

## Article

# Interspecies Recombination-Led Speciation of a Novel Geminivirus in Pakistan

Aamir Lal <sup>1,2,3</sup>, Eui-Joon Kil <sup>2</sup>, Thuy T. B. Vo <sup>1</sup>, I Gusti Ngurah Prabu Wira Sanjaya <sup>1</sup>, Muhammad Amir Qureshi <sup>1</sup>, Bupi Nattanong <sup>1</sup>, Muhammad Ali <sup>4</sup>, Malik Nawaz Shuja <sup>5</sup> and Sukchan Lee <sup>1,\*</sup>

<sup>1</sup> Department of Integrative Biotechnology, Sungkyunkwan University, Suwon 16419, Korea

<sup>2</sup> Department of Plant Medicals, College of Life Sciences, Andong National University, Andong 36729, Korea

<sup>3</sup> Agricultural Science and Technology Research Institute, Andong National University, Andong 36729, Korea

<sup>4</sup> Department of Life Sciences, School of Science, University of Management and Technology (UMT), Lahore 54770, Pakistan

<sup>5</sup> Department of Microbiology, Kohat University of Science and Technology, Kohat 26000, Pakistan

\* Correspondence: cell4u@skku.edu

**Abstract:** Recombination between isolates of different virus species has been known to be one of the sources of speciation. Weeds serve as mixing vessels for begomoviruses, infecting a wide range of economically important plants, thereby facilitating recombination. *Chenopodium album* is an economically important weed spread worldwide. Here, we present the molecular characterization of a novel recombinant begomovirus identified from *C. album* in Lahore, Pakistan. The complete DNA- A genome of the virus associated with the leaf distortion occurred in the infected *C. album* plants was cloned and sequenced. DNA sequence analysis showed that the nucleotide sequence of the virus shared 93% identity with those of the rose leaf curl virus and the duranta leaf curl virus. Interestingly, this newly identified virus is composed of open reading frames (ORFs) from different origins. Phylogenetic networks and complementary recombination detection methods revealed extensive recombination among the sequences. The infectious clone of the newly detected virus was found to be fully infectious in *C. album* and *Nicotiana benthamiana* as the viral DNA was successfully reconstituted from systemically infected tissues of inoculated plants, thus fulfilling Koch's postulates. Our study reveals a new speciation of an emergent ssDNA plant virus associated with *C. album* through recombination and therefore, proposed the tentative name 'Chenopodium leaf distortion virus' (CLDV).

**Keywords:** *Chenopodium album*; recombination; Chenopodium leaf distortion virus; begomovirus



**Citation:** Lal, A.; Kil, E.-J.; Vo, T.T.B.; Wira Sanjaya, I.G.N.P.; Qureshi, M.A.; Nattanong, B.; Ali, M.; Shuja, M.N.; Lee, S. Interspecies Recombination-Led Speciation of a Novel Geminivirus in Pakistan. *Viruses* **2022**, *14*, 2166. <https://doi.org/10.3390/v14102166>

Academic Editor: K. Andrew White

Received: 12 July 2022

Accepted: 27 September 2022

Published: 30 September 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Interspecific interactions play a primary role in the diversification and organization of life [1]. Numerous recurrent formations of allopolyploid species have been reported in the plant and animal kingdoms [2,3]. Genetic recombination is a major source of genetic variability in viruses and creates new opportunities for the viruses to overcome selection pressures [4–7]. The expansion of viral host ranges, alteration of transmission vector specificities, and increases in virulence and pathogenesis are associated with recombination [8–10]. Recombination between isolates of different species as a source of speciation has been reported widely [11]. Among the DNA viruses, the role of recombination in geminiviruses (family: *Geminiviridae*) in the formation of new DNA virus species is well-documented [12–14] and, therefore, plays an essential role in geminivirus diversification and evolution [15–17].

The family *Geminiviridae* includes some of the most damaging plant pathogens that affect a wide host range and cause economic losses throughout the world [18,19]. These plant-infecting viruses have very compact monopartite (DNA-A) or bipartite (DNA-A and DNA-B) genomes [20–22]. Geminiviruses infect both monocots and dicots [23] and are

transmitted by either whiteflies, leafhoppers, treehoppers, or aphids [24–26]. Based on the genome organization, host range, and insect vectors, geminiviruses are classified into fourteen genera: *Becurtovirus*, *Begomovirus*, *Capulavirus*, *Citlodavirus*, *Curtovirus*, *Eragrovirus*, *Grablovirus*, *Maldovirus*, *Mastrevirus*, *Mulcrileovirus*, *Opunovirus*, *Topilevirus*, *Topocuvirus*, *Turncurtovirus*, [13] and a few unclassified genera, such as *Olea europaea geminivirus* [27] and *Fraxinus symptomless virus* [28]. Begomoviruses are the largest plant virus genera within the family *Geminiviridae* having single-stranded circular DNA genome, either monopartite or bipartite components, ranging from 2.5 to 3.0 kb in size and are transmitted by whiteflies, i.e., *Bemisia tabaci* [18,24,29]. DNA-A of the New World and Old World begomoviruses contains five and six protein-encoding open reading frames (ORFs), respectively [22]. The DNA-A virion-sense strand encodes a coat protein (*AVI/VI*) and a movement protein (*AV2/V2*, only in the Old World begomoviruses) [22,30,31], whereas its complementary sense strand encodes a replication initiator protein (*AC1/C1*), a transcriptional activator protein (*AC2/C2*), a replication enhancer protein (*AC3/C3*), and the *AC4/C4* protein (*AC4/C4*) [32–34]. *AC1* (replication-associated protein; Rep) initiates viral DNA replication by binding to iterons within the intergenic region and creates a nick into the conserved TAATATT↓AC sequence [35]. *AC4*, the smallest ORF embedded within the coding region of the Rep protein, is required for monopartite geminivirus infection [36,37]. The DNA-B of bipartite begomoviruses encodes a nuclear shuttle protein (*BV1*) on the virion-sense strand and a movement protein (*BC1*) on the complementary-sense strand [12].

*Chenopodium album* (family: *Amaranthaceae*) is an erect, branched (occasionally unbranched) cosmopolitan weed, which is widely distributed in Canada, Europe, India, Mexico, New Zealand, Pakistan, and South Africa and is ranked among the six most harmful weeds of the world [38]. Except for extreme desert climates, *C. album* is found in all inhabited areas of the world, where it thrives on all soil types and over a wide range of pH [39]. It grows most vigorously in temperate and subtemperate regions; however, it can also potentially affect almost all summer- and winter-sown crops in the tropics and subtropics [40]. *C. album* is responsible for important economic losses in agriculture around the world. In Pakistan, weed species have been identified to be responsible for causing monetary loss worth 3 billion USD annually [41]. Despite some toxic effects, *C. album* is used as a vegetable and a medicinal plant in some regions of the world. A mixed infection of the tomato yellow leaf curl virus (TYLCV) and the tomato yellow leaf curl Sardinia virus (TYLCSV) has been reported to occur in *C. album*. Another begomovirus reported infecting the genus *Chenopodium* is the chenopodium leaf curl virus found in *Chenopodium ambrosioides* [14]. *C. album* is often used as a herbaceous test plant in plant virology.

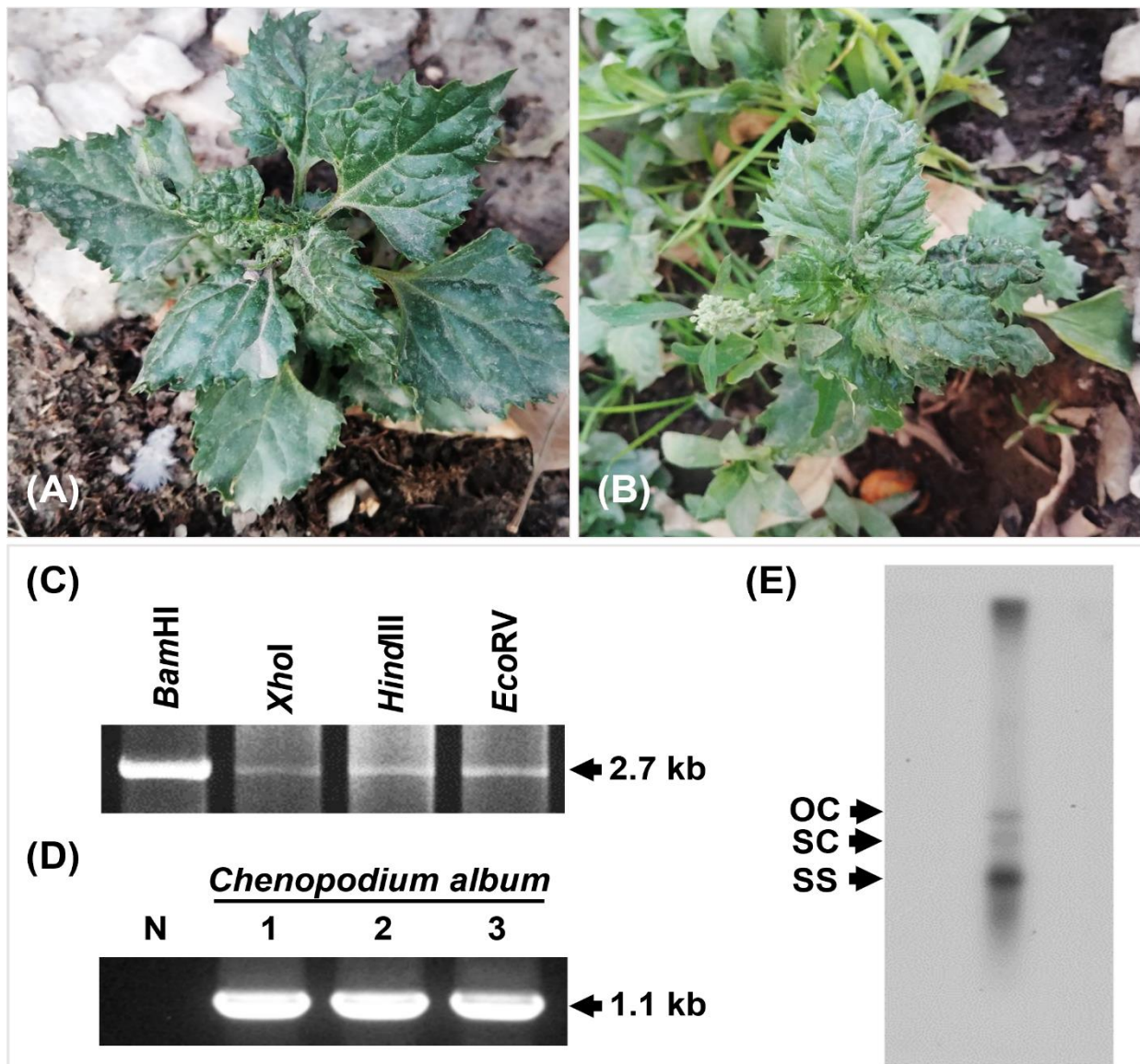
In this study, we characterized a new recombinant begomovirus species in *C. album* that has emerged from the begomovirus interspecies recombination events.

