Brazilian Journal of Biology®

ISSN 1519-6984 (Print)
ISSN 1678-4375 (Online)

Original Article

Alternanthera yellow vein virus (AYVV); a betasatellite independent begomovirus infecting *Sonchus palustris* in Pakistan

Alternanthera yellow vein virus (AYVV); um begomovírus independente de betassatélites infectando *Sonchus palustris* L. no Paquistão

M. S. Nawaz-Ul-Rehman[a], I. Liaqat[b], N. Nahid[c], F. Saleem[a], S. Alkahtani[d], A. Al Qahtani[e], J. Ye[f] and M. Mubin[a,f*]

[a]University of Agriculture, Centre of Agricultural Biochemistry and Biotechnology, Faisalabad, Pakistan
[b]GC University Lahore, Department of Zoology, Lahore, Pakistan
[c]GC University Faisalabad, Department of Bioinformatics and Biotechnology, Faisalabad, Pakistan
[d]King Saud University, College of Science, Department of Zoology, Riyadh, Saudi Arabia
[e]King Faisal Specialist Hospital and Research Center, Research Center, Department of Infection and Immunity, Riyadh, Saudi Arabia
[f]Chinese Academy of Sciences, Institute of Microbiology, Laboratory of Vector-Borne Diseases, State Key Laboratory of Plant Genomics, Beijing, PR China

Abstract

Satellites associated begomoviruses are the most diverse group of plant viruses in tropical and subtropical regions. In Pakistan, during field surveys in 2019-2020, *Sonchus palustris* (a weed plant) was observed showing begomovirus symptoms i.e., vein yellowing and mosaic patterns on leaves. Rolling circle amplification from total isolated DNA of symptomatic leaves was performed to amplify circular viral genomes. Subsequent cloning and sequencing showed that a new strain of *Alternanthera yellow vein virus* (AlYVV) is associated with vein yellowing disease of *S. palustris*. The identity percentage analysis through BLAST search and SDT analysis showed that the new strain is 94-98% identical to AlYVV isolates reported from Pakistan, India and China. In phylogenetic tree, it clustered with AlYVV-[PK:E prostrata:15-KX710155], AlYVV-[PK:E prostrata:13]-KX906697] and AlYVV-[PK:E prostrata:11]-KX906694] previously reported from Pakistan. There was no detectable level of betasatellite or any other satellite molecule in the samples studied here. Phylogenetic analysis of *Rep* and *CP* genes of AlYVV with corresponding genes of closely related viruses circulating in Southeast Asia showed intra-specific recombination involving both complementary and virion sense region of virus. Relaxed clock and Bayesian Skyline Plot analysis based on *CP* gene sequences indicated slight higher substitution rates (4.75×10^{-3} substitutions/nucleotide/year). In the Indian subcontinent satellite-associated monopartite begomoviruses predominately infect crops and non-crop plants. But AlYVV is found infecting mostly non-crop plants independent of satellite molecules. We hypothesize here that AlYVV evolved as a true monopartite begomovirus in the Indian sub-continent and could be a great threat to introduced crops under suitable conditions. Such studies are crucial to understand probable future epidemics of begomoviruses in the region.

Keywords: begomovirus, phylogeny, recombination, Bayesian analysis, mutation.

Resumo

Os begomovírus associados aos satélites são o grupo mais diversificado de vírus de plantas encontrado em regiões tropicais e subtropicais. No Paquistão, durante pesquisas de campo entre 2019 e 2020, a espécie *Sonchus palustris* L. (uma planta daninha) foi observada apresentando sintomas de begomovírus, ou seja, amarelecimento das veias e padrões de mosaico nas folhas. A Amplificação em Círculo Rolante (ACR) a partir de DNA isolado total de folhas sintomáticas foi realizada para amplificar genomas virais circulares. A clonagem e sequenciamento subsequentes mostraram que uma nova cepa de *Alternanthera yellow vein virus* (AlYVV) está associada à doença do amarelecimento das veias *de S. palustris*. A análise da porcentagem de identidade por meio de pesquisa BLAST e análise SDT mostrou que a nova cepa é 94-98% idêntica aos isolados de AlYVV relatados no Paquistão, Índia e China. Na árvore filogenética, essa cepa se agrupou com AlYVV-[PK:E prostrata:15-KX710155], AlYVV-[PK:E prostrata:13]-KX906697] e AlYVV-[PK:E prostrata:11]-KX906694] relatada anteriormente de Paquistão. Não houve nível detectável de betassatélite ou qualquer outra molécula satélite nas amostras estudadas aqui. A análise filogenética de genes *Rep* e *CP* de AlYVV com genes correspondentes de vírus intimamente relacionados que estão circulando no Sudeste Asiático mostrou recombinação intraespecífica envolvendo a região complementar e de sentido viral do vírus. Relógio molecular relaxado e análise de Bayesian Skyline Plot (BSP) com base nas sequências do gene *CP* indicaram taxas

*e-mail: mmubin@uaf.edu.pk
Received: March 19, 2022 – Accepted: May 13, 2022

de substituição ligeiramente mais altas (4,75 x 10-3 substituições/nucleotídeo/ano). No subcontinente indiano, os begomovírus monopartidos associados aos satélites infectam predominantemente culturas e plantas não cultivadas. Mas o AlYVV é encontrado infectando principalmente plantas não cultivadas, independentemente de moléculas satélites. Desenvolveu-se a hipótese de que o AlYVV evoluiu como um verdadeiro begomovírus monopartido no subcontinente indiano e pode ser uma grande ameaça às culturas introduzidas em condições adequadas. Tais estudos são cruciais para entender prováveis e futuras epidemias de begomovírus na região.

Palavras-chave: begomovírus, filogenia, recombinação, análise Bayesiana, mutação.

1. Introduction

Whitefly-transmitted begomoviruses of *Geminiviridae* family are emerging as major pathogens on food and fiber crops in the Indian subcontinent (Mansoor et al., 2006). These viruses have ~2.6kb to 5.2kb circular single-stranded genomes and are transmitted by whitefly (Zerbini et al., 2017). The emergence of more diverse begomovirus species with broader host range could be the result of multiple infections, recombination, component capture/exchange and emergence of different biotypes of whiteflies (Mansoor et al., 2006; Czosnek et al., 2017). The natural sources of resistance are limited and often prone to breakdown due to emergence of highly virulent strains (Amrao et al., 2010).

Weeds serve as alternative host for crop-infecting begomoviruses, when main cropping season is not there. There are many reports of begomoviruses and associated satellites i.e., alphasatellite and betasatellite (Briddon and Stanley, 2006) infecting weeds, like *Sonchus arvensis, Eclipta prostrata* and *Alternanthera* (Mubin et al., 2009; Guo and Xhou, 2005; Mubin et al., 2010; Ghulam et al., 2017).

The genomes of begomoviruses consist of either two genomic components of approx. 2.8kb, known as DNA A and DNA B i.e., New World (NW) begomoviruses or of a single component homologous to the DNA A component of the bipartite viruses i.e., Old World (OW) monopartite begomoviruses (Briddon, 2003; Mansoor et al., 2006). Betasatellites are pathogenicity determinant molecules while alphasatellites are diverse in nature and do not play a significant role in the development of disease symptoms. Although *Alpha-Rep* protein was found to be a suppressors of gene silencing, but its downstream pathways are not understood (Nawaz-ul-Rehman et al., 2010). In the Old World, alphasatellites and betasatellites molecules are associated mostly with monopartite begomoviruses (Mansoor et al., 2001; Briddon and Stanley, 2006). We know more about begomoviruses infecting field crops as compared to overall begomovirus diversity in agriculture ecosystem. Based on available data of begomoviruses diversity, it can be hypothesized that a large number of unknown begomoviruses reside within non-cultivated plant species (Haider et al., 2007; Iram et al., 2005; Murtaza et al., 2018; Mubin et al., 2009, 2012).

