Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Emergence of novel begomovirus species through recombination in chilli leaf curl disease epidemics in southern India

E. Santhoshinii

santhoshiniielango@gmail.com

University of Agricultural Sciences, Dharwad

Shamarao Jahagirdar

University of Agricultural Sciences, Dharwad

Krishnaraj P.U.

University of Agricultural Sciences, Dharwad

Kambrekar D.N.

University of Agricultural Sciences, Dharwad

Pragadeesh Ayyamuthu Rajarathinam Uma

Western Sydney University

Gurudatta Hegde

University of Agricultural Sciences, Dharwad

Premchand Udavatha

University of Agricultural Sciences, Dharwad

Shankarappa K.S.

Univeristy of horticultural sciences, Bagalkot

Research Article

Keywords: Chilli leaf curl disease, begomovirus, betasatellite, alphasatellite, recombination, novel species, Begomovirus dindigulense, Begomovirus karurense

Posted Date: September 16th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10914768/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Emergence of novel begomovirus species through recombination in chilli leaf curl disease epidemics in southern India

Santhoshinii E.[1]* (ORCID: 0000-0002-9941-8508), **Shamarao Jahagirdar[1]*** (ORCID: 0000-0002-7590-5395), **Krishnaraj P. U.[2]**, **Kambrekar D. N.[3]**, **Pragadeesh Ayyamuthu Rajarathinam Uma[4]** (ORCID: 0009-0009-4092-0681), **Gurudatta Hegde[1]** (ORCID: 0000-0002-5340-1401), **Premchand Udavatha[1]** (ORCID: 0000-0001-5991-5975), **Shankarappa K. S.[5]*** (ORCID: 0000-0003-4894-0280)

[1] Department of Plant Pathology, College of Agriculture, Dharwad, University of Agricultural Sciences, Dharwad 580005, India

[2] Department of Microbiology, College of Agriculture, Dharwad, University of Agricultural Sciences, Dharwad 580005, India

[3] Department of Entomology, College of Agriculture, Dharwad, University of Agricultural Sciences, Dharwad 580005, India

[4] Hawkesbury Institute for the Environment, Western Sydney University, Penrith, New South Wales, Australia

[5] Department of Plant Pathology, College of Horticulture, University of Horticultural Sciences, Bengaluru 560065, India

*Corresponding author: Santhoshinii E.
Email: santhoshiniielango@gmail.com

Abstract

Chilli leaf curl disease (ChiLCD), caused by whitefly-transmitted monopartite begomoviruses, severely constrains chilli (*Capsicum annuum* L.) production in India and worldwide. During the 2023–2024 cropping season, surveys across 15 major chilli-growing districts of Karnataka, Tamil Nadu, and Andhra Pradesh recorded disease incidence of 30.5–76.5%. All 50 symptomatic samples collected during the survey tested positive for begomovirus DNA-A and were negative for DNA-B, confirming monopartite genomic organization. Full-genome sequencing of 16 representative isolates via rolling-circle amplification and Sanger sequencing, coupled with SDT analysis, revealed that 12 isolates belong to Chilli leaf curl virus (ChiLCV) (96.6–99.5% identity), two to Pepper leaf curl Bangladesh virus (PepLCBV) (94.6–95.8% identity), and two represent previously undescribed species (<91% identity with any recognised begomovirus): Chilli leaf curl Dindigul virus (ChiLCDinV; PQ141073**)** **(***Begomovirus dindigulense***)** and Chilli leaf curl Karur virus (ChiLCKarV; PQ141074**)** **(***Begomovirus karurense***)**. Among the 16 begomovirus isolates analyzed, 13 begomovirus isolates were associated with betasatellites. Of the 13 betasatellite isolates, all were found to be strains of previously described South Asian beta satellite strains (>94.5%) rather than novel betasatellites.Two isolates additionally harboured alphasatellites (94.5% inter se identity) forming a new species most closely related to Papaya leaf curl alphasatellite (PaLCA) and Papaya leaf curl virus-associated alphasatellite (PaLCVA) (92.5–92.8% identity). Maximum-likelihood phylogenies confirmed taxonomic assignments and revealed strong geographic clustering with South Asian (India, Pakistan, and Sri Lanka) begomovirus–satellite complexes.

Recombination analysis using RDP analysis identified 58 significant events in begomovirus genomes (mainly in Rep and CP regions) and limited recombination in betasatellites, but none in alphasatellites. At the survey level, the 16 isolates showed a slightly negative Tajima's D (-0.724), indicating a stable, non-expanding local population, whereas at the global level, all available isolates revealed a strong positive Tajima's D values evidencing the deep interspecies divergence. Gene-wise (dN/dS analysis) further exhibited strong purifying selection on the coat protein and replication-associated genes (ω = 0.127, 0.166 respectively), in contrast with diversifying selection on AC4 (ω = 1.618). This study reports two novel begomovirus species species underpinning ongoing ChiLCD epidemics in southern India, highlighting the rapid evolution of begomovirus–satellite complexes and the pressing need for durable resistance breeding and integrated disease management in chilli cultivation.

Keywords: Chilli leaf curl disease, begomovirus, betasatellite, alphasatellite, recombination, novel species, *Begomovirus dindigulense*, *Begomovirus karurense*

1. Introduction

Chilli (*Capsicum annuum* L.), a member of the *Solanaceae* family, is one of the most important spice crops globally and the leading spice commodity in India [1,2]. This annual herb belongs to the genus *Capsicum*, which comprises five major domesticated species, viz., *C. annuum, C. baccatum, C. chinense, C. frutescens,* and *C. pubescens* [2,3]. Among these, *C. annuum* is the most widely cultivated species in India, valued for both pungent (hot pepper) and non-pungent (bell pepper, paprika) fruit types [4]. The crop is grown across tropical, subtropical, and temperate regions worldwide, with optimal growth occurring at 20–30 °C; prolonged exposure to temperatures above 35 °C or below 15 °C severely impairs development and yield [5,6].

Originating in South and Central America, with Mexico as the primary centre of domestication and Peru, Guatemala, and Bulgaria as secondary centres [7,8,9,10]. Chilli was introduced to India during the 16th century and has since become integral to its cuisine and economy for its intense colour and pungency [11]. India's diverse agro-climatic zones enable year-round cultivation establishing the country as the world's largest producer, consumer, and exporter of chilli, contributing approximately 42% of national spice export volume [12,13]. In 2022, India produced 1.98 million tonnes, accounting for 43% of global output [14]. During 2024–25, Andhra Pradesh recorded 269.75 thousand hectares under chilli cultivation with a production of 1,212.90 thousand metric tonnes and a productivity of 4.50 MT/ha, followed by Telangana and Madhya Pradesh [15].

Chilli fruits are rich in vitamin A, β-carotene, ascorbic acid, flavonoids, phenolic compounds, and capsaicinoids, which confer potent antioxidant and anticancer properties [16]. Capsaicin activates TRPV1 receptors to produce pungency [17], whereas capsanthin is responsible for the intense red colour of ripe fruits and serves as a strong antioxidant [18]. Paprika oleoresin extracted from red peppers is widely used in food, pharmaceutical, and cosmetic industries for its high carotenoid content [19].

Despite its economic significance, chilli is highly susceptible to numerous pathogens, particularly viruses, resulting in substantial yield losses [20]. Approximately 75 distinct viruses have been reported in chilli worldwide [21,22], among which begomoviruses causing chilli leaf curl disease (ChiLCD) are the most devastating in tropical and subtropical regions [20,23,24]. First documented incidence of ChiLCD in India was by [25]; it can cause up to 100% yield loss under severe epidemics [26].

Begomoviruses (family *Geminiviridae*) possess circular single-stranded DNA genomes (2.5–5.2 kb) encapsidated in twinned icosahedral particles and are transmitted in a persistent, circulative, non-propagative manner by the whitefly *Bemisia tabaci* cryptic species complex [27]. The genus *Begomovirus* currently includes over 440 ICTV-recognised species, representing ~86% of all geminiviruses [28]. Most Old World begomoviruses are monopartite, often associated with symptom-modulating DNA satellites: betasatellites (family *Tolecusatellitidae*) and alphasatellites (family *Alphasatellitidae*)[29]. Betasatellites encode a single multifunctional βC1 protein that suppresses host silencing and enhances virulence [30], whereas alphasatellites are self-replicating but depend on helper viruses for encapsidation and transmission [31,32].

In India, monopartite Chilli leaf curl virus (ChiLCV) and its associated satellites dominate ChiLCD epidemics [33]. Rapid evolution driven by recombination, pseudorecombination, and satellite capture has led to the emergence of new variants and species, complicating disease management [34]. Given the persistently high incidence of ChiLCD in southern India and the dynamic nature of begomovirus–satellite complexes, comprehensive molecular characterization is critical to elucidate pathogen diversity and evolution [21,35]. The present study investigates the genetic diversity, phylogenetic relationships, recombination patterns, and satellite associations of begomoviruses causing ChiLCD in major chilli-growing regions of southern India during the 2023–2024 cropping season.

2. Materials and methods

2.1. Disease Survey, Sample Collection, and Molecular Characterization of Begomovirus Isolates

A field survey was conducted during the 2023–2024 cropping season to assess the incidence of chilli leaf curl disease (ChiLCD) in major chilli-growing regions of southern India, covering 15 districts across Karnataka, Tamil Nadu, and Andhra Pradesh. Disease incidence was estimated by visual assessment of characteristic begomovirus symptoms, including upward and downward leaf curling, vein thickening, leaf distortion, plant stunting, and reduced fruit size, within randomly selected 25 × 25 m quadrats at each location. A total of 50 symptomatic leaf samples representing different symptom expression and crop growth stages were collected from cultivated chilli fields and stored individually in sterile polythene bags. Samples were labelled CLC-1 to CLC-50, transported on ice, and stored at −80 °C at the Department of Plant Pathology, College of Agriculture, University of Agricultural Sciences, Dharwad, Karnataka, India, until further analysis.

Total genomic DNA was extracted from the 50 symptomatic samples and one asymptomatic healthy chilli leaf sample using the cetyltrimethylammonium bromide (CTAB) method [36]. DNA concentration and purity were determined using a NanoDrop™ 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA), and DNA integrity

was verified by electrophoresis on 0.8% agarose gels. Initial detection of begomoviruses was carried out using degenerate primers specific for DNA-A [37] and DNA-B [38]. PCR-positive DNA-A amplicons from all symptomatic samples were subsequently subjected to preliminary sequence analysis to assess their genetic relatedness before selecting representative isolates for complete genome characterization.

2.2. PCR Detection of Betasatellites and Alphasatellites

All 50 symptomatic leaf samples and one asymptomatic control were screened for the presence of betasatellites and alphasatellites using specific PCR assays. Detection of betasatellite was performed using the universal primer pair Beta01/Beta02 [39] : Beta01 (5′-GGTACCACTACGCTACGCAGCAGCC-3′) and Beta02 (5′-GGTACCTACCCTCCCAGGGGTACAC-3′) alphasatellite was performed using the primer pair DNA101/DNA102 [40] : DNA101 (5′-CTGCAGATAATGTAGCTTACCAG-3′) and DNA102 (5′-CTGCAGATCCTCCACGTGTATAG-3′) respectively. The PCR reaction mixture (25 µL) contained 1× PCR buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl), 1.5 mM $MgCl_2$, 0.2 mM each dNTP, 0.4 µM each primer, 1.25 U of *Taq* DNA polymerase (Takara Bio Inc., Japan), and approximately 50 ng of template DNA. Amplification was performed under the following conditions: initial denaturation at 94 °C for 3 min; 35 cycles of 94 °C for 30 s, 55–57 °C (depending on primer set) for 45 s, and 72 °C for 1 min; and a final extension at 72 °C for 10 min. Expected amplicon sizes were approximately 1.3 kb for betasatellites and 1.3–1.4 kb for alphasatellites (Table 1). Among the 50 symptomatic samples, betasatellite amplicons were detected in 13 of the 16 begomovirus isolates selected for full-genome characterization (CLC-6, CLC-12, CLC-16, CLC-21, CLC-26, CLC-32, CLC-36, CLC-39, CLC-40, CLC-42, CLC-44, CLC-47 and CLC-49). Alphasatellites were detected in two isolates (CLC-21 and CLC-32). All positive amplicons were gel-purified, cloned into pGEM-T Easy vector (Promega, Madison, Wisconsin, USA), and bidirectionally Sanger-sequenced.

Table 1. PCR primers and amplification conditions for detection of Begomovirus and DNA satellites from infected leaf curl samples of chilli

Primers	Oligonucleotide (5'-3')	Length (nt)	PCR conditions						Remarks	References
			Initial denaturation	No. of cycles	Denaturation	Annealing	Extension	Final extension		
Degenerative begomovirus primers										
PAL1v1978	GCATCTGCAGGCCCACATYGT CTTYCCNGT	30	94 ºC for 3 min	30	94 ºC for 1 min	55 ºC for 2 min	72 ºC for 2 min	72 ºC for 10 min	DNA-A 1.2 kb	[37]
PAR1c496	AATACTGCAGGGCTTYCTRTA CATRGG	27								
ToLCBD2081F	GCGTACTCWACGC GCTCAGATTG	23	94 ºC for 3 min	35	94 ºC for 1 min	55 ºC for 45 sec	72 ºC for 90s	72 ºC for 10 min	DNA-B 1.2 kb	[38]
ToLCBD656R	GTGTTTCACAGATTTCCTTACGCG	24								
Betasatellite primers										
BETA01	GGTACCACTACGCTACGCAGC AGCC	25	95 ºC for 5 min	35	94 ºC for 45 sec	57 ºC for 45 sec	72 ºC for 1.30 min	72 ºC for 10 min	1.3 kb	[39]
BETA02	GGTACCTACCCTCCCAGGGGTA CAC	25								
Alphasatellite primers										
DNA101	CTGCAGATAATGTAGCTTAC CAG	23	94 ºC for 5 min	35	94 ºC for 45 sec	55 ºC for 45 sec	72 ºC for 1.30 min	72 ºC for 10 min	1.3 kb	[40]
DNA102	CTGCAGATCCTCCACGTGTATAG	23								

2.3. Full-Genome Characterization Using Rolling Circle Amplification (RCA)

Based on PCR confirmation and preliminary DNA-A sequence analysis, sixteen representative begomovirus isolates (CLC-1, CLC-6, CLC-10, CLC-12, CLC-16, CLC-21, CLC-25, CLC-26, CLC-32, CLC-36, CLC-39, CLC-40, CLC-42, CLC-44, CLC-47, and CLC-49) were selected for complete genome characterization. The selection strategy was designed to ensure both geographical and biological representation of the surveyed population. One representative isolate was selected from each surveyed district to capture the distribution of begomoviruses across southern India. The selected isolates represented the two major crop growth stages encountered during the survey, namely the vegetative and fruiting stages. In Coimbatore district, two representative isolates were included because symptomatic plants were collected independently from both vegetative- and fruiting-stage crops, enabling comparison of viral populations associated with different crop developmental stages within the same geographical region. Thus, the sixteen isolates represented the diversity of geographical locations, crop growth stages, and preliminary sequence variation observed among the fifty field samples.

