJAB for publication.

Ethics Approval

The present research does not involve any animal as experimental organism therefore approval from ethical committee was not needed.

Data Availability

The full genome sequences of clone IA6 (CLCuBuV PK: Sin: CCRI: Gossy: 18) is available under accession number MW166879 and alphasatellite clones IA31 and IA32 under accession numbers MW166934 and MW166935, respectively in Genbank database.

References

Akhtar KP, S Haidar, MKR Khan, M Ahmad, N Sarwar, MA Murtaza, M Aslam (2010). Evaluation of Gossypium species for resistance to cotton leaf curl Burewala virus. *Ann Appl Biol* 157:135–147

Akram A, K Hussain, N Nahid, Mahmood-ur-Rahman, A Nasim, S Shaheen (2017). Molecular characterization of a begomovirus and associated satellites from cotton (Gossypium hirsutum) from Dera Ghazi Khan district of Pakistan. *J Anim Plant Sci* 27:1245–1255

Amin I, K Hussain, R Akbergenov, JS Yadav, J Qazi, S Mansoor, T Hohn, CM Fauquet, RW Briddon (2011). Suppressors of RNA silencing encoded by the components of the cotton leaf curl begomovirus-betasatellite complex. *Mol Plant Microbe Interact* 24:973–983

Amrao L, S Akhter, MN Tahir, I Amin, RW Briddon, S Mansoor (2010). Cotton leaf curl disease in Sindh province of Pakistan is associated with recombinant begomovirus components. *Virus Res* 153:161–165

Anwar I, HA Bukhari, N Nahid, K Rashid, I Amin, S Shaheen, K Hussain, S Mansoor (2020). Association of cotton leaf curl Multan betasatellite and Ageratum conyzoides symptomless alphasatellite with tomato leaf curl New Delhi virus in Luffa cylindrica in Pakistan. *Aust Plant Pathol* 49:25–29

Biswas KK, UK Bhattacharyya, S Palchoudhury, N Balram, A Kumar, R Arora, SK Sain, P Kumar, RK Khetarpal, A Sanyal, PK Mandal (2020). Dominance of recombinant cotton leaf curl Multan-Rajasthan virus associated with cotton leaf curl disease outbreak in northwest India. *PLoS One* 15:e0231886

Briddon RW, SE Bull, I Amin, S Mansoor, ID Bedford, N Rishi, SS Siwatch, Y Zafar, AM Abdel-Salam, PG Markham (2004). Diversity of DNA 1: A satellite-like molecule associated with monopartite begomovirus–DNA β complexes. *Virology* 324:462–474

Bull SE, RW Briddon, PG Markham (2003). Universal primers for the PCR-mediated amplification of DNA 1: A satellite-like molecule associated with begomovirus-DNA β complexes. *Mol Biotechnol* 23:83–86

Doyle JJ, JL Doyle (1990). Isolation of plant DNA from fresh tissue. *Focus* 12:13–15

Hamza M, MN Tahir, R Mustafa, H Kamal, MZ Khan, S Mansoor, RW Briddon, I Amin (2018). Identification of a dicot infecting mastrevirus along with alpha-and betasatellite associated with leaf curl disease of spinach (Spinacia oleracea) in Pakistan. *Virus Res* 256:174–182

Kumar S, G Stecher, K Tamura (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33:1870–1874

Mahmood-ur-Rahman, K Hussain, MA Khan, A Bakhsh, A Q Rao (2012). An insight of cotton leaf curl virus: A devastating plant pathogenic begomovirus. *Pure Appl Biol* 1:52–58

Mansoor S, M Hussain, SH Khan, A Bashir, AB Leghari, GA Panwar, WA Siqqiqui, Y Zafar, KA Malik (1998). Polymerase chain reaction-based detection of cotton leaf curl and other whitefly-transmitted geminiviruses from Sindh. *Pak J Biol Sci* 1:39–43

Muhire BM, A Varsani, DP Martin (2014). SDT: A virus classification tool based on pairwise sequence alignment and identity calculation. *PLoS One* 9:e108277

Nawaz-ul-Rehman MS, CM Fauquet (2009). Evolution of geminiviruses and their satellites. *FEBS Lett* 583:1825–1832

Nawaz-ul-Rehman MS, N Nahid, S Mansoor, RW Briddon, CM Fauquet (2010). Post-transcriptional gene silencing suppressor activity of two non-pathogenic alphasatellites associated with a begomovirus. *Virology* 405:300–308

Nawaz-ul-Rehman MS, S Shakir, S Ijaz, M Mubin, N Nahid (2017). Status and Diversity of Begomoviruses in Pakistan. *In Begomoviruses: Occurrence and Management in Asia and Africa*, pp:255–269. Springer, Singapore

Sanz AI, A Fraile, F García-Arenal, X Zhou, DJ Robinson, S Khalid, T Butt, BD 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

Sattar MN, A Kvarnheden, M Saeed, RW Briddon (2013). Cotton leaf curl disease–an emerging threat to cotton production worldwide. *J Gen Virol* 94:695–710

Siddiqui K, S Mansoor, RW Briddon, I Amin (2016). Diversity of alphasatellites associated with cotton leaf curl disease in Pakistan. *Virol Rep* 6:41–52

Simón B, JL Cenis, F Beitia, S Khalid, IM Moreno, A Fraile, F García-Arenal (2003). Genetic structure of field populations of begomoviruses and of their vector *Bemisia tabaci* in Pakistan. *Phytopathology* 93:1422–1429

Varsani A, P Roumagnac, M Fuchs, J Navas-Castillo, E Moriones, A Idris, RW Briddon, R Rivera-Bustamante, FM Zerbini, DP Martin (2017). Capulavirus and Grablovirus: Two new genera in the family Geminiviridae. *Arch Virol* 162:1819–1831

Yasmeen A, S Kiani, A Butt, AQ Rao, F Akram, A Ahmad, IA Nasir, T Husnain, S Mansoor, I Amin, S Aftab (2016). Amplicon-based RNA interference targeting V2 gene of cotton leaf curl Kokhran Virus-Burewala strain can provide resistance in transgenic cotton plants. *Mol Biotechnol* 58:807–820

Zaidi SS, M Shafiq, I Amin, BE Scheffler, JA Scheffler, RW Briddon, S Mansoor (2016). Frequent occurrence of Tomato leaf curl New Delhi virus in cotton leaf curl disease affected cotton in Pakistan. *PLoS One* 11:e0155520

Zhang Z, S Schwartz, L Wagner, W Miller (2000). A greedy algorithm for aligning DNA sequences. *J Comput Biol* 7:203–214

Zubair M, SS Zaidi, S Shakir, M Farooq, I Amin, JA Scheffler, BE Scheffler, S Mansoor (2017). Multiple begomoviruses found associated with cotton leaf curl disease in Pakistan in early 1990 are back in cultivated cotton. *Sci Rep* 7:680