The *Sonchus palustris* belonging to family *Asteraceae* is an annual herb found in subtropical and tropical regions of the world including Pakistan. *S. palustris* can grow and thrive in the varied geographical conditions, mainly found around the wet crop fields and moist places. Objective of this study is to uncover the component viruses, possibility of variation in the disease complex and understand the phylogeny of begomoviruses infecting *S. palustris*.

During a survey conducted in 2019-20, natural occurrence of vein yellowing disease was observed on *S. palustris*. We analyzed symptomatic plants for the possible presence of components of begomovirus disease complex using rolling circle amplification-based protocol, which has improved our ability to identify circular DNA viruses in large numbers. As expected, vein-yellowing disease was associated with AlYVV while no betasatellites or alphasatellites were detected from the plants analyzed. Phylogenetic analysis of AlYVV and coat protein and replication associated genes was also analyzed. This data will help in understanding the diversity of begomoviruses in weed hosts may aid in devising control strategies against begomoviruses.

2. Materials and Methods

2.1. Virus sources

Ten *Sonchus palustris* symptomatic plants showing vein yellowing were collected from farmer's fields in Faisalabad, Punjab. Two types of leaf samples; green leaves with no visible symptoms as control (Figure 1A) and leaves showing vein yellowing (Figure 1B), were collected. Young leaves were collected, labeled and transported on ice to lab and stored at -80°C. Total DNA was extracted from leaf samples by CTAB method described by Doyle and Doyle (1990).

2.2. Cloning and sequencing of viral molecules

Total DNA extracted from infected leaves of *S. palustris* was subjected to rolling circle amplification (RCA) using ϕ29 DNA polymerase (Blanco et al., 1989). The RCA product of *S. palustris* DNA was restricted using different restriction enzymes i.e., *Sal*I, *Bgl*II, *Hin*dIII, *Kpn*I and *Eco*RI. The 2.8kb sized molecules (equal to the size of begomovirus) were generated by restriction digestion with *Eco*RI enzyme. The restricted product (2.8kb) was gel eluted and cloned into the pre-digested pTZ57R vector at *Eco*RI restriction site (Fermentas) and two clones out of 23 clones were completely sequenced to generate clone names pVirol901-2.8 and pVirol902-2.8.

2.3. Sequence analysis and recombination detection

Sequences were assembled and analyzed by the Lasergene DNA analysis package (v8; DNAStar Inc., Madison, WI, USA). Phylogenetic trees were generated, first by aligning the molecules using CLUSTAL-W, followed by Neighbor joint method of phylogenetic tree construction in MEGA7 program (Kumar et al., 2016). Sequences of AlYVV isolated from *S. palustris* are submitted to GenBank (waiting for accession numbers). The other viral sequences were

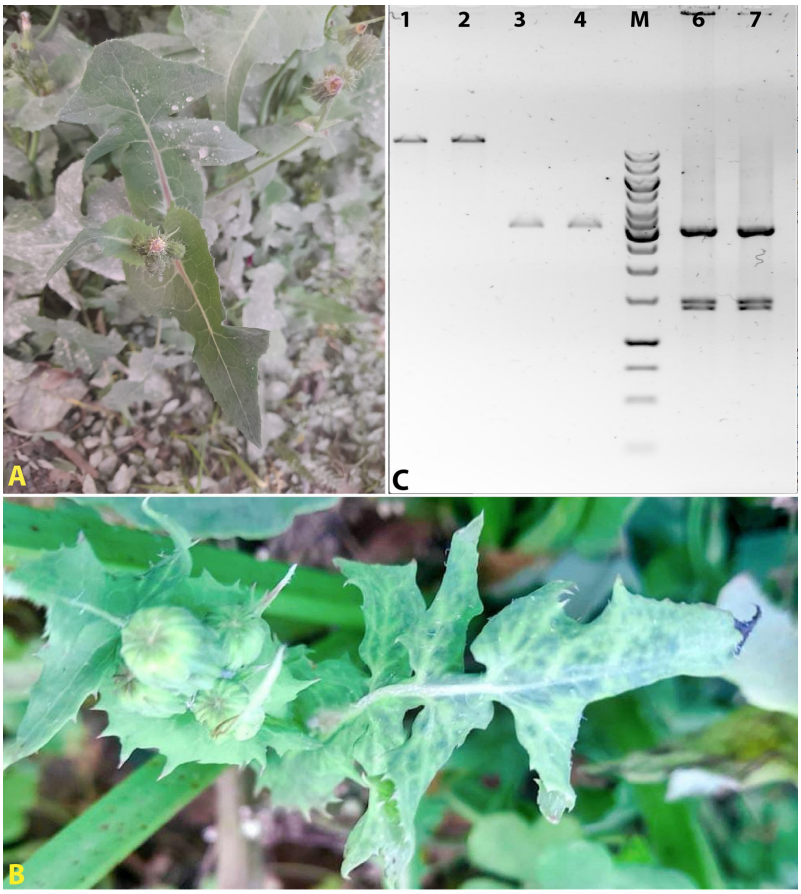

Figure 1. *Sonchus Palustris* plant and cloning of begomovirus: (A) Asymptomatic plant showing vein yellowing symptoms; (B) Symptomatic plant; (C) Amplification using Phi29 DNA polymerase, restriction and cloning of viral molecules; (lane 1-2) Amplified RCA product, (lane 3-4) restriction of RCA product resulted in 2.8kb fragment using *Eco*RI, (lane 6-7) Confirmation of cloning of virus 2.8kb in pTZ57R vector using *Eco*RI and *HindIII,* lane 5 shows 1Kb marker. Size of pTZ57R vector is also 2.88 Kb.

downloaded from GenBank and virus abbreviations are used as described by Zerbini et al. (2017). After the sequence confirmation, detailed recombination analysis was conducted for viral molecules through RDP-4 (Martin et al., 2015). Prior to recombination analysis the sequences were aligned by CLUSTAL-W in MEGA7 DNA analysis software (Kumar et al., 2016) followed by recombination detection through RDP-4 program (Martin et al., 2015).

2.4. SDT and phylogenetic analysis

As a requirement of begomovirus taxonomy (Zerbini et al., 2017), further confirmation of virus was performed using Multiple sequence comparison by log-expectation (MUSCLE) alignment in sequence demarcation tool (SDT) (Muhire et al., 2014). Separate files for coat protein genes and replication-associated genes of all AlYVV isolates in the GenBank and in this study were also generated manually for Bayesian evolutionary analysis by sampling trees (BEAST) analysis (Drummond and Rambaut, 2007). For these two datasets (*Rep* and *CP*) nexus files were generated after aligning sequences in MEGA-7 as described above. The general time reverse (GTR+E) substitution was chosen as the best-fit model and the nucleotide dataset was partitioned into 3 sets (codon positions 1, 2 and 3). Coalescent Bayesian Skyline was chosen in the tree panel as a prior and each dataset was run for a chain length of $4x10_7$ to ensure an adequate sample size in the MCMC panel of the BEAUTi module in BEAST (Drummond and Rambaut, 2007).

