3 Biotech (2021) 11:14
https://doi.org/10.1007/s13205-020-02499-5

ORIGINAL ARTICLE

Interaction of eukaryotic proliferating cell nuclear antigen (PCNA) with the replication-associated protein (Rep) of cotton leaf curl Multan virus and pedilanthus leaf curl virus

Sara Shakir[1,2,4] · Georg Jander[2] · Nazia Nahid[3] · Muhammad Mubin[1] · Ayesha Younus[5] · Muhammad Shah Nawaz-ul-Rehman[1]

Received: 7 March 2019 / Accepted: 19 October 2020

Abstract

The replication-associated (Rep) proteins of pathogenic begomoviruses, including cotton leaf curl Multan virus (CLCuMuV) and pedilanthus leaf curl virus (PeLCV), interact with the DNA replication machinery of their eukaryotic hosts. The analysis of Rep protein sequences showed that there is 13–28% sequence variation among CLCuMuV and PeLCV isolates, with phylogenetic clusters that can separated at least in part based on the country of origin of the respective viruses. To identify specific host factors involved in the virus replication cycle, we conducted yeast two-hybrid assays to detect possible interactions between the CLCuMuV and PeLCV Rep proteins and 30 protein components of the *Saccharomyces cerevisiae* DNA replication machinery. This showed that the proliferating cell nuclear antigen (PCNA) protein of *S. cerevisiae* interacts with Rep proteins from both CLCuMuV and PeLCV. We used the yeast PCNA sequence in BLAST comparisons to identify two PCNA orthologs each in *Gossypium hirsutum* (cotton), *Arabidopsis thaliana* (Arabidopsis), and *Nicotiana benthamiana* (tobacco). Sequence comparisons showed 38–40% identity between the yeast and plant PCNA proteins, and > 91% identity among the plant PCNA proteins, which clustered together in one phylogenetic group. The expression of the six plant PCNA proteins in the yeast two-hybrid system confirmed interactions with the CLCuMuV and PeLCV Rep proteins. Our results demonstrate that the interaction of begomovirus Rep proteins with eukaryotic PCNA proteins is strongly conserved, despite significant evolutionary variation in the protein sequences of both of the interacting partners.

Keywords Cotton · Rep · PCNA · Begomovirus

Introduction

Cotton leaf curl Multan virus (CLCuMuV), a *Begomovirus* species in the *Geminiviridae* family, was a major virus associated with the first epidemic of cotton leaf curl disease (CLCuD) in Pakistan (1988–1999) and still is a component of this disease complex (Datta et al. 2017; Zubair et al. 2017). CLCuMuV was also a donor parent for the evolution of a recombinant virus, cotton leaf curl Kokhran virus-Burewala strain (CLCuKoV-Bur) that was primarily involved in the second leaf curl disease epidemic in Pakistan and India (2002–2014) (Amrao et al. 2010a, b; Kumar et al. 2010). On the other hand, pedilanthus leaf curl virus (PeLCV) is an emerging *Begomovirus* that infects several plant species including weeds, vegetables, and ornamental plant species in Pakistan and India (Ilyas 2010; Ismail et al. 2017; Shakir et al. 2018; Tahir et al. 2009; Zaidi et al. 2016).

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s13205-020-02499-5) contains supplementary material, which is available to authorized users.

✉ Muhammad Shah Nawaz-ul-Rehman
msnawazulrehman@uaf.edu.pk

1 Virology Lab, Center for Agricultural Biochemistry and Biotechnology, University of Agriculture, Faisalabad 38000, Pakistan

2 Boyce Thompson Institutute, Ithaca, NY 14853, USA

3 Department of Bioinformatics and Biotechnology, GC University, Faisalabad, Pakistan

4 Present Address: Plant Genetics, Lab, Agro-Bio Tech, University of Liège, Gembloux, Belgium

5 Laser Matter Interaction and Nano-Sciences Lab, Department of Physics, University of Agriculture, Faisalabad, Pakistan

Published online: 02 January 2021

Begomoviruses, the most abundant viruses in the family *Geminiviridae*, have single stranded (ss) DNA genomes and replicate in the nuclei of infected plant cells by recombination-mediated and rolling circle replication (Saunders et al. 1991; Stenger et al. 1991). As these viruses have a small, ~ 2.8 kb genome with limited coding capacity, they do not encode any DNA polymerases and are largely dependent on the host DNA replication machinery to drive their replication. However, begomoviruses express two replication-associated proteins. One of these is the replication-associated (Rep) protein that initiates the replication process by site-specific cleavage at a conserved nonamer (TAATAT TAC) in the viral replication origin, and then rejoins the nascent DNA. The other protein is the replication enhancer protein (REn), which promotes the replication process and viral DNA accumulation (Hanley-Bowdoin et al. 2000; Settlage et al. 2005).

Several studies have demonstrated that *Begomovirus* Rep and REn proteins interact with host proteins. For example, Indian mung bean yellow mosaic virus (IMYMV) Rep and tomato leaf curl virus (TYLCV) Rep and REn proteins interact with plant proliferating cell nuclear antigen (PCNA) (Bagewadi et al. 2004; Castillo et al. 2003). Proliferating cell nuclear antigen is an integral component of DNA replisome, acting as sliding clamp during DNA synthesis (Maga and Hübscher 2003). The mung bean yellow mosaic India virus (MYMIV) Rep protein interacts with RAD51 and RAD54, which are recombinational repair proteins that work in conjunction to each other and other host factors (Petukhova et al. 1999). Tomato golden mosaic virus (TGMV) Rep interacts with retinoblastoma protein (RBR), mitotic kinesin, histone H3, and Ser/Thr kinase of *Nicotiana benthamiana* to alter the cell cycle in nondividing host cells (Kong and Hanley-Bowdoin 2002; Kong et al. 2000). However, begomoviruses that are abundantly found in the Indian subcontinent are unexplored in terms of their interaction with host factors.

The molecular interactions of agriculturally important viruses in Pakistan and India with proteins in their host plants were a missing link in begomoviruses research. To address this issue, we used yeast two-hybrid assays to identify interactions of CLCuMuV and PeLCV Rep proteins with eukaryotic proteins. Our experiments exploited the replication machinery of the eukaryotic model organism, *Saccharomyces cerevisiae* (yeast), which consists of only thirty proteins, thereby avoiding the gene redundancy problems that are inherent in plant systems. Our identification of PCNA as a Rep-interacting protein was generalized in follow-up experiments using PCNA homologs from three plant species.

