Virology Journal

RESEARCH

Open Access

Virome analysis of field-collected chilli samples reveals diverse viruses

Vamsidhar Reddy Netla[1†], Shridhar Hiremath[1,7†], Mantesh Muttappagol[1,7], H. D. Vinay Kumar[1], S. Koti Prasanna[2], T. L. Mohan Kumar[3], C. R. Jahir Basha[1], V. Venkataravanappa[4], K. S. Shankarappa[5], M. N. Maruthi[6] and C. N. Lakshminarayana Reddy[1*]

Abstract

Background Chilli (*Capsicum annuum* L.), an important spice crop, is susceptible to diverse viral infections. Traditional detection methods including PCR and its variants had difficulty in identifying the complete spectrum of viruses, especially in mixed infections. High-throughput sequencing (HTS) has emerged as a successful tool for comprehensive virome analyses, enabling the identification of the known and novel viruses in the infected samples. Using HTS, we investigated virome analyses to identify known and novel viruses in chilli.

Methods In 2021–22, 19 leaf samples were collected from chili plants in farmer fields in Karnataka, India, showing symptoms such as leaf curling, vein banding, mosaic, mottling, filiform, leathery, dull-colored, and bunchy leaves. Total RNA was extracted, pooled at equimolar concentrations, and subjected to virome profiling. rRNA-depleted RNA was used to prepare mRNA and sRNA libraries, which were sequenced on the Illumina NovaSeq 6000 platform. Bioinformatics tools were used to analyze the sequencing data and identify plant viruses.

Results Viral disease incidences varied from 26.6 to 47.5% in the farmer fields surveyed. Virome analyses revealed complete/ near-complete genomes of six different viruses: chilli leaf curl virus (ChiLCV), cucumber mosaic virus (CMV), groundnut bud necrosis orthotospovirus (GBNV), pepper cryptic virus-2 (PCV-2), pepper vein yellows virus (PeVYV) and bell pepper alphaendornavirus (BPEV). The viral copy number of ChiLCV was found to be the highest (45.36%) and had the least mutational frequency (SNPs) and was also associated with five satellites. Recombination breakpoints were observed in ChiLCV (coat protein and AC4 regions), CMV RNA2 (2a protein) and PeVYV (P0, P3 and P5 proteins), indicating their origins from intra- and interspecific recombination events. Identified viruses in the pooled RNA sample were confirmed by PCR. Further, novel loop-mediated isothermal amplification (LAMP) diagnostic assays were developed for diagnosing the identified viruses for future use. Among the six viruses identified in chilli, PeVYV and BPEV are the first reports from India.

[†]Vamsidhar Reddy Netla and Shridhar Hiremath contributed equally to this work and share first authorship.

*Correspondence:
C. N. Lakshminarayana Reddy
cnl.reddy@uasbangalore.edu.in; cnlreddy@gmail.com

Full list of author information is available at the end of the article

Conclusions This study presents the first virome profiling of chili using HTS and identified known and previously unreported viruses in farmer fields of Karnataka, India. Understanding viral diversity provides insights for developing diagnostic tools and effective management strategies.

Keywords Chilli viruses, High-throughput sequencing (HTS), Virome analysis, PeVYV, BPEV, LAMP

Introduction

Chilli (*Capsicum annuum* L.) is a widely cultivated ethnopharmacological spice crop, thriving across tropical and subtropical zones globally [1]. In India, it plays a crucial role as a cash crop, supporting both domestic markets and international exports [2]. Green chilli is particularly valued for its high vitamin C content, as well as its significant contributions of vitamins A and B. It also contains trace amounts of proteins, fats, carbohydrates, and essential minerals such as phosphorus, iron, and calcium [3, 4]. Despite India's contribution of 30% to the global chilli production, the crop faces increasing challenges due to viral infections, which are responsible for substantial economic losses [5, 6]. Chilli plants are vulnerable to a diverse range of DNA and RNA plant viruses, with the number of viral infections rising sharply over the past decade [7]. Currently, around 75 viruses are recognized to infect chilli worldwide, 37 of which have been classified by the International Committee on Taxonomy of Viruses (ICTV), while six others remain as tentative species [5, 8].