3. Results and Discussion

To avoid future epidemics and design control strategies for plant viral diseases, identification of initial viral inoculum is important. Begomoviruses infecting important crop plants, are one of the rapidly emerging groups of plant viruses, which can be attributed to various factors, including increased insect vector populations, presence of alternative hosts and/or rapid manifestation of insect vectors (Mansoor et al., 2006). These begomoviruses probably hibernate on alternate hosts when the main crop is not present in the field. A higher diversity of economically important viruses is found in weeds (Maliano et al., 2021). But unfortunately, these reservoirs are often neglected;

partly because many weed species display non-significant viral symptoms and partly because cost and effort required acquiring sequence data from significant number of weed plants is huge. Pakistan is home for begomoviruses and there is a high disease incidence as well as high diversity of viruses infecting crops and non-crops like *Zinnia elegans*, *Solanum nigrum*, *Ageratum conyzoides* (Haider et al., 2007), *Duranta erecta* (Iram et al., 2005), Chili pepper, Tomato (Hussain et al., 2004; Shih et al., 2003; Mansoor et al., 1997) *Croton bonplandianus* (Amin et al., 2002) Okra, Watermelon and Radish (Mansoor et al., 2001; Mansoor et al., 2000a, b) *Vigna aconitifolia* (Qazi et al., 2006) Mungbean (Hameed and Robinson, 2004), *Eclipta prostrata* (Murtaza et al., 2018) papaya (Nadeem et al., 1997) *Digera arvensis*, *Sonchus arvensis* and *Xanthium strumarium* (Mubin et al., 2009; 2012; Mubin et al., 2009). Old World begomoviruses infecting economically important field crops are believed to be originated from the weeds/non-crop plants usually grow in or around the fields of these crops (Ndunguru et al., 2005; Nawaz-ul-Rehman et al., 2012). Resistance breaking begomoviruses, Cotton Leaf Curl Burewala Virus (CLCuBuV), which is a recombinant molecule of Cotton Leaf Curl Multan Virus (CLCuMuV) and Cotton leaf Curl Khokhran Virus (CLCuKV) (Amrao et al., 2010), was isolated from a weed known as *X. strumarium* L. found inside and around the cotton fields (Mubin et al., 2012).

3.1. Symptomatology and characterization of begomoviruses

Sonchus palustris is a perennial weed and found around water channels and in crop fields. In 2019-2020 during a field survey, symptomatic leaves of *S. palustris,* a common weed, showing vein yellowing were collected from different farmers fields growing cotton and vegetables in Punjab (Figure 1B). Vein yellowing was so conspicuous that the infected plants showed clear distinction from the non-infected ones (Figure 1A). Rolling circle amplification (RCA) (Blanco et al., 1989) was used to amplify all circular ssDNA molecules from isolated DNA of infected plant samples. Restriction with *Eco*RI enzyme yielded 2.8kb fragment i.e., size of begomovirus was cloned in pTZ57R (Figure 1C). The restriction digestion with *Eco*RI and *Hin*dIII enzyme confirmed the 2.8kb sized bands in the gel (Figure 1C). Sequencing of 2.8kb sized bands confirmed the presence of begomovirus. We were unable to amplify DNA B, betasatellites or alphasatellites from all samples with universal primers (Briddon et al., 2002, 2004; Rojas et al., 1993) using the dilution of same RCA dilution as a template in PCR reaction. The absence of betasatellites confirmed our previous results (Murtaza et al., 2018) showing that betasatellite might not be associated with AlYVV in *S. palustris*. For the first time AlYVV was amplified from *Alternanthera* plants in Hainan province of China during 2004 as new species of begomoviruses (Guo and Xhou, 2005). Since then AlYVV has been reported from *E. prostrata* and several other weeds in different Asian countries (Mubin et al., 2010; Murtaza et al., 2018). AlYVV seems to be prevalent in Pakistan, India and China and none of the hosts in these countries have shown to be associated with DNA-B component, suggesting it as a monopartite begomovirus. In China, no satellites have been reported to be associated with AlYVV, from *E. prostrata* or *Alternanthera* plants (Guo and Xhou, 2005), (Huang et al., 2006) (Table 1). Similarly in Vietnam, there are two reports of AlYVV infecting *Z. elegance* and *E. prostrate* (Ha et al., 2008). In case of *Z. elegance* AlYVV is associated with *Alternanthera* yellow vein betasatellite (AlYVB) but infectivity analysis showing trans-replication was not performed while in case of *E. prostrata* no satellite was found (Ha et al., 2008). In Pakistan, AlYVV has been found to be associated with satellites, like *Ageratum* yellow vein betasatellite (AYVB) and potato leaf curl alphasatellite (PotLCuA) in *Sonchus arvensis* (Mubin et al., 2010). Later on, AlYVV was found to be associated with new species of alphasatellite, *Alternanthera* yellow vein alphasatellite (EcYVA), but no betasatellite, in *Eclipta prostrata* (Murtaza et al., 2018). Infectivity analysis showed that when AlYVV was co-inoculated with Cotton leaf curl Multan betasatellite (CLCuMB) and betasatellite was not transreplicated by AlYVV. This gives an indication that for AlYVV to interact with betasatellite there is a need of suitable hosts. Or in other words certain host plants do not support betasatellites replication. AlYVV is among the few begomoviruses that are found in both China and Pakistan, as virus distribution is usually restricted by the Himalayan Mountain range which serves as a natural geographical barrier. In contrast, the viruses found in India, Pakistan and Bangladesh are often related due to their geographical proximity with no natural barriers. This is the first report of a begomovirus found in China, Vietnam, and Pakistan. The trade by land route may have disseminated the virus in China, Pakistan, and India.

3.2. Sequence and phylogenetic analysis of AlYVV

Total of 23 full-length begomoviruses were cloned from 10 different samples of *S. palustris*. Partial sequencing showed that all molecules show maximum sequence homology to AlYVV (data not shown). Two molecules were completely sequenced. Begomoviruses isolated from *S. palustris* showed highest sequence homology with AlYVV though sampling was done from different areas. Total of 25 full-length sequences of ALYVV are present in the GenBank so far and mostly sequences were isolated from weed plants rather crops. The phylogenetic tree was constructed using 25 sequences of AlYVV (Table 1) along with AlYVV-[PK:S Palustris:20]-1 and AlYVV-[PK:S Palustris:20]-2 isolated from *S. palustris* and closely related begomoviruses. Sequence comparisons showed the begomovirus to be closely related to isolates of AlYVV (99.3 to 99.7% nucleotide sequence identity to 25 AlYVV sequences available in the databases) with the highest to isolates from Pakistan i.e., AlYVV-[PK:E prostrata:13]-KX906697, AlYVV-[PK:E prostrata:15]-KX710155] and AlYVV-[PK:E prostrata:11]-KX906694 from Pakistan (Figure 2). This indicates that the virus isolated from *S. palustris* is an isolate of AlYVV for which we propose the isolate descriptor for two AlYVV sequences as AlYVV-[PK:S Palustris:20]-1 and AlYVV-[PK:S Palustris:20]-2. The homology difference was observed as a point mutation throughout the genome as compared to recombination. In the phylogenetic tree (Figure 3) it was evident that

Table 1. Geographical distribution and host range of Alternanthera yellow vein virus (AlYVV) and associated satellites.