Materials and methods

Microorganisms, plasmids and general methods

Manipulation of *S. cerevisiae* and *Escherichia coli* nucleic acids was done using standard molecular methods (Sambrook et al. 1989). *Escherichia coli* strain Top10 was used for subcloning. Gateway compatible two-hybrid plasmids pGADT7-GW and pGBKT7-GW (Clontech, https://www.clontech.com) were used to prepare prey and bait constructs, respectively, for yeast two-hybrid assays.

Saccharomyces cerevisiae strain, AH109 (*MATa, trp1-901, leu2-3, 112, ura3-52, his3-200, gal4Δ, gal80Δ, LYS2:: GAL1* $_{UAS}$*—GAL1*$_{TATA}$ *-HIS3, GAL2*$_{UAS}$*- GAL2*$_{TATA}$*-ADE2, URA3:: MEL1*$_{UAS}$*- MEL1*$_{TATA}$*-lacZ*), a derivative of strain PJ69-2A that contains *ADE2* and *HIS3* markers (James et al. 1996), was used in the yeast two-hybrid experiment.

Amplification and cloning of replication-associated genes from *S. cerevisiae*

Thirty genes have been reported to be involved in *S. cerevisiae* DNA replication (Fragkos et al. 2015) (Supplementary Table S1). Based on the available sequences at the *Saccharomyces* genome database (SGD; https://www.yeastgenome.org), specific primers for all thirty genes were designed using the online tool, Primer3 (Untergasser et al. 2012). Primer sequences are listed in Supplementary Table S2. Candidate genes for the prey library were amplified with high fidelity Phusion (*pfu*) polymerase using *S. cerevisiae* DNA as the template and processed for Gateway recombination-based cloning in pGADT7-GW, as described by the manufacturer (New England Biolabs). All of the cloned genes were confirmed by restriction digestion using *Bgl*II and *Xho*I, PCR using T7 promoter forward and gene-specific reverse primers, and DNA sequencing to verify the absence of any mutations. To prepare the bait constructs, coding regions of PeLCV-Rep and CLCuMuV-Rep proteins were amplified from the full-length PeLCV clone (Accession no: MF135486) and the full-length CLCuMuV clone (Accession no: EU365616), respectively, and were cloned in the pGBKT7-GW plasmid vector. As a positive control in the yeast two-hybrid assay, replication enhancer protein (REn) was amplified from the CLCuMuV clone (Accession no: EU365616) and cloned in both pGADT7-GW and pGBKT7-GW plasmids. All of the cloned genes were confirmed by PCR using T7 promoter forward and gene-specific reverse primers, followed by sequencing to verify the absence of any mutations. Correct reading frames and intact amino acid sequences were confirmed using BLAST and alignment with the respective reference genome.

Amplification and cloning of cotton, tobacco, and Arabidopsis PCNA cDNAs

Primers for *Gossypium hirsutum* PCNAs (GhPCNA1 and GhPCNA2) and *Nicotiana benthamiana* PCNAs (NbPCNA1 and NbPCNA2) were designed based on reference genomes available at Cottongen (https://www.cottongen.org) and the Solanaceae Genomics Network (https://www.solgenomics.net), respectively. Total RNA was isolated from *G. hirsutum* and *N. benthamiana* plants grown in a greenhouse at 28–30 °C, using the SV total RNA isolation kit (Promega). Single-stranded cDNA was synthesized using a first strand cDNA-synthesis kit (SMART-MMLV Reverse Transcriptase-Clontech). Amplification of GhPCNA1 and GhPCNA2 was performed with primers CotAD-37582 and CotAD-04662, respectively (Supplementary Table S2). Tobacco PCNA1 and PCNA2 cDNAs were amplified from synthesized cDNA using *pfu* polymerase, using primers Nben-14 and Nben-8, respectively (Supplementary Table S2), cloned in a Gateway-compatible two-hybrid plasmid (pGADT7-GW), and confirmed by PCR, restriction digestion, and Sanger sequencing. Arabidopsis cDNA clones for PCNA1 and PCNA2 (AtPCNAs) were purchased from ABRC/TAIR (https://www.arabidopsis.org) and cloned in pGADT7-GW using the Gateway cloning system (New England Biolabs).

Sequence and phylogenetic analysis of Arabidopsis, cotton, tobacco and yeast PCNAs

The Lasergene (V.8) package (DNASTAR, Madison, Wisconsin) was used to assemble and analyze DNA sequences. The sequences obtained here were compared with the reference genomes of each host, which are available in GenBank. Amino acid sequences of PCNA from *G. hirsutum, A. thaliana*, and *N. tabacum*, as well as other begomoviruses-susceptible plant species whose genomes are available in the GenBank, were aligned using MUSCLE alignment in the Sequence Demarcation Tool (SDT) (Muhire et al. 2014), and identity scores were calculated. Furthermore, PCNA cDNA sequences were aligned using MEGA6 (Tamura et al. 2013) by applying the CLUSTAL-W algorithm (Edgar 2004) and a neighbor joining phylogenetic tree was constructed with 1000 bootstrap replication values.

Diversity and molecular phylogeny of replication-associated protein of CLCuMuV and PeLCV

To investigate the diversity of CLCuMuV and PeLCV Rep, we used the amino acid sequences of Rep protein from two virus isolates used here, along with the selected isolates available in the National Center for Biotechnology Information (NCBI) database. All selected Rep sequences were aligned using the MUSCLE algorithm in SDT (Muhire et al. 2014) and pairwise identity percentages were calculated. The Rep phylogenetic tree was constructed using the neighbor-joining method based on a CLUSTAL-W alignment (Edgar 2004) in MEGA6 (Tamura et al. 2013).

Yeast two-hybrid assay

The Gateway compatible yeast two-hybrid plasmids, pGBKT7-GW and pGADT7-GW (https://www.clontech.com) were used to express the bait and prey proteins, respectively. Replication-associated proteins of two begomoviruses (CLCuMuV and PeLCV) were used as bait in this study. Self-interaction of CLCuMuV REn protein (AC3 × AC3), as reported previously (Settlage et al. 2001), was used as the positive control. Empty plasmids (pGADT7-GW × pGBKT7-GW) were used as the negative control in the initial yeast two-hybrid assay. In the subsequent experiments, yeast DNA Primase (YIR008C, PRI1) protein was used as a negative control against CLCuMuV-Rep protein.