## 2. Materials and Methods

### 2.1. Sample Collection and Virus Detection

In June 2018, during a routine survey to record begomoviruses other than the cotton-infecting viruses in Pakistan, leaf distortion was observed in *C. album* weed grown in a residence area, Lahore, Pakistan (Figure 1A,B). Leaf tissues from three plants were collected and stored at  $-20^{\circ}\text{C}$  until processing. Whiteflies, the insect vectors for begomoviruses, were observed in all symptomatic plants. Total nucleic acid was extracted from the samples using the Viral Gene-spin Viral DNA/RNA Extraction Kit (iNtRON Biotechnology, Inc., Seongnam, Korea) following the manufacturer's instructions. Circular DNA was amplified using the extracted total DNA as a template through rolling circle amplification (RCA) (TempliPhi Amplification Kit; GE Healthcare Life Sciences, Uppsala, Sweden) before being digested with the restriction enzymes *Bam*HI, *Xho*I, *Hind*III, and *Eco*RV (TaKaRa Bio, Inc., Shiga, Japan) [42,43]. All the amplified products digested by all the above-mentioned restriction enzymes were visualized using gel electrophoresis and determined to be approximately 2.7 kb in size. Along with RCA amplification, the presence of begomoviruses was confirmed through PCR amplification of coat protein (CP)

and replication enhancer protein (REn) using begomovirus-specific primers, Beg-F (5'-CCGTGCTGCTGCCCCATTGTCCGCGTCAC-3') and Beg-R (5'-CTGCCACAACCATGATTACACGCACAGGG-3') with target size about 1.1 kb [44]. All these amplicons from both RCA and PCR processing were cloned into the pGEM-3Zf (+) vector (Promega Corporation, WI, USA) and then sequenced by a commercial sequencing service, Macrogen (Seoul, Korea). Sequence contigs were assembled and analyzed using BLASTn and BLASTx [45]. We also attempted to detect satellite molecules, i.e., alphasatellite and betasatellite, and DNA-B through PCR using universal primers [46,47]. Southern hybridization analysis was conducted to confirm the replication of the newly detected virus in the samples using the modified method by Southern et al. [48,49].



**Figure 1.** Amplification of virus DNA from the symptomatic *Chenopodium album* plant and confirmation of viral infection; (A,B) symptomatic samples with leaf distortion symptoms collected and processed. (C) Digestion of RCA product using restriction enzymes *Bam*HI, *Xho*I, *Hind*III, and *Eco*RV. (D) PCR processing using begomovirus-specific primers of *C. album* samples. N: Negative control, Lane 1–3: *C. album* samples processed. (E) Southern blot hybridization of one symptomatic *C. album* sample to confirm virus infection. OC: open circular, SC: supercoiled, SS: single-stranded.



## 2.2. Infectious Clone Construction and Infectivity

An infectious clone (1.1 mer) of the detected virus was constructed to check its infectivity in the host plants (Supplementary Figure S1). Two partial genomes containing restriction sites at the edges, i.e., *SpeI/BamHI* and *BamHI/XbaI* respectively were amplified using primer sets based on the extracted sequence and were ligated into the pGEM-T Easy vector (Promega, Madison, WI, USA), using the TA cloning technique according to the manufacturer's instructions. This was followed by sequencing (Macrogen, Seoul, Korea) and restriction digestion using specific enzymes. The two partial genomes were introduced into the pCAMBIA1303 vector and first transformed into competent *Escherichia coli* strain DH5 $\alpha$  using the heat shock method and then into the GV3101 *Agrobacterium* strains. GV3101 *Agrobacterium* strains (both transformed and untransformed) were cultured in LB broth in the presence of a pCAMBIA1303 selection antibiotic, such as kanamycin (50 mg/L), and strain-specific selection antibiotics, such as gentamycin and rifampicin (50 mg/L), at 28 °C with agitation for 30 h (until the OD value at 600 nm was 0.8–1.0). The untransformed GV3101 *Agrobacterium* with no plasmid was used as the mock. Agro-inoculation was performed by the pin-pricking method [50] in approximately 4- and 6-week-old *N. benthamiana* and *C. album* plants, respectively. Leaf tissue samples were collected from mock and infected plants 28 days post-inoculation (dpi) to check infectivity through PCR processing using the primers: CLDV-IC1/F (5'-ACTAGTTTTGGCAATCGGTGTCTCAC-3') and CLDV-IC1/R (5'-GGATCCACACTCGTTTACATCC-3') specifically designed to amplify IC1 (intergenic region, movement protein and coat protein) of the infectious clone of CLDV with a target size of 1.4 kb. The infection caused by CLDV in *N. benthamiana* and *C. album* plants was further confirmed by Southern blot hybridization.

## 2.3. Target-Specific Primer Construction and PCR Processing

To explore the intriguing genomic composition of the detected virus, i.e., Chenopodium leaf distortion virus (CLDV), primers were designed based on one ORF sequence of the recombinant viruses rather than their identical ORFs in the new virus genomic composition. The ageratum enation virus (AEV) was found to share the highest identity with the Rep and C4 proteins. New primers were designed to target the gene encoding CP of AEV to confirm whether the detected virus was a separate entity. Similarly, primers were designed to target the genes encoding Rep proteins of the rose leaf curl virus (RLCV), duranta leaf curl virus (DLCV), papaya leaf crumple virus (PaLCrV), and catharanthus yellow mosaic virus (CYMV) (Table 1).

**Table 1.** Target-specific primers were constructed to confirm the presence of the viruses in a single host, *C. album*.

Virus	Primer	Sequences 5'-3'	Product Size (bps)	Targeted Area
AEV	AEVCP-F	TGGTCCCCAGACAAACAAC	349	CP
	AEVCP-R	TGGGCTGTCTCGAAGTTGAGAC		
RLCV	RLVREP-F	GTTCCCTAATGACTCTAAGAGC	384	Rep
	RLVREP-R	AGAAGAAGCCCTCTATCAATTAC		
CYMV	CMREP-F	GCTAAAGCTGCGTCAGCAGA	375	Rep
	CMREP-R	AAAGGAGCAAATGCTCGAACTC		
PaLCrV	PaLRep-F	CAGGATGTACAGGATGTATAGGAG	336	Rep
	PaLRep-R	GTGCTGGGCTCATTATCAAACA		
DLCV	DLCREP-F	TAAAGCTGCTTCAGCTGAACC	365	Rep
	DLCREP-R	GAGCAAATGCTCGAACTCCTTA		

## 2.4. Nucleotide Sequences and Phylogenetic Analysis

A total of 56 full-length DNA-A sequences of AEV (30), RLCV (5), CYMV (6), PaLCrV (9), and DLCV (5) were used in this study including the sequence of CLDV (Supplementary Table S1). BLASTn analysis of CLDV showed the sequence identity of these selected viruses though the percentage varies. All sequences were retrieved from the GenBank database

([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov) (accessed on 23 November 2021)) and were aligned at the nicking site in the nonanucleotide motif at the origin of replication (5'-TAATATT//AC-3'). All multiple-sequence alignments were constructed using the MUSCLE method as implemented in MEGA X [51] and manually corrected as well. Phylogenetic tree construction was performed using Mr. Bayes software version 3.2.7a provided by the CIPRES server [52]. In addition, distances were corrected with the best fit model estimated with jModelTest v2.1.6 on XSEDE on the CIPRES Gateway [52–54]. Visualization and editing of phylogenetic trees were carried out using a Newick file generated through FigTree in iTOL [55]. The full-length genome sequences of these top hits aligned with the MUSCLE algorithm were subjected to pairwise comparison using SDT v1.2 [56].

### 2.5. Recombination Analysis

Putative recombinants and major and minor “parents” within the datasets were determined using the RDP, GeneConv, Bootscan, MaxChi, Chimaera, SiScan, and 3Seq recombination detection methods implemented on the RDP4 v4.100 suite [57]. In RDP4, the major parent and minor parent are the presumed parent contributing the larger fraction of the sequence and the presumed parent contributing the smaller fraction of the sequence, respectively. Alignments for all methods were performed using default settings and by *p*-value cutoff of 0.05.

### 2.6. Nucleotide Diversity and Haplotype Variability Indices

The average pairwise number of nucleotide differences per site (nucleotide diversity,  $\pi$ ) was estimated using DnaSP version 6.12.03 (Librado and Rozas 2009, Universitat de Barcelona, Barcelona, Spain). DnaSP version 6.12.03 was also used to calculate the number of haplotypes (H), the number of segregating sites (S), haplotype diversity (Hd), testing Tajima D [58], and Fu and Li’s F [59]. Nucleotide diversity estimates the average pairwise differences among sequences while haplotype diversity refers to the frequency and number of haplotypes in the population. Tajima’s D test is based on the differences between the number of segregating sites and the average number of nucleotide differences whereas, Fu and Li’s F test is based on the differences between the number of singletons and the average number of nucleotide differences between every pair of sequences. The statistically significant differences among the mean nucleotide diversity from all datasets were estimated and represented using GraphPad Prism version 8.0. (Harvey Motulsky 1989, Dotmatics, CA, USA).

### 2.7. Estimation of Gene Genealogies through TCS

The method of Templeton, Crandall, and Sing (TCS) resulting genealogical networks identifies both the relationship between the different sequences as well as the number of nucleotide substitutions connecting them. All the sequences of the AEV, RLCV, CYMV, moieties identified in the novel CLDV recombinant begomovirus along with highly identical DLCV and PaLCrV were analyzed using statistical parsimony with the program TCS (v.1.21) implemented in the software Population Analysis with Reticulate Trees (POPART) [60,61].