Serial No	Virus Accession/Year	Host name	Country	Associated satellites	Reference
1	AlYVV-AM050736/2005	*Alternanthera philoxeroide*	China	No	(Guo and Xhou, 2005)
2	AlYVV- EF544601/2005	*Eclipta prostrata*	China	No	Unpublished
3	AlYVV-DQ375456/2006	*Eclipta prostrata*	China	No	(Guo and Xhou, 2005)
4	AlYVV- AJ965540/2006	*Ludwigia hyssopifolia*	China	No	(Huang et al., 2006)
5	AlYVV- DQ641704/2006	*Eclipta prostrata*	Vietnam	No	(Ha et al., 2008)
6	AlYVV- DQ641703/2006	*Zinnia elegans*	Vietnam	AlYVB	(Ha et al., 2008)
7	AlYVV- EF544604/2006	*Eclipta prostrata*	China	No	Unpublished
8	AlYVV- EF544603/2006	*Eclipta prostrata*	China	No	Unpublished
9	AlYVV-EF544602/2006	*Eclipta prostrata*	China	No	Unpublished
10	AlYVV-EU286798/2007	*Eclipta prostrata*	China	No	Unpublished
11	AlYVV- EU286797/2007	*Eclipta prostrata*	China	No	Unpublished
12	AlYVV-FJ712190/2008	*Alternanthera philoxeroide*	China	No	Unpublished
13	AlYVV- FJ015062/2008	*Eclipta prostrata*	China	No	Unpublished
14	AlYVV- FN432361/2009	*Sonchus arvensis*	Pakistan	AYVB and CLCuMuB	(Mubin et al., 2010)
15	AlYVV- LN795903/2013	*Rumex nepalensis*	India	No	Unpublished
16	AlYVV-KT717678/2013	*Picrorhiza kurrooa*	India	No	Unpublished
17	AlYVV-KX885031/2015	*Eclipta prostrata*	China	No	Unpublished
18	AlYVV-KX710155/2015	*Eclipta prostrata*	Pakistan	No	Unpublished
19	AlYVV-KX906694/2013	*Eclipta prostrata*	Pakistan	TYLCCNA^{Ecl} and ECYVA	(Murtaza et al., 2017)
20	AlYVV-MG686552/2017	*Synedrella sp*	India	No	Unpublished
21	AlYVV-LC316183/2012	*Eclipta prostrata*	India	No	Unpublished
22	AlYVV-LC316182/2012	*Eclipta prostrata*	India	No	Unpublished
23	AlYVV-KX906695/2013	*Eclipta prostrata*	Pakistan	TYLCCNA^{Ecl} and ECYVA	(Murtaza et al., 2018)
24	AlYVV-KX906696/2013	*Eclipta prostrata*	Pakistan	TYLCCNA^{Ecl} and ECYVA	(Murtaza et al., 2018)
25	AlYVV-KX906697/2013	*Eclipta prostrata*	Pakistan	TYLCCNA^{Ecl} and ECYVA	(Murtaza et al., 2018)

sequences of the AlYVV had a high level of sequence similarity with the viruses isolated from Pakistan and clustered with AlYVV-[PK:E prostrata:13]-KX906697, AlYVV-[PK:E prostrata:15]-KX710155] and AlYVV-[PK:E prostrata:11]-KX906694. The virus cloned from *S. palustris* has the typical genome organization of monopartite begomoviruses (or DNA-A component of bipartite begomoviruses) with two ORFs in the virion-sense (encoding the V2 protein and coat protein [CP]) and four in the complementary sense (encoding the replication associated protein [Rep], the transcriptional activator protein [TrAP], the replication enhancer protein [REn] and the C4 protein). A phylogenetic tree, based on an alignment of the complete nucleotide sequence of the begomovirus isolated from *S. palustris* with selected begomovirus genome (or DNA A component) sequences is shown in Figure 3. This shows the sequence isolated from *S. palustris* to segregate with isolates of AlYVV, confirming it as an isolate of this species. Tree showed that this conservation in the sequence of AlYVV was not only consistent with respect to the geographical positions rather it was also consistent over the period of time. The sequence of the virus isolated from Pakistan in 2019-2020 had 98% to 100% similarity level with each other. Similarly, the full length AlYVV isolated from China in 2006 had least difference in the nucleotide sequence and had a higher level of homology with each other and these viruses were seemed to be the strains of each other without having considerable difference over a

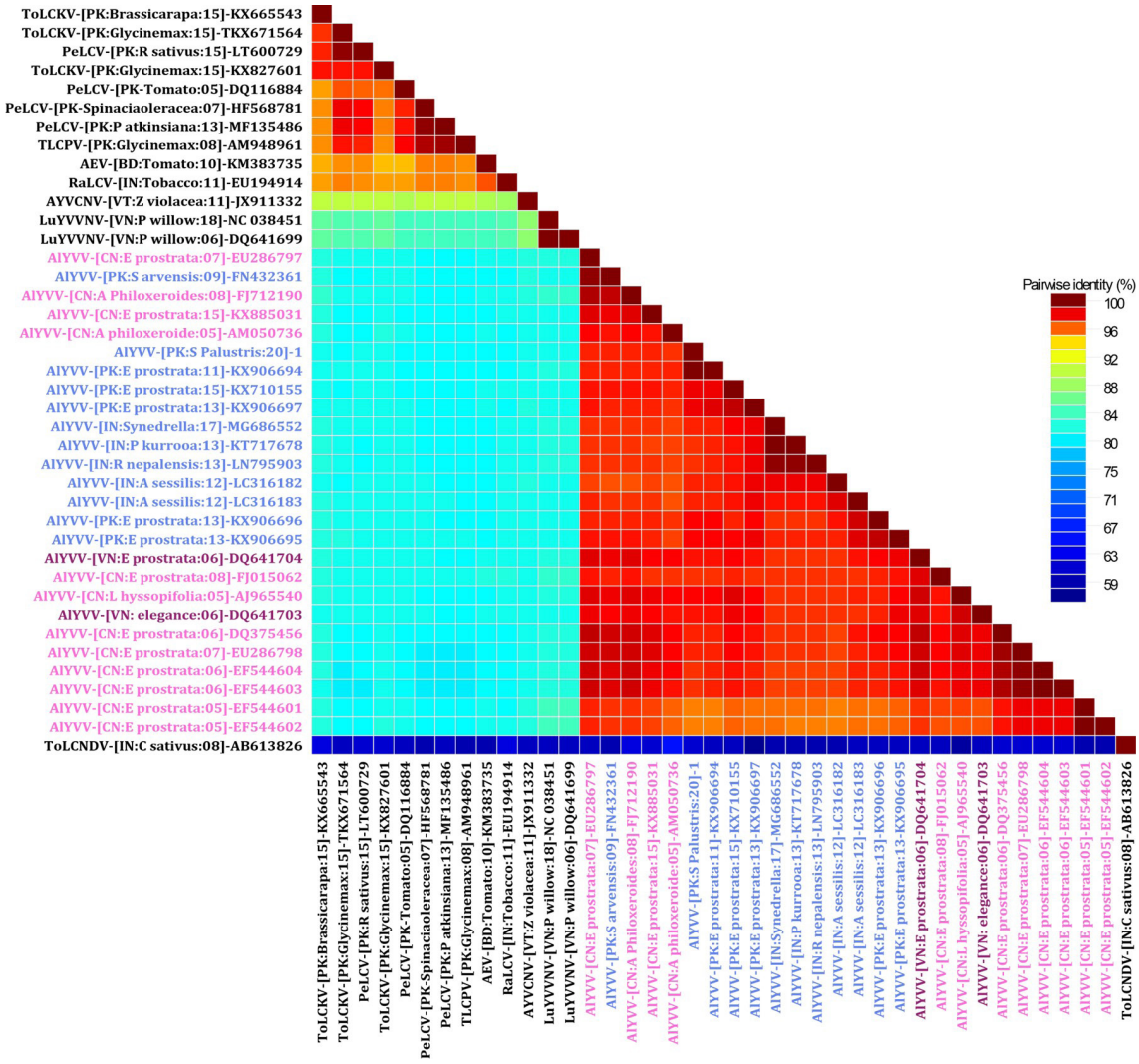

Figure 2. Color-coded matrix of pairwise nucleotide identity inferred from alignment of full-length AlYVV present in data bank. The matrix uses a discontinuous range of three shades of color (red, green and blue) differentiating two cut-off values representing the strain (93-94%, brown red) and the species (90-91%, yellow green) demarcation thresholds of begomoviruses. Identities were calculated with SDT v. 1.2.

period of 8 to 10 years. Higher level of sequence similarity also revealed that apparently no recombination of AlYVV occurred with any other begomovirus. Sequence analysis of AlYVV reported so far from different countries revealed that the full-length viral molecules (~ 2.8kb) were found to be highly conserved and level of sequence homology was 92 to 98%. The homology difference could be observed as a point mutation throughout the genome as compared to recombination.