For the yeast two-hybrid screening, bait proteins were first checked for autoactivation by transformation into yeast using PEG/LiAc/ss carrier DNA method (Gietz and Woods 2006). After observing no autoactivation in the yeast two-hybrid assay, we used the bait plasmids for co-transformation with individual prey plasmids from *S. cerevisiae* replication machinery using the same PEG/LiAc/ss carrier DNA method. Transformants were first plated onto low stringency yeast synthetic dropout medium (SD) Leu^-, Trp^- to recover the transformed cells with prey and bait plasmids. Transformants were then allowed to grow at 28 °C for 3 days, diluted in 10 µl of dH_2O, streaked to low stringency SD medium/Leu^-, Trp^- media, and incubated at 28 °C for 3 days. Colonies were then streaked on SD medium/Leu^-, Trp^- media supplemented with X-α-gal and incubated again at 28 °C for 3 days. Colonies were further streaked on high stringency SD medium/Trp^-, Leu^-, His^- supplemented with 2 mM 3-amino-1,2, 4 triazole (3AT) and X-α-gal (Bartel et al. 1993) to prevent the appearance of false positives. Plates were incubated at 28 °C for 3–7 days for yeast growth. Plant orthologs of PCNA were co-transformed with CLCuMuV and PeLCV Rep individually in yeast cells and screened on high stringency SD medium, Trp^-, Leu^-, His^- supplemented with 2 mM of 3AT and X-α-gal, using the method described above.

Results

Diversity and molecular phylogeny of replication-associated proteins (Rep) of begomoviruses

Begomoviruses, which have compromised proofreading activity because of their single-stranded DNA genome organization, exhibit a high mutation frequency and a fast evolutionary rate. Based on the calculations of pairwise percent identity, we found great diversity within and among Rep proteins from different CLCuMuV and PeLCV isolates. CLCuMuV-Rep is quite different from PeLCV-Rep, sharing only 72–87% sequence identity at the protein level. Thus, our phylogenetic analysis separates CLCuMuV-Rep (group a) from PeLCV-Rep (group b) proteins (Fig. 1). The PeLCV Rep protein that was

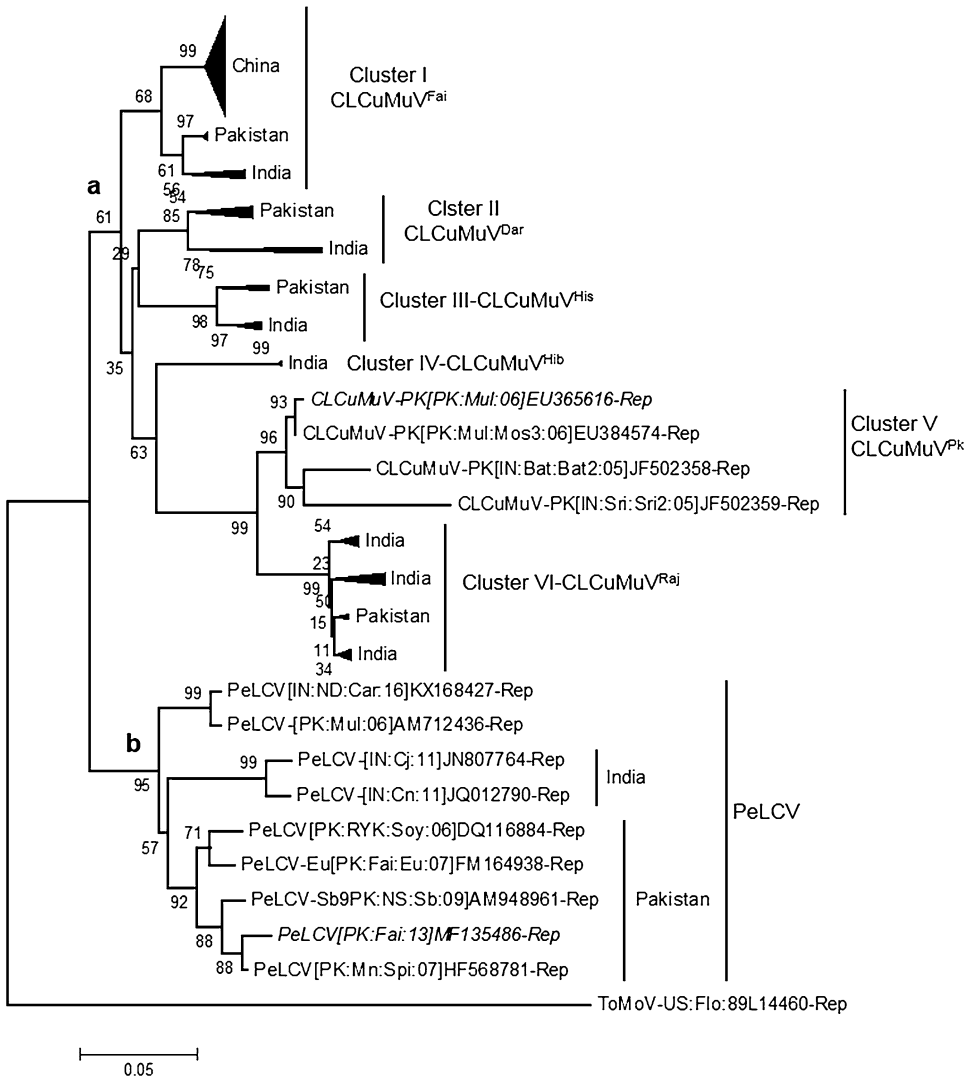

Fig. 1 Phylogenetic analysis for PeLCV-Rep and CLCuMuV-Rep. A neighbor-joining phylogenetic tree of replication-associated (Rep) proteins of cotton leaf curl Multan virus (CLCuMuV) and pedilanthus leaf curl virus (PeLCV) was generated in Mega6 (Tamura et al. 2013) based on a CLUSTAL-W alignment using the CLCuMuV and PeLCV Rep isolates from this study and all related sequences available in GenBank. Isolates of CLCuMuV-Rep (group a) segregate separately into different clade than PeLCV-Rep isolates (group b), indicative of the high level of amino acid sequence variation (13–28%). Moreover, sequence variation also exists within PeLCV-Rep and CLCuMuV-Rep isolates based on the specific strain and country of origin. Rep sequences of PeLCV isolates segregated into two clear clusters based on their origins in Pakistan and India, respectively. Similarly, CLCuMuV segregated into five clusters based on the strain and country of origin. Different clusters have been named according to their strains or country of origin and are compressed for better representation. Rep sequences of PeLCV and CLCuMuV that were used as bait plasmids in this study are marked with italic letters. The Rep protein of tomato mottle virus (ToMoV) was taken as outgroup. A number at each node represent the percentage of 1000 bootstrap replication values

used here (amplified from a full-length clone of accession MF135486), shares 91–99% sequence identity with other PeLCV isolates from Pakistan and India, thus constituting a monophyletic group in the analysis.