## 3. Results

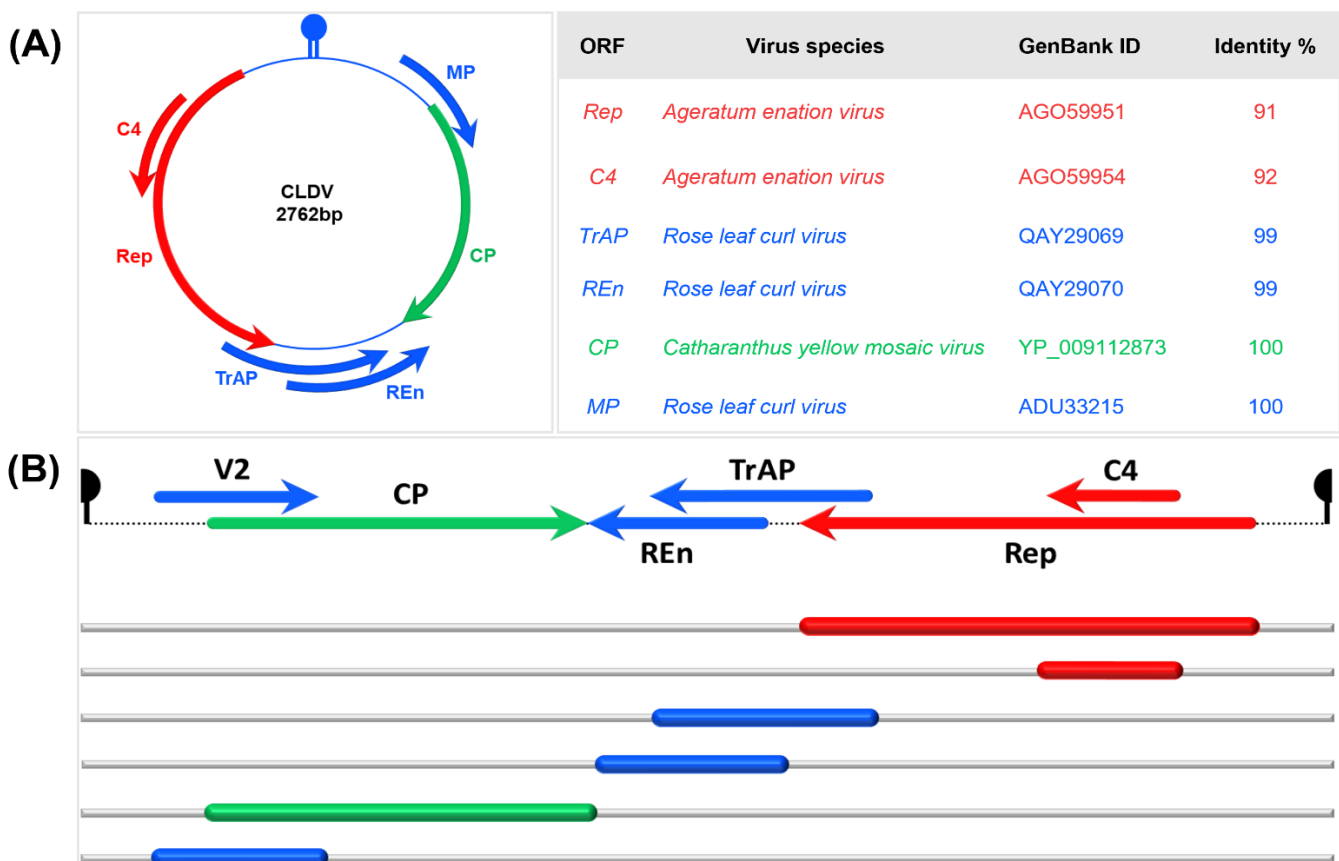
### 3.1. Detection of a New Virus in *C. album*

Sequencing of cloned DNA fragments following the digestion of RCA product with *Bam*HI confirmed the presence of a new begomovirus species, i.e., chenopodium leaf distortion virus (CLDV) in all samples (Figure 1C). The complete nucleotide sequence of DNA- A of CLDV was deposited in GenBank under the accession number MN423112. The expected amplicon size of 1.1 kb targeting CP and Ren regions was observed during PCR processing in all symptomatic samples (Figure 1D), yielding the same sequencing results as MN423112. NCBI basic local alignment search tool (BLASTn) analysis revealed a 93% identity of the nucleotide sequence of the newly identified virus (2.7 kb) with that of the rose leaf curl virus (RLCV; MN746285) and the duranta leaf curl virus (DLCV; MN537564)

respectively. No satellite molecules or DNA-B were detected in association with the DNA-A using universal primers. Southern blot hybridization assay confirms that the viral DNA was integrated into the genome of the *C. album* by producing a noticeable specific band of the expected full-length genome size (Figure 1E).

### 3.2. Genome Organization and Homology Analysis of Genes

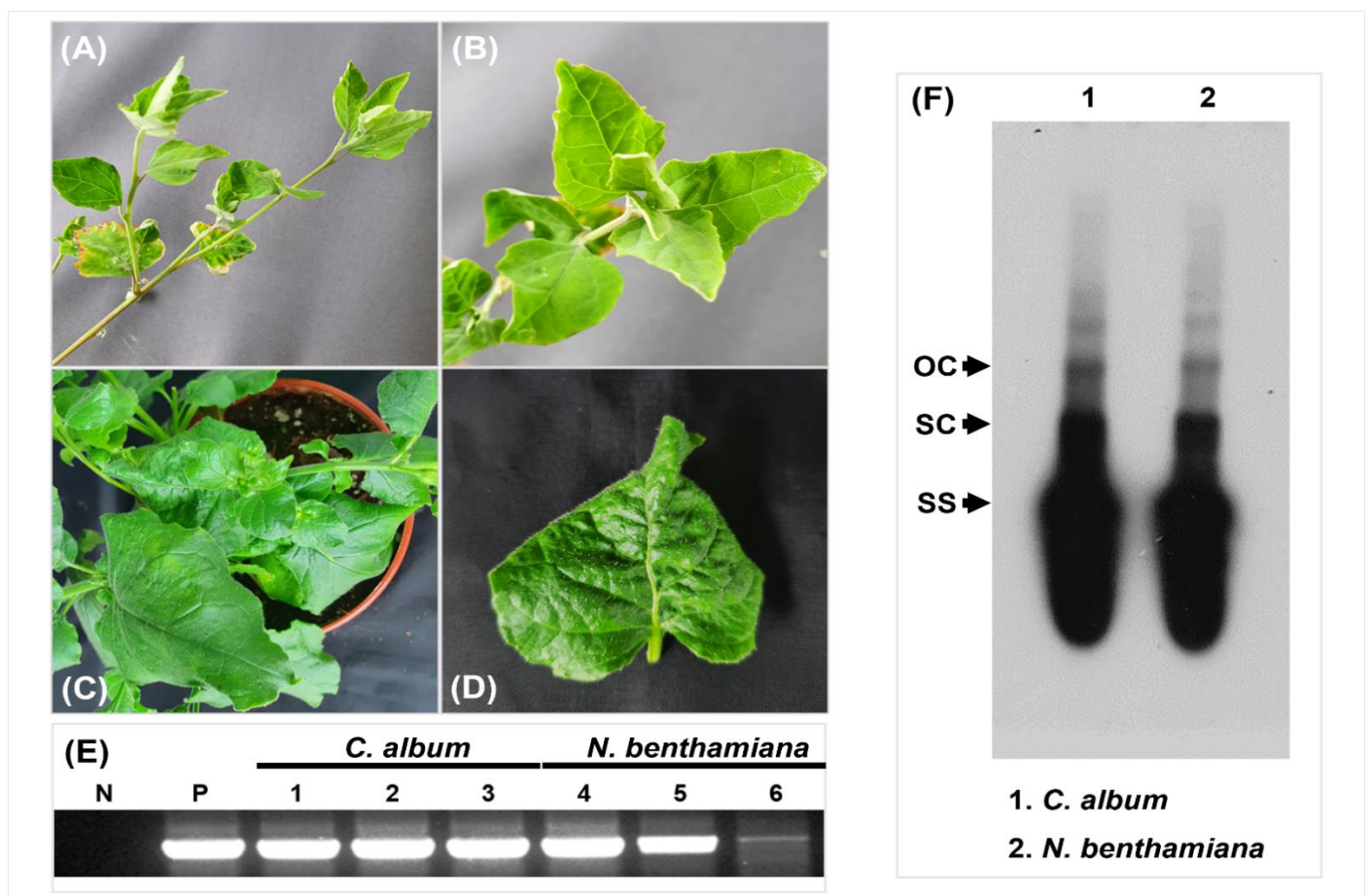
DNA-A of CLDV contained six ORFs following the Old World organization. Sequence analysis (amino acid level) of each ORF using BLASTX showed that AC1, which encodes Rep protein, showed 91% identity with AC1 of AEV (AGO59951) and AC4 showed 92% identity with AC4 of AEV (AGO59954). Similarly, AC2 (TrAP) and AC3 (REn) were 99% identical to the AC2 and AC3 of the RLCV (QAY29069) and (QAY29070), respectively. The ORF AVI (CP) was 100% similar to AV1 of CYMV (YP\_009112873), and AV2 showed 100% identity with AV2 of RLCV (ADU33215) (Figure 2A,B). (Detailed comparison of the sequence homology of each ORF of CLDV with the ORFs of other viruses is provided in Supplementary Figure S2 and Supplementary Table S2). In the overlapping regions: CP/MP ( $\Delta 1$ ) and Rep/TrAP ( $\Delta 2$ ), RLCV was found more identical to CLDV than CYMV and AEV respectively (Supplementary Figure S3). The similarities and differences in the DNA-A of the newly identified virus with the reference sequences are highly noticeable and make a strong case for proving the identity of CLDV as a separate new species.



**Figure 2.** Genome organization of CLDV DNA-A. (A) genomic circular map of DNA-A of Chenopodium leaf distortion virus (CLDV) along with open reading frame (ORF) annotation. Each ORF is represented by a specific color-coated line according to its location in the genome. Along with that, the identical ORFs with their GenBank accession numbers and the percentage of identity at the amino acids level have been noted. The genome composition of CLDV depicts ORFs/IR originating from three different viruses, i.e., AEV (Rep, C4), RLCV (TrAP, REn, MP, and IR), and CP (CYMV). (B) Linear map of CLDV genome with ORFs representation.

### 3.3. Infectivity through Infectious Clone Inoculation

Mild symptoms were observed in both CLDV infected plant groups, i.e., *N. benthamiana* and *C. album*. Leaf tissues were harvested and analyzed by PCR to investigate viral replication ability; the virus was detected in the samples, which confirmed its presence (Figure 3A–E). CLDV (1.4 kb amplicon) was successfully detected through PCR in all three *C. album* samples and three *N. benthamiana* samples. Sample no. 3 of *N. benthamiana* (Lane 6 in Figure 2E) shows a faint band which might be due to poor inoculation or any other experimental error resulting in less virus titer but still, CLDV was reconstituted in this sample when it was sequenced. The virus reconstituted in *C. album* and *N. benthamiana* maintained the exact nucleotide sequence of the original clone. PCR using vector-specific primers showed negative results, which backs the virus detection on its own instead of containing the virus in different parts of the plant. Southern blot hybridization data also confirms the viral infection (Figure 3F).