In India, several chilli viral diseases are characterized based on their symptomatology, host range, particle morphology, and molecular detection techniques such as serology and nucleic acid-based assays. These include chilli leaf curl virus (ChiLCV), chilli veinal mottle virus (ChiVMV), cucumber mosaic virus (CMV), capsicum chlorosis virus (CaCV), tobacco mosaic virus (TMV), potato virus Y (PVY), groundnut bud necrosis virus (GBNV), and tobacco streak virus (TSV) [9–14]. Furthermore, a variety of other viruses have been reported to infect chilli in India and different global regions, including Begomoviruses, Potyviruses, Cucumoviruses, Tospoviruses, Poleroviruses, Criniviruses, Alfamoviruses, Tobamoviruses, Tymoviruses, Ilarviruses, Tombusvirus, Cryptoviruses, Nepoviruses, and Endornaviruses [7] (Table 1).

Earlier, the identification of plant viruses relied on conventional diagnostic methods, including symptomatology, biological assays, electron microscopy, nucleic acid hybridization, and serological and molecular techniques such as ELISA (enzyme-linked immunosorbent assay) and PCR (polymerase chain reaction) with its various adaptations [15–17]. However, these traditional methods, such as ELISA and PCR, are limited to detecting only known viruses, leaving the identification of unknown or novel viruses unaddressed due to their dependence on prior viral knowledge [18, 19]. In field conditions, the synergistic interactions between multiple viruses in mixed infections can generate a broad spectrum of symptom phenotypes or even lead to the emergence of new viral strains [20]. Symptoms may include leaf curling, mosaic patterns, chlorosis, vein banding, leaf crumpling, reduced leaf size, leaf distortion, stunting, mottling, and vein swelling [7]. These complexities make it challenging for traditional diagnostic methods to capture the full array of viruses present in plants [21]. Therefore, to comprehensively identify all viruses both known and novel it is essential to study the virome, the total viral population associated with infected plant samples [22].

Recent advancements in high-throughput sequencing (HTS) technology have revolutionized the identification of a wide array of viruses across diverse ecosystems, providing valuable insights into viral diversity, interactions, evolution, and epidemiology [23]. HTS methods offer a comprehensive snapshot of the virome in infected plants without the need for prior knowledge of the viruses present, enabling the detection of both known and novel viruses [24]. HTS has been successfully applied to virome sequencing in numerous plant species, including grapes [25, 26], sweet potato [27], tomato [28], garlic [29], apple [30], pepper [31], citrus [32], and various forest plants.

The objective of our study was to characterize the virome associated with the chilli crop in Karnataka, India, with the aim of identifying both known and unknown viruses. In addition, we developed loop-mediated isothermal amplification (LAMP) diagnostic assays for the viruses identified in chilli, facilitating rapid and specific detection.

Materials and methods

Collection of plant samples with viral disease symptom phenotypes

A survey was conducted in 2021-22 to assess the prevalence of viral diseases in chilli crops across key growing regions in Bengaluru (Mavallipura and Muthagadahalli) and Kolar (Rajakallahalli, Madderi Mallandahalli) districts of Karnataka, India (Fig. 1a). The incidence of viral diseases was quantified using the percentage of disease incidence (PDI) method [12]. PDI was calculated by dividing the number of plants exhibiting virus-like symptoms by the total number of plants assessed, then multiplying the result by 100 to express it as a percentage (PDI = Number of symptomatic plants / Total plants × 100). A total of 19 chilli leaf samples showing a range of virus-like symptoms (Supplementary Table S1) were

Table 1 List of viruses identified in Chilli in India compared with viruses identified in Chilli world wide

Country	Virus name	Reference
India	Chili leaf curl Joydebpur virus, chili leaf curl Palampur virus, chilli leaf curl vellanad virus, pepper cryptic virus-2, pepper mild mottle virus and large cardamom chirke virus	[78–81]
China	Chilli ringspot virus and paprika mild mottle virus	[82, 83]
Costa Rica	Tomato chlorosis virus	[84]
Cuba	Tomato chlorotic spot virus	[85]
Czech Republic	Broad bean wilt virus-2	[86]
Dominican Republic	Tomato spotted wilt virus	[87]
Italy	Pepper vein yellows virus	[88]
Korea	Tobacco mild green mosaic virus	[89]
Mali	Pepper veinal mottle virus	[90]
Mexico	Chitelpin yellow mosaic virus	[91]
Pakistan	Chilli leaf curl Multan virus and pepper leaf curl Lahore virus	[92]
Republic of Macedonia	Alfalfa mosaic virus	[93]
Saudi Arabia	Pepper leaf roll chlorosis virus	[94]
Slovakia	Bell pepper alphaendornavirus	[95]
Spain	Broad bean wilt virus-1 and parietaria mottle virus	[96, 97]
Srilanka	Chilli leaf curl Sri Lanka virus	[98]
United States	Pepper cryptic virus-1, pepper cryptic virus-2 and moroccan pepper virus	[99, 100]

collected from all surveyed locations for virome analysis (Fig. 1b). The samples were immediately snap-frozen in liquid nitrogen in the field and stored at -80°C for further processing.