3.3. Estimation of nucleotide substitution rates

The mean nucleotide substitution rates for the *CP* gene of AlYVV were determined using recombination free datasets with the relaxed clock and Bayesian Skyline Plot (BSP) method. For each dataset, the sequences were partitioned into the 3 codons positions. The mean substitute rates for *CP* gene of ALYVV was considerably higher (4.75 x 10^{-3}) substitutions/nucleotide/year, respectively) than those for the CLCuKoV CP (2.706 x10^{-4}) (Nawaz-ul-Rehman et al., 2012). This high substitution rate in *CP* gene is closer to substitution rate estimated for the bipartite begomovirus, *East African cassava mosaic virus CP* (1.37x10^{-3} subst./nt/year) and *Tomato yellow leaf curl virus CP* (4.6x10-4 subst./nt/year) (Nawaz-ul-Rehman et al., 2012). This may indicate that these viruses face the same evolutionary pressure, despite infecting different hosts in different parts of the world. To further estimate the selection pressure, we used the 3-position clock model in BSP analysis. Codon positions 1, 2 and 3 showed normal behavior for AlYVV-CP (0.965, 0.783 and 1.253).

3.4. SDT and recombination analysis

Sequences were aligned with begomovirus sequences available in databases using MUSCLE and pairwise identity

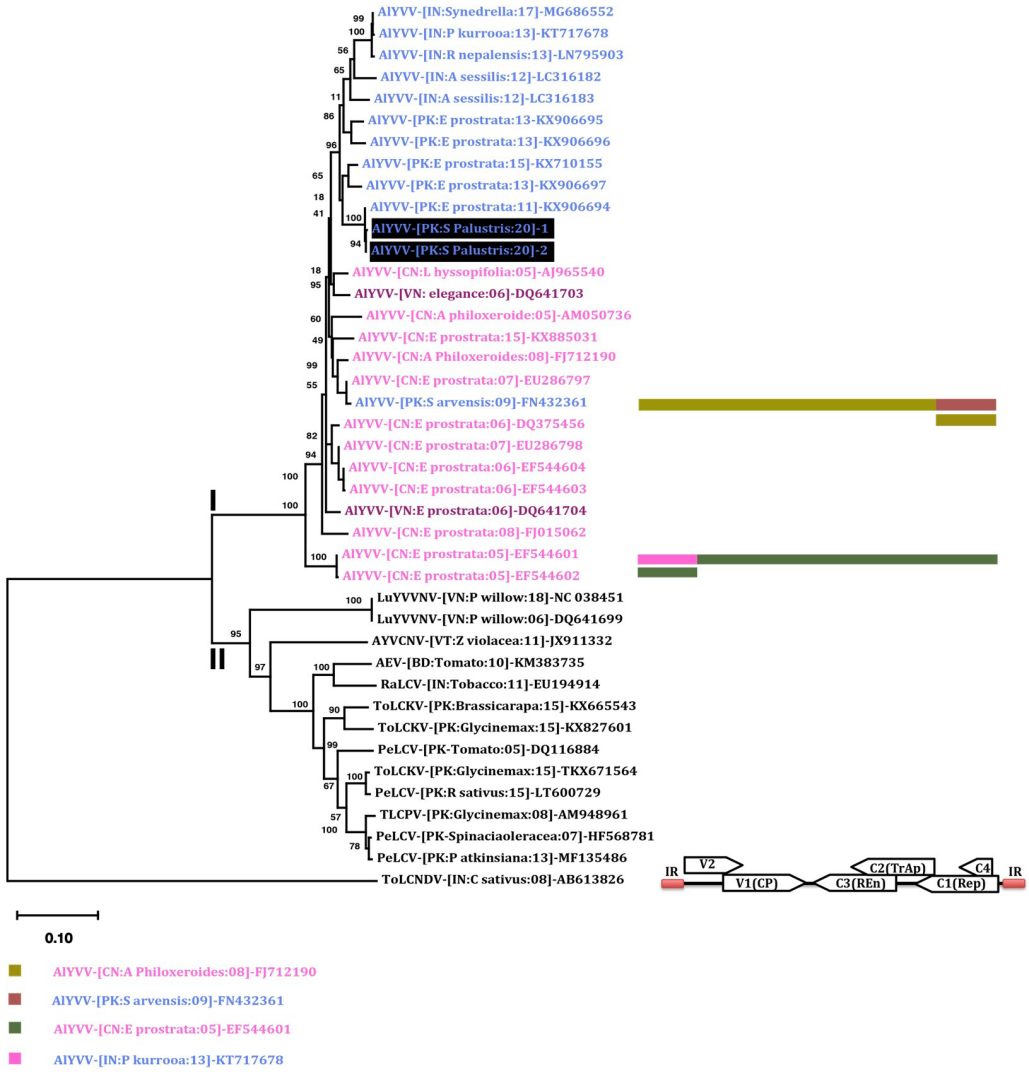

Figure 3. Phylogenetic relationships and intra specific recombination patterns among different AlYVV molecules. The maximum-likelihood phylogenetic tree contains 27 known complete genomes of AlYVV from databank and two complete genomes determined in this study (indicated in black boxes). The tree was rooted on ToLCNDV (AB613826) as an out-group. The schematic representation of recombination events detected by RDP4. Arrows and blocks at the bottom correspond respectively to open reading frames (ORFs) and intergenic regions: pre-coat protein (*AV2*), coat protein (*CP*), replication-associated proteins (*Rep* and *REn*), transcriptional protein (TrAP), and *AC4* region. AlYVV from different countries were colored differently. The colors of blocks represent the different ALYVV species and strains. Numbers at nodes indicate bootstrap confidence scores (1000 replicates).