**Figure 3.** Infectivity of CLDV infectious clones in *C. album* and *N. benthamiana* samples. The results of inoculation assays of CLDV clones observed in (A,B) *C. album* and (C,D) *N. benthamiana* 28 dpi. Mild symptoms were induced by CLDV in both *C. album* and *N. benthamiana* samples. (E) Successful detection of CLDV (1.4 kb amplicon) was done through PCR in all three *C. album* samples and three *N. benthamiana* samples. The mock plant was used as the negative control whereas DNA extracted from one of the infected plants from Pakistan was used as the template for positive control in PCR processing. N: negative control, P: positive control, Lane 1–3: *C. album* samples, Lane 4–6: *N. benthamiana* samples processed. (F) Southern blot hybridization data for infection confirmation purpose, Lane 1: *C. album* sample, Lane 2: *N. benthamiana* sample. OC: open circular, SC: supercoiled, SS: single-stranded.



### 3.4. Target-Specific Primer Construction and PCR Analysis

To explore the intriguing genomic composition of the detected virus, primers were constructed based on one of the ORF sequences of the recombinant viruses rather than their identical ORFs in the new virus genomic composition (Table 1). In all cases, the results remained negative showing the presence of only CLDV in the host sample (Supplementary Figure S4).

### 3.5. Phylogenetic and Recombination Analysis

To assess the standing evolutionary relatedness among these populations, we performed molecular phylogenetic analysis of CLDV and closely associated viruses, using full-genome sequences (sequences included in the analysis were added in Supplementary Table S1). CLDV was found to share a clade with RLCV (Figure 4A). Analysis through Sequence demarcation tool version 1.2 (SDT v1.2.) showed the sequence comparison among the viruses with the revelation of CLDV maximum identity of 93% with RLCV and DLCV isolates (Figure 4B). Recombination analysis detected recombination events in the main genome of CLDV. RDP analysis revealed that CLDV is probably a recombinant genome resulting from a recombination event and originated through recombination between the isolates of the CYMV (GenBank MH643737; 86% similar), AEV (KC795968; 88% similar), and the RLCV (GenBank GQ478342; 93% similar). The recombinant nucleotide coordinates are 2280–1059, spanning the AC1, AC4, IR, V2, and CP genes. The recombination event was validated by the lower  $p$ -value of  $2.98 \times 10^{-15}$ , maximum recombination methods, i.e., RGMCS3, and an acceptable R score of 0.47 (Figure 5 and Supplementary Table S3).

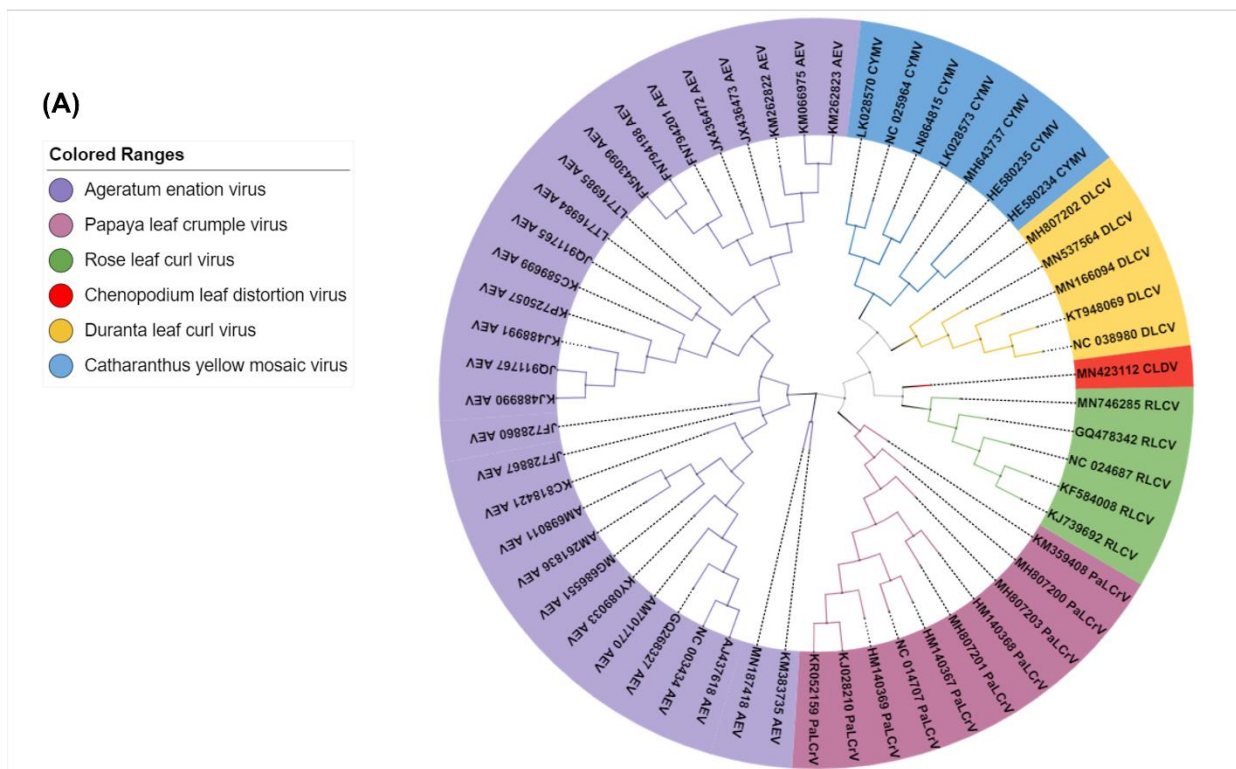
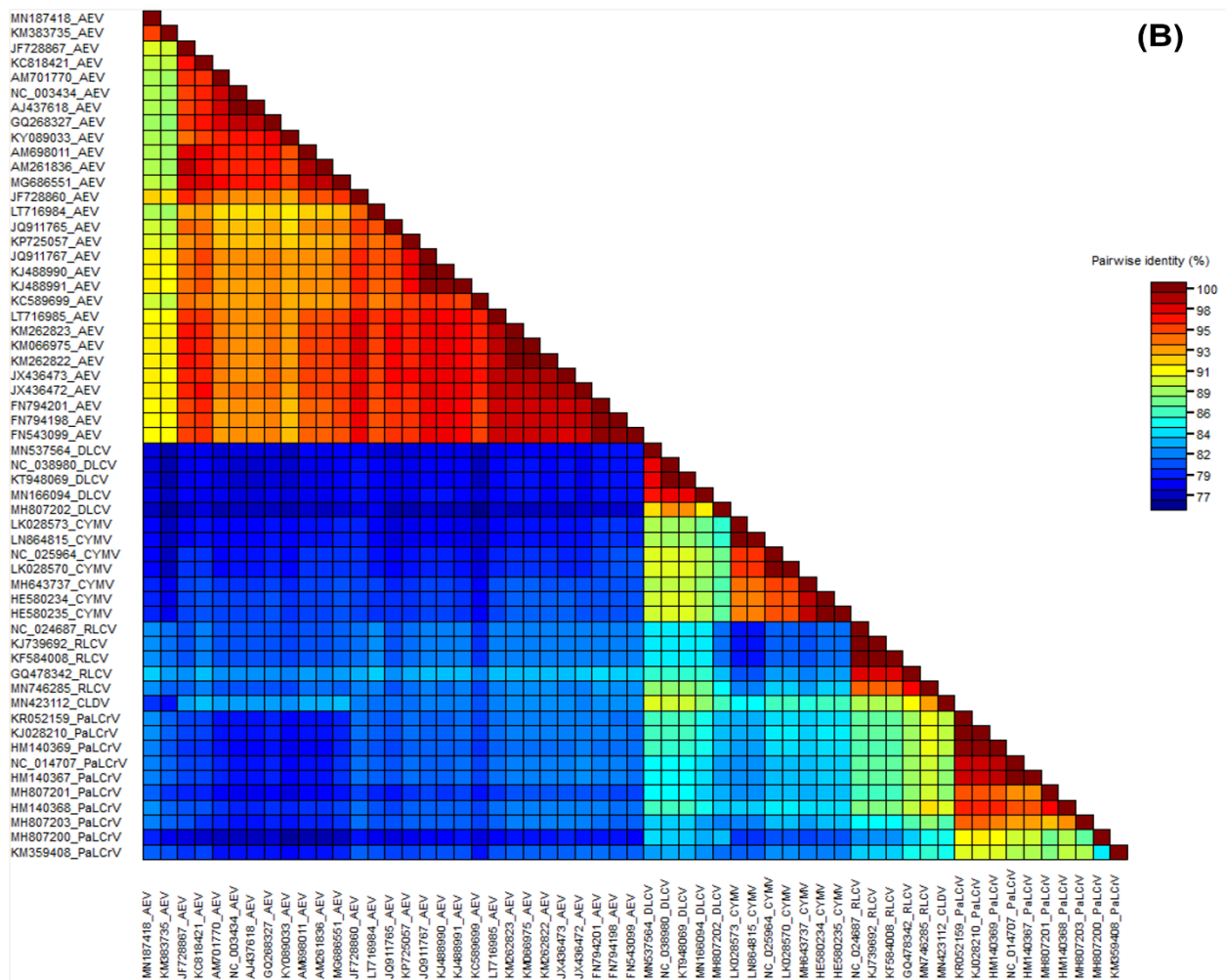


Figure 4. Cont.





**Figure 4.** (A) Phylogenetic tree generated based on an alignment of the sequences of genomic components (DNA-A) of all viruses with a high percentage identity to CLDV. (B) Pairwise complete nucleotide sequence comparison of begomoviruses. Color-coded pairwise identity matrix was generated using the sets of begomovirus sequences by using the SDT v1.2.