High-throughput sequencing (HTS)

Total RNA was extracted from the 19 chilli leaf samples using the phenol-chloroform and lithium-chloride (LiCl) method [33, 34]. The RNA was then quantified using a Nanodrop (Thermo Fisher Scientific, USA) and assessed for purity and integrity with an Agilent 4150 Tapestation system. The total RNA from the 19 leaf samples, collected from all surveyed locations, was pooled at equimolar concentrations to create two representative RNA samples. One pooled sample was used to construct an mRNA library, while the other was used for constructing an sRNA library. Ribosomal RNA (rRNA) was removed from both pooled samples using the Ribo-Zero rRNA removal kit (Illumina, San Diego, CA, USA). mRNA and sRNA libraries were then prepared using the Illumina TrueSeq stranded mRNA library preparation kit (Illumina, San Diego, CA, USA) and TruSeq small RNA sample preparation protocol, respectively. Library quantification was performed using a Qubit 4.0 fluorometer (Thermo Fisher Scientific, USA) with a DNA HS assay kit (Thermo Fisher Scientific, USA). The insert sizes of the libraries were verified using D1000 ScreenTapes (Agilent) on the Tapestation 4150 system (Agilent). Finally, the libraries were sequenced on an Illumina NovaSeq 6000 platform by the Neuberg Center for Genomic Medicine, Ahmedabad, India, to generate 150-bp paired-end reads.

De Novo assembly

The quality of the raw sequencing reads was assessed using the FastQC Version 0.11.9 tool [35]. Sequence adapters and low-quality reads were trimmed using Trim Galore version 0.6.5, applying a Phred quality score 30 [36]. High-quality mRNA and sRNA reads were selected for de novo assembly. A novel approach was employed to comprehensively characterize the chilli virome, incorporating individual assembly steps for both mRNA and sRNA reads. First, de novo assembly of mRNA reads was performed using Trinity version 2.13.2 [37] with default parameters, followed by a separate de novo assembly of sRNA reads using Velvet version 1.2.10 [38]. Subsequently, the mRNA and sRNA reads were combined for a unified de novo assembly using SPAdes version 3.13.1 [39]. The resulting contigs from each individual assembly were clustered using the cdhit-est version 4.7 tool [40] to eliminate redundant sequences, ensuring the retention of only the unique contigs.

Identification of viruses

The unique contigs generated from the individual de novo assemblies were subjected to homology search using standalone blastn version 2.12.0 [41] to identify closely related sequences (MEGABLAST) with an e-value threshold of 1e-5 and a query coverage of ≥ 90% in the virus database, which includes complete viral reference sequences from the National Center for Biotechnology Information (NCBI) GenBank. Only contigs longer than 200 nucleotides (for mRNA and whole transcriptome (mRNA + sRNA) data) and 50 nucleotides (for sRNA) were included in the homology search. Virus species were identified and named in accordance with the latest classification guidelines from the International Committee