scores were calculated with SDT (species demarcation tool). Viral sequence isolated in this study i.e., AlYVV-[PK:S Palustris:20]-1 shared 98% and 100% nucleotide identity, with isolates from Pakistan i.e., AlYVV-[CN:E prostrata:07]-EU286797, AlYVV-[PK:S arvensis:09]-FN432361 and AlYVV-[CN:A philoxeroide:05]-AM050736 (Figure 2). Nucleotide sequence homology of two begomoviruses with the most closely related begomoviruses was above the 97%, threshold for species demarcation, thus confirming that the begomoviruses found infecting *S. palustris* in Pakistan are isolates not new species. To find out any recombination, phylogenies of full-length AlYVV (Figure 3) AlYVV-Rep (Fig, 4A) and AlYVV-CP genes (Figure 4B) were studied. Interestingly trees have two major clades. Clade I contains AlYVV and clade II encompasses closely related viruses like LuYVVNV, RaLCV and AEV. This shows no interspecific recombination. But when phylogenetic trees generated for *Rep* and *CP* genes were analyzed along with full-length AlYVV tree, it showed interesting results. Clade I in *CP* tree and AlYVV tree distribution is uniform but in *Rep* tree differs in distribution considerably from *CP* and AlYVV trees. Not surprisingly the phylogenetic tree generated for *Rep* gene sequences shows AlYVV-[CN:A philoxeroide:05]-AM050736 segregates from rest of the ALYVV sequences. While for *CP* tree, AlYVV-[CN:A philoxeroide:05]-AM050736 forms group with AlYVV-[PK:E prostrata:13]-KX906695 and AlYVV-[VN:Hue:Z elegance:06]-DQ641703. Similarly for *CP* tree, AlYVV-[CN:E

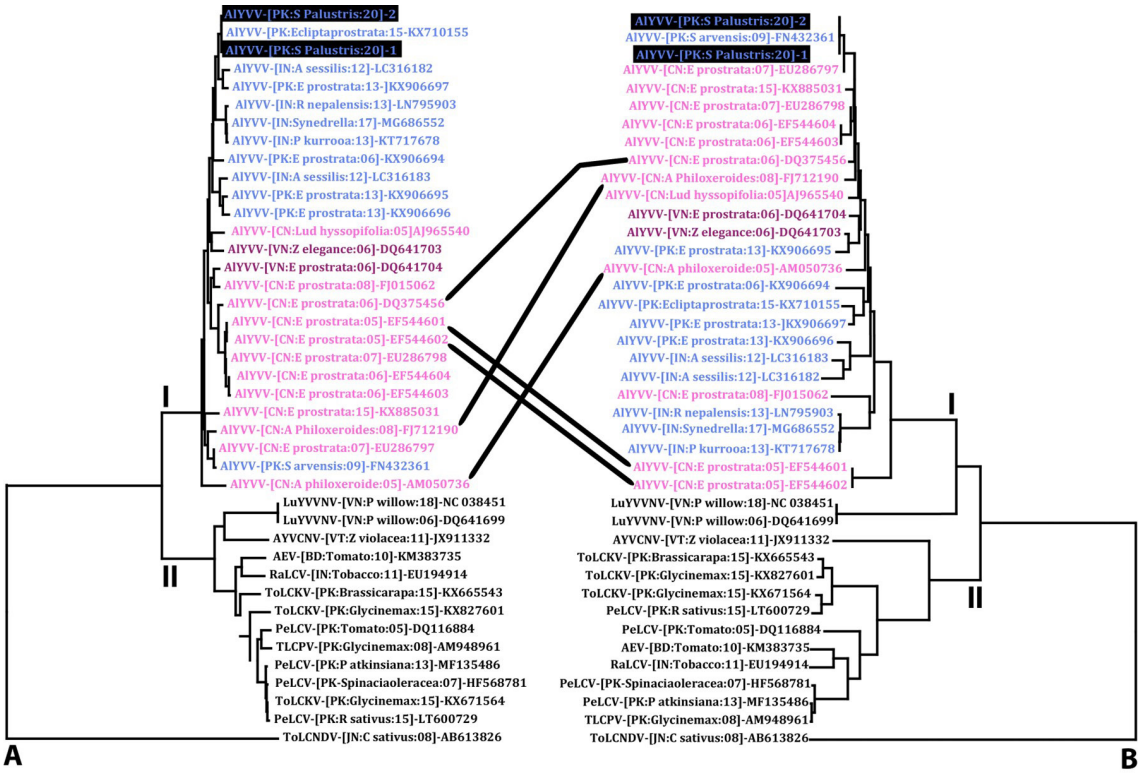

Figure 4. Comparison of the phylogenetic trees of the *Rep*-gene (A) and *CP*-gene (B) of AlYVV for recombination analysis. AlYVV from different countries were colored differently. Lines show assortment of Rep gene and *CP* genes in different clusters showing intraspecific recombination. *Rep* and *CP* gene from AlYVV determined in this study are indicated in black boxes. Clade I contains AlYVV and clade II encompasses closely related viruses like LuYVVNV, RaLCV and AEV. The trees were arbitrarily rooted on the sequences of the *Tomato leaf curl New Delhi virus* (AB613826). Numbers at nodes indicate bootstrap confidence scores (1000 replicates).

prostrata:05]-EF544601 and AlYVV-[CN:E prostrata:05]-EF544602 make a monophyletic group while for *Rep* these two group together with AlYVV-[CN:E prostrata:07]-EU286797 and other AlYVV from China. Indeed, the phylogenetic trees of *CP* and full-length molecules are alike. Where, the isolates of Indian sub-continent and Chinese origin cluster separately, indicating their geographical origin. However, the Phylogenetic tree for *Rep* gene shows segregation of Chinese isolates with the Indian sub-continent. The isolate AlYVV-[PK:S arvensis:09]-FN432361, which segregates with Chinese isolates in the full-length sequences tree is an indication of its recent introduction to the Indian subcontinent. However, its *Rep* gene is positioned along with the AlYVV molecules isolated in this study. The segregation of Chinese isolates with the Indian sub-continents isolates is a clear indication of recombination within the AlYVV isolates. Together these finding suggest that there might be no inter specific recombination in AlYVV but there is a definite intraspecific recombination on virion sense and complementary sense strand in AlYVV. AlYVV-[PK:S arvensis:09]-FN432361 shows recombination at complementary sense strand as *Rep* portion of it shows sequence homology with AlYVV-[CN:A Philoxeroides:08]-FJ712190 with *p-value* of P-value 2.441×10^{-2}. In *CP* tree and in *Rep* tree distribution of AlYVV-[CN:A Philoxeroides:08]-FJ712190 and AlYVV-[PK:S arvensis:09]-FN432361 supports this hypothesis. In full-length AlYVV tree, AlYVV-[CN:E prostrata:05]-EF544601 and AlYVV-[CN:E prostrata:05]-EF544602 cluster with AlYVV-[CN:E prostrata:08]-FJ015062 and AlYVV-[VN:E prostrata:06]-DQ641704 just like in *Rep* tree but in *CP* tree these two cluster with AlYVV-[IN:P kurrooa:13]-KT717678, AlYVV-[IN:Synedrella:17]-MG686552 and AlYVV-[IN:R nepalensis:13]-LN795903. This shows intraspecific recombination at virion sense strand with *p-value* of 1.328×10^{-2}. In contrast recombination analysis for CLCuBuV showed a distinct recombination pattern. The analysis showed CLCuBuV to consist of the virion-sense sequences of CLCuKoV and the complementary-sense sequences of CLCuMuV, so showing interspecific recombination (Nawaz-ul-Rehman et al., 2012). We need more sequencing data and experiments to understand recombination patterns in different begomoviruses.