Circular viral genomes and associated satellite molecules were amplified using the TempliPhi™ Rolling Circle Amplification Kit (Cytiva, Marlborough, MA, USA) according to the manufacturer's instructions. RCA products (2 µL) were digested with BamHI and KpnI restriction endonucleases to release monomeric DNA-A (~2.7 kb) and satellite DNA molecules (~1.3–1.4 kb). The resulting fragments were gel-purified, ligated into BamHI/KpnI-digested pBluescript KS (+) vector, and transformed into *Escherichia coli* DH5α competent cells [41]. Plasmid DNA from three independent positive clones of each isolate was purified by the alkaline lysis method and sequenced bidirectionally by primer walking at Medauxin Pvt. Ltd., Bengaluru, Karnataka, India, using an ABI 3730xl DNA Analyzer (Applied Biosystems, Foster City, CA, USA).

2.4. Genome Sequencing, Sequence Analysis, Phylogenetic Reconstruction, and Recombination Detection

Raw Sanger sequences were quality-trimmed and vector-contaminated regions were removed using VecScreen (NCBI). Contigs were assembled *de novo* and edited using BioEdit v7.2.5 [42]. Initial identification of sequence similarities was performed by BLASTn against the NCBI non-redundant nucleotide database. Reference sequences of begomoviruses (n = 59), alphasatellites (n = 23), and betasatellites (n = 32) exhibiting the highest BLAST scores were retrieved from GenBank. Multiple sequence alignments were generated using MAFFT v7.490 [43] with default parameters.

Species demarcation and intergenomic similarity were determined using SDT analysis for complete genome, intergenic and ORFs (AV2,AV1,AC1,AC2,AC3,AC4) applying the ICTV-approved 91% threshold for begomovirus species demarcation, 78% threshold for betasatellites and 83–88% threshold for alphasatellites. Open reading frames (ORFs) were predicted using NCBI ORFfinder and verified manually.

Maximum-likelihood (ML) phylogenetic trees were constructed with IQ-TREE v2.2.2.6 [44]. The best-fit nucleotide substitution models, selected by ModelFinder, were TIM3+F+I+G4 (DNA-A), GTR+F+I+G4 (betasatellites), and TPM3u+F+I+G4 (alphasatellites). Branch support was assessed with 10,000 ultrafast bootstraps and SH-aLRT tests.

Evidence of recombination was first visualized by constructing phylogenetic networks in FastTree v2.1.11 [45] using neighbor-joining trees (60% consensus cutoff, 1000 bootstraps). Putative recombination events were then evaluated using seven algorithms (RDP, GENECONV, BootScan, MaxChi, Chimaera, SiScan, and 3Seq) implemented in RDP5 v5.5 [46]. Only events detected by ≥4 methods with Bonferroni-corrected $P \leq 0.01$ and supported by phylogenetic evidence were considered significant**.** Default settings were used except for a sliding window of 10 variable sites.

2.5. Population Genetics and Selection Pressure Analysis of DNA-A Genes

DNA polymorphism and neutrality statistics were computed for two datasets: (i) the 16 chilli leaf curl virus (ChiLCV) study isolates alone, representing the intraspecific ChiLCD population sampled in this study, and (ii) the full 75-sequence reference-inclusive alignment. For each dataset, sites containing gaps or missing data in any sequence were excluded (complete-deletion criterion). Number of segregating sites (S), total mutations (Eta), haplotype number (h) and diversity (Hd), nucleotide diversity (π), Watterson's theta (θW), Tajima's D [47], Fu and Li's D* and F*[47], and Fu's Fs[48] were calculated using EGGLIB v3.6.1 [50], with multi-hit–aware mutation counting enabled. Statistical significance of neutrality statistics was assessed against approximate critical value thresholds ($|D|$, $|D|$, $|F| \approx 2.0$; two-tailed $\alpha = 0.05$)

Selection pressure acting on each of the six DNA-A-encoded open reading frames (AC1/Rep, AC2/TrAP, AC3/REn, AC4, AV1/CP, and AV2) was estimated by calculating the ratio of nonsynonymous to synonymous

substitution rates (ω = dN/dS). Coding sequences for each gene were verified for correct reading frame and coding-strand orientation, translated, aligned as amino acid sequences using MAFFT v7 (L-INS-i), and back-translated to codon-aware nucleotide alignments following the PAL2NAL approach [51]. Pairwise dN and dS values were then calculated for all isolate pairs within each gene using the counting method [52] implemented in Biopython v1.88 [53] and averaged to obtain a gene-wide ω estimate. Genes with ω < 1 were interpreted as evolving under purifying selection and ω > 1 as showing evidence of diversifying selection or relaxed functional constraint.

3. Results

3.1. Disease Incidence and Symptomatology of Chilli Leaf Curl Disease

Field surveys conducted during the 2023–2024 cropping season across 50 commercial chilli fields in 15 districts of Karnataka, Tamil Nadu, and Andhra Pradesh revealed widespread occurrence of chilli leaf curl disease (ChiLCD), with incidence ranging from 30.5% to 76.5% (Table.2). Infected plants displayed characteristic begomovirus symptoms, including upward or downward leaf curling, vein banding and thickening, leaf puckering and crinkling, yellowing of leaf margins, shortened internodes and petioles, apical bunching, severe stunting, flower abortion, and production of few, small, deformed fruits (Fig. 1). Fifty symptomatic leaf samples exhibiting severe leaf curl were collected and designated CLC-1 to CLC-50 (Table.2) for subsequent molecular analysis.

Table 2. Locations of chilli leaf curl isolates collected from chilli and representative isolates used for full length genome sequencing.

S.No.	State	District	Taluk	Village	Per cent Disease Incidence	Name assigned
1	Karnataka	Belagavi	Belagavi	kudremani	47.6	CLC-1
2	Karnataka	Belagavi	Belagavi	Aste	44.3	CLC-2
3	Karnataka	Belagavi	Athani	Athani	39.5	CLC-3
4	Karnataka	Bellary	Bellary	Asundi	38.7	CLC-4
5	Karnataka	Bellary	Bellary	Kudathini	45.4	CLC-5
6	Karnataka	Bellary	Bellary	Kudathini	52.3	CLC-6
7	Karnataka	Bellary	Bellary	Beivinahalli	49.3	CLC-7
8	Karnataka	Gadag	Gadag	Harlapur	53.4	CLC-8
9	Karnataka	Gadag	Mundargi	Mundargi	58.5	CLC-9
10	Karnataka	Gadag	Mundargi	Dindur	60.3	CLC-10
11	Karnataka	Dharwad	Kalghati	Gambyapur	64.5	CLC-11
12	Karnataka	Dharwad	Kalghati	Dhummawad	76.5	CLC-12
13	Karnataka	Dharwad	Dharwad	Dharwad	53.5	CLC-13
14	Karnataka	Dharwad	Khundgol	Devanur	55.6	CLC-14
15	Karnataka	Dharwad	Navalkund	Dundur	62.3	CLC-15
16	Karnataka	Haveri	Byadgi	Byadgi	72.3	CLC-16
17	Karnataka	Haveri	Byadgi	Montebennur	70.5	CLC-17
18	Karnataka	Haveri	Haveri	Devihosur	67.6	CLC-18
19	Karnataka	Haveri	Shiggaon	Adavisomapur	57.4	CLC-19
20	Karnataka	Bagalkot	Bagalkot	Kaladgi	47.5	CLC-20
21	Karnataka	Bagalkot	Kerur	Kerur	53.6	CLC-21
22	Karnataka	Davengere	Harapanahalli	Bennihalli	36.6	CLC-22

23	Karnataka	Koppal	Koppal	Bevinhalli	42.0	CLC-23
24	Karnataka	Koppal	Koppal	Mattur	44.6	CLC-24
25	Karnataka	Davengere	Harihara	Haraganahalli	39.4	CLC-25
26	Karnataka	Vijayanagara	Hagaribommanahalli	Hagaribommanahalli	63.2	CLC-26
27	Karnataka	Vijayanagara	Hagaribommanahalli	Hagaribommanahalli	30.5	CLC-27
28	Karnataka	Vijayanagara	Kudligi	Haralu	61.9	CLC-28
29	Karnataka	Vijayanagara	Kotturu	Rampura	59.8	CLC-29
30	Karnataka	Vijayanagara	Harpanahalli	Ananthahalli	60.4	CLC-30
31	Tamil Nadu	Dindugal	Dindugal	Kamachipuram	70.6	CLC-31
32	Tamil Nadu	Dindugal	Ottanchatram	Reddiapatti	75.7	CLC-32
33	Tamil Nadu	Dindugal	Ottanchatram	Pottikampatti	76.5	CLC-33
34	Tamil Nadu	Dindugal	Ottanchatram	Pottikampatti	69.5	CLC-34
35	Tamil Nadu	Karur	Kulithalai	Poyyamani	43.6	CLC 35
36	Tamil Nadu	Karur	Kulithalai	Pulitheri	62.5	CLC 36
37	Tamil Nadu	Dharmapuri	Karimangalam	Elumichanahalli	60.7	CLC-37
38	Tamil Nadu	Dharmapuri	Palacode	Nallur	58.8	CLC-38
39	Tamil Nadu	Dharmapuri	Pennagaram	Chinnampalli	54.5	CLC-39
40	Tamil Nadu	Salem	Omalur	Darapuram	52.5	CLC-40
41	Tamil Nadu	Salem	Attur	Arasanatham	50.2	CLC-41
42	Tamil Nadu	Coimbatore	Coimbatore	Ikkaraibooluvampatti	50.2	CLC-42
43	Tamil Nadu	Coimbatore	Thondamuthur	Irutupallam	48.5	CLC-43
44	Tamil Nadu	Coimbatore	Thondamuthur	Vellimallaipattinam	42.6	CLC-44
45	Tamil Nadu	Krishnagiri	Pochampalli	Damodarhalli	40.5	CLC-45
46	Tamil Nadu	Krishnagiri	Veppanahalli	Doddaganamma	52.0	CLC-46
47	Tamil Nadu	Krishnagiri	Veppanahalli	Sigaramayanapalli	53.5	CLC-47
48	Andhra Pradesh	Prakasham	Markapur	Sivarampuram	48.5	CLC-48
49	Andhra Pradesh	Prakasham	Markapur	Sivarampuram	59.5	CLC-49
50	Andhra Pradesh	Prakasham	Podili	Mallavaram	42.5	CLC-50

3.2. Molecular Detection of Begomoviruses, Betasatellites, and Alphasatellites

Total genomic DNA extracted from all 50 symptomatic samples yielded the expected ~1.2-kb amplicon with degenerate DNA-A-specific primers, confirming begomovirus infection in every case. No amplification was observed with DNA-B-specific primers, indicating that all detected begomoviruses were monopartite. Based on geographical representation and symptom severity, 16 isolates (Table.3) were selected for full-genome characterization.

PCR screening further revealed the presence of betasatellites in 13 of the 16 selected isolates (CLC-6, CLC-12, CLC-16, CLC-21, CLC-26, CLC-32, CLC-36, CLC-39, CLC-40, CLC-42, CLC-44, CLC-47, and CLC-49), as evidenced by specific ~1.3-kb amplicons. Alphasatellites were detected in two isolates (CLC-21 and CLC-32) using

alphasatellite-specific primers, also producing ~1.3-kb products. All diagnostic amplicons were cloned and Sanger-sequenced bidirectionally for confirmation.

Table. 3. Leaf curl isolates collected from chilli used for full length genome sequencing

S.No.	State	District	Taluk	Village	Name assigned
1	Karnataka	Belagavi	Belagavi	Kudremani	CLC-1
2	Karnataka	Bellary	Bellary	Kudathini	CLC-6
3	Karnataka	Gadag	Mundargi	Dindur	CLC-10
4	Karnataka	Dharwad	Kalghati	Dhummawad	CLC-12
5	Karnataka	Haveri	Byadgi	Byadgi	CLC-16
6	Karnataka	Bagalkot	Kerur	Kerur	CLC-21
7	Karnataka	Davengere	Harihara	Haraganahalli	CLC-25
8	Karnataka	Vijayanagara	Hagaribommanahalli	Hagaribommanahalli	CLC-26
9	Tamil Nadu	Dindugal	Ottanchatram	Reddiapatti	CLC-32
10	Tamil Nadu	Karur	Kulithalai	Pulitheri	CLC 36
11	Tamil Nadu	Dharmapuri	Pennagaram	Chinnampalli	CLC-39
12	Tamil Nadu	Salem	Omalur	Darapuram	CLC-40
13	Tamil Nadu	Coimbatore	Coimbatore	Ikkaraibooluvampatti	CLC-42
14	Tamil Nadu	Coimbatore	Thondamuthur	Vellimallaipattinam	CLC-44
15	Tamil Nadu	Krishnagiri	Veppanahalli	Sigaramayanapalli	CLC-47
16	Andhra Pradesh	Prakasham	Markapur	Sivarampuram	CLC-49

3.3. Genome Organization and Taxonomic Classification of Begomoviruses and Associated Satellites

Full-length DNA-A sequences of the 16 selected monopartite begomoviruses ranged from 2737 to 2773 nucleotides (GenBank accessions PQ141071–PQ141086). All isolates exhibited the typical Old World monopartite begomovirus genome organization, encoding two virion-sense ORFs (AV1/CP and AV2) and four complementary-sense ORFs (AC1/Rep, AC2, AC3/Ren, and AC4) (Fig. 2; Table 1). Betasatellites associated with 13 isolates were 1367–1390 nt in length (GenBank PQ057804–PQ057816), whereas the two alphasatellites associated with CLC-21 and CLC-32 were each 1353 nt and named A-21 and A-32, respectively. Betasatellites contained a single βC1 ORF, an A-rich region, and a satellite conserved region (SCR), while alphasatellites encoded a replication-associated protein (Rep) and an A-rich region (Tables 4,5).