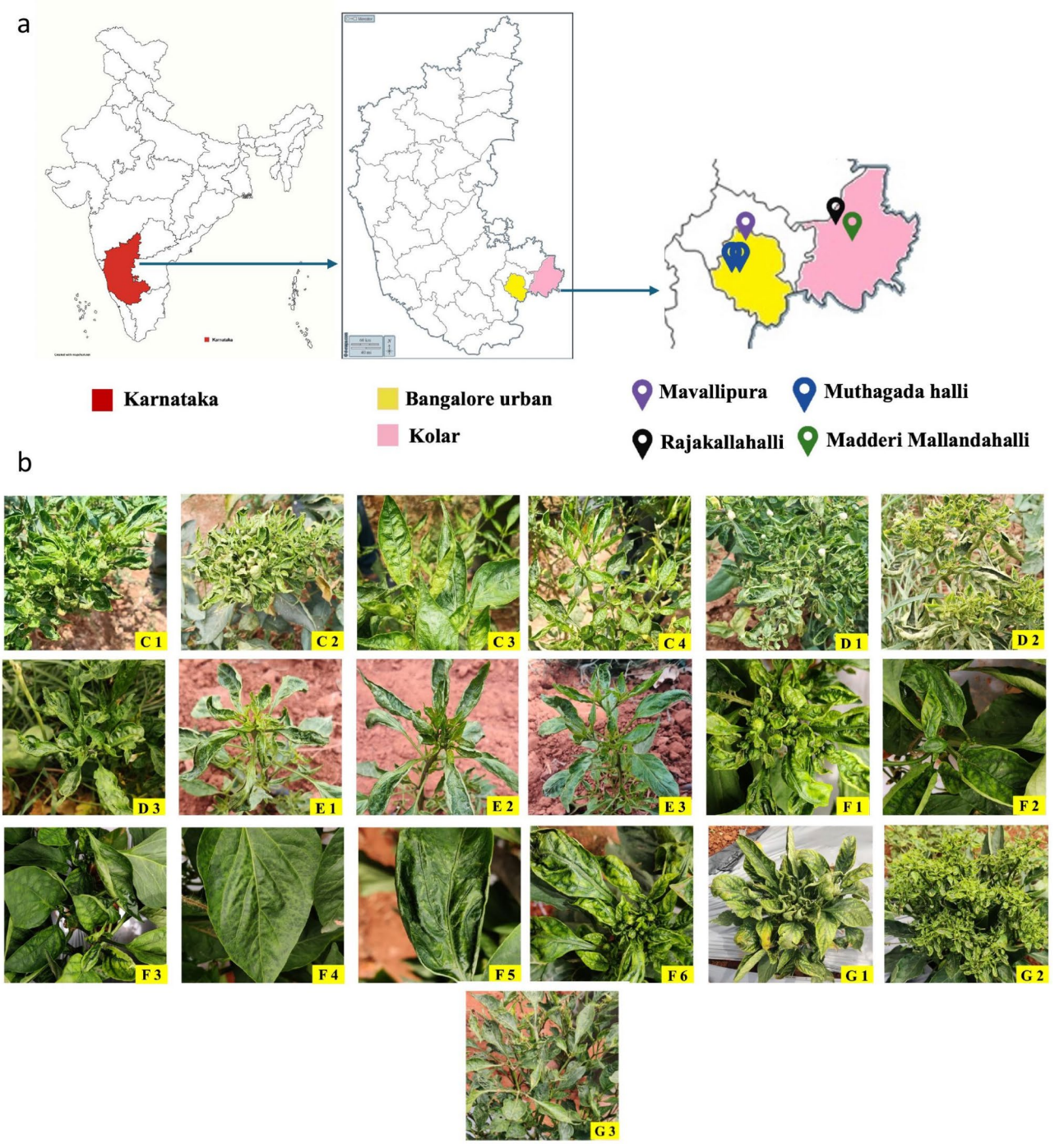

Fig. 1 **a**. Map showing the sampling locations in Karnataka, India. The collection sites included were Bangalore Urban and Kolar districts in Karnataka, India. The samples were collected from Mavallipura (one location, four samples) and Muthagada halli (two locations, three samples per colaction) villages from Bnagalore Urban district and Rajakallahalli (one location, six samples) and Madderi Malllandahalli (one location, three samples) from Kolar districts. **b.** Chilli plants showing different viral disease symptoms under natural conditions. Symptoms observed in each sample are as follows. C1) leaf curling and mosaic, C2) leaf curling, mosaic and bunchy appearance of leaves, C3) mottling and dark green vein banding, C4) leathery leaves and vein banding, D1) leaf curling, D2) leaf curling and mosaic, D3) mottling and mosaic of leaves, E1) leaf curling and mosaic, E2) mosaic and rat tailed leaves, E3) mosaic, F1) leaf curling, mottling and vein banding, F2) mottling and vein banding, F3) vein banding, F4) mosaic, F5) mottling and mosaic, F6) mottling and vein banding, G1) leaf curling and vein banding, G2) leaf curling and small leaves and G3) mottling, mosaic and dull coloured leaves

on Taxonomy of Viruses (ICTV) [42], which incorporate nucleotide and amino acid identity thresholds for species and strain demarcation, as detailed in the ICTV guidelines (Supplementary Table S2).

Reconstruction of viral genomes

Longer contigs (ChiLCV, CMV, PCV-2, PeVYV, and BPEV, >1000 nucleotides; GBNV, >200 nucleotides), assembled by SPAdes and Trinity, were compared to the sequences from the NCBI-designated virus database. The best-hit sequence was selected as the reference seed sequence. Using this reference seed sequence and the contigs of varying lengths, multiple sequence alignment was performed using Clustal X version 2.0 [43] to identify overlapping regions and extend the fragmented virome. Further reconstruction of the virome with maximum coverage was accomplished using BioEdit version 7.2 [44]. From the assembled complete or near-complete viral genomes, all predicted open reading frames (ORFs) were identified using the NCBI ORF Finder.