4. Conclusions

Usually in Old World, begomoviruses are associated with betasatellites but in this study no betasatellite was found. AlYVV was found to be associated with satellites in some hosts but not proven through infectivity analysis. There are a number of reports showing multiple begomovirus

infections in some hosts therefore it might be possible that these reported satellites were associated with some unamplified begomovirus. In present study, AlYVV was found to be recombination free across the infected samples collected from Punjab, which again is an interesting observation. The aim of this study was to understand the begomovirus complex responsible for the disease in *S. palustris* and subsequently phylogenetic and mutational analysis of AlYVV prevailing in the region. Findings of these studies were very interesting and novel as *S. palustris* proved to be a new reservoir for begomoviruses. So, there is always a chance that whitefly feed on different weed hosts and mix up these components. Crop plants showed specificity to specific viruses and components and thus only specific viruses and their components propagated in crop plants. Weeds act as reservoir of these viral components and recombination vessels where all the components are present together with greater chance of recombination. The host range of AlYVV is increasing which can be attributed to various factors, including increased insect vector populations, mutation, inter-specific recombination and/or rapid manifestation of insect vectors. In future such viruses could jump from weeds to main crops and cause economic loss.

Acknowledgements

We acknowledge all members of virology lab for help and support to perform experiments for this work. Higher education commission of Pakistan (HEC) under the grant numbers 1682 and 6427 funds the research in Virology lab. This work was partly funded by Researchers Supporting Project number (RSP-2021/26), King Saud University, Riyadh, Saudi Arabia.

References

AMIN, I., MANSOOR, S., IRAM, S., KHAN, M.A., HUSSAIN, M., ZAFAR, Y., BULL, S.E., BRIDDON, R.W. and MARKHAM, P.G., 2002. Association of a monopartite begomovirus producing subgenomic DNA and a distinct DNA beta on Croton bonplandianus showing yellow vein symptoms in Pakistan. *Plant Disease*, vol. 86, no. 4, pp. 444-444. http://dx.doi.org/10.1094/PDIS.2002.86.4.444B. PMid:30818738.

AMRAO, L., AKHTER, S., TAHIR, M.N., AMIN, I., BRIDDON, R.W. and MANSOOR, S., 2010. Cotton leaf curl disease in Sindh province of Pakistan is associated with recombinant begomovirus components. *Virus Research*, vol. 153, no. 1, pp. 161-165. http://dx.doi.org/10.1016/j.virusres.2010.07.003. PMid:20621137.

BLANCO, L., BERNAD, A., LAZARO, J., MARTIN, G., GARMENDIA, C. and SALAS, M., 1989. Highly efficient DNA synthesis by the phage phi 29 DNA polymerase. Symmetrical mode of DNA replication. *The Journal of Biological Chemistry*, vol. 264, no. 15, pp. 8935-8940. http://dx.doi.org/10.1016/S0021-9258(18)81883-X. PMid:2498321.

BRIDDON, R. and STANLEY, J., 2006. Subviral agents associated with plant single-stranded DNA viruses. *Virology*, vol. 344, no. 1, pp. 198-210. http://dx.doi.org/10.1016/j.virol.2005.09.042. PMid:16364750.

BRIDDON, R., BULL, S.E., MANSOOR, S., AMIN, I. and MARKHAM, P., 2002. Universal primers for the PCR-mediated amplification of DNA beta: a molecule associated with some monopartite begomoviruses. *Molecular Biotechnology*, vol. 20, no. 3, pp. 315-318. http://dx.doi.org/10.1385/MB:20:3:315. PMid:11936260.

BRIDDON, R.W., 2003. Cotton leaf curl disease, a multicomponent begomovirus complex. *Molecular Plant Pathology*, vol. 4, no. 6, pp. 427-434. http://dx.doi.org/10.1046/j.1364-3703.2003.00188.x. PMid:20569402.

BRIDDON, R.W., BULL, S.E., AMIN, I., MANSOOR, S., BEDFORD, I.D., RISHI, N., SIWATCH, S.S., ZAFAR, Y., ABDEL-SALAM, A.M. and MARKHAM, P.G., 2004. Diversity of DNA 1: a satellite-like molecule associated with monopartite begomovirus–DNA β complexes. *Virology*, vol. 324, no. 2, pp. 462-474. http://dx.doi.org/10.1016/j.virol.2004.03.041. PMid:15207631.

CZOSNEK, H., HARITON-SHALEV, A., SOBOL, I., GOROVITS, R. and GHANIM, M., 2017. The incredible journey of begomoviruses in their whitefly vector. *Viruses*, vol. 9, no. 10, pp. 273. http://dx.doi.org/10.3390/v9100273. PMid:28946649.

DOYLE, J.J. and DOYLE, J.L., 1990. Isolation of plant DNA from fresh tissue. *Focus (San Francisco, Calif.)*, vol. 12, no. 1, pp. 13-15.

DRUMMOND, A.J. and RAMBAUT, A., 2007. BEAST: bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, vol. 7, no. 1, pp. 214. http://dx.doi.org/10.1186/1471-2148-7-214. PMid:17996036.

GHULAM, M., MUBIN, M., NAWAZ-UL-REHMAN, M.S. and AMRAO, L., 2017. Alternanthera yellow vein virus (AYVV) infecting Eclipta prostrata plant is associated with two novel alphasatellites in Pakistan. *Pakistan Journal of Life and Social Sciences*, vol. 15, no. 2, pp. 84-89.

GUO, X.A. and XHOU, Z., 2005. Molecular characterization of Alternanthera yellow vein virus: a new Begomovirus species infecting Alternanthera philoxeroides. *Journal of Phytopathology*, vol. 153, no. 3, pp. 694-696. http://dx.doi.org/10.1111/j.1439-0434.2005.01039.x.

HA, C., COOMBS, S., REVILL, P., HARDING, R., VU, M. and DALE, J., 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. *The Journal of General Virology*, vol. 89, no. 1, pp. 312-326. http://dx.doi.org/10.1099/vir.0.83236-0. PMid:18089756.

HAIDER, M.S., TAHIR, M., SAEED, A., SHAH, A., RASHID, N., JAVED, M. and IQBAL, J., 2007. Vinca minor: another host of a tomato infecting begomovirus in Pakistan. In *Proceedings of the African Crop Science Conference Proceedings*, 2007, Kampala, Uganda. Kampala, Uganda: African Crop Science Society, pp. 905-907.

HAMEED, S. and ROBINSON, D.J., 2004. Begomoviruses from mungbeans in Pakistan: epitope profiles, DNA A sequences and phylogenetic relationships. *Archives of Virology*, vol. 149, no. 4, pp. 809-819. http://dx.doi.org/10.1007/s00705-003-0256-9. PMid:15045567.

HUANG, J.F., JIANG, T. and ZHOU, X.P., 2006. Molecular characterization of Begomoviruses infecting Ludwigia hyssopifloia. *Journal of Plant Pathology*, vol. 88, no. 1, pp. 83-88.

HUSSAIN, M., MANSOOR, S., IRAM, S., ZAFAR, Y. and BRIDDON, R.W., 2004. First report of Tomato leaf curl New Delhi virus affecting chilli pepper in Pakistan. *Plant Pathology Journal*, vol. 53, no. 6, pp. 794-794. http://dx.doi.org/10.1111/j.1365-3059.2004.01073.x.

IRAM, S., AMRAO, L., MANSOOR, S., MALIK, K.A., BRIDDON, R.W. and ZAFAR, Y., 2005. First report of a begomovirus associated with leaf curl disease of Duranta erecta in Pakistan. *Plant Pathology*,

vol. 54, no. 2, pp. 260-260. http://dx.doi.org/10.1111/j.1365-3059.2005.01129.x.

KUMAR, S., STECHER, G. and TAMURA, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, vol. 33, no. 7, pp. 1870-1874. http://dx.doi.org/10.1093/molbev/msw054. PMid:27004904.