Table 4. Genome organization of betasatellites of begomoviruses isolated from chilli

Betasatellite		B-6	B-12	B-16	B-21	B-26	B-32	B-36	B-39	B-40	B-42	B-44	B-47	B-49
Nucleotide length (nt)		1367	1370	1367	1367	1370	1371	1366	1370	1367	1372	1367	1370	1390
Begomovirus isolate		CLC 6	CLC 12	CLC 16	CLC 21	CLC 26	CLC 32	CLC 36	CLC 39	CLC 40	CLC 42	CLC 44	CLC 47	CLC 49
BC1	SS0	180-584	563-201	180-584	180-584	563-201	565-200	563-201	565-200	180-584	563-201	180-584	562-200	583-221
	PO	405	363	405	405	363	366	363	366	405	363	405	363	363
	PP	134	120	134	134	120	121	120	121	134	120	134	120	120
	PWP	15.03	13.96	15.03	15.03	13.96	13.97	13.96	14.01	15.03	13.96	15.03	13.96	14.03

Note : NT: Nucleotide (nt); SSC: Start–Stop codon (nt); PO: Predicted size of ORFs (nt); PP: Predicted size of protein (no. of amino acids); PWP: Predicted weight of protein (kDa). BC1- pathogenicity factor

Table 5. Genome organization of alphasatellites of begomoviruses isolated from chilli

Alphasatellite		A-21	A-32
Nucleotide length (nt)		1353	1353
Begomovirus isolate		CLC-21	CLC-32
Rep	SS0	65-1012	65-1012
	PO	948	948
	PP	315	315
	PWP	36.5	36.5

Note : NT: Nucleotide (nt); SSC: Start–Stop codon (nt); PO: Predicted size of ORFs (nt); PP: Predicted size of protein (no. of amino acids); PWP: Predicted weight of protein (kDa). Rep - Replicase associated protein

3.3.1. Begomoviruses infecting chilli

Independent SDT pairwise nucleotide identity analysis of the 16 full-length DNA-A sequences, together with 59 selected Old World begomovirus references, revealed four distinct taxonomic affiliations (Fig. 3; Supplementary Table. S2). Twelve isolates **(CLC-1, CLC-6, CLC-10, CLC -12, CLC-16, CLC-21, CLC-25, CLC-26, CLC-39, CLC-44, CLC-47, CLC-49)** shared 96.6–99.9% nucleotide identity with established Chilli leaf curl virus (ChiLCV) isolates and >94% identity among themselves, placing them unequivocally within the species ChiLCV (ICTV species demarcation threshold ≥91%). Within this group, CLC-1 was slightly more divergent (76.8-96.6 % with most ChiLCV isolates).

Two isolates, **CLC-40 and CLC-42**, shared their highest identity to Pepper leaf curl Bangladesh virus (PepLCBV), sharing 95.3% and 96.4% nucleotide identity respectively with established PepLCBV-IN:TN:Coi:Chi:18-HM007096 and 96.40% identity with each other. These values are well within the ICTV species demarcation threshold for PepLCBV (≥91%) confirming their classification as PepLCBV rather than ChiLCV or a novel species.

The remaining two isolates, **CLC-32 and CLC-36**, exhibited maximum pairwise identities of only 88.1% and 89.6%, respectively, with any currently recognized begomovirus species (CLC-32 with ChiLCV-SL:Ban:Tom:03-AF274349; CLC-36 with ChiLCV-PK:SP:Chi:24-MW615736). Since both values are below the 91% ICTV species demarcation threshold, these isolates represent two previously undescribed monopartite begomovirus species, provisionally designated Chilli leaf curl Dindigul virus (ChiLCDinV; CLC-32) **(*Begomovirus dindigulense*)** and Chilli leaf curl Karur virus (ChiLCKarV; CLC-36) **(*Begomovirus karurense*)**.

Maximum-likelihood phylogenetic analysis (IQ-TREE 2, TIM3+F+I+G4 model, log-likelihood: –152865.06) strongly corroborated the SDT taxonomic assignments (Fig. 4). The unrooted phylogenetic tree of the 75-sequence DNA-A dataset resolved 12 major clades radiating from a central polytomy, consistent with the SDT-based groupings. The largest clade comprised the core ChiLCV group together with reference ChiLCV sequences, formed a highly supported monophyletic clade (100% ultrafast bootstrap / 100% SH-aLRT) together with previously described ChiLCV sequences from India, Pakistan, and Bangladesh. A distinct, well-separated clade contained CLC-32, CLC-36, and their closest reference matches (ChiLCSV, ChiLCVV, ChiLCSLV), confirming that these two isolates are deeply divergent lineages basal to the ChiLCV–PepLCBV complex, each with strong nodal support (≥98% bootstrap). CLC-40 and CLC-42 fell within a separate clade alongside PepLCBV, ChiLCNV, and PepLCLV reference sequences. The remaining reference taxa (MYMV, PepGMV, CabLCV, PepHYVV, and other divergent Old World begomoviruses) each formed short, isolated single- or two-taxon branches at the periphery of the topology, reflecting their substantial evolutionary distance from all chilli-infecting isolates.

Nucleotide sequence analysis of the AV2 (pre-coat protein) gene highlighted maximum pairwise identities of 94.9-100% for the core ChiLCV-clade isolates. Among them, 6 isolates (CLC-1, CLC-10, CLC-12, CLC-16, CLC-25 and CLC-47) showed highest pairwise identity of 99.4-99.7 % against Eclipta yellow vein virus (ELCCV) rather

than the ChiLCV reference, indicating that these isolates is more broadly conserved across the old world begomoviruses than the rest of the genome. Despite the genome wide divergence of CLC-32, CLC-36, the AV2 gene in these isolates showed high pairwise similarity 99.4 % and 98.2 % respectively, evidencing the AV2 gene is the more conserved gene in the DNA-A ORF (Supplementary Table. S3).

Nucleotide sequence analysis of the AV1 (coat protein) gene was highly conserved across the dataset, with maximum identities spanning 96.9-99.9% to ChiLCV references the core isolates. CLC-40 (96.8%) and CLC-44 (96.8%) shared highest AV1 identity with PepLCBV. Meanwhile, isolates CLC-32, CLC-36 showed 98.7% and 96.9 % identities within the ChiLCV species range (Supplementary Table. S4).

Nucleotide sequence analysis of the AC1 (replication associated) gene revealed a diverse pattern of pairwise identity with related begomoviruses. Ten isolates (CLC-1, CLC-6, CLC-10, CLC-16, CLC-21, CLC-25, CLC-26, CLC-39, CLC-47, CLC-49) shared 98.3-99.9 % identity with ChiLCV references. CLC-32 showed the lowest AC1 nucleotide pairwise identity (84.3 % against ChiLCV references), indicating that AC1 is the principal locus of gneome-wide divergence in the isolate (Supplementary Table. S5).

Nucleotide sequence analysis of the AC2 (transcriptional activator) gene across the dataset revealed a clear pattern of conservation with related begomoviruses. Among the core isolates clade, nine isolates exhibited 98.5 -100 % AC2 identity with ChiLCV references. CLC-32 showed reduced similarity in the pairwise identity, show genome divergence in the region. CLC-36 (97.8%), CLC-40 (99.3%) and CLC-42 (99.5%) shared their highest AC2 identity with PepLCLC or PepLCBV references, while CLC-44 (98.3%) aligned most closely with ChiLCAV (Supplementary Table. S6).

Nucleotide sequence analysis of the AC3 (replication enhancer) gene across the dataset revealed AC3 pairwise identities were highly conserved, with 10 isolates reaching more than 98 % identity to ChiLCV. CLC-32 and CLC-36 showed their lowest identities at this gene level relative to AV1/AV2, though both remained above the AC1/AC2 (Supplementary Table. S7).

Nucleotide sequence analysis of the AC4 gene across isolates revealed a clear pattern of conservation with related begomoviruses. The core isolates clade showed 97.6 – 100 % similarity with ChiLCV references. In contrast. CLC-12 (94.9%), CLC- 32 (93.9%) and CLC-36 (95.2%) shared their highest AC4 gene identity with Tomato leaf curl Gujarat virus (ToLCGV) rather than ChiLCV, suggesting AC4 has a distinct evolutionary history in these three isolates (Supplementary Table. S8).

Recombination analysis using RDP5 identified **58 putative recombination events** with Bonferroni-corrected P-values ≤ 0.05 and confirmation by at least four methods (Supplementary Table. S1). Breakpoints were predominantly located in the Rep (C1/AC1) gene (approximately nt 950–1,950) and the CP (AV1) gene

(approximately nt 1,950–2,650), consistent with established recombination hotspots in Old World monopartite begomoviruses.

Among the 16 CLC isolates, strong and unambiguous recombination signals (supported by ≥4 methods and highest P-values < 10^{-20}) were detected in four isolates belonging to the divergent or novel taxa. **Event 1** (P = 1.42 × 10^{-59}) identified CLC-32 as recombinant; major parent CLC-36, minor parent unknown (PepLCLV-IN:Kar:Ram:Chi:17, MW978695); breakpoints nt 1,209–1,906 (Rep region). **Event 5** (P = 1.44 × 10^{-50}) detected CLC-40 and CLC-42 as recombinants; major parent CLC 1 (ChiLCV), minor parent Croton yellow vein mosaic virus (CYVMV-IN:LK:Jat:07, EU727086); breakpoints nt 1,914–2,660 (Rep/CP overlap). **Event 8** (P = 4.38 × 10^{-56}) recorded CLC 36 as recombinant; major parent CLC 1 (and other ChiLCV isolates), minor parent Chilli leaf curl Sri Lanka virus (ChiLCSLV-SL:Noc:Chi:19, JN555601); breakpoints nt 143–1,276 (intergenic region to Rep). **Event 10** (P = 1.02 × 10^{-28}) identified CLC 1 as recombinant; major parent ChiLCV-IN:Nar:Chi:07 (EU939533), minor parent Tomato leaf curl Gujarat virus (ToLCGV-PK:Tob:Cot:15, LN794215); breakpoints nt 2,265–2,531 (CP region).

Parental sequences in the confirmed events were exclusively South Asian begomoviruses (ChiLCV, PepLCBV, CYVMV, ChiLCSLV, ToLCGV), demonstrating that inter-species recombination continues to shape the genetic diversity of chilli-infecting begomoviruses in the region. The strongest recombination signatures were concentrated in the two novel species (CLC-32 and CLC-36) and the two PepLCBV isolates (CLC-42 and CLC-44), underscoring recombination as a primary mechanism generating taxonomic and biological novelty.

3.3.2. Betasatellites

Thirteen betasatellites (B-6, B-12, B-16, B-21, B-26, B-32, B-36, B-39, B-40, B-42, B-44, B-47, B-49)**,** associated with CLC-6, CLC-12, CLC-16, CLC-21, CLC-26, CLC-32, CLC-36, CLC-39, CLC-40, CLC-42, CLC-44, CLC-47, and CLC-49, ranged from 1345 to 1361 nt in length were analysed. SDT pairwise nucleotide identity analysis revealed substantial diversity among them (39.6–99.3% identity; Supplementary Table S9). Ten betasatellites (B-6, B-12, B-16, B-21, B-26, B-39, B-40, B-42, B-47, and B-49**)** shared 99.3–99.9 % identity with each other and with reference Tomato leaf curl Bangladesh betasatellite (ToLCBB) (Fig.5).

The remaining three betasatellites, B-32 (associated with CLC-32), B-36 (associated with CLC-36), and B-44 (associated with CLC-44) were highly divergent, with B2 sharing 95.3% similarity with CLCSLB, B-36 sharing 94.8% similarity with ChiLCB and B-44 sharing 96.1% similarity with TOLCBB. All isolates clearly exceed the current ICTV species demarcation threshold for betasatellites (>78% identity) and are therefore classified as strains instead of species.

Maximum-likelihood phylogenetic analysis strongly supported geographic and taxonomic clustering of the betasatellites (Fig. 6). Along the reference sequences, the phylogram resolved into 10 clades. Ten betasatellites (B-6, B-12, B-16, B-21, B-26, B-39, B-40, B-42, B-47, and B-49) formed a single large, branching clade together with

reference ToLCBB and with other related ChiLCB reference sequences, confirms their close relationship. B-32 and B-36 grouped together with a separate well-defined clade alongside CLCSLB and ChiLCB reference sequences, rather than ToLCBB. The remaining reference betasatellites formed a distinct, more peripheral clade.

RDP5 analysis identified eight putative recombination events, but only one event unambiguously involved a newly characterized betasatellite as the recombinant. **Event 3** (supported by 4 methods, $P = 2.45 \times 10^{-5}$) identified B-26 as recombinant; major parent ToLCBB-[IN:UP:Chi:23] (OR180506); minor parent unknown (BYVIB-[IN:Bih:Sam:Okra:15], KT390401); breakpoint interval 1,075–1,123 nt (49-nt segment within the βC1 ORF). No unique, robust recombination breakpoints were detected within the ten ToLCBB variants or the three novel betasatellites.