Viral copy number Estimation and single nucleotide polymorphism (SNP) analysis

The copy number of the identified viruses was assessed using two methods. The first approach involved determining the total number of virus-associated contigs by performing BLAST analysis of the assembled reads against the viral database. In the second approach, the full-length or near-full-length viral genomes were used as reference sequences, and the raw reads were mapped to these genomes. The number of Fragments Per Kilobase of transcript per Million (FPKM) was then calculated to quantify the reads mapping to each viral genome. Single nucleotide polymorphisms (SNPs) in complete or near-complete viral genomes were identified by mapping raw reads to the assembled virome sequences using the BWA aligner [45] with default parameters. The resulting Binary Alignment Map (BAM) files were indexed and visualized using the Tablet program [46] to detect the presence of SNPs.

Sequence comparison, phylogenetic and recombination analyses

The assembled complete or near-complete genomes of the identified viruses, along with their respective reference genomes retrieved from NCBI GenBank, were utilized for sequence comparison, phylogenetic analysis, and recombination detection (Supplementary Tables S19-S26). Pairwise sequence identity scores were calculated by aligning the assembled viral genomes with reference sequences from NCBI using ClustalW within the Sequence Demarcation Tool (SDT) version 1.2 [47]. Phylogenetic trees were constructed using the Neighbor-Joining (NJ) method with the Kimura 2-parameter (K2P) model with 1000 bootstrap replicates in MEGA 11 [48]. Recombination analysis was performed using RDP5 [49], which integrates six algorithms: Recombination Detection Program (RDP), GENECONV, MaxChi, Chimaera, SiScan, and 3Seq. Recombination events detected by at least three of these algorithms in the assembled viral genomes were considered significant. The overall workflow for virome analysis in chilli is depicted in Fig. 2.

Confirmation of HTS results

The presence of the identified viruses was validated through PCR assays. Specific primers targeting the coat protein (for ChiLCV, CMV, PCV-2, and PeVYV), movement protein (for GBNV), and polyprotein (for BPEV) genes were designed (Supplementary Table S3).

PCR

DNA was extracted from the 19 leaf samples using the CTAB method and pooled into a single representative sample to confirm the presence of DNA viruses. PCR assays were carried out with an initial denaturation step at 94 °C for 5 min, followed by 35 cycles of denaturation at 94 °C for 30 s, annealing at 50 °C for 30 s, and extension at 72 °C for 45 s, with a final extension step at 72 °C for 10 min. PCR reactions were performed on a ProFlex PCR system (Applied Biosystems, Thermo Fisher Scientific, USA), in a 20 µL reaction volume containing 2 µL of DNA (50 ng/µL) or 5 µL of cDNA (20 ng/µL) template, 10 µL of EmeraldAmp® GT PCR Master Mix (Takara Bio, USA), and 1 µL (10 µM) of each forward and reverse primer. Amplification products were resolved on a 1% agarose gel (electrophoresed at 80 V for 1 h), stained with ethidium bromide (10 mg/mL), and visualized using the Gel Doc XR Imaging System (Bio-Rad, USA).

RT-PCR

For RNA viruses, RNA was extracted from the same 19 leaf samples, as described earlier, and similarly pooled into one representative sample. cDNA synthesis was performed on 1000 ng of pooled RNA using 10 µM of virus-specific reverse primers for each reconstructed RNA virus and RevertAid Reverse Transcriptase (Thermo Fisher Scientific, USA) according to the manufacturer's instructions. PCR assays were carried out with an initial denaturation step at 94 °C for 5 min, followed by 35 cycles of denaturation at 94 °C for 30 s, annealing at 50 °C for 30 s, and extension at 72 °C for 45 s, with a final extension step at 72 °C for 10 min. PCR reactions were performed on a ProFlex PCR system (Applied Biosystems, Thermo Fisher Scientific, USA), in a 20 µL reaction volume containing 5 µL of cDNA (20 ng/µL) template, 10 µL of EmeraldAmp® GT PCR Master Mix (Takara Bio, USA), and 1 µL (10 µM) of each forward and reverse primer. Amplification products were resolved on a 1% agarose gel