MALIANO, M.R., MACEDO, M.A., ROJAS, M.R. and GILBERTSON, R.L., 2021. Weed-infecting viruses in a tropical agroecosystem present different threats to crops and evolutionary histories. *PLoS One*, vol. 16, no. 4, pp. e0250066. http://dx.doi.org/10.1371/journal.pone.0250066. PMid:33909644.

MANSOOR, S., KHAN, S., SAEED, M., BASHIR, A., ZAFAR, Y., MALIK, K.A. and MARKHAM, P.G., 1997. Evidence for the association of a bipartite geminivirus with tomato leaf curl disease in Pakistan. *Plant Disease*, vol. 81, no. 8, pp. 958. http://dx.doi.org/10.1094/PDIS.1997.81.8.958C. PMid:30866394.

MANSOOR, S., KHAN, S., HUSSAIN, M., ZAFAR, Y., PINNER, M., BRIDDON, R.W., STANLEY, J. and MARKHAM, P.G., 2000a. Association of a begomovirus and nanovirus-like molecule with Ageratum yellow vein disease in Pakistan. *Plant Disease*, vol. 84, no. 1, pp. 101. http://dx.doi.org/10.1094/PDIS.2000.84.1.101A. PMid:30841203.

MANSOOR, S., MUKHTAR, S., HUSSAIN, M., AMIN, I., ZAFAR, Y., MALIK, K.A. and MARKHAM, P.G., 2000b. Widespread occurrence of Cotton leaf curl virus on radish in Pakistan. *Plant Disease*, vol. 84, no. 7, pp. 809. http://dx.doi.org/10.1094/PDIS.2000.84.7.809B. PMid:30832124.

MANSOOR, S., AMIN, I., HUSSAIN, M., ZAFAR, Y., BULL, S.E., BRIDDON, R.W. and MARKHAM, P.G., 2001. Association of a disease complex involving a begomovirus, DNA 1 and a distinct DNA beta with leaf curl disease of okra in Pakistan. *Plant Disease*, vol. 85, no. 8, pp. 922. http://dx.doi.org/10.1094/PDIS.2001.85.8.922B. PMid:30823073.

MANSOOR, S., ZAFAR, Y. and BRIDDON, R.W., 2006. Geminivirus disease complexes: the threat is spreading. *Trends in Plant Science*, vol. 11, no. 5, pp. 209-212. http://dx.doi.org/10.1016/j.tplants.2006.03.003. PMid:16616578.

MARTIN, D.P., MURRELL, B., GOLDEN, M., KHOOSAL, A. and MUHIRE, B., 2015. RDP4: detection and analysis of recombination patterns in virus genomes. *Virus Evolution*, vol. 1, no. 1, pp. vev003. http://dx.doi.org/10.1093/ve/vev003. PMid:27774277.

MUBIN, M., BRIDDON, R.W. and MANSOOR, S., 2009. Diverse and recombinant DNA betasatellites are associated with a begomovirus disease complex of Digera arvensis, a weed host. *Virus Research*, vol. 142, no. 1-2, pp. 208-212. http://dx.doi.org/10.1016/j.virusres.2009.01.020. PMid:19428755.

MUBIN, M., SHAHID, M., TAHIR, M., BRIDDON, R.W. and MANSOOR, S., 2010. Characterization of begomovirus components from a weed suggests that begomoviruses may associate with multiple distinct DNA satellites. *Virus Genes*, vol. 40, no. 3, pp. 452-457. http://dx.doi.org/10.1007/s11262-010-0470-y. PMid:20306125.

MUBIN, M., AKHTAR, S., AMIN, I., BRIDDON, R.W. and MANSOOR, S., 2012. Xanthium strumarium: a weed host of components of begomovirus–betasatellite complexes affecting crops. *Virus Genes*, vol. 44, no. 1, pp. 112-119. http://dx.doi.org/10.1007/s11262-011-0662-0. PMid:21969121.

MUHIRE, B.M., VARSANI, A. and MARTIN, D.P.2014. SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. *PLoS One*, vol. 9, no. 9, pp. e108277. http://dx.doi.org/10.1371/journal.pone.0108277. PMid:25259891.

MURTAZA, G., MUBIN, M., NAWAZ-UL-REHMAN, M.S. and AMRAO, L., 2018. Genetic analysis of Alternanthera Yellow vein virus (AYVV) infecting Eclipta prostrata plant in Pakistan. *Pakistan Journal of Agricultural Sciences*, vol. 3, no. 51, pp. 505-512. http://dx.doi.org/10.21162/PAKJAS/18.6663.

NADEEM, A., MEHMOOD, T., TAHIR, M., KHALID, S. and XIONG, Z., 1997. First report of papaya leaf curl disease in Pakistan. *Plant Disease*, vol. 81, no. 11, pp. 1333. http://dx.doi.org/10.1094/PDIS.1997.81.11.1333B. PMid:30861752.

NAWAZ-UL-REHMAN, M.S., BRIDDON, R.W. and FAUQUET, C.M.. 2012. A melting pot of Old World begomoviruses and their satellites infecting a collection of Gossypium species in Pakistan. *PLoS One*, vol. 7, no. 8, pp. e40050. http://dx.doi.org/10.1371/journal.pone.0040050. PMid:22899988.

NAWAZ-UL-REHMAN, M.S., NAHID, N., MANSOOR, S., BRIDDON, R.W. and FAUQUET, C.M., 2010. Post-transcriptional gene silencing suppressor activity of two non-pathogenic alphasatellites associated with a begomovirus. *Virology*, vol. 405, no. 2, pp. 300-308. http://dx.doi.org/10.1016/j.virol.2010.06.024. PMid:20598726.

NDUNGURU, J., LEGG, J.P., AVELING, T., THOMPSON, G. and FAUQUET, C.M., 2005. Molecular biodiversity of cassava begomoviruses in Tanzania: evolution of cassava geminiviruses in Africa and evidence for East Africa being a center of diversity of cassava geminiviruses. *Virology Journal*, vol. 2, no. 1, pp. 21. http://dx.doi.org/10.1186/1743-422X-2-21. PMid:15784145.

QAZI, J., MANSOOR, S., AMIN, I., AWAN, M., BRIDDON, R.W. and ZAFAR, Y., 2006. First report of Mungbean yellow mosaic India virus on mothbean in Pakistan. *Plant Pathology*, vol. 55, no. 6, pp. 818-818. http://dx.doi.org/10.1111/j.1365-3059.2006.01475.x.

ROJAS, M.R., GILBERTSON, R.L., RUSSELL, D.R. and MAXWELL, D.P., 1993. Use of degenerate primers in the polymerase chain reaction to detect whitefly-transmitted geminiviruses. *Plant Disease*, vol. 77, no. 4, pp. 340-347. http://dx.doi.org/10.1094/PD-77-0340.

SHIH, S., TSAI, W., GREEN, S., KHALID, S., AHMAD, I., REZAIAN, M. and SMITH, J., 2003. Molecular characterization of tomato and chili leaf curl begomoviruses from Pakistan. *Plant Disease*, vol. 87, no. 2, pp. 200. http://dx.doi.org/10.1094/PDIS.2003.87.2.200A. PMid:30812928.

ZERBINI, F.M., BRIDDON, R.W., IDRIS, A., MARTIN, D.P., MORIONES, E., NAVAS-CASTILLO, J., RIVERA-BUSTAMANTE, R., ROUMAGNAC, P., VARSANI, A. and ICTV REPORT CONSORTIUM, 2017. ICTV virus taxonomy profile: geminiviridae. *The Journal of General Virology*, vol. 98, no. 2, pp. 131-133. http://dx.doi.org/10.1099/jgv.0.000738. PMid:28284245.