3.3.3. Alphasatellites

Two alphasatellites, **A-21 (associated with CLC-21)** and **A-32 (associated with CLC- 32)**, were 1368–1372 not in length. SDT analysis revealed 98.3% nucleotide identity between them, placing both well above the current ICTV species demarcation threshold for alphasatellites (<83–88%) and classifying them as strains of the same, previously undescribed species (Supplementary Table. S10, Fig.7). The highest pairwise nucleotide identities of A-21 and A-32 were 93% with Papaya leaf curl alphasatellite (PaLCA) and 92.8% with Papaya leaf curl virus associated alphasatellite (PaLCVA) respectively. On this basis, they can be best classified as strains of PaLCA and PaLCVa, respectively, rather than as novel alphasatellite species.

Maximum-likelihood phylogenetic analysis placed A-21 and A-32 in a strongly supported two-taxon branch pair nested within a large clade containing PaLCA, PaLCVA (Fig. 8). The direct topological proximity of A-21 and A-32 to PaLCA and PaLCVA within this radiating clade, rather than a isolated deeply lineage provides evidence for the independent phylogenetic support. Recombination analysis detected 33 putative events across the dataset, but **none involving A-21 or A-32 were supported by three or more methods with $P < 0.05$**. SplitsTree4 network analysis confirmed widespread recombination in the dataset (Φ test, P = 0.0), but Delta scores for A-21 (0.260) and A-32 (0.269) were among the lowest recorded, indicating predominantly tree-like evolution.

3.4. DNA polymorphism and neutrality tests

Analysis of the 16 ChiLCV survey level isolates alone (2,720 unambiguously aligned sites) revealed extensive intraspecific polymorphism, with 838 segregating sites and 942 total inferred mutations. All 16 isolates represented distinct haplotypes (h = 16; Hd = 1.000), and nucleotide diversity was high ($\pi = 0.0775$; $\pi(JC) = 0.0818$). Tajima's D (−0.724), Fu and Li's D* (−0.205) and F* (−0.544), and Fu's Fs (+0.297) were all negative or near zero (Table 4). When the same analysis was extended to the full 75-sequence reference-inclusive alignment (2,408 gap-free sites), polymorphism was substantially higher (S = 1,809; Eta = 3,403), resolving 73 unique haplotypes among the 75 sequences (Hd = 0.999; two exact-duplicate pairs were identified: ChiLCV-AF274349/ToLCSV-NC004647 and ChiLCV-NC055131/ChiLCSLV-JN555601). Nucleotide diversity was markedly elevated relative to the 16-

isolate subset ($\pi = 0.1990$; $\pi(JC) = 0.2312$), reflecting the inclusion of multiple distinct begomovirus species. Tajima's D was positive (D = +1.032), and Fu and Li's D* is −0.570 and F* is −0.949. Fu's Fs was strongly negative (−2.268) (Table.6). Because this 75-sequence dataset pools multiple distinct begomovirus species rather than a single interbreeding population.

Table. 6. DNA polymorphism, haplotype diversity, and neutrality test statistics for the 16-isolate ChiLCV study population and the full 75-sequence reference-inclusive DNA-A dataset.

Statistic	N = 16 (ChiLCV isolates)	N = 75 (full reference-inclusive dataset)
Sites analysed (bp)	2,720	2,408
Segregating sites (S)	838	1,809
Total mutations (Eta)	942	3,403
Number of haplotypes (h)	16	73
Haplotype diversity (Hd)	1.000	0.999
Avg. pairwise differences (k)	210.858	479.065
Nucleotide diversity (π)	0.0775	0.1990
Nucleotide diversity $\pi(JC)$	0.0818	0.2312
Theta-W (per site)	0.0929	0.1537
Theta (per site, from Eta)	0.1044	0.2891
Tajima's D	−0.7242	+1.0322
Fu & Li's D*	−0.2049	−0.5699
Fu & Li's F*	−0.5435	−0.9485
Fu's Fs	+0.2970	−2.2680

3.5. Gene-wise selection pressure (dN/dS)

Pairwise dN/dS analysis of the six DNA-A-encoded genes (2,080–2,556 pairwise comparisons per gene, depending on the number of sequences retaining an unambiguous open reading frame) revealed markedly heterogeneous selective constraints (Table.7; Fig. 9). All genes except AC4 showed ω substantially below 1, consistent with predominant purifying selection. The coat protein gene (AV1) showed the strongest purifying selection observed ($\omega = 0.127$), followed by the replication-associated protein gene (AC1/Rep; $\omega = 0.166$), consistent with strong functional constraint on virion structure, whitefly transmissibility, and rolling-circle replication initiation, respectively. The precoat/movement protein gene (AV2) and replication enhancer gene (AC3) showed intermediate purifying constraint ($\omega = 0.249$ and 0.344), while the transcriptional activator gene (AC2) showed the weakest purifying signal among the negatively selected genes ($\omega = 0.668$). AC4 was the only gene for which ω exceeded 1 ($\omega = 1.618$), indicating an excess of nonsynonymous over synonymous substitutions consistent with diversifying selection and/or relaxed purifying constraint.

Table.7. Mean pairwise dN, dS, and ω (dN/dS) per DNA-A-encoded gene.

Gene	Protein	N	Pairwise comparisons	Mean dN	Mean dS	ω (dN/dS)	Selection regime
AC1	Rep	72	2,556	0.163	0.984	0.166	Strong purifying
AC2	TrAP	72	2,556	0.235	0.352	0.668	Moderate purifying
AC3	REn	69	2,346	0.182	0.528	0.344	Purifying
AC4	—	72	2,556	0.374	0.231	1.618	Diversifying/relaxed

AV1	CP	73	2,415	0.081	0.643	0.127	Strongest purifying
AV2	Precoat/MP	65	2,080	0.141	0.568	0.249	Purifying

4. Discussion

Chilli leaf curl disease (ChiLCD) continues to pose a severe threat to *Capsicum annuum* production across South Asia. The present study provides a high-resolution genomic snapshot of the monopartite begomovirus–satellite complexes responsible for ongoing epidemics in southern India during the 2023–2024 cropping season.

4.1. Disease incidence and symptomatology

Field surveys across 15 districts in Karnataka, Tamil Nadu, and Andhra Pradesh revealed ChiLCD incidence ranging from 30.5% to 76.5%, consistent with previous reports of widespread and often devastating outbreaks in the region [26,54,55]. Symptom expressions were uniform and indistinguishable across locations, reinforcing the diagnostic challenge posed by mixed infections of closely related begomoviruses [56,57].

4.2. Dominance of monopartite begomoviruses and complex satellite associations

All 50 symptomatic samples tested positive for begomovirus DNA-A and negative for DNA-B, confirming the exclusive involvement of monopartite begomoviruses [27,35]. Full-genome sequencing of 16 representative isolates revealed extensive satellite co-infection: 11 isolates (68.75%) were associated only with betasatellites, while two isolates (12.5%) harboured both betasatellites and alphasatellites. Such multi-component complexes dramatically enhance symptom severity and viral DNA accumulation through synergistic interactions [39, 58,59].

4.3. Taxonomic diversity and emergence of novel begomovirus species

SDT and phylogenetic analyses identified four distinct taxonomic groups among the 16 isolates. ChiLCV remained the predominant species (12 isolates, 96.6–99.5% nt identity), underscoring its entrenched role as the primary driver of ChiLCD in India and Pakistan [20,21,33]. Two isolates belonged to PepLCBV, extending the documented range of this species within southern India [24].

Most significantly, isolates CLC-32 and CLC-36 exhibited <91% genome-wide identity with any currently recognised begomovirus species and are formally proposed as two new species: **Chilli leaf curl Dindigul virus (ChiLCDinV; PQ141073) (*Begomovirus dindigulense*)** and **Chilli leaf curl Karur virus (ChiLCKarV; PQ141074) (*Begomovirus karurense*)** . Their deep phylogenetic divergence and basal position relative to the ChiLCV–PepLCBV clade suggest a long period of cryptic circulation followed by recent host adaptation, probably driven by the high recombination and genetic variability characteristic of begomoviruses infecting solanaceous crops in South Asia [60,61].

4.4. Recombination as the primary driver of begomovirus evolution

Recombination analysis uncovered 58 significant events concentrated in the Rep (AC1) and CP (AV1) coding regions, well-established recombination hotspots [60,62,63]. The strongest signals were associated with the two novel species and the PepLCBV isolates, with parental sequences drawn exclusively from South Asian begomoviruses. The notably high Tajima's D values found in the begomovirus populations show an excess of intermediate-frequency variations, suggesting a departure from neutral evolution. This pattern may arise from population bottlenecks, balancing selection, or population mixing. Along with the extensive reticulations and homoplasy seen in the phylogenetic network, these results indicate that recombination has significantly influenced the genetic structure and diversity of the virus populations [64,65,66].

4.5. Diversity and evolutionary constraints of associated betasatellites

Ten of the 13 betasatellites were variants of ToLCBB, consistent with its near-ubiquitous association with severe monopartite begomovirus diseases across the Indian subcontinent [39,67,68]. Recombination in betasatellites was limited and largely inconclusive, with only short, weakly supported segments detected within the functionally critical βC1 ORF [27,39]. This pattern reflects strong purifying selection and functional constraint on βC1 [20,60] the primary virulence determinant that suppresses transcriptional and post-transcriptional gene silencing [59,69].

4.6. Limited role and evolutionary stability of alphasatellites

Alphasatellites were detected in only two isolates and belonged to a single, previously uncharacterised species most closely related to PaLCA and related South Asian [58,67]. Their low prevalence, absence of robust recombination signals, and strongly positive Tajima's D (11.15) are consistent with a facultative, transmission-dependent lifestyle subject to frequent stochastic loss [70, 71]. Despite evidence demonstrating the ability of alphasatellites to decrease symptoms and limit helper virus accumulation under laboratory settings, their inconsistent occurrence in natural populations indicates that they may possess a selective advantage only in certain scenarios [73,74,75].

4.7. Population genetics and selection pressure corroborate recombination-driven diversification

The population genetic analyses of the chilli-infecting begomoviruses revealed striking contrasts between the 16 ChiLCV survey isolates and the broader 75-sequence reference-inclusive dataset. Within the study population, all isolates represented unique haplotypes (Hd = 1.000), with high nucleotide diversity ($\pi = 0.0775$) and negative or near-zero neutrality statistics (Tajima's D = −0.724; Fu's Fs = +0.297). These values are consistent with purifying selection on a genetically diverse but relatively cohesive population, where deleterious mutations are removed while maintaining haplotype richness. In contrast, the expanded dataset incorporating multiple begomovirus species showed

elevated polymorphism (S = 1,809; $\pi = 0.1990$) and a positive Tajima's D (+1.032), reflecting population subdivision and the pooling of divergent taxa rather than a single interbreeding population. The strongly negative Fu's Fs (−2.268) in this dataset further supports historical population expansion and recombination-driven diversification across begomovirus lineages.

Gene-wise dN/dS analysis further corroborated this interpretation. Most DNA-A encoded genes (AV1, AV2, AC1, AC2, AC3) were under strong to moderate purifying selection, consistent with functional constraints on coat protein-mediated transmission, replication initiation, and transcriptional regulation. However, AC4 exhibited a dN/dS ratio >1 ($\omega = 1.618$), indicating diversifying selection or relaxed constraint. This suggests AC4 may act as a hotspot for adaptive evolution, potentially linked to host defense suppression or pathogenicity modulation. The coexistence of strong purifying selection in structural and replication genes with diversifying selection in AC4 highlights a dual evolutionary strategy: conservation of essential functions alongside adaptive innovation in host–virus interactions.

These patterns align closely with the recombination signals detected in the dataset. Breakpoints were concentrated in the Rep (AC1) and CP (AV1) genes both regions under strong purifying selection, evidencing that recombination provides a mechanism to shuffle conserved modules while maintaining functional integrity. The strongest recombination signatures were observed in the two novel species (CLC-32, CLC-36) and PepLCBV isolates (CLC-40, CLC-42), underscoring recombination as a primary driver of taxonomic novelty. The population genetic evidence of high haplotype diversity, neutrality test deviations, and gene-specific selection pressures corroborates the recombination-driven diversification revealed by RDP5 analysis. Together, these findings demonstrate that begomovirus evolution in chilli is shaped by a dynamic interplay of purifying selection, localized diversifying pressure, and recurrent recombination events, generating both genetic stability and innovation within the ChiLCV–PepLCBV complex .

In summary, the ChiLCD pathosystem in southern India is characterised by a highly dynamic, recombination-driven begomovirus population dominated by ChiLCV, near-universal ToLCBB-like betasatellites, frequent emergence of novel viral and satellite species, and only sporadic alphasatellite involvement. This extraordinary evolutionary plasticity enables rapid adaptation to resistant cultivars and changing vector populations, explaining the persistent failure of single-gene resistance strategies [21,76,77]. The identification of two novel begomovirus species highlights the critical need for continuous genomic surveillance, deployment of durable broad-spectrum resistance by gene pyramiding and RNAi-based approaches, and coordinated regional vector management to protect chilli production in South Asia.

5. Conclusion

The results of this study demonstrate that chilli leaf curl disease (ChiLCD) across southern India is driven by a complex interplay of high genetic diversity, recombination, and heterogeneous selection pressures. Field surveys revealed widespread incidence with severe symptomatology, while molecular characterization confirmed the predominance of monopartite begomoviruses, often accompanied by betasatellites and occasionally alphasatellites. Genome-wide analyses identified both established ChiLCV lineages and divergent isolates, including two novel species (ChiLCDinV and ChiLCKarV), underscoring the ongoing emergence of new viral taxa. Population genetic analyses highlighted extensive haplotype diversity and strong purifying selection acting on structural and replication-associated genes, ensuring functional stability. In contrast, the AC4 gene exhibited diversifying selection, suggesting adaptive roles in host–virus interactions. Recombination signals concentrated in conserved genomic regions further corroborate that recombination is a major evolutionary force generating taxonomic novelty and shaping begomovirus diversity. Together, these findings establish that the evolutionary trajectory of chilli-infecting begomoviruses is defined by a balance between purifying selection, localized diversifying pressure, and recurrent recombination. This dynamic process not only maintains essential viral functions but also drives diversification, enabling begomoviruses to adapt to new hosts and environments. The identification of novel species and recombinant lineages emphasizes the need for continuous surveillance and molecular monitoring to mitigate the growing threat of ChiLCD to chilli cultivation in India.