The Rep protein sequences of the two viruses can be further divided into subclades that are separated based on the strains and their geographic origins. The PeLCV-Rep group segregated into two clades, with Rep proteins clustering together according to their origins from Pakistan or India, with the exception of two isolates, PeLCV [IN:ND:Car:16] KX168427-Rep and PeLCV [PK:Mul:06] AM712436-Rep, that constituted a third cluster. Similarly, CLCuMuV-Rep segregated into six clusters, mainly according to the strain and country of origin: Cluster I represents the isolates that belong to Faisalabad strain of CLCuMuV reported from China, India and Pakistan; Cluster II represents the isolates of the Darwinii strain of CLCuMuV, which was reported from Pakistan and India; Cluster III consists of Pakistani and Indian isolates of the Hisar strain of CLCuMuV; Cluster IV represents CLCuMuV isolates reported only from India and belongs to the hibiscus viral group; Cluster V represents the Pakistan strain of CLCuMuV isolates reported from either Pakistan or India; and Cluster VI represents the Rajasthan stain of CLCuMuV, isolated from Pakistan and India. All of the Rep protein sequences used to construct the phylogenetic tree have been provided in Supplementary File S1. Rep protein alignments from representative strains of CLCuMuV and PeLCV are presented in Supplementary Fig. S1.

Replication-associated proteins (Rep) of CLCuMuV and PeLCV show significant interaction with *S. cerevisiae* PCNA protein

The DNA replication machinery of *S. cerevisiae*, which consists of only 30 essential proteins (Fragkos et al. 2015) (Supplementary Table S1), was used to identify cellular proteins involved in *Begomovirus* replication. To screen the potential candidates, we cloned the thirty *S. cerevisiae* replication-associated genes into a two-hybrid plasmid, pGADT7-GW, which has a GAL4 activation domain. All cloned constructs were confirmed to have intact open reading frames by DNA sequencing. BLAST comparisons to the *S. cerevisiae* genome database (https://www.yeastgenome.org/) showed 99–100% identity with the reference genome. We selected genes encoding CLCuMuV (Accession MF135486) and PeLCV (Accession EU365616) Rep proteins as bait, fusing them to the GAL4 binding domain in the two-hybrid plasmid pGBKT7-GW. BLAST DNA sequence analysis of the bait constructs, PeLCV-Rep and CLCuMuV-Rep, showed 100% sequence identity with their respective viral clones.

In yeast two-hybrid assays, *S. cerevisiae* replication-associated genes were screened as prey, with CLCuMuV and PeLCV Rep proteins as bait. Positive colonies on low stringency SD medium were further screened for expression of two reporter genes, *His*3 and beta galactosidase on high stringency SD medium (Leu-/Trp-/His-) supplemented with 2 mM 3-Amino-1,2,4-triazole (3-AT) and X-α-gal. 3-AT is a competitive inhibitor of imidazole glycerol-phosphate dehydratase (HIS3), https://www.en.wikipedia.org/wiki/Imidazoleglycerol-phosphate_dehydratase and high affinity binding between two candidate prey and bait proteins allows the expression of *HIS3* reporter gene from the plasmid, resulting in yeast growth on minimal medium containing 2 mM 3-AT. This shows true interactions between candidate proteins in the yeast two-hybrid system.

One out of the sixty prey-bait combinations, only YBR088C, the cell nuclear antigen of *S. cerevisiae* (ScPCNA), allowed growth on medium stringent SD medium (Leu^-/Trp^-), along with the positive control (Fig. 2a–c). Interaction of ScPCNA with the Rep proteins of both CLCuMuV and PeLCV was confirmed by sequencing, re-transformation in yeast, and selection on high stringency SD medium (Fig. 2d–g).

Molecular phylogeny of PCNA from *S. cerevisiae* and selected plant hosts

Based on BLAST searches in GenBank, we identified two PCNA proteins each in *N. benthamiana, G. hirsutum*, and *A. thaliana* genomes. These plant proteins shared 38–40% amino acid sequence identity with the 258-amino acid *S. cerevisiae* PCNA (ScPCNA) (Table 1, Supplementary Fig. S2), thus making the *S. cerevisiae* protein a separate branch radiating from the plant PCNA clade in the phylogenetic tree (Fig. 3). The NbPCNA1 and NbPCNA2 proteins contain 264 amino acids each, the GhPCNA1 and GhPCNA2 proteins are comprised of 266 amino acids, and the AtPCNA1 and AtPCNA2 proteins have 263 and 264 amino acids, respectively. All PCNA orthologs from begomoviruses-susceptible plant hosts studied here, as well as other begomoviruses-susceptible plants whose genomes are available in the GenBank, shared 91–99% sequence identity with one another and thus constitute a monophyletic group (Fig. 3). All sequences that were used to construct the phylogeny are listed in fasta format in Supplementary File S2. The nucleotide sequences of all PCNA clones that were used in this study are available in the NCBI database under the accession numbers MK424839-MK424845.

CLCuMuV and PeLCV Rep proteins interact with different plant PCNAs

As CLCuMuV and PeLCV do not infect *S. cerevisiae*, we determined whether the Rep proteins interact with PCNA from *G. hirsutum*, an agriculturally relevant host of CLCuMuV, as well as *N. benthamiana* and *A. thaliana*, which