### 3.6. Nucleotide Diversity and Haplotype Variability Indices

We analyzed all datasets comprising AEV (30 sequences), CYMV (6 sequences), DLCV (5 sequences), PaLCrV (9 sequences) and RLCV (5 sequences) to compare the standing molecular diversity. Despite the difference in sample numbers, we were able to calculate average pairwise nucleotide diversities ( $\pi$ ) for the aforementioned datasets. The average pairwise nucleotide differences were higher for PaLCrV ( $\pi = 0.0651$ ), followed by AEV ( $\pi = 0.0494$ ), DLCV ( $\pi = 0.0442$ ), CYMV ( $\pi = 0.0424$ ), and RLCV ( $\pi = 0.0257$ ), respectively (Figure 6a). The average number of segregating sites ( $\theta_w$ ) was remarkably higher in the case of PaLCrV ( $\theta_w = 0.0819$ ), while these values were lower for AEV ( $\theta_w = 0.0635$ ), DLCV ( $\theta_w = 0.0480$ ), CYMV ( $\theta_w = 0.0344$ ), and RLCV ( $\theta_w = 0.0255$ ) respectively (Figure 6b). Tajima's D values for PaLCrV ( $-1.4063$ ), AEV ( $-1.1259$ ), DLCV ( $-0.8438$ ), and RLCV ( $-0.0109$ ) were found negative whereas CYMV (1.1617) was interestingly found positive (Figure 6c). Likewise, we also found Fu and Li's F values for each virus: PaLCrV ( $-1.5289$ ), AEV ( $-1.3594$ ), DLCV ( $-0.8903$ ), and RLCV ( $-0.0024$ ) were found negative whereas CYMV (1.2396) was found positive (Figure 6d and Supplementary Table S4).

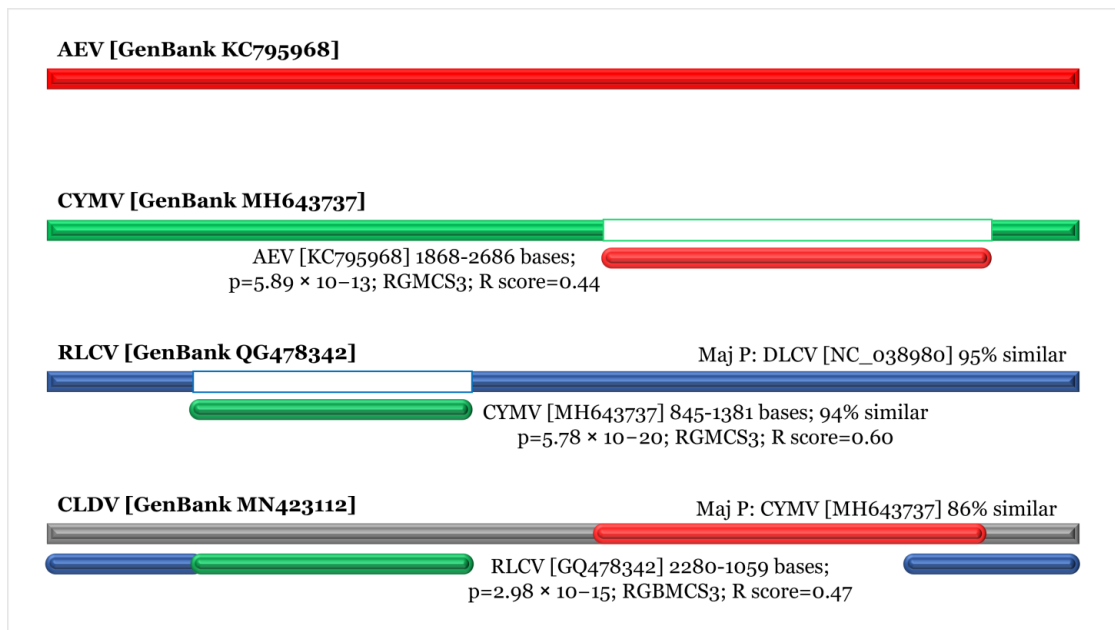


Figure 5. Recombination analysis of recombinant CLDV identified in this study. Recombination events were identified and found between the isolates of the CYMV, AEV, and RLCV.

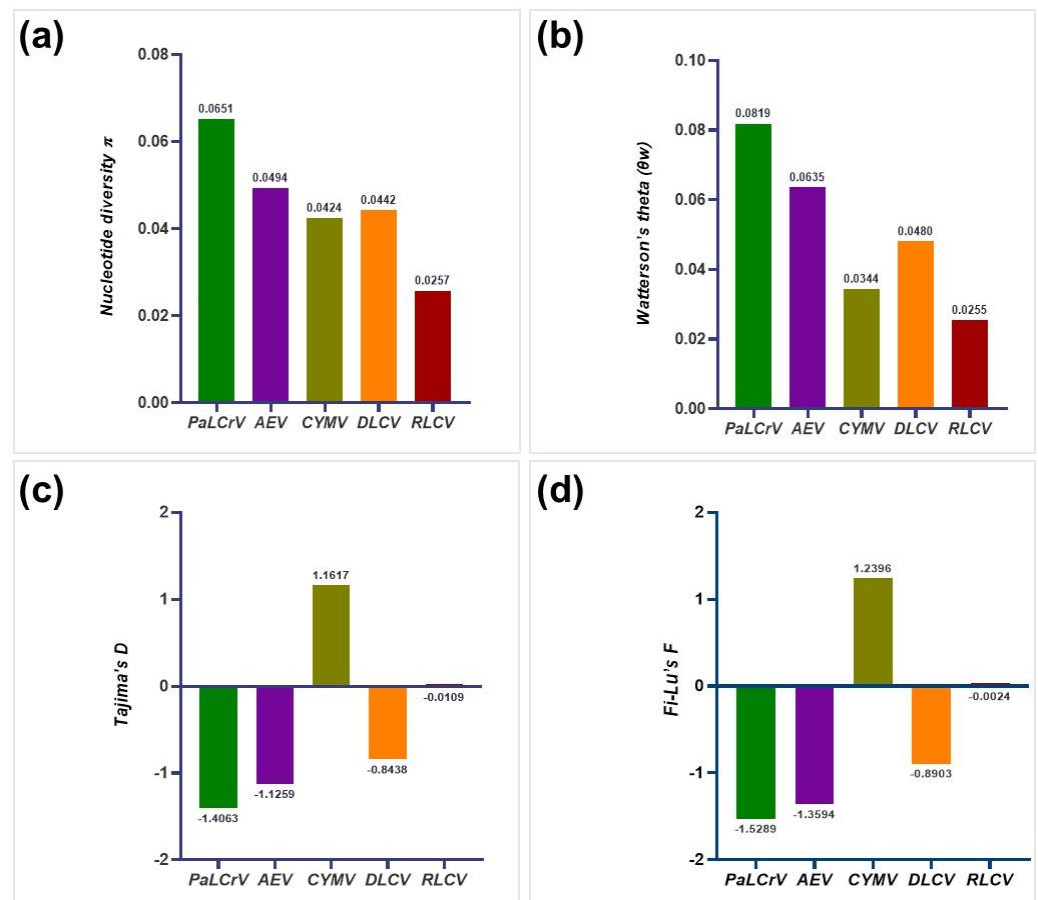
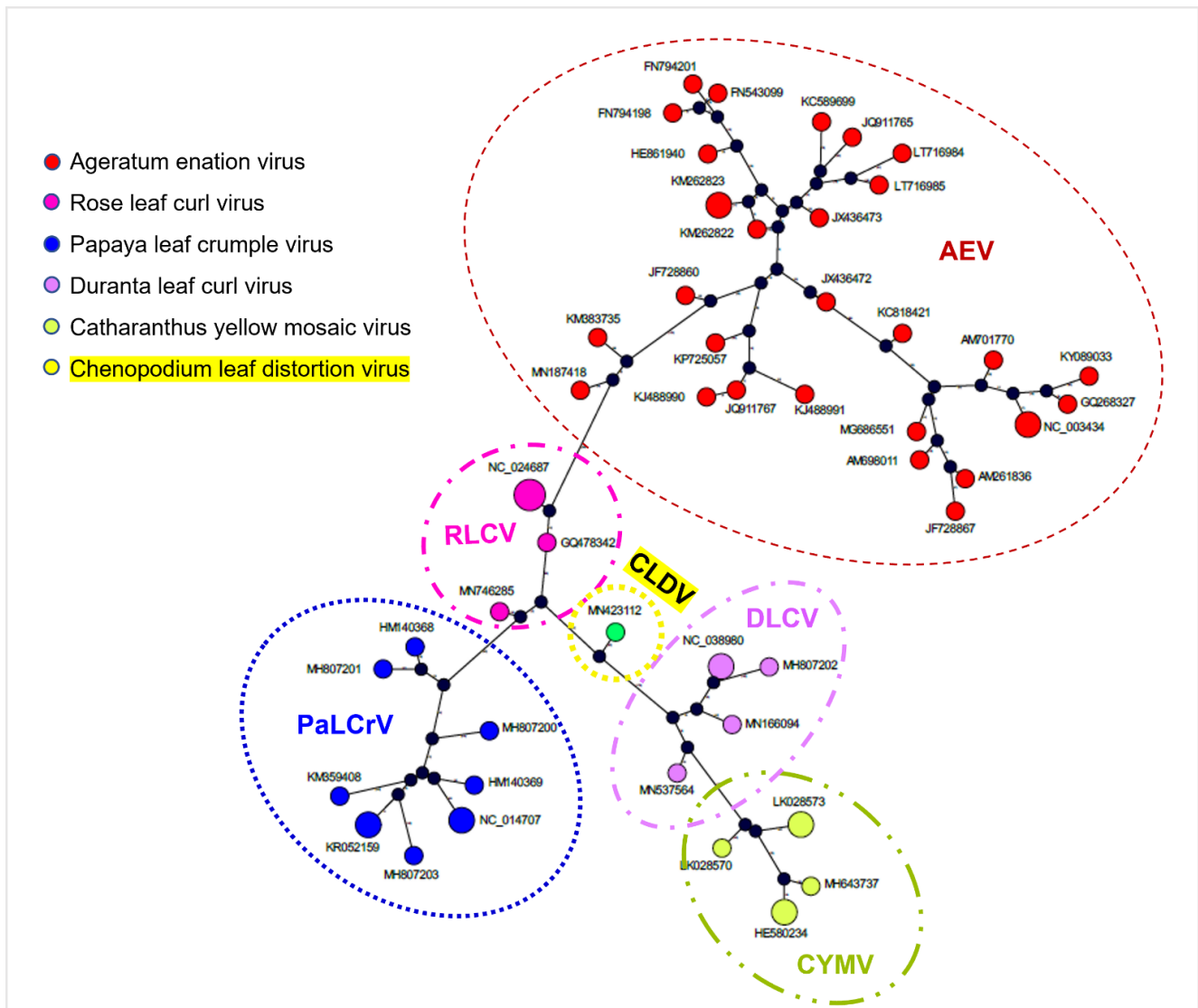


Figure 6. Estimation of genetic diversity was performed for AEV, DLCV, PaLCrV, RLCV, and CYMV populations. The calculated population genetic parameters include (a) nucleotide diversity ( $\pi$ ), (b) Watterson's theta ( $\theta_w$ ), (c) Tajima's D, and (d) Fi-Lu's F.