Data availability

The genome sequences reported in this paper have been deposited in GenBank under accession numbers **PQ141071–PQ141086** (begomoviruses), **PQ057804 - PQ057816** (betasatellites), and **PQ037111, PQ057803** (alphasatellites). Supplementary materials are available online.

Funding Declaration

The work was supported by Department of Biotechnology (DBT), Government of India, under the DBT-Boost to University Interdisciplinary Life Science Departments for Education and Research (BUILDER) programme at University of Agricultural Sciences Dharwad, awarded to P.U. Krishnaraj (Principal Investigator).

Author contributions

S.E. conceived and designed the study, conducted the experiments, performed molecular characterization and bioinformatic analyses, and prepared the initial manuscript. S.J., K.S.S. and G.H. provided guidance and critical input during the research. D.N.K. provided guidance and supervision throughout the research. P.U.K. provided financial support for the study. P.A.R.U. and P.U. contributed to data analysis and interpretation. S.E. prepared the figures and tables and wrote the main manuscript text.

Competing interests

The authors declare no competing interests.

Supplementary Information

Supplementary Table S1. Locations of chilli leaf curl isolates collected from chilli and representative isolates used for full-length genome sequencing.

Supplementary Table S2. Leaf curl isolates collected from chilli used for full-length genome sequencing.

Supplementary Table S3. Recombination analysis of breakpoints and putative parental sequences of begomoviruses isolated from chilli (58 events).

Supplementary Table S4. VIRIDIC pairwise nucleotide analysis of the predicted complete sequences of begomoviruses isolated from chilli with other begomoviruses.

Supplementary Table S5. VIRIDIC pairwise nucleotide analysis of the predicted intergenic sequences of begomoviruses isolated from chilli with other begomoviruses.

Supplementary Table S6. VIRIDIC pairwise nucleotide analysis of the predicted AV2 gene product of begomoviruses isolated from chilli with other begomoviruses

Supplementary Table S7. VIRIDIC pairwise nucleotide analysis of the predicted AV1 gene product of begomoviruses isolated from chilli with other begomoviruses

Supplementary Table S8 VIRIDIC pairwise nucleotide analysis of the predicted AC1 gene product of begomoviruses isolated from chilli with other begomoviruses

Supplementary Table S9. VIRIDIC pairwise nucleotide analysis of the predicted AC2 gene product of begomoviruses isolated from chilli with other begomoviruses

Supplementary Table S10. VIRIDIC pairwise nucleotide analysis of the predicted AC3 gene product of begomoviruses isolated from chilli with other begomoviruses

Supplementary Table S11. VIRIDIC pairwise nucleotide analysis of the predicted AC4 gene product of begomoviruses isolated from chilli with other begomoviruses

Supplementary Table S12. VIRIDIC pairwise nucleotide analysis of isolated betasatellites and other betasatellites.

Supplementary Table S13. VIRIDIC pairwise nucleotide analysis of isolated alphasatellites and other alphasatellites.

Supplementary Figures S1–S6. (As provided in main manuscript: Figs 1–6).

References

[1]. Madhavi Reddy K, Ponnam N, Barik S, Rakesh Reddy V, Saha K, Lakshamana Reddy DC, Sujatha K. Chillies and capsicums in India. In: [Editors]. Handbook of spices in India: 75 years of research and development. Singapore: Springer Nature Singapore; 2023. p. 2033-2084. https://doi.org/10.1007/978-981-19-3728-6_28

[2]. Tripodi P, Kumar S. The Capsicum crop: an introduction. In: The capsicum genome. Cham: Springer International Publishing; 2019. p. 1-8. https://doi.org/10.1007/978-3-319-97217-6_1

[3].Pickersgill B. Chile peppers (Capsicum spp.). In: Ethnobotany of Mexico: interactions of people and plants in Mesoamerica. New York, NY: Springer New York; 2016. p. 417-437. https://doi.org/10.1007/978-1-4614-6669-7_17

[4]. Reddy UK, Almeida A, Abburi VL, Alaparthi SB, Unselt D, Hankins G, Park M, Choi D, Nimmakayala P. Identification of gene-specific polymorphisms and association with capsaicin pathway metabolites in Capsicum annuum L. collections. PLoS One. 2014;9(1):e86393. https://doi.org/10.1371/journal.pone.0086393

[5]. Erickson AN, Markhart AH. Flower developmental stage and organ sensitivity of bell pepper (Capsicum annuum L.) to elevated temperature. Plant Cell Environ. 2002;25(1):123-30. https://doi.org/10.1046/j.0016-8025.2001.00807.x

[6]. Oh SY, Koh SC. Fruit development and quality of hot pepper (*Capsicum annuum* L.) under various temperature regimes. Hortic Sci Technol. 2019;37(3):313-21. https://doi.org/10.7235/HORT.20190032

[7]. Bosland PW, Votava EJ, Votava EM. Peppers: vegetable and spice capsicums. 2nd ed. Wallingford: CABI; 2012.

[8]. Misra S, Lal RK, Darokar MP, Khanuja SPS. Genetic variability in germplasm accessions of *Capsicum annuum* L. Am J Plant Sci. 2011;2(5):629-35.

[9]. Kraft KH, Brown CH, Nabhan GP, Luedeling E, Luna Ruiz JD, Coppens d'Eeckenbrugge G, Hijmans RJ, Gepts P. Multiple lines of evidence for the origin of domesticated chili pepper, *Capsicum annuum*, in Mexico. Proc Natl Acad Sci U S A. 2014;111(17):6165-70. https://doi.org/10.1073/pnas.1308933111

[10]. Thakur H, Jindal SK, Sharma A, Dhaliwal MS. A monogenic dominant resistance for leaf curl virus disease in chilli pepper (Capsicum annuum L.). Crop Prot. 2019;116:115-20. https://doi.org/10.1016/j.cropro.2018.10.007

[11]. Shah NC. Origin, introduction, and cultural history of capsicum in India. Indian J Hist Sci. 2023;58(4):312-9. https://doi.org/10.1007/s43539-023-00104-1

[12]. Bansal R, Kumar M, Shaikh AS. Growth and export trends of major spices in India: An analytical study. Gujarat J Ext Educ. 2022;33(2):145-50.

[13]. Kiruthika N. Economics of Mundu Chilli Cultivation in Ramanathapuram District of Tamil Nadu, India. Asian J Agric Ext Econ Sociol. 2024;42(1):44-8. https://doi.org/10.9734/AJAEES/2024/v42i12345

[14]. Food and Agriculture Organization of the United Nations (FAO). FAOSTAT statistical database. Rome: FAO; 2024. Available from: https://www.fao.org/faostat/. Accessed 30 Aug 2026.

[15]. Spices Board. Indian spice statistics. Kochi: Ministry of Commerce and Industry, Government of India; 2025. Available from: https://www.indianspices.com. Accessed 30 Aug 2026.

[16]. Pandit MK, Pandit R, Bairagi S. Chili. In: Ethnopharmacological investigation of Indian spices. Hershey, PA: IGI Global; 2020. p. 253-68.

[17]. He W, Liang L, Zhang Y. Pungency perception and the interaction with basic taste sensations: an overview. Foods. 2023;12(12):2317. https://doi.org/10.3390/foods12122317

[18]. Mohd Hassan N, Yusof NA, Yahaya AF, Mohd Rozali NN, Othman R. Carotenoids of capsicum fruits: pigment profile and health-promoting functional attributes. Antioxidants. 2019;8(10):469. https://doi.org/10.3390/antiox8100469

[19]. Pérez-Gálvez A, Jarén-Galán M, Isabel M. Processing of red pepper fruits (*Capsicum annuum* L.) for production of paprika and paprika oleoresin. In: Handbook of fruits and fruit processing. 2006. p. 565. https://doi.org/10.1002/9780470277737

[20]. Nalla MK, Schafleitner R, Pappu HR, Barchenger DW. Current status, breeding strategies and future prospects for managing chilli leaf curl virus disease and associated begomoviruses in Chilli (Capsicum spp.). Front Plant Sci. 2023;14: 1223982. https://doi.org/10.3389/fpls.2023.1223982

[21].Jayanthi P, Geetanjali AS. Molecular characterization to study the genetic diversity of begomoviruses occurring in the major chilli growing areas of Tamil Nadu state of India. Int Microbiol. 2025;28(4):863-76. https://doi.org/10.1007/s10123-024-00580-0

[22]. Netla VR, Hiremath S, Muttappagol M, Vinay Kumar HD, Koti Prasanna S, Kumar TM, Basha CJ, Venkataravanappa V, Shankarappa KS, Maruthi MN, Lakshminarayana Reddy CN. Virome analysis of field-collected chilli samples reveals diverse viruses. Virol J. 2025;22(1):116. https://doi.org/10.1186/s12985-025-02713-x

[23]. Kenyon L, Kumar S, Tsai WS, Hughes JD. Virus diseases of peppers (Capsicum spp.) and their control. In: Advances in virus research. Vol. 90. Amsterdam: Academic Press; 2014. p. 297-354. https://doi.org/10.1016/B978-0-12-801246-8.00006-8

[24]. Krishnan N, Kumari S, Kumar RV, Sakthivel K, Dubey V, Reddy R, Singh AK, Kumar R. Prevalence of chilli leaf curl virus and tomato leaf curl New Delhi virus with chilli leaf curl disease in India. Physiol Mol Biol Plants. 2025;31(8):1363-74. https://doi.org/10.1007/s12298-025-01570-w

[25]. Vasudeva RS. Report of the Division of Mycology and Plant Pathology. Scientific Report of the Agricultural Research Institute, New Delhi. 1954-1955. p. 87-101.

[26]. Thakur H, Jindal SK, Sharma A, Dhaliwal MS. Chilli leaf curl virus disease: a serious threat for chilli cultivation. J Plant Dis Prot. 2018;125(3):239-49. https://doi.org/10.1007/s41348-018-0146-8

[27]. Fiallo-Olivé E, Lett JM, Martin DP, Roumagnac P, Varsani A, Zerbini FM, Navas-Castillo J, ICTV Report Consortium. ICTV virus taxonomy profile: Geminiviridae 2021. J Gen Virol. 2021;102(12):001696. https://doi.org/10.1099/jgv.0.001696

[28]. International Committee on Taxonomy of Viruses (ICTV). Virus taxonomy: 2024 release. Available from: https://ictv.global/taxonomy. Accessed 28 Aug 2026.

[29]. Kumar S, Gupta N, Chakraborty S. Geminiviral betasatellites: critical viral ammunition to conquer plant immunity. Arch Virol. 2023;168(7):196. https://doi.org/10.1007/s00705-023-05776-9

[30]. Zhao SX, Liu YM, Wang SD, Wang XW, Liu SS, Liu YQ, Pan LL. The βC1 protein encoded in betasatellites promotes begomovirus-whitefly coexistence by subverting vector infestation-induced plant antiviral defenses. PLoS Pathog. 2026;22(1):e1013907. https://doi.org/10.1371/journal.ppat.1013907

[31]. Zhou X. Advances in understanding begomovirus satellites. Annu Rev Phytopathol. 2013;51(1):357-81. https://doi.org/10.1146/annurev-phyto-082712-102234

[32]. Nawaz-ul-Rehman MS, Nahid N, Hassan M, Mubin M. Betasatellites and deltasatellites (Tolecusatellitidae). In: Encyclopedia of virology. 3rd ed. 2021. p. 239-46. https://doi.org/10.1016/B978-0-12-809633-8.21562-8

[33]. Shingote PR, Wasule DL, Parma VS, Holkar SK, Karkute SG, Parlawar ND, Senanayake DM. An overview of chili leaf curl disease: molecular mechanisms, impact, challenges, and disease management strategies in Indian subcontinent. Front Microbiol. 2022;13:899512. https://doi.org/10.3389/fmicb.2022.899512

[34]. Devaraj, Ashwathappa KV, Madhu GS, Venkataravanappa V, Manthesh V, Reddy MK, Lakshminarayana Reddy CN. Cross species transmission and recombination dynamics of chilli leaf curl virus: implications for host range expansion and LAMP-based detection. 3 Biotech. 2025;15(8):260. https://doi.org/10.1016/j.tim.2021.05.007

[35]. Naveen M, Shanmugam PS, Srinivasan R, Shanthi M, Murugan M, Sivagnanapazham K, Angappan K, Oliva R, Karthikeyan G, Boopathi NM. Recombination, phylogenetic insights and co-existence dynamics of begomoviruses and satellite DNAs: implications for tomato viral disease management. Molecular Biology Reports. 2025;52(1):390. https://doi.org/10.1007/s11033-025-10499-y

[36]. Lodhi MA, Ye GN, Weeden NF, Reisch BI. A simple and efficient method for DNA extraction from grapevine cultivars and Vitis species. Plant Mol Biol Report. 1994;12(1):6-13. . https://doi.org/10.1007/BF02668658

[37]. Rojas MR, Gilbertson RL, Russell DR, Maxwell DP. Use of degenerate primers in the polymerase chain reaction to detect whitefly-transmitted geminiviruses. Plant Dis. 1993;77(4):340-7. https://doi.org/10.1094/PD-77-0340

[38]. Venkataravanappa V, Lakshminarayana Reddy CN, Jalali S, Krishna Reddy M. Molecular characterization of distinct bipartite begomovirus infecting bhendi (Abelmoschus esculentus L.) in India. Virus Genes. 2012;44(3):522-35. https://doi.org/10.1007/s11262-012-0732-y

[39]. Briddon RW, Bull SE, Mansoor S, Amin I, Markham PG. Universal primers for the PCR-mediated amplification of DNA β: a molecule associated with some monopartite begomoviruses. Mol Biotechnol. 2002;20(3):315-8. https://doi.org/10.1385/MB:20:3:315

[40]. Bull SE, Briddon RW, Markham PG. Universal primers for the PCR-mediated amplification of DNA 1: a satellite-like molecule associated with begomovirus-DNA β complexes. Mol Biotechnol. 2003;23(1):83-6. https://doi.org/10.1385/MB:23:1:83

[41]. Sambrook J, Russell DW. Molecular cloning: a laboratory manual. 3rd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 2001.