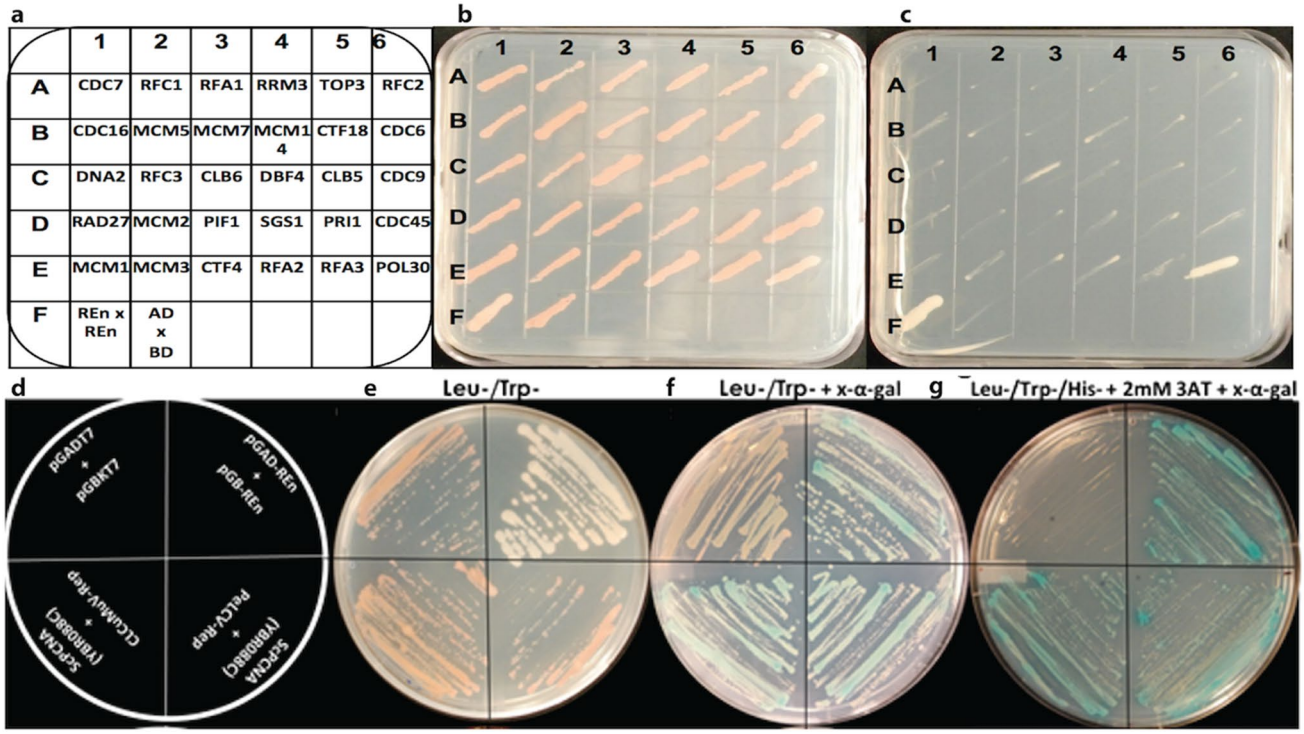

Fig. 2 CLCuMuV and PeLCV Rep interact with yeast YBR088C (PCNA) in a yeast two-hybrid system. **a** Sector diagram illustrating the combinations used against CLCuMuV-Rep in the yeast two-hybrid system. **b** AH109 yeast cells transformed with the combinations described in panel A were recovered on synthetic dropout (SD) Leu^-/Trp^- media. **c** Growth of transformed yeast cells on high stringency SD medium ($Leu^-/Trp^-/His^-$), where interaction of CLCuMuV-AC1, with YBR088/POL30 (E6) and CLCuMuV REn self-interaction as positive control (F1) can be observed in panel B. F2 is the negative control, where empty plasmids pGADT7 and pGBKT7) were co-transformed. Interaction of CLCuMuV-Rep and PeLCV-Rep with yeast PCNA was confirmed by growing yeast cells on low stringency SD Leu^-/Trp^- media (**e**), medium stringency SD $Leu^-/Trp^- + X$-α-gal media (**f**) and high stringency SD $Leu^-/Trp^-His^- + X$-α-gal media supplemented with 2 mM 3-amino-1,2, 4-triazole (3-AT) (**g**) after transformation with the construct combination described in **d**. Growth of yeast in the form of blue colored colonies in the presence of X-α-gal in **f**, **g** indicates protein – protein interactions, whereas yeast growth in the form orange colonies in the presence of X-α-gal (**f**) and no growth on high stringency medium (**g**) represent negative controls with no interaction

are commonly used for begomoviruses studies in laboratory settings. cDNA sequences from *G. hirsutum* (GhPCNA1 and GhPCNA2), *N. benthamiana* (NbPCNA1 and NbPCNA2) and *A. thaliana* (AtPCNA1 and AtPCNA2) were amplified and cloned. DNA sequencing showed 100% identity with reference genomes of the respective species. We performed yeast two-hybrid assays with each bait plasmid, CLCuMuV-Rep and PeLCV-Rep, co-transformed separately with GhPCNA1, GhPCNA2, NbPCNA1, NbPCNA2, AtPCNA1, AtPCNA2. This demonstrated that both viral Rep proteins interact with each of the six plant PCNA proteins in the yeast two-hybrid system (Fig. 4).

Discussion

As the small genomes of begomoviruses do not allow them to encode their own DNA polymerases and other accessary factors for DNA replication, they are largely dependent on the host cellular machinery for their replication. Begomoviruses infect terminally differentiated plant cells (Nagar et al. 1995) that have exited the cell division cycle and are left with little or none of the machinery required for DNA replication. Therefore, the viruses reprogram the plant cell cycle through interaction of the viral Rep protein with plant PCNA, thereby releasing the E2F transcription factor to revert plant cells from resting G0 phase to dividing S phase. As a result, begomoviruses are able to replicate in mature plant cells (Egelkrout et al. 2001, 2002). It has been reported that Rep proteins of tomato yellow leaf curl Sardinia virus (TYLCSV) and Indian mung bean yellow mosaic virus (IMYMV) interact with tomato and pea nuclear PCNA, respectively (Bagewadi et al. 2004; Castillo et al. 2003). Rep also interacts with 32 kDa subunit of plant replication protein A, which binds to ssDNA, and the larger subunit of replication factor C, which loads PCNA onto DNA (Luque et al. 2002; Singh et al. 2007). Geminivirus dsDNA forms 11–12 nucleosome-based minichromosomes; thereby, benefiting

Table 1 Nucleotide sequence identity percentages for pairwise comparisons of PCNA from different species

	PCNA-*S. cerevisiae*	PCNA1-*A. thaliana*	PCNA2-*A thaliana*	PCNA1-*N. benthamiana*	PCNA2-*N. benthamiana*	PCNA1-*G. hirsutum*	PCNA2-*G. hirsutum*	PCNA-*S. lycopersicum*	PCNA-*C. annuum*	PCNA-*N. tomentosiformis*	PCNA-*N. sylvestris*	PCNA-*N. tabacum*	PCNA-*N. attenuata*
PCNA-*S. cerevisiae*	100	40	40	39	39	38	38	39	39	40	40	40	40
PCNA1-*A. thaliana*		100	97	91	91	91	91	91	91	91	91	91	91
PCNA2-*A. thaliana*			100	91	91	91	91	91	92	92	91	92	92
PCNA1-*N. benthamiana*				100	99	95	93	97	97	99	100	100	99
PCNA2-*N. benthamiana*					100	95	93	97	97	99	99	99	99
PCNA1-*G. hirsutum*						100	96	95	95	95	95	95	95
PCNA2-*G. hirsutum*							100	92	92	93	93	93	93
PCNA-*S. lycopersicum*								100	97	98	97	97	97
PCNA-*C. annuum*									100	98	98	98	98
PCNA-*N. tomentosiformis*										100	99	99	99
PCNA-*N. sylvestris*											100	99	99
PCNA-*N. tabacum*												100	99
PCNA-*N. attenuata*													100