### 3.7. Estimation of Genealogies through TCS

As the TCS method provides an important tool for dealing with species or genes at the population level and has proved to be a valuable tool in DNA analysis, TCS calculations revealed that most of the isolates sustained a significant number of mutations compared with each other. CLDV is a species that has arisen from the recombination between RLCV, DLCV, and CYMV. Despite CLDV nucleotide sequences having high identity with those of the RLCV, the CLDV genome was found localized in between RLCV and DLCV as shown by the genealogical network analysis (Figure 7).



**Figure 7.** Estimation of gene genealogies through TCS network. CLDV was found located in between RLCV and DLCV.

## 4. Discussion

Geminiviruses have the ability to adapt and evolve quickly as a result of genome-associated changes and recombination events [62,63]. These recombination events have also been documented among members of the genus *Begomovirus* enhancing their virulence [64–67]. These recombination and recurrent mutations can occur in all plant viruses. In this study, we characterized a new recombinant virus (a name suggested as CLDV) with ORFs originating from three different viruses AEV, RLCV, and CYMV (Figure 2). There are no known begomovirus species with such diverse origins. *C. album* samples with leaf distortion symptoms processed in this study were collected from Lahore, Pakistan from regions where RLCV, DLCV, AEV,

and CYMV had been detected previously in various hosts [44,68–70]. As all viruses that constitute CLDV were reported in the same region, it compelled us to think of the three possible scenarios of CLDV speciation: (i) collection of all aforementioned viruses from different hosts by whiteflies and transmitted to the *C. album* plant, which might act as a mixing vessel to facilitate interspecific recombination, (ii) these viruses might have been intermixed in any other host and then transmitted to *C. album* by whiteflies, (iii) the interspecific recombination could have been carried out inside the insect vector, which in this case is the whitefly. To investigate the first scenario, we designed the target-specific primers based on one of the ORF sequences of the recombinant viruses rather than their identical ORFs in CLDV (Table 1). PCR amplification showed no positive results which dismisses the possibility of interspecific recombination in CLDV (Supplementary Figure S3). In the case of the second scenario, we collected a lot of samples i.e., *Vinca rosea*, *Ficus virens*, *Duranta repens*, *Rosa indica*, *Cestrum nocturnum*, etc., from the surrounding areas of the location of CLDV infected *C. album* samples and processed through PCR amplification by using begomovirus specific primers as well as target-specific primers as mentioned above but could not find CLDV in any of the cases (data not shown). There might have been a possibility of missing any other host of CLDV during sample collection. Attempt to detect CLDV from the insect vector, i.e., whiteflies did not succeed either and needs further investigation but the presence of whitefly on *C. album* plants substantiated CLDV as a whitefly vectored begomovirus.

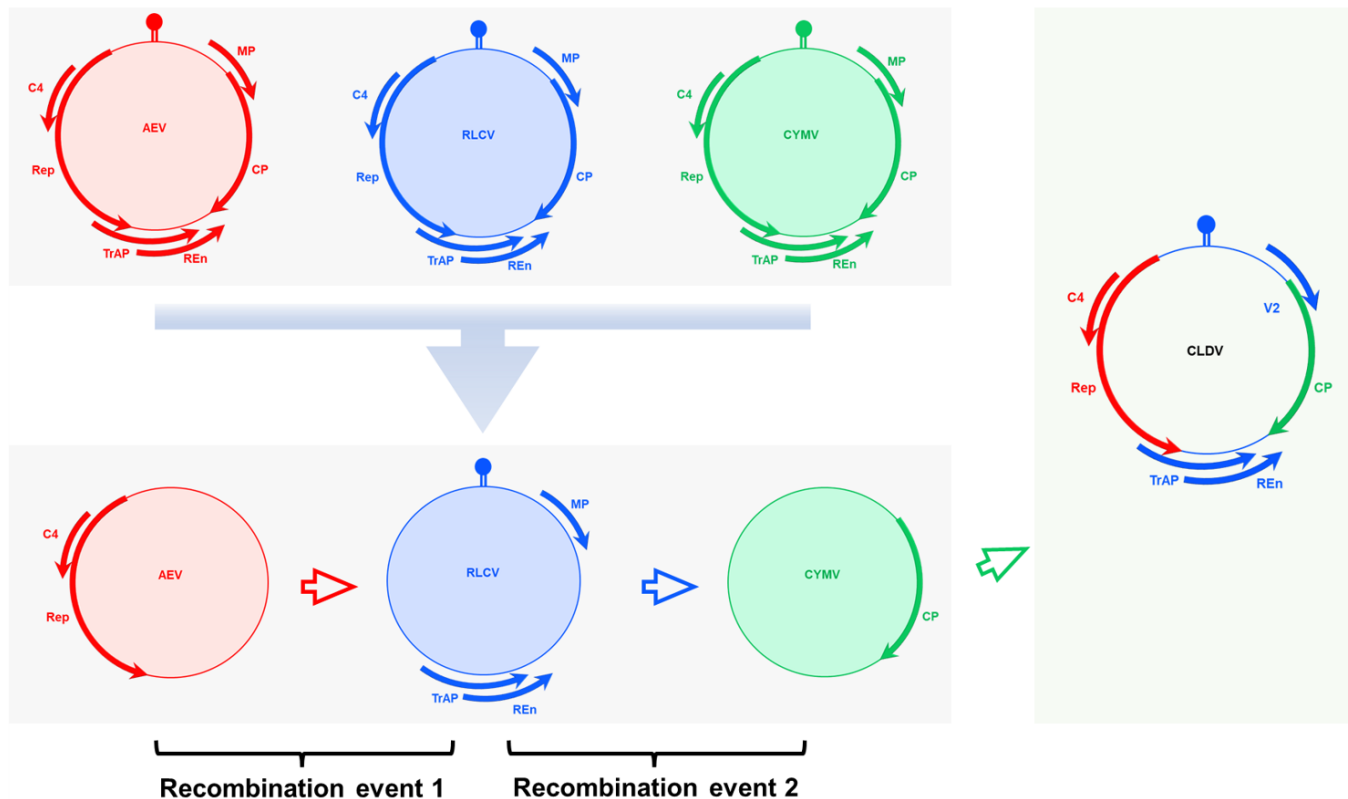
Since genetic variation influences viral emergence, evolution, and vector transmission [4], we investigated the existing genetic diversity of each related virus, i.e., AEV, RLCV, CYMV, PaLCrV, DLCV to determine the extent of genomic variations in these datasets. The average pairwise number of nucleotide differences per site (nucleotide diversity,  $\pi$ ), the number of haplotypes (H), the number of segregating sites (S), haplotype diversity (Hd), Tajima's D value [58], and Fu and Li's F value [59] were calculated for the aforementioned viruses using DnaSP version 6.12.03 (Librado and Rozas 2009, Universitat de Barcelona, Barcelona, Spain). as shown in Figure 6 and Supplementary Table S4. Though we successfully found genetic diversity among these viruses respectively but due to the difference in numbers of sequences (due to the scarcity of viruses in the numbers reported in NCBI GenBank) we cannot conclude that the result data is trustworthy, specifically in regard to CLDV. At least, we observed genetic diversity in all of these CLDV-related viruses which emphasizes the possibility of CLDV existence in the derived class from these viruses.

The possibility of virus existence in the current composition cannot be considered naturally original as clearly recombination events were identified during recombination analysis (Figure 5). Phylogenies are really useful tools to establish genealogical relationships among organisms or their parts (e.g., genes) [60]. Phylogenetic analysis through Mr. Bayes highlighted the evolutionary relatedness among the viruses with the revelation of CLDV localization in between RLCV and DLCV (Figure 4A,B). Along with this traditional method of phylogenetic analysis, an alternative approach TCS [60] was used to provide accurate estimates of gene genealogies at the population level which also showed the same results (Figure 7). On the basis of all these evaluations, we believe that the RLCV is a parent virus here and possesses the CP of CYMV and Rep and C4 of AEV through recombination events respectively as shown in Figure 8. Based on the absence of the other component, i.e., the DNA-B, the novel begomovirus CLDV can be considered a monopartite begomovirus with no associated satellite molecules identified and proved through Koch's postulates by successfully reconstituting the virus from the host after agro-inoculation (Figure 3).

Based on general ICTV demarcation criteria the newly detected virus (CLDV) should be categorized as a new isolate of RLCV (93% sequence identity between them), but the ICTV report clearly demonstrates exceptions in the case of recombinant viruses [29] such as tomato yellow leaf curl Malaga virus and tomato yellow leaf curl Axarquia virus, which have  $\geq 91\%$  identity to both parental viruses (tomato yellow leaf curl virus and tomato yellow leaf curl Sardinia virus) causing both parental species to merge into a single species, despite the fact that all isolates of the parental viruses have  $< 91\%$  identity. Following



this rule, CLDV which shows >91% identity to both RLCV and DLCV is categorized as a new species with the proposal of merging these parental species (88% identical) into a single species.