[42]. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In: Nucleic acids symposium series. Vol. 41. 1999. p. 95-8.

[43]. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30(4):772-80. https://doi.org/10.1093/molbev/mst010

[45]. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol. 2020;37(5):1530-4. https://doi.org/10.1093/molbev/msaa015

[45]. Price MN, Dehal PS, Arkin AP. FastTree 2: approximately maximum-likelihood trees for large alignments. PLoS One. 2010;5(3):e9490. https://doi.org/10.1371/journal.pone.0009490

[46]. Martin DP, Varsani A, Roumagnac P, Botha G, Maslamoney S, Schwab T, Kelz Z, Kumar V, Murrell B. RDP5: a computer program for analyzing recombination in, and removing signals of recombination from, nucleotide sequence datasets. Virus Evol. 2021;7(1). https://doi.org/10.1093/ve/veaa087

[47]. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 1989;123(3):585-95. https://doi.org/10.1093/genetics/123.3.585

[48]. Fu YX, Li WH. Statistical tests of neutrality of mutations. Genetics. 1993;133(3):693-709. https://doi.org/10.1093/genetics/133.3.693

[49]. Fu YX. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics. 1997;147(2):915-25. https://doi.org/10.1093/genetics/147.2.915

[50]. De Mita S, Siol M. EggLib: processing, analysis and simulation tools for population genetics and genomics. BMC Genet. 2012;13(1):27. https://doi.org/10.1186/1471-2156-13-27

[51]. Suyama M, Torrents D, Bork P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. 2006;34(suppl_2):W609-12 https://doi.org/10.1093/nar/gkl315

[52]. Nei M, Gojobori T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol. 1986;3(5):418-26. https://doi.org/10.1093/oxfordjournals.molbev.a040410

[53]. Cock PJ, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, Friedberg I, Hamelryck T, Kauff F, Wilczynski B, De Hoon MJ. Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics. 2009;25(11):1422. https://doi.org/10.1093/bioinformatics/btp163

[54].Chaithanya BH, Srinivas T, Reddy BB. Survey and molecular identification of begomovirus associated with chilli leaf curl disease in Kurnool district of Andhra Pradesh. Indian J Agric Res. 2023;57(6):836-40. https://doi.org/10.18805/IJARe.A-6028

[55]. Yadav RK, Reddy KM, Ashwathappa KV, Kumar M, Naresh P, Reddy MK. Screening of Capsicum germplasm and inheritance of resistance to chilli leaf curl virus. Indian Phytopathol. 2022;75(4):1129-36. https://doi.org/10.1007/s42360-022-00530-8

[56]. Senanayake DM, Varma A, Mandal B. Virus–vector relationships, host range, detection and sequence comparison of Chilli leaf curl virus associated with an epidemic of leaf curl disease of chilli in Jodhpur, India. Journal of Phytopathology. 2012;160(3):146-55. https://doi.org/10.1111/j.1439-0434.2011.01876.x

[57]. Kumar RA, Kumar VI, Kadiri S, Palicherla SR. Epidemiology and diagnosis of chilli leaf curl virus in central India, a major chilli growing region. Indian Phytopathol. 2016a;69(4s):61-4.

[58]. Mishra M, Verma RK, Marwal A, Sharma P, Gaur RK. Biology and interaction of the natural occurrence of distinct monopartite begomoviruses associated with satellites in Capsicum annuum from India. Front Microbiol. 2020; 11:512957. https://doi.org/10.3389/fmicb.2020.512957

[59].Saeed A, Arif M, Rafiq M, Song C, Albaqami M, Abdelbacki AM. Molecular characterization, genetic divergence, expression of encapsidiation and synergism by a bipartite begomovirus; Tomato leaf curl Palampur virus (ToLCPMV) infecting bitter gourd (Momordica charantia). Microb Pathog. 2024;196:106953. https://doi.org/10.1016/j.micpath.2024.106953

[60]. Mishra M, Verma RK, Pandey V, Srivastava A, Sharma P, Gaur R, Ali A. Role of diversity and recombination in the emergence of chilli leaf curl virus. Pathogens. 2022;11(5):529. https://doi.org/10.3390/pathogens11050529

[61]. Devendran R, Kumar M, Ghosh D, Yogindran S, Karim MJ, Chakraborty S. Capsicum-infecting begomoviruses as global pathogens: host-virus interplay, pathogenesis, and management. Trends Microbiol. 2022;30(2):170-84. https://doi.org/10.1016/j.tim.2021.05.007

[62]. Padidam M, Sawyer S, Fauquet CM. Possible emergence of new geminiviruses by frequent recombination. Virology. 1999;265(2):218-25. https://doi.org/10.1006/viro.1999.0056

[63]. Lefeuvre P, Lett JM, Varsani A, Martin DP. Widely conserved recombination patterns among single-stranded DNA viruses. J Virol. 2009;83(6):2697-707. https://doi.org/10.1128/jvi.02152-08

[64]. Duffy S, Holmes EC. Validation of high rates of nucleotide substitution in geminiviruses: phylogenetic evidence from East African cassava mosaic viruses. J Gen Virol. 2009;90(6):1539-47. https://doi.org/10.1099/vir.0.009266-0

[65]. Lima AT, Silva JC, Silva FN, Castillo-Urquiza GP, Silva FF, Seah YM, Mizubuti ES, Duffy S, Zerbini FM. The diversification of begomovirus populations is predominantly driven by mutational dynamics. Virus Evol. 2017;3(1):vex005. https://doi.org/10.1093/ve/vex005

[66]. Sahu AK, Pandey V, Srivastava A, Verma RK, Mishra NS, Akhtar A, Gaur RK. Molecular characterization of a bipartite begomovirus associated with leaf crinkle and curling symptoms of Duranta erecta in India. Indian Phytopathol. 2024;77(1):191-201. https://doi.org/10.1007/s42360-024-00718-0

[67]. Kumar S, Srivastava A, Jaidi M, Chauhan PS, Raj SK. Molecular characterization of a begomovirus, α-satellite, and β-satellite associated with leaf curl disease of Parthenium hysterophorus in India. Plant Dis. 2016b;100(11):2299-305.

[68]. Jeyaraj G. Infectivity analysis of chilli leaf curl Ahmedabad virus and its associated tomato leaf curl Bangladesh betasatellite complex in chili. Physiol Mol Plant Pathol. 2023;128:102167. https://doi.org/10.1016/j.pmpp.2023.102167

[69]. Cui X, Li G, Wang D, Hu D, Zhou X. A begomovirus DNAβ-encoded protein binds DNA, functions as a suppressor of RNA silencing, and targets the cell nucleus. J Virol. 2005;79(16):10764-75. https://doi.org/10.1128/jvi.79.16.10764-10775.2005

[70]. Bornancini VA, Irazoqui JM, Flores CR, Vaghi Medina CG, Amadio AF, López Lambertini PM. Reconstruction and characterization of full-length begomovirus and alphasatellite genomes infecting pepper through metagenomics. Viruses. 2020;12(2):202. https://doi.org/10.3390/v12020202

[71]. Idris AM, Shahid MS, Briddon RW, Khan AJ, Zhu JK, Brown JK. An unusual alphasatellite associated with monopartite begomoviruses attenuates symptoms and reduces betasatellite accumulation. J Gen Virol. 2011;92(3):706-17. https://doi.org/10.1099/vir.0.025288-0

[72]. Abbas Q, Amin I, Mansoor S, Shafiq M, Wassenegger M, Briddon RW. The Rep proteins encoded by alphasatellites restore expression of a transcriptionally silenced green fluorescent protein transgene in Nicotiana benthamiana. VirusDis. 2019;30(1):101-5. . https://doi.org/10.1007/s13337-017-0413-5

[73]. Hassan I, Orílio AF, Fiallo-Olivé E, Briddon RW, Navas-Castillo J. Infectivity, effects on helper viruses and whitefly transmission of the deltasatellites associated with sweepoviruses (genus Begomovirus, family Geminiviridae). Sci Rep. 2016;6(1):30204. https://doi.org/10.1038/srep30204

[74]. Abbas Q, Amin I, Mansoor S, Shafiq M, Wassenegger M, Briddon RW. The Rep proteins encoded by alphasatellites restore expression of a transcriptionally silenced green fluorescent protein transgene in Nicotiana benthamiana. VirusDisease. 2019.30(1):101-5. https://doi.org/10.1007/s13337-017-0413-5

[75]. Fondong VN. DNA satellites impact begomovirus diseases in a virus-specific manner. Int J Mol Sci. 2025;26(12):5814. https://doi.org/10.3390/ijms26125814

[76]. Varma A, Singh MK. The role of satellites in the evolution of begomoviruses. Viruses. 2024;16(6):970. https://doi.org/10.3390/v16060970

[77]. Jain H, Singh I, Chahal S, Kaur R, Siwach P. Phylogenetic and recombination analysis of begomoviruses associated with cotton leaf curl disease and in silico analysis of viral-host protein interactions. Microb Pathog. 2024;186:106504. https://doi.org/10.1016/j.micpath.2023.106504



Figure 1. Different symptoms of leaf curl disease in chilli. a) Shortening of internodes along with petioles b) Upward leaf curling with yellow leaf margin c) leaf puckering d) upward leaf curling e) vein banding f) Bunching of leaves along with stunting of plants

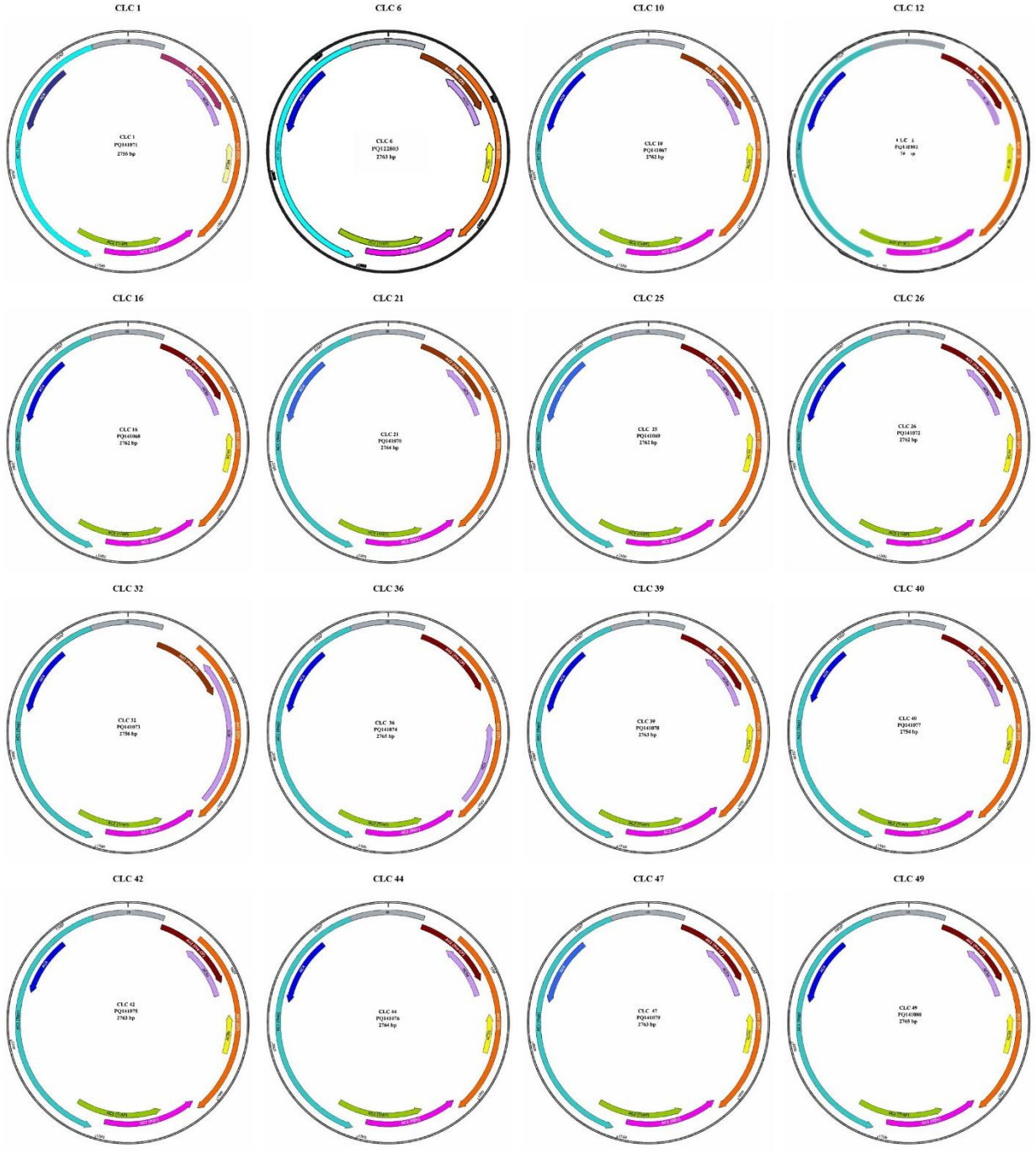


Figure 2. Genome organization of DNA-A component of isolated begomoviruses from chilli