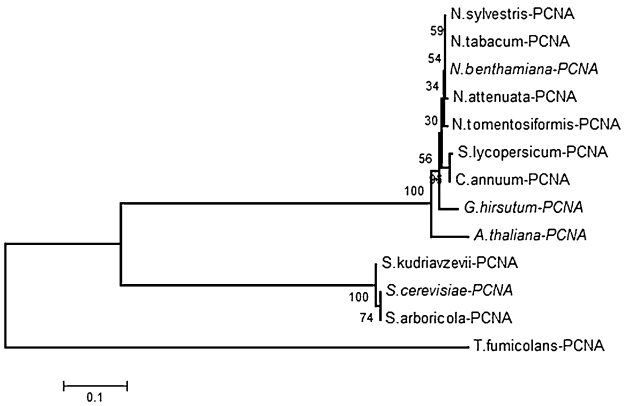

Fig. 3 Molecular phylogeny of proliferating cell nuclear antigen (PCNAs). A neighbor-joining phylogenetic tree of PCNA proteins was constructed in Mega6 (Tamura et al. 2013) based on a CLUSTAL-W alignment of selected PCNA sequences from yeast, begomovirus-susceptible plants used in this study, and related plant species. The PCNA proteins from different plant species have high protein sequence identity (91–99%) and constitute a monophyletic clade. In contrast, due to a lower level of protein sequence identity (38–40%) with plant PCNAs, PCNA from yeast is in a separate branch radiating from plant clade. PCNA sequences that were cloned in prey plasmids for yeast two-hybrid analysis are marked with italic letters. The PCNA protein from an Archaea, *Thermococcus fumicolans* was used as an outgroup to validate the tree. Numbers at each node represent the percentage of 1000 bootstrap replication values

from Rep interaction with the host minichromosome maintenance protein 2 (Suyal et al. 2013). However, the exact mechanism of these interactions is unknown.

Begomoviruses encode two replication-associated proteins, Rep and REn, which are not DNA polymerases, but interact with plant host proteins to drive viral replication (Bagewadi et al. 2004; Settlage et al. 2001). Although both PeLCV and CLCuMuV are abundant begomoviruses in Pakistan, their dependence on the interaction of Rep with host proteins was largely unexplored. Protein sequence and phylogenetic analysis of Rep proteins from CLCuMuV and PeLCV showed great diversity, with 13–28% differences in the amino acid sequences. This level of variation allowed the differentiation of two branches in the phylogenetic tree according to the full-length virus stains and country from which they were isolated (Fig. 1).

We used the yeast two-hybrid system to demonstrate that one of the thirty yeast candidate proteins, PCNA, interacts with the replication-associated protein of CLCuMuV and PeLCV. The high level of functional conservation among PCNA proteins from different species likely explains the yeast PCNA interaction with CLCuMuV and PeLCV Rep protein in the yeast two-hybrid system. Consistent with the observed Rep interactions, our sequence analysis of PCNA from yeast and plants shows that this protein is highly conserved. Previous studies with human PCNA showed that it can complement the function of *Schizosaccharomyces*