**Figure 8.** Proposed schematic diagram of CLDV construction through recombination. Two recombination events were hypothesized which results in the existence of CLDV.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/v14102166/s1>, Supplementary Figure S1: Schematic diagram of IC construction of CLDV. 1.1 mer IC was constructed by the addition of restriction enzyme sites i.e., *SpeI* at the start of IC1 and *XbaI* at the end of IC2. *BamHI* is the common point of digestion (end of IC1; start of IC2) existing naturally in the sequence. The sequences of both IC1 and IC2 have been shown as well in the box on the right side. Restriction enzymes in the sequences (IC1 and IC2) are shown in bold letters. Both IC1 and IC2 are ligated with digested pCambia-1303 followed by the transformation into *Agrobacterium* strain GV3101. Supplementary Figure S2. Alignment of Each ORF of CLDV with their respective identical ORF. AEV was aligned with CLDV in the cases of AC1 and AC4. Alignment of RLCV with CLDV was shown in the cases of AC2, AC3, and AV2. In the case of AV1, CYMV was aligned with CLDV. Supplementary Figure S3. Analysis of the common region between the ORFs: ( $\Delta 1$ ) coat protein-movement protein and ( $\Delta 2$ ) replication protein-transcriptional activator protein. The common region of CLDV ORFs was compared with their respective contendents i.e., RLCV and CYMV in the case of  $\Delta 1$ ; RLCV and AEV in the case of  $\Delta 2$ . Analysis was done on both i.e., nucleotide and amino acid levels. In the case of  $\Delta 1$ , RLCV showed more identity with CLDV than CYMV. In the case of  $\Delta 2$ , RLCV showed more identity to CLDV than AEV. The asterisk in red color (\*) shows the variations in the nucleotides whereas the asterisk in black color (\*) shows the variations in the comparison at the amino acid level. Supplementary Figure S4. The exploration of the intriguing genomic composition of the newly detected virus (CLDV). The primers were constructed based on one of the ORF sequences of the recombinant viruses rather than their identical ORFs in the new virus genomic composition. (a) RLCV, (b) CYMV, (c) PaLCrV, (d) DLCV, and (e) AEV couldn't be detected in the infected samples using the specific ORF-based primers. P: positive control (only for DLCV and AEV as didn't have positive control for RLCV, CYMV, PaLCrV), N: Negative control, Lanes 1-3:

infected samples (*C. album*). Supplementary Table S1. Sequences downloaded from NCBI GenBank and used in this study. Supplementary Table S2. The sequence homology of each ORFs of CLDV with the genes of its composition. Supplementary Table S3. Recombination analysis of recombinant CLDV identified in this study. Supplementary Table S4. Estimation of molecular diversity among PaLCrV, AEV, CYMV, DLCV and RLCV populations.

**Author Contributions:** A.L., E.-J.K. and S.L. outlined and conceptualized the review theme. A.L. wrote the first draft of the manuscript. M.A., T.T.B.V., I.G.N.P.W.S., M.A.Q., B.N., M.N.S., E.-J.K. and S.L. contributed to manuscript preparation and revision, and also read and approved the submitted version. All authors contributed to the article and approved the submitted version. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Research Program (Grant #Z-1543086-2017-21-01) for the Exportation Support of Agricultural Products at the Animal and Plant Quarantine Agency of the Republic of Korea.

**Data Availability Statement:** All data are available and present either in the publication or in the indicated databases.

**Conflicts of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

1. Maeso, I.; Roy, S.W.; Irimia, M. Widespread recurrent evolution of genomic features. *Genome Biol. Evol.* **2012**, *4*, 486–500. [[CrossRef](#)] [[PubMed](#)]
2. Blasio, F.; Prieto, P.; Pradillo, M.; Naranjo, T. Genomic and Meiotic Changes Accompanying Polyploidization. *Plants* **2022**, *11*, 125. [[CrossRef](#)] [[PubMed](#)]
3. Mason, A.S.; Wendel, J.F. Homoeologous Exchanges, Segmental Allopolyploidy, and Polyploid Genome Evolution. *Front. Genet.* **2020**, *11*, 1014. [[CrossRef](#)] [[PubMed](#)]
4. Sanjuán, R.; Domingo-Calap, P. Genetic Diversity and Evolution of Viral Populations. *Encycl. Virol.* **2021**, *1*, 53–61.
5. Rubio, L.; Guerri, J.; Moreno, P. Genetic variability and evolutionary dynamics of viruses of the family Closteroviridae. *Front. Microbiol.* **2013**, *4*, 151. [[CrossRef](#)]
6. Watt, W.B. Intragenic Recombination as a Source of Population Genetic Variability. *Am. Nat.* **1972**, *106*, 737–753. [[CrossRef](#)]
7. Duffy, S.; Shackelton, L.A.; Holmes, E.C. Rates of evolutionary change in viruses: Patterns and determinants. *Nat. Rev. Genet.* **2008**, *9*, 267–276. [[CrossRef](#)]
8. Stapley, J.; Feulner, P.G.D.; Johnston, S.E.; Santure, A.W.; Smadja, C.M. Recombination: The good, the bad and the variable. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2017**, *372*, 20170279. [[CrossRef](#)]
9. Heyer, W.-D.; Kanaar, R. Recombination Mechanisms: Fortieth Anniversary Meeting of the Holliday Model. *Mol. Cell* **2004**, *16*, 1–9.
10. Posada, D.; Crandall, K.A.; Holmes, E.C. Recombination in Evolutionary Genomics. *Annu. Rev. Genet.* **2002**, *36*, 75–97. [[CrossRef](#)]
11. Fiallo-Olivé, E.; Trenado, H.P.; Louro, D.; Navas-Castillo, J. Recurrent speciation of a tomato yellow leaf curl geminivirus in Portugal by recombination. *Sci. Rep.* **2019**, *9*, 1332. [[CrossRef](#)] [[PubMed](#)]
12. Rojas, M.R.; Hagen, C.; Lucas, W.J.; Gilbertson, R.L. Exploiting Chinks in the Plant's Armor: Evolution and Emergence of Geminiviruses. *Annu. Rev. Phytopathol.* **2005**, *43*, 361–394. [[CrossRef](#)] [[PubMed](#)]
13. Fiallo-Olivé, E.; Lett, J.-M.; Martin, D.P.; Roumagnac, P.; Varsani, A.; Zerbini, F.M.; Navas-Castillo, J. ICTV Virus Taxonomy Profile: Geminiviridae 2021. *J. Gen. Virol.* **2021**, *102*, 001696. [[CrossRef](#)]
14. Ng, T.F.F.; Duffy, S.; Polston, J.E.; Bixby, E.; Vallad, G.E.; Breitbart, M. Exploring the Diversity of Plant DNA Viruses and Their Satellites Using Vector-Enabled Metagenomics on Whiteflies. *PLoS ONE* **2011**, *6*, e19050. [[CrossRef](#)] [[PubMed](#)]
15. Briddon, R.W.; Bedford, I.D.; Tsai, J.H.; Markham, P.G. Analysis of the nucleotide sequence of the treehopper-transmitted geminivirus, tomato pseudo-curly top virus, suggests a recombinant origin. *Virology* **1996**, *219*, 387–394. [[CrossRef](#)] [[PubMed](#)]
16. Sanz, A.I.; Fraile, A.; García-Arenal, F.; Zhou, X.; Robinson, D.J.; Khalid, S.; Butt, T.; Harrison, B.D. Multiple infection, recombination and genome relationships among begomovirus isolates found in cotton and other plants in Pakistan. *J. Gen. Virol.* **2000**, *81*, 1839–1849. [[CrossRef](#)] [[PubMed](#)]
17. Padidam, M.; Sawyer, S.; Fauquet, C.M. Possible Emergence of New Geminiviruses by Frequent Recombination. *Virology* **1999**, *265*, 218–225. [[CrossRef](#)]
18. Mansoor, S.; Briddon, R.W.; Zafar, Y.; Stanley, J. Geminivirus disease complexes: An emerging threat. *Trends Plant Sci.* **2003**, *8*, 128–134. [[CrossRef](#)]
19. Varma, A.; Malathi, V.G. Emerging geminivirus problems: A serious threat to crop production. *Ann. Appl. Biol.* **2003**, *142*, 145–164. [[CrossRef](#)]