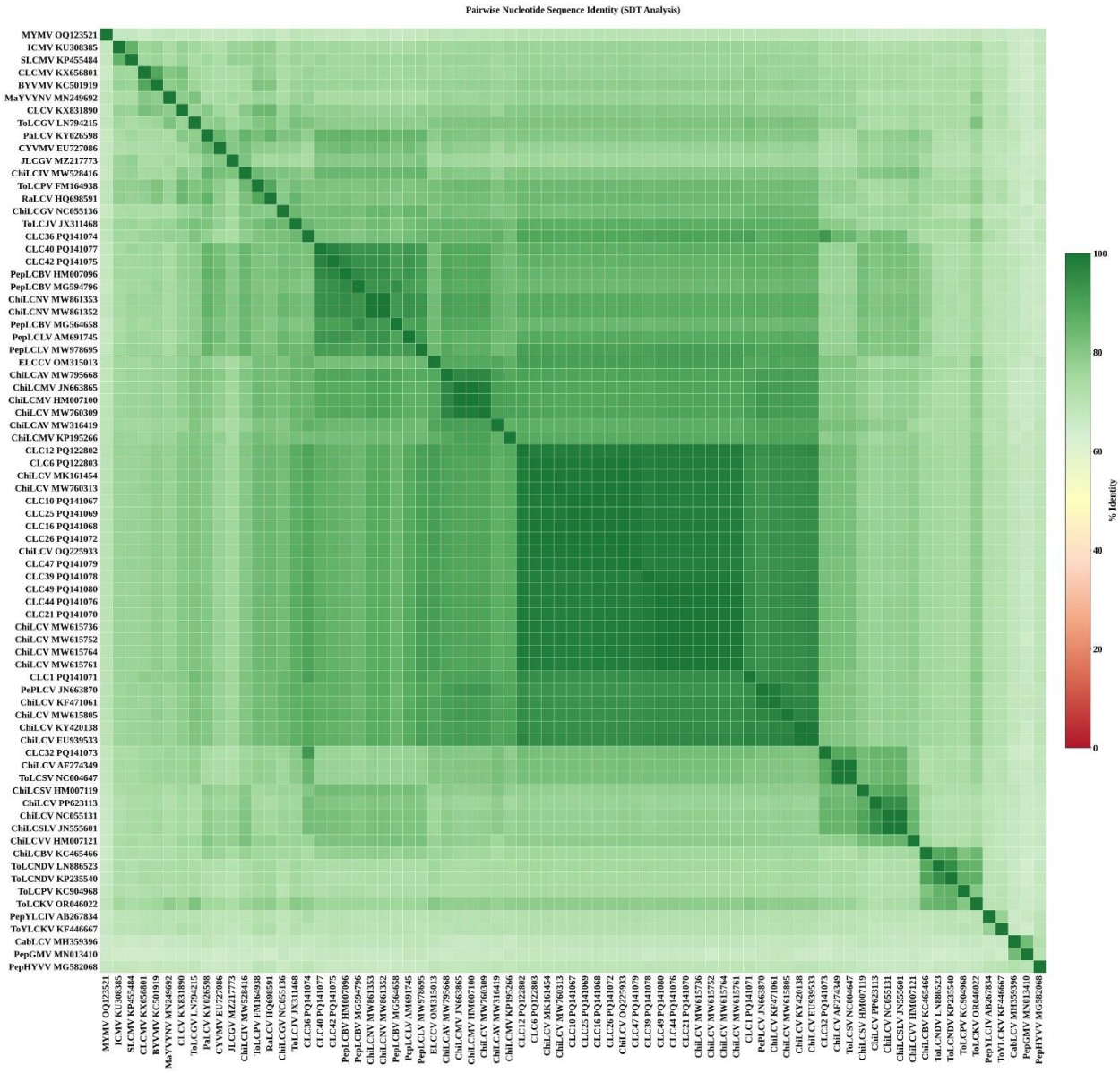


Figure 3. SDT pairwise nucleotide identity analysis of isolated begomoviruses and other begomoviruses from NCBI database

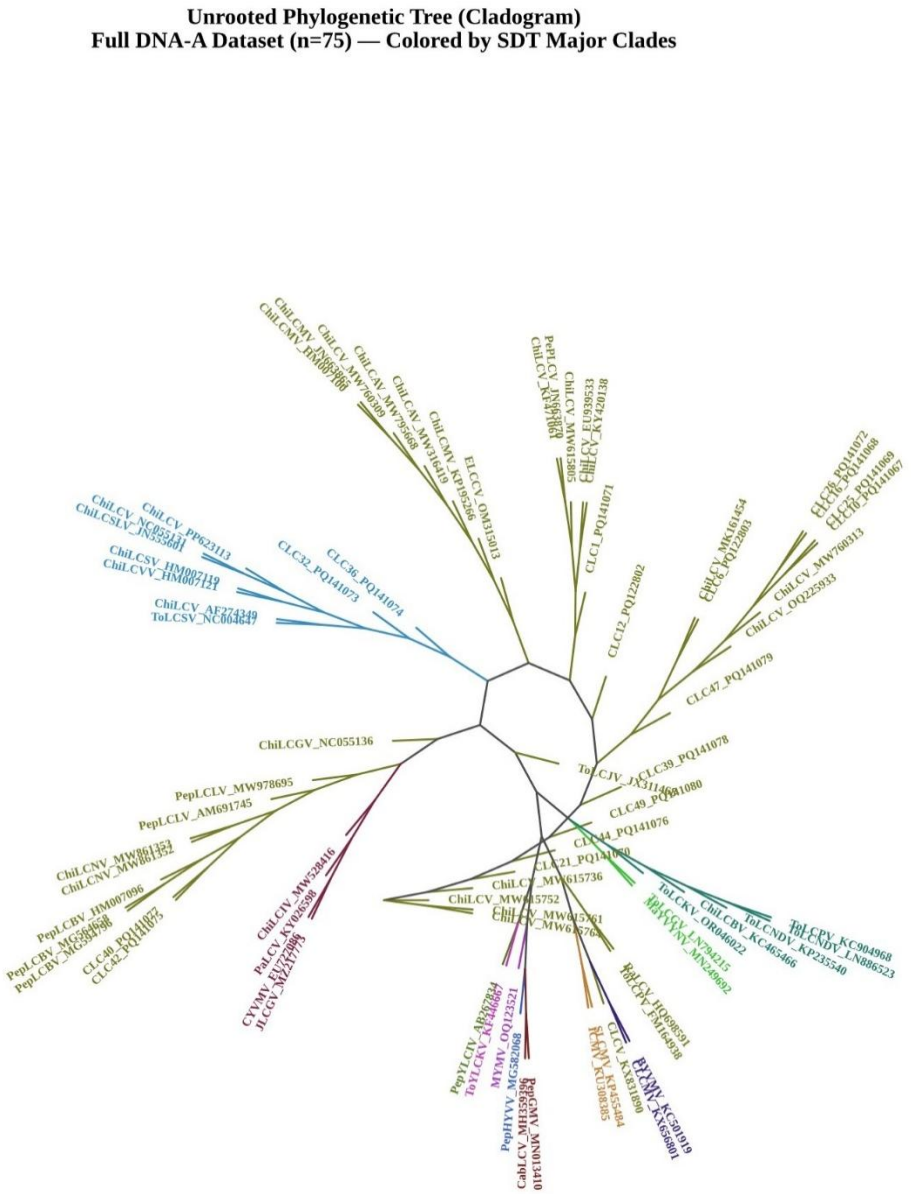


Figure 4. Phylogenetic analysis of isolated begomoviruses and other begomoviruses from NCBI database

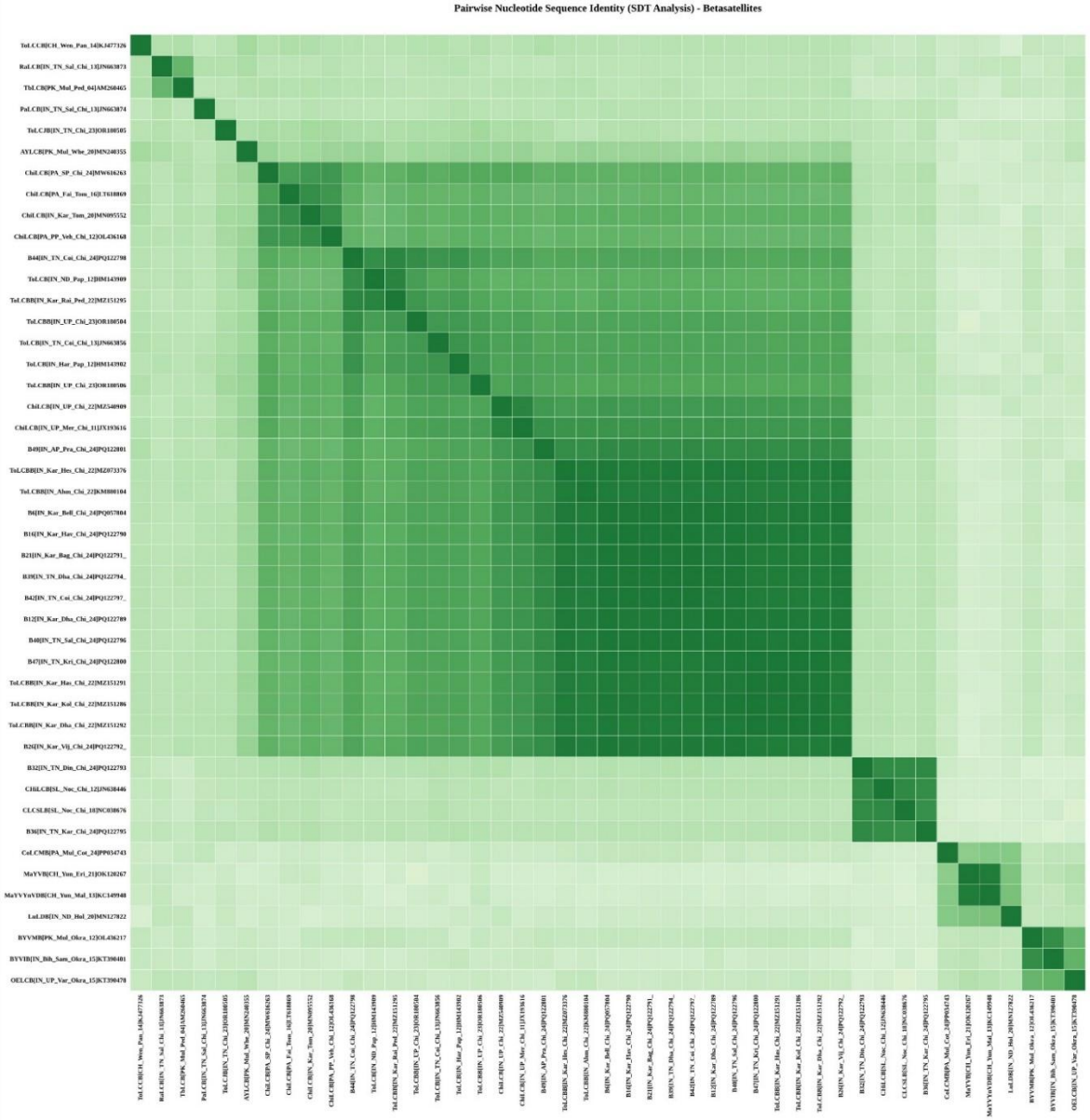

Pairwise Nucleotide Sequence Identity (SDT Analysis) - Betasatellites
ToLCCB[CH_Wen_Pan_14]KJ477126
RaLCB[IN_TN_Sal_Chi_13]JN663873
TbLCB[PK_Mul_Pod_04]AM260465
PaLCB[IN_TN_Sal_Chi_13]JN663874
TeLCB[IN_TN_Chi_23]OR180505
AYLCB[PK_Mul_Whe_20]MN240355
ChiLCB[PA_SP_Chi_24]MW616763
ChiLCB[PA_Fai_Tom_16]LT618869
ChiLCB[IN_Kar_Tom_20]MN095552
ChiLCB[PA_PP_Veh_Chi_12]OL436168
B44[IN_TN_Cui_Chi_24]PQ122798
TeLCB[IN_ND_Pop_12]HM143909
TeLCBB[IN_Kar_Rai_Pod_22]MZ151295
TeLCBB[IN_UP_Chi_23]OR180504
TeLCB[IN_TN_Cui_Chi_13]JN663856
TeLCB[IN_Har_Pop_12]HM143902
TeLCBB[IN_UP_Chi_23]OR180506
ChiLCB[IN_UP_Chi_22]MZ540909
ChiLCB[IN_UP_Mer_Chi_11]JX193616
B49[IN_AP_Pra_Chi_24]PQ122801
TeLCBB[IN_Kar_Hos_Chi_22]MZ073376
TeLCBB[IN_Abus_Chi_22]KM880104
B4[IN_Kar_Bell_Chi_24]PQ057804
B16[IN_Kar_Hav_Chi_24]PQ122790
B21[IN_Kar_Bag_Chi_24]PQ122791_
B39[IN_TN_Dha_Chi_24]PQ122794_
B4[IN_TN_Cui_Chi_24]PQ122797_
B12[IN_Kar_Dha_Chi_24]PQ122789
B40[IN_TN_Sal_Chi_24]PQ122796
B47[IN_TN_Kri_Chi_24]PQ122800
TeLCBB[IN_Kar_Hav_Chi_22]MZ151291
TeLCBB[IN_Kar_Kol_Chi_22]MZ151286
TeLCBB[IN_Kar_Dha_Chi_22]MZ151292
B26[IN_Kar_Vij_Chi_24]PQ122792_
B32[IN_TN_Din_Chi_24]PQ122793
ChiLCB[SL_Noc_Chi_12]JN638446
CLCSLB[SL_Noc_Chi_18]NC038676
B36[IN_TN_Kar_Chi_24]PQ122795
CoLCMB[PA_Mul_Cot_24]PP034743
MaYVB[CH_Yun_Fri_21]OK128267
MaYVYoVDB[CH_Yun_Mal_13]KC149948
LuLDB[IN_ND_Hol_20]MN127822
BYVMB[PK_Mol_Okra_12]OL436217
BYVIB[IN_Bih_Sam_Okra_15]KT390401
OELCB[IN_UP_Var_Okra_15]KT390478
% Identity
100
80
60
40
20
0

Figure 5. SDT pairwise nucleotide identity analysis of isolated betasatellites and other betasatellites from NCBI database