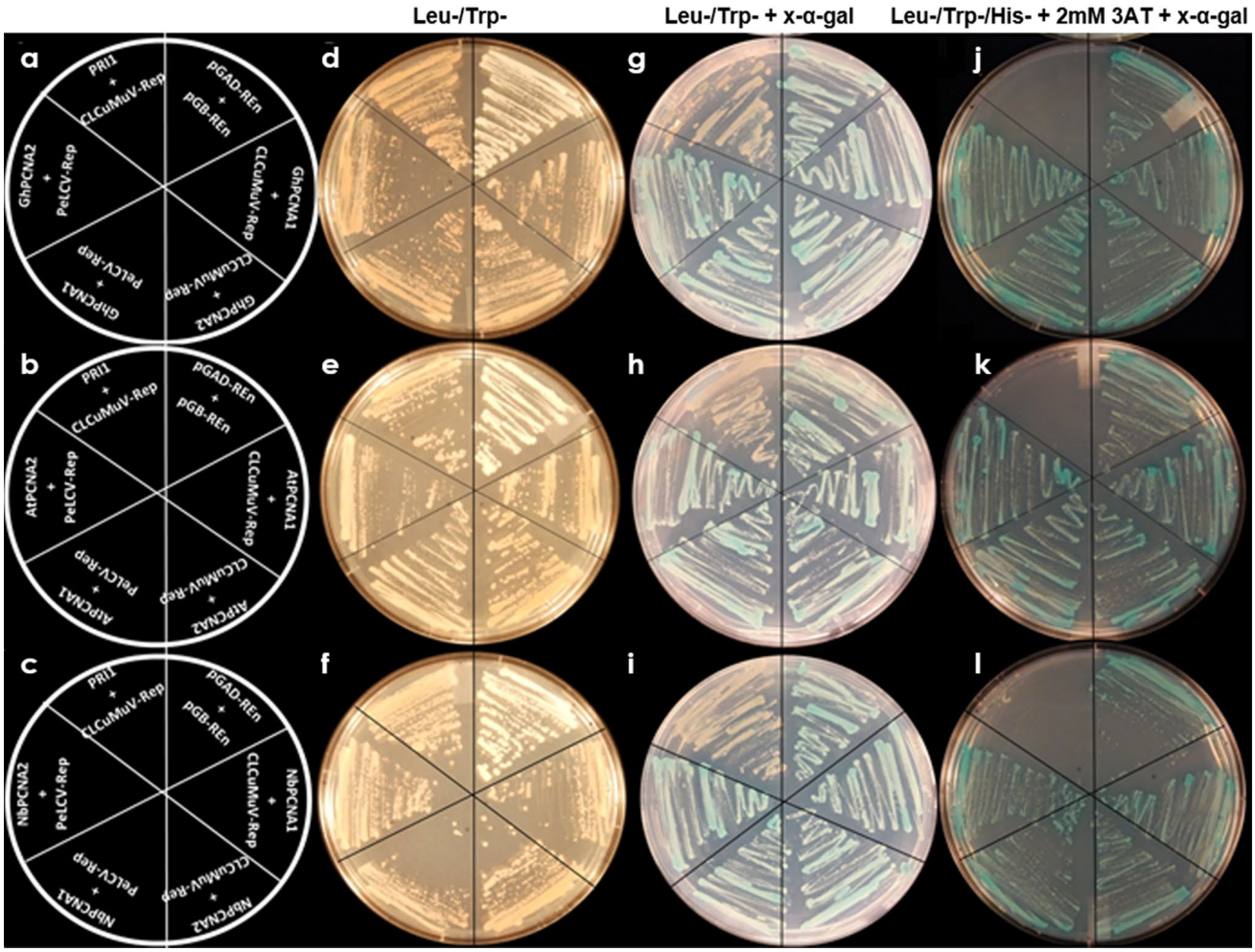

Fig. 4 Interaction of cotton, Arabidopsis, and tobacco PCNA orthologs with CLCuMuV and PeLCV Rep in a yeast two-hybrid system. **a** Illustrates the combinations of constructs transformed in yeast strain AH109C for cotton PCNAs. **b–d** AH109 yeast cells transformed with a combination of constructs (as described in panel **a**) grown on synthetic dropout (SD) Leu^{-}/Trp^{-} medium (low stringency), low stringency SD medium supplemented X-α-gal (medium stringency) and SD $Leu^{-}/Trp^{-}/His^{-}$ medium supplemented with 2 mM 3-amino-1,2, 4-triazole (3-AT) and X-α-gal (high stringency), respectively. **e** describes the combinations of constructs transformed in yeast for Arabidopsis PCNAs. **f–h** Yeast cells transformed with the construct combinations (as described in **e**) grown on low, medium, and high stringency SD medium, respectively. **i** Illustrates the combinations of constructs transformed in yeast for tobacco PCNAs. **j–l** show the interaction verification by growing transformed yeast cells with the construct's combinations (as described in **i**) grown on low, medium, and high stringency medium, respectively. The growth of yeast in the form of blue colonies on medium and high stringency medium indicates protein–protein interactions, whereas orange colonies on medium stringency medium containing X-α-gal (**c**, **g**, **k**) and no growth on high stringency medium (**d**, **h**, **l**) represent negative controls with no interaction

pombe PCNA (Waseem et al. 1992). Another report, which showed that African cassava mosaic virus (ACMV) Rep induced re-replication in *S. pombe* (Kittelmann et al. 2009) by interacting with PCNA, is also consistent with a high level of functional conservation.

In the follow-up experiments, we used plant homologs to generalize the yeast PCNA interaction with virus Rep proteins. For one of the plant proteins, we choose cotton because it is the third most important crop in the world and CLCuMuV is one of the most pathogenic and prevalent begomoviruses in Asia (Zubair et al. 2017; Datta et al. 2017). The other two plants, *A. thaliana* and *N. tabacum*, were chosen because they are important models for begomovirus research in the laboratory. Further studies will be needed to determine the specific protein domains that are responsible for interactions between virus Rep and plant PCNA proteins.

As Rep interaction with PCNA already has been studied in detail for some other begomoviruses (Bagewadi et al. 2004; Castillo et al. 2003; Luque et al. 2002), we can propose the same role for *G. hirsutum* PCNA in CLCuMuV replication during infection of cotton plants. Rolling circle replication is initiated by begomoviruses Rep following its site-specific nicking/closing and ATPase activity, but both

abilities get lost upon Rep interaction with PCNA, thereby limiting the copy number of virus DNA in the infected cell. This loss is overcome by begomoviruses Rep interaction with another host factor, RFC (Luque et al. 2002), that helps to move PCNA away from Rep to the 3′ end of the replication fork. In this way, Rep is released to perform its ATPase and nicking-closing activity at the origin of replication, which is necessary for rolling circle replication. RFC later moves PCNA to the 3′OH end of the nicked DNA at the replication fork, where it acts as a primer for begomoviruses replication and strand synthesis by DNA polymerases.

In conclusion, we explored the use of yeast for investigating begomovirus—host interactions and show that yeast PCNA protein has a conserved structure that allows interaction with begomovirus Rep protein in a yeast two-hybrid system. The virus Rep interaction with yeast PCNA offered a good model system for studying the role of PCNA in virus replication in more detail. Furthermore, this system can be used to study the interaction and functional characterization of other host factors involved in the virus infection process. For instance, cyclin-dependent kinases (CDKs) are important regulators of cell cycle and signal transduction, and a growing body of evidence suggests the interaction of such kinases with multiple geminiviruses proteins (Garnelo Gómez et al. 2019; Mei et al. 2018; Shen et al. 2012). Yeast, due to the simplicity of its genome, offers a good system for elucidating the molecular mechanisms of such host–pathogen interactions.

Our study not only confirmed the previous findings of Rep interaction with Arabidopsis and tobacco PCNA but also presented the role of cotton PCNA in CLCuMuV replication, which opens another field of research for the scientific community. In further experiments, it will be interesting to determine how this Rep-PCNA interaction affects the replication of virus and host chromosomes. One hypothesis is that Rep interaction with PCNA will prevent its interaction with other cellular proteins such as p57, p21 and cyclin D; thereby, regulating the chromosomal replication and cell-cycle progression.

Acknowledgements The authors are thankful to all lab members for their support in conducting research experiments. The research was funded by the International Foundation for Science (IFS grant no, C/5434-1) and the Higher Education Commission (HEC) of Pakistan (project no 1682). Sara Shakir was supported by the HEC indigenous scholarship program and the International Research Support Initiative Program (IRSIP).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest related to this work.

References

Amrao L, Akhter S, Tahir MN, Amin I, Briddon RW, Mansoor S (2010) Cotton leaf curl disease in Sindh province of Pakistan is associated with recombinant begomovirus components. Virus Res 153:161–165. https://doi.org/10.1007/s00705-010-0760-7

Amrao L, Amin I, Shahid MS, Briddon RW, Mansoor S (2010) 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. https://doi.org/10.1016/j.virusres.2010.06.019

Bagewadi B, Chen S, Lal SK, Choudhury NR, Mukherjee SK (2004) PCNA interacts with Indian mung bean yellow mosaic virus Rep and downregulates Rep activity. J Virol 78:11890–11903

Bartel P, Chien C-T, Sternglanz R, Fields S (1993) Elimination of false positives that arise in using the two-hybrid system. Biotechniques 14:920–924

Castillo AG, Colinet D, Deret S, Kashoggi A, Bejarano ER (2003) Dual interaction of plant PCNA with geminivirus replication accessory protein (REn) and viral replication protein (Rep). Virology 312:381–394. https://doi.org/10.1016/S0042-6822(03)0023-4

Datta S et al (2017) Rebound of Cotton leaf curl Multan virus and its exclusive detection in cotton leaf curl disease outbreak, Punjab (India), 2015. Sci Rep 7:17361. https://doi.org/10.1038/s41598-017-17680-9

Edgar R (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinform 5:113

Egelkrout EM, Mariconti L, Settlage SB, Cella R, Robertson D, Hanley-Bowdoin L (2002) Two E2F elements regulate the proliferating cell nuclear antigen promoter differently during leaf development. Plant Cell 14:3225–3236

Egelkrout EM, Robertson D, Hanley-Bowdoin L (2001) Proliferating cell nuclear antigen transcription is repressed through an E2F consensus element and activated by geminivirus infection in mature leaves. Plant Cell 13:1437–1452

Fragkos M, Ganier O, Coulombe P, Méchali M (2015) DNA replication origin activation in space and time. Nat Rev Mol Cell Biol 16:360. https://doi.org/10.1038/nrm4002