20. Fauquet, C.M.; Bisaro, D.; Briddon, R.; Brown, J.; Harrison, B.; Rybicki, E.; Stenger, D.; Stanley, J. Revision of taxonomic criteria for species demarcation in the family Geminiviridae, and an updated list of begomovirus species. *Arch. Virol.* **2003**, *148*, 405–420. [[CrossRef](#)]
21. Zhang, W.; Olson, N.H.; Baker, T.S.; Faulkner, L.; Agbandje-McKenna, M.; Boulton, M.I.; Davies, J.W.; McKenna, R. Structure of the Maize streak virus geminate particle. *Virology* **2001**, *279*, 471–477. [[CrossRef](#)] [[PubMed](#)]
22. Lazarowitz, S.G.; Shepherd, R. Geminiviruses: Genome structure and gene function. *Crit. Rev. Plant Sci* **1992**, *11*, 327–349. [[CrossRef](#)]
23. Fauquet, C.M.; Mayo, M.A.; Maniloff, J.; Desselberger, U.; Ball, L.A. *Virus Taxonomy: VIIIth Report of the International Committee on Taxonomy of Viruses*; Academic Press: Cambridge, MA, USA, 2005.
24. Brown, J. Current status of Bemisia tabaci as a plant pest and virus vector in agroecosystems worldwide. *FAO Plant Prot. Bull.* **1994**, *42*, 3–32.
25. Roumagnac, P.; Granier, M.; Bernardo, P.; Deshoux, M.; Ferdinand, R.; Galzi, S.; Fernandez, E.; Julian, C.; Abt, I.; Filloux, D.; et al. Alfalfa Leaf Curl Virus: An Aphid-Transmitted Geminivirus. *Virology* **2015**, *89*, 9683–9688. [[CrossRef](#)]
26. Zerbini, F.M.; Briddon, R.W.; Idris, A.; Martin, D.P.; Moriones, E.; Navas-Castillo, J.; Rivera-Bustamante, R.; Roumagnac, P.; Varsani, A. ICTV Virus Taxonomy Profile: Geminiviridae. *J. Gen. Virol.* **2017**, *98*, 131–133. [[CrossRef](#)] [[PubMed](#)]
27. Chiumenti, M.; Greco, C.; De Stradis, A.; Loconsole, G.; Cavalieri, V.; Altamura, G.; Zicca, S.; Saldarelli, P.; Saponari, M. A Novel Bipartite Geminivirid Infecting Olive Trees. *Viruses* **2021**, *13*, 481. [[CrossRef](#)]
28. Lal, A.; Kim, Y.-H.; Vo, T.T.B.; Wira Sanjaya, I.G.N.P.; Ho, P.T.; Byun, H.-S.; Choi, H.-S.; Kil, E.-J.; Lee, S. Identification of a Novel Geminivirus in Fraxinus rhynchophylla in Korea. *Viruses* **2021**, *13*, 2385. [[CrossRef](#)]
29. Brown, J.K.; Zerbini, F.M.; Navas-Castillo, J.; Moriones, E.; Ramos-Sobrinho, R.; Silva, J.C.F.; Fiallo-Olivé, E.; Briddon, R.; Hernández-Zepeda, C.; Idris, A.; et al. Revision of Begomovirus taxonomy based on pairwise sequence comparisons. *Arch Virol.* **2015**, *160*, 1593–1619. [[CrossRef](#)]
30. Harrison, B.D.; Swanson, M.M.; Fargette, D. Begomovirus coat protein: Serology, variation and functions. *Physiol. Mol* **2002**, *60*, 257–271. [[CrossRef](#)]
31. Hehnle, S.; Wege, C.; Jeske, H. Interaction of DNA with the movement proteins of geminiviruses revisited. *J. Virol.* **2004**, *78*, 7698–7706. [[CrossRef](#)]
32. Fondong, V.N. Geminivirus protein structure and function. *Mol. Plant Pathol.* **2013**, *14*, 635–649. [[CrossRef](#)] [[PubMed](#)]
33. Argüello-Astorga, G.R.; Guevara-González, R.G.; Herrera-Estrella, L.R.; Rivera-Bustamante, R.F. Geminivirus Replication Origins Have a Group-Specific Organization of Iterative Elements: A Model for Replication. *Virology* **1994**, *203*, 90–100. [[CrossRef](#)] [[PubMed](#)]
34. Sunter, G.; Stenger, D.C.; Bisaro, D.M. Heterologous Complementation by Geminivirus AL2 and AL3 Genes. *Virology* **1994**, *203*, 203–210. [[CrossRef](#)] [[PubMed](#)]
35. Laufs, J.; Traut, W.; Heyraud, F.; Matzeit, V.; Rogers, S.G.; Schell, J.; Gronenborn, B. In vitro cleavage and joining at the viral origin of replication by the replication initiator protein of tomato yellow leaf curl virus. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 3879–3883. [[CrossRef](#)]
36. Fontes, E.P.; Eagle, P.A.; Sipe, P.S.; Luckow, V.A.; Hanley-Bowdoin, L. Interaction between a geminivirus replication protein and origin DNA is essential for viral replication. *Int. J. Biol. Chem.* **1994**, *269*, 8459–8465. [[CrossRef](#)]
37. Fondong, V.N. The Ever-Expanding Role of C4/AC4 in Geminivirus Infection: Punching above Its Weight? *Mol. Plant* **2019**, *12*, 145–147. [[CrossRef](#)]
38. Altenhofen, L.M.; Dekker, J. Complex regulation of Chenopodium album seed germination. *Appl. Ecol. Environ. Sci.* **2013**, *1*, 133–142. [[CrossRef](#)]
39. Scheepens, P.C.; Kempenaar, C.; Andreasen, C.; Eggert, T.H.; Netland, J.; Vurro, M. Biological control of the annual weed Chenopodium album, with emphasis on the application of Ascochyta caulina as a microbial herbicide. *J. Integr. Pest Manag.* **1997**, *2*, 71–76. [[CrossRef](#)]
40. Eslami, S.V. Comparative germination and emergence ecology of two populations of common lambsquarters (*Chenopodium album*) from Iran and Denmark. *Weed Sci.* **2011**, *59*, 90–97. [[CrossRef](#)]
41. Matloob, A.; Safdar, M.E.; Abbas, T.; Aslam, F.; Khaliq, A.; Tanveer, A.; Rehman, A.; Chadhar, A.R. Challenges and prospects for weed management in Pakistan: A review. *Crop Prot.* **2020**, *134*, 104724. [[CrossRef](#)]
42. Shepherd, D.N.; Martin, D.P.; Lefeuvre, P.; Monjane, A.L.; Owor, B.E.; Rybicki, E.P.; Varsani, A. A protocol for the rapid isolation of full geminivirus genomes from dried plant tissue. *J. Virol. Methods* **2008**, *149*, 97–102. [[CrossRef](#)]
43. Crespo-Bellido, A.; Hoyer, J.S.; Dubey, D.; Jeannot, R.B.; Duffy, S.; Simon, A.E. Interspecies Recombination Has Driven the Macroevolution of Cassava Mosaic Begomoviruses. *J. Virol.* **2021**, *95*, e00541-21. [[CrossRef](#)] [[PubMed](#)]
44. Lal, A.; Kil, E.-J.; Vo, T.T.B.; Fadhila, C.; Ho, P.T.; Shuja, M.N.; Ali, M.; Lee, S. First Report of Duranta leaf curl virus Infecting Ficus virens Showing Leaf Curl Symptoms in Pakistan. *Plant Dis.* **2020**, *104*, 2034. [[CrossRef](#)]
45. Altschul, S.F.; Madden, T.L.; Schäffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402. [[CrossRef](#)]
46. Briddon, R.W.; Bull, S.E.; Mansoor, S.; Amin, I.; Markham, P.G. Universal primers for the PCR-mediated amplification of DNA  $\beta$ . *Mol. Biotechnol.* **2002**, *20*, 315–318. [[CrossRef](#)]
47. Rojas, M. Use of degenerate primers in the polymerase chain reaction to detect whitefly-transmitted geminiviruses. *Plant Dis.* **1993**, *77*, 340–347. [[CrossRef](#)]
48. Southern, E.M. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **1975**, *98*, 503–517. [[CrossRef](#)]

49. Kil, E.-J.; Kim, S.; Lee, Y.-J.; Byun, H.-S.; Park, J.; Seo, H.; Kim, C.-S.; Shim, J.-K.; Lee, J.-H.; Kim, J.-K. Tomato yellow leaf curl virus (TYLCV-IL): A seed-transmissible geminivirus in tomatoes. *Sci. Rep.* **2016**, *6*, 19013. [[CrossRef](#)]
50. Seol, E.; Jung, Y.; Lee, J.; Cho, C.; Kim, T.; Rhee, Y.; Lee, S. In planta transformation of *Notocactus scopia* cv. Soonjung by *Agrobacterium tumefaciens*. *Plant Cell Rep.* **2008**, *27*, 1197–1206. [[CrossRef](#)]
51. Kumar, S.; Stecher, G.; Tamura, K. Evolution, MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [[CrossRef](#)]
52. Miller, M.A.; Pfeiffer, W.; Schwartz, T. The CIPRES science gateway: A community resource for phylogenetic analyses. In Proceedings of the 2011 TeraGrid Conference: Extreme digital discovery, Salt Lake City, Utah, USA, 18–21 July 2011; pp. 1–8.
53. Nylander, J. *MrModeltest, a Program to Evaluate the Fit of Several Models of Evolution to a Given Data and Unrooted Tree (Version 2.2)*; Program Distributed by the Author. Evolutionary Biology Centre, Uppsala University: Uppsala, Sweden, 2004. Available online: <http://www.abc.se/~jnylander> (accessed on 21 January 2022).
54. Darriba, D.; Taboada, G.L.; Doallo, R.; Posada, D. jModelTest 2: More models, new heuristics and parallel computing. *Nat. Methods* **2012**, *9*, 772. [[CrossRef](#)] [[PubMed](#)]
55. Letunic, I.; Bork, P. Interactive Tree Of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **2019**, *47*, W256–W259. [[CrossRef](#)] [[PubMed](#)]
56. Muhire, B.M.; Varsani, A.; Martin, D.P. SDT: A Virus Classification Tool Based on Pairwise Sequence Alignment and Identity Calculation. *PLoS ONE* **2014**, *9*, e108277. [[CrossRef](#)] [[PubMed](#)]
57. Martin, D.P.; Murrell, B.; Golden, M.; Khoosal, A.; Muhire, B. RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evol.* **2015**, *1*, vev003. [[CrossRef](#)]
58. Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **1989**, *123*, 585–595. [[CrossRef](#)]
59. Fu, Y.X.; Li, W.H. Statistical tests of neutrality of mutations. *Genetics* **1993**, *133*, 693–709. [[CrossRef](#)]
60. Clement, M.; Posada, D.; Crandall, K.A. TCS: A computer program to estimate gene genealogies. *Mol. Ecol.* **2000**, *9*, 1657–1659. [[CrossRef](#)]
61. Leigh, J.W.; Bryant, D. popart: Full-feature software for haplotype network construction. *Methods Ecol. Evol.* **2015**, *6*, 1110–1116. [[CrossRef](#)]
62. Zhu, Z.; Meng, K.; Meng, G. Genomic recombination events may reveal the evolution of coronavirus and the origin of SARS-CoV-2. *Sci. Rep.* **2020**, *10*, 21617. [[CrossRef](#)]
63. Lefevre, P.; Moriones, E. Recombination as a motor of host switches and virus emergence: Geminiviruses as case studies. *Curr. Opin. Virol.* **2015**, *10*, 14–19. [[CrossRef](#)]
64. Patil, B.L.; Fauquet, C.M. Cassava mosaic geminiviruses: Actual knowledge and perspectives. *Mol. Plant Pathol.* **2009**, *10*, 685–701. [[CrossRef](#)] [[PubMed](#)]
65. Monci, F.; Sánchez-Campos, S.; Navas-Castillo, J.; Moriones, E. A Natural Recombinant between the Geminiviruses Tomato yellow leaf curl Sardinia virus and Tomato yellow leaf curl virus Exhibits a Novel Pathogenic Phenotype and Is Becoming Prevalent in Spanish Populations. *Virology* **2002**, *303*, 317–326. [[CrossRef](#)] [[PubMed](#)]
66. Lima, A.T.M.; Silva, J.C.F.; Silva, F.N.; Castillo-Urquiza, G.P.; Silva, F.F.; Seah, Y.M.; Mizubuti, E.S.G.; Duffy, S.; Zerbini, F.M. The diversification of begomovirus populations is predominantly driven by mutational dynamics. *Virus Evol.* **2017**, *3*, vex005. [[CrossRef](#)] [[PubMed](#)]
67. Lapidot, M.; Gelbart, D.; Gal-On, A.; Sela, N.; Anfoka, G.; Ahmed, F.H.; Abou-Jawada, Y.; Sobh, H.; Mazyad, H.; Aboul-Ata, A.-A.E.; et al. Frequent migration of introduced cucurbit-infecting begomoviruses among Middle Eastern countries. *Virology* **2014**, *11*, 181. [[CrossRef](#)]
68. Rashid, K.; Tariq, M.; Briddon, R.W.; Kotta-Loizou, I.; Hussain, K. Identification and molecular characterization of rose leaf curl virus in ornamental pomegranate (*Punica granatum* L.). *Australas. Plant Pathol.* **2021**, *50*, 353–356. [[CrossRef](#)]
69. Ilyas, M.; Nawaz, K.; Shafiq, M.; Haider, M.S.; Shahid, A.A. Complete nucleotide sequences of two begomoviruses infecting Madagascar periwinkle (*Catharanthus roseus*) from Pakistan. *Arch. Virol.* **2013**, *158*, 505–510. [[CrossRef](#)]
70. Tahir, M.; Amin, I.; Haider, M.S.; Mansoor, S.; Briddon, R.W. Ageratum enation virus—A Begomovirus of Weeds with the Potential to Infect Crops. *Viruses* **2015**, *7*, 647–665. [[CrossRef](#)]