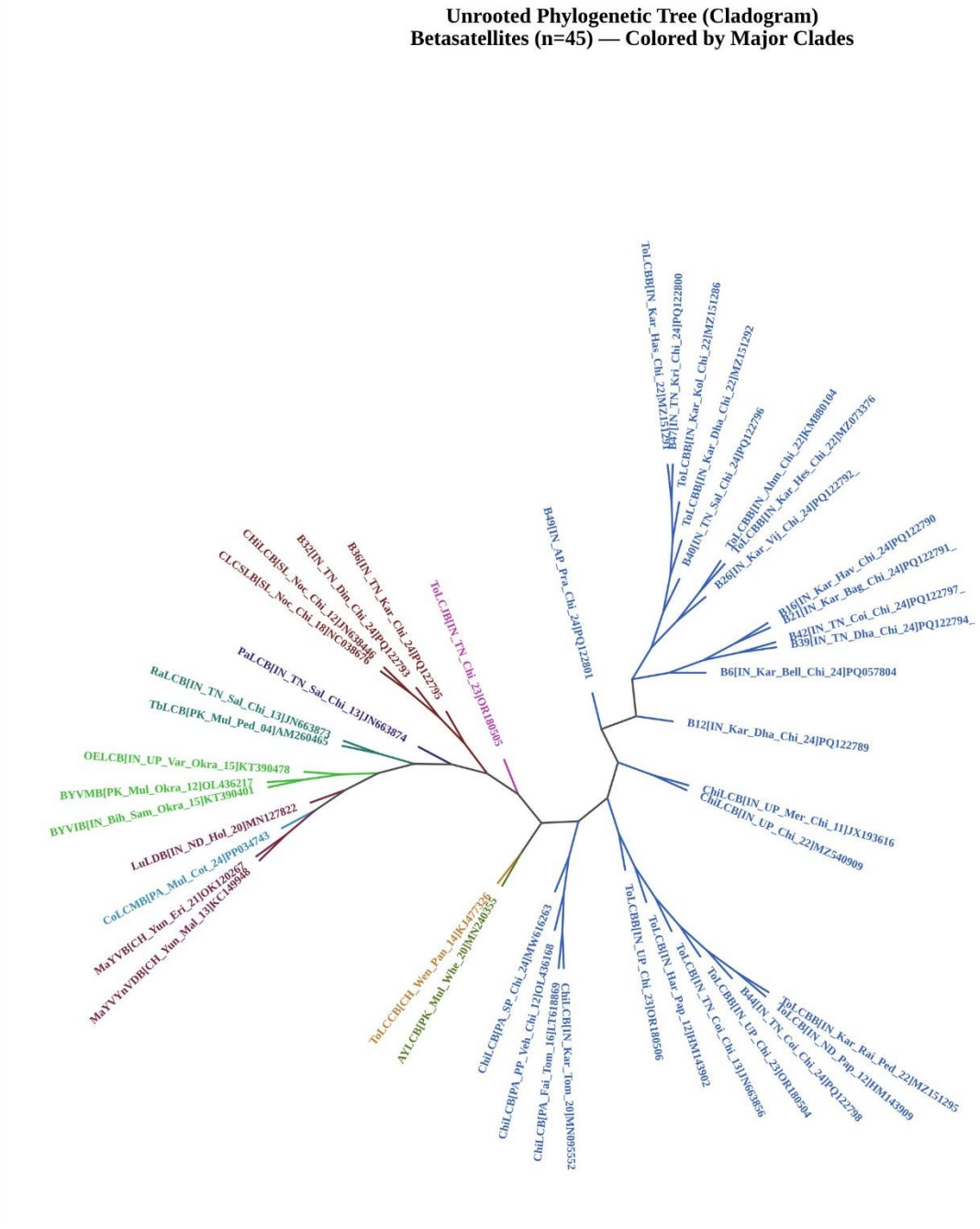


Figure 6. Phylogenetic analysis of isolated Betasatellites and other betasatellites from NCBI database

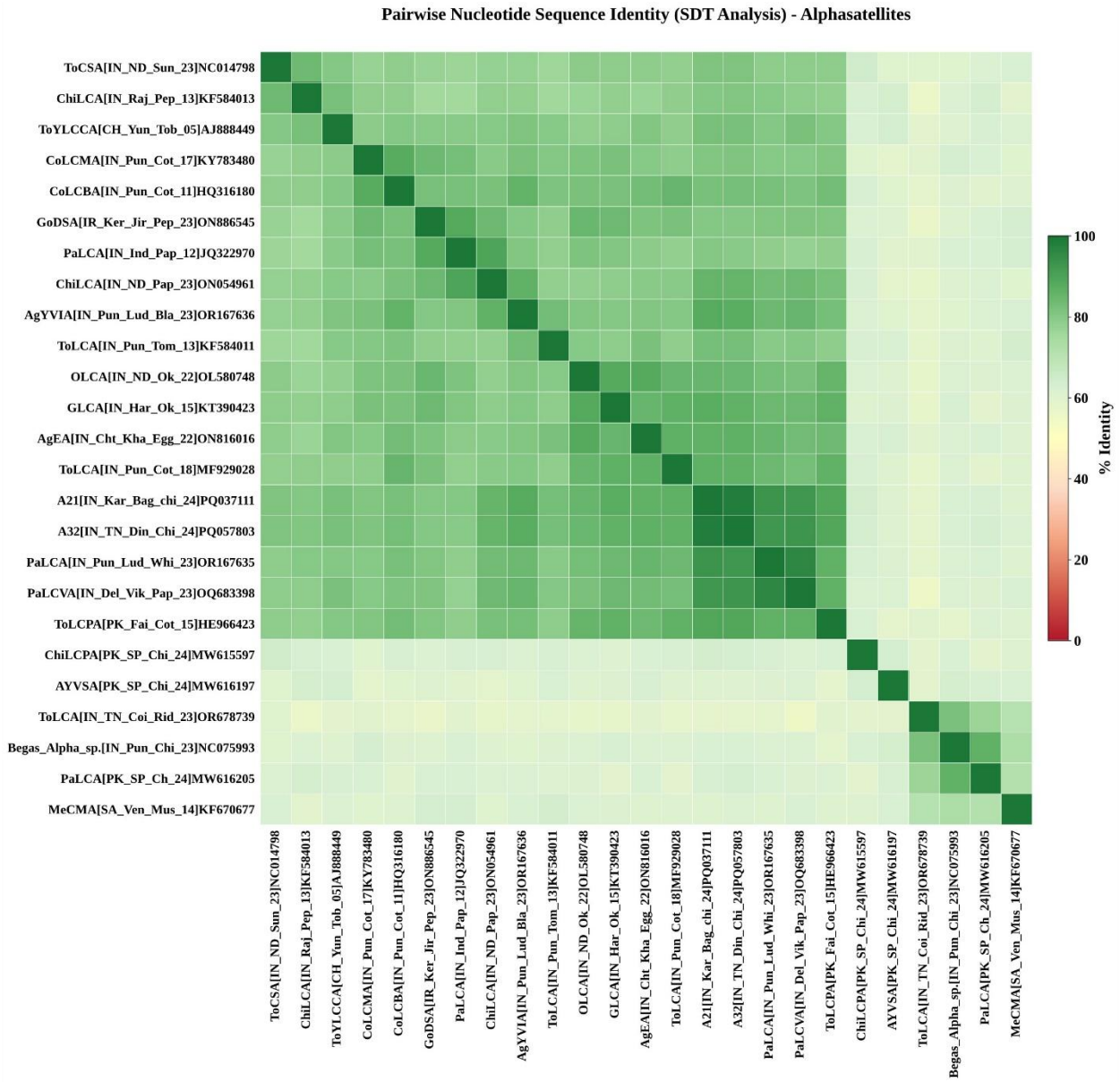


Figure 7. SDT pairwise nucleotide identity analysis of isolated alphasatellites and other alphasatellites from NCBI database

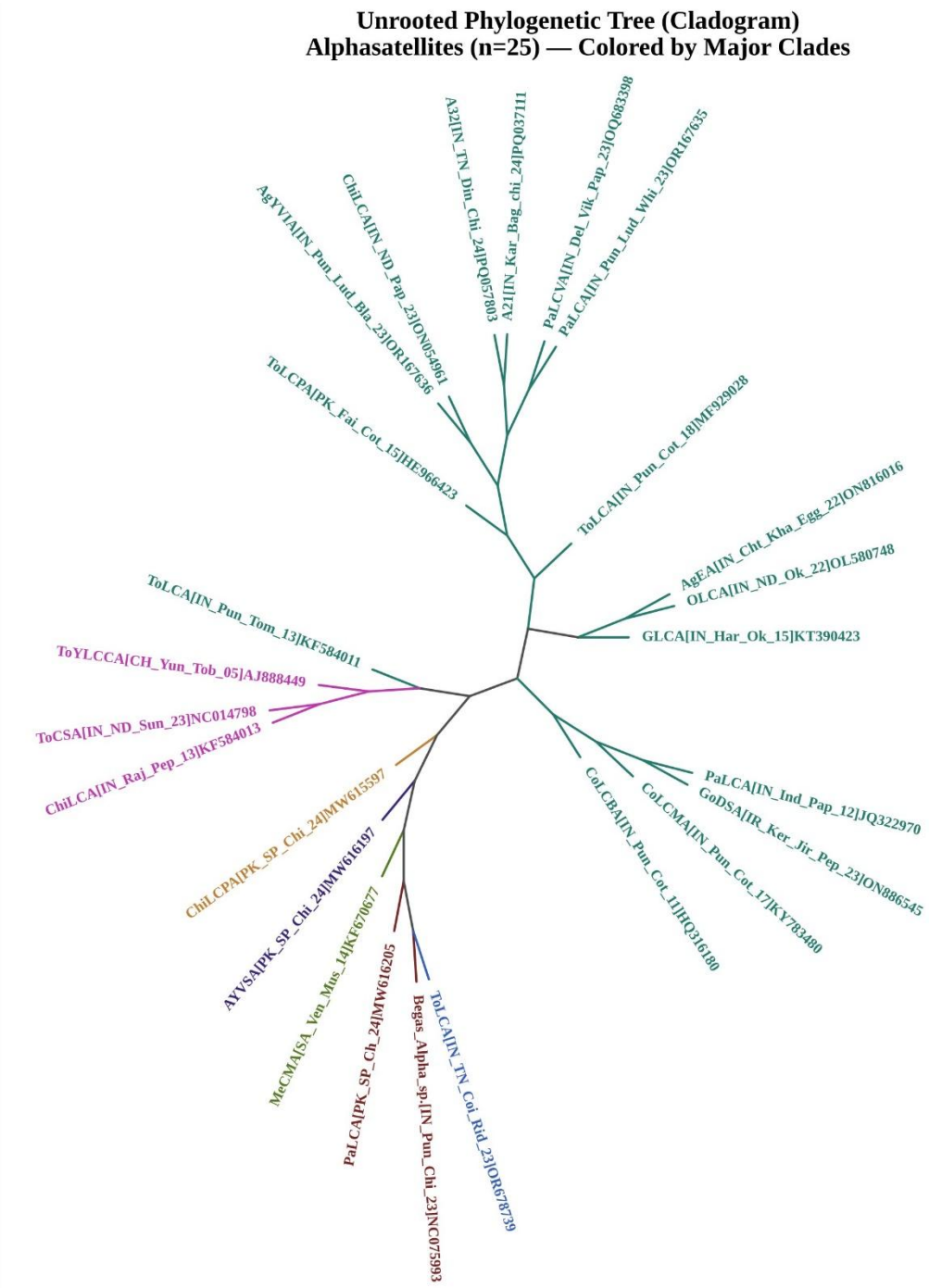


Figure 8. Phylogenetic analysis of isolated alphasatellites and other alphasatellites from NCBI database

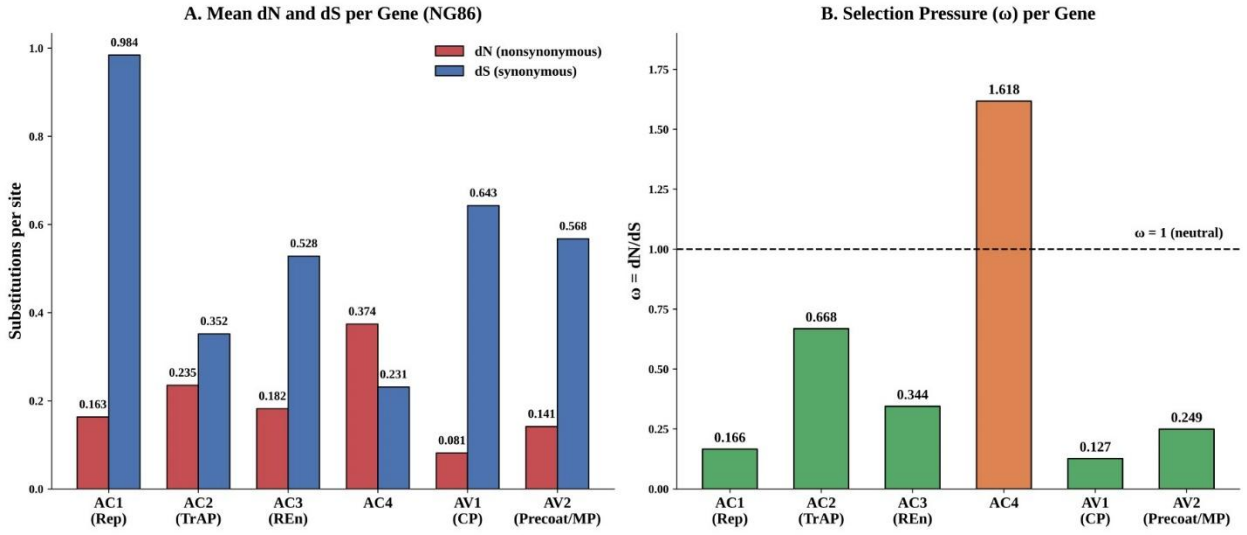


Figure. 9. Mean pairwise dN, dS, and ω (dN/dS) per DNA-A-encoded gene, showing strong purifying selection on AV1 (CP) and AC1 (Rep), and diversifying/relaxed selection on AC4.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarytables.docx