Garnelo Gómez B, Zhang D, Rosas-Díaz T, Wei Y, Macho AP, Lozano-Durán R (2019) The C4 protein from tomato yellow leaf curl virus can broadly interact with plant receptor-like kinases. Viruses. https://doi.org/10.3390/v11111009

Gietz RD, Woods RA (2006) Yeast transformation by the LiAc/SS carrier DNA/PEG method. Yeast Protoc 313:107–120

Hanley-Bowdoin L, Settlage SB, Orozco BM, Nagar S, Robertson D (2000) Geminiviruses: models for plant DNA replication, transcription, and cell cycle regulation. Crit Rev Biochem Mol Biol 35:105–140

Ilyas M (2010) Legume-infecting begomoviruses: diversity and host interaction vol Ph.D. Quaid-i-Azam University, Islamabad

Ismail H et al (2017) First report of Pedilanthus leaf curl virus, Tobacco leaf curl betasatellite and Guar leaf curl alphasatellite infecting radish (*Raphanus sativus*) in Pakistan. Plant Dis. https://doi.org/10.1094/PDIS-09-16-1265-PDN

James P, Halladay J, Craig EA (1996) Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. Genetics 144:1425–1436

Kittelmann K, Rau P, Gronenborn B, Jeske H (2009) Plant geminivirus rep protein induces rereplication in fission yeast. J Virol 83:6769–6778. https://doi.org/10.1128/jvi.02491-08

Kong L-J, Hanley-Bowdoin L (2002) A geminivirus replication protein interacts with a protein kinase and a motor protein that display different expression patterns during plant development and infection. Plant Cell 14:1817–1832

Kong LJ et al (2000) A geminivirus replication protein interacts with the retinoblastoma protein through a novel domain to determine symptoms and tissue specificity of infection in plants. EMBO J 19:3485–3495

Kumar A, Kumar J, Khan J (2010) Sequence characterization of cotton leaf curl virus from Rajasthan: phylogenetic relationship with other members of geminiviruses and detection of recombination. Virus Genes 40:282–289

Luque A, Sanz-Burgos AP, Ramirez-Parra E, M.M. C, Gutierrez C, (2002) Interaction of geminivirus Rep protein with replication factor C and its potential role during geminivirus DNA replication. Virology 302:83–94

Maga G, Hübscher U (2003) Proliferating cell nuclear antigen (PCNA): a dancer with many partners. J Cell Sci 116:3051–3060

Mei Y, Yang X, Huang C, Zhang X, Zhou X (2018) Tomato leaf curl Yunnan virus-encoded C4 induces cell division through enhancing stability of Cyclin D 11 via impairing NbSKη -mediated phosphorylation in Nicotiana benthamiana. PLOS Pathog 14:e1006789. https://doi.org/10.1371/journal.ppat.1006789

Muhire BM, Varsani A, Martin DP (2014) SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. PLoS ONE 9:e108277. https://doi.org/10.1371/journal.pone.0108277

Nagar S, Pedersen TJ, Carrick KM, Hanleybowdoin L, Robertson D (1995) A geminivirus induces expression of a host dna-synthesis protein in terminally differentiated plant-cells. Plant Cell 7:705–719

Petukhova G, Van Komen S, Vergano S, Klein H, Sung P (1999) Yeast Rad54 promotes Rad51-dependent homologous DNA pairing via ATP hydrolysis-driven change in DNA double helix conformation. J Biol Chem 274:29453–29462

Sambrook J, Frisch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, New York

Saunders K, Lucy A, Stanley J (1991) DNA forms of the geminivirus african cassava mosaic-virus consistent with a rolling circle mechanism of replication. Nucleic Acids Res 19:2325–2330. https://doi.org/10.1093/nar/19.9.2325

Settlage SB, Miller AB, Gruissem W, Hanley-Bowdoin L (2001) Dual Interaction of a geminivirus replication accessory factor with a viral replication protein and a plant cell cycle regulator. Virology 279:570–576. https://doi.org/10.1006/viro.2000.0719

Settlage SB, See RG, Hanley-Bowdoin L (2005) Geminivirus C3 protein: replication enhancement and protein interactions. J Virol 79:9885–9895

Shakir S, Nawaz-ul-Rehman MS, Mubin M, Ali Z (2018) Characterization, phylogeny and recombination analysis of Pedilanthus leaf curl virus-Petunia isolate and its associated betasatellite. Virol J 15:134. https://doi.org/10.1186/s12985-018-1047-y

Shen Q, Bao M, Zhou X (2012) A plant kinase plays roles in defense response against geminivirus by phosphorylation of a viral pathogenesis protein. Plant Signal Behav 7:888–892. https://doi.org/10.4161/psb.20646

Singh DK, Islam MN, Choudhury NR, Karjee S, Mukherjee SK (2007) The 32 kDa subunit of replication protein A (RPA) participates in the DNA replication of Mung bean yellow mosaic India virus (MYMIV) by interacting with the viral Rep protein. Nucleic Acids Res 35:755–770. https://doi.org/10.1093/nar/gkl1088

Stenger DC, Revington GN, Stevenson MC, Bisaro DM (1991) Replicational release of geminivirus genomes from tandemly repeated copies: evidence for rolling-circle replication of a plant viral DNA. Proc Natl Acad Sci USA 88:8029–8033

Suyal G, Mukherjee SK, Srivastava PS, Choudhury NR (2013) *Arabidopsis thaliana* MCM2 plays role(s) in mungbean yellow mosaic India virus (MYMIV) DNA replication. Arch Virol 158:981–992. https://doi.org/10.1007/s00705-012-1563-9

Tahir M, Haider MS, Iqbal J, Briddon RW (2009) Association of a distinct begomovirus and a betasatellite with leaf curl symptoms in *Pedilanthus tithymaloides*. J Phytopathol 157:188–193. https://doi.org/10.1111/j.1439-0434.2008.01464.x

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729. https://doi.org/10.1093/molbev/mst197

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3—new capabilities and interfaces. Nucleic Acids Res 40:e115–e115

Waseem N, Labib K, Nurse P, Lane D (1992) Isolation and analysis of the fission yeast gene encoding polymerase delta accessory protein PCNA. EMBO J 11:5111

Zaidi SS, Amin I, Iqbal Z, Akhtar KP, Scheffler BE, Mansoor S (2016) *Sesbania bispinosa*, a new host of a begomovirus–betasatellite complex in Pakistan. Can J Plant Pathol 38:107–111. https://doi.org/10.1080/07060661.2015.1128980

Zubair M et al (2017) Multiple begomoviruses found associated with cotton leaf curl disease in Pakistan in early 1990 are back in cultivated cotton. Sci Rep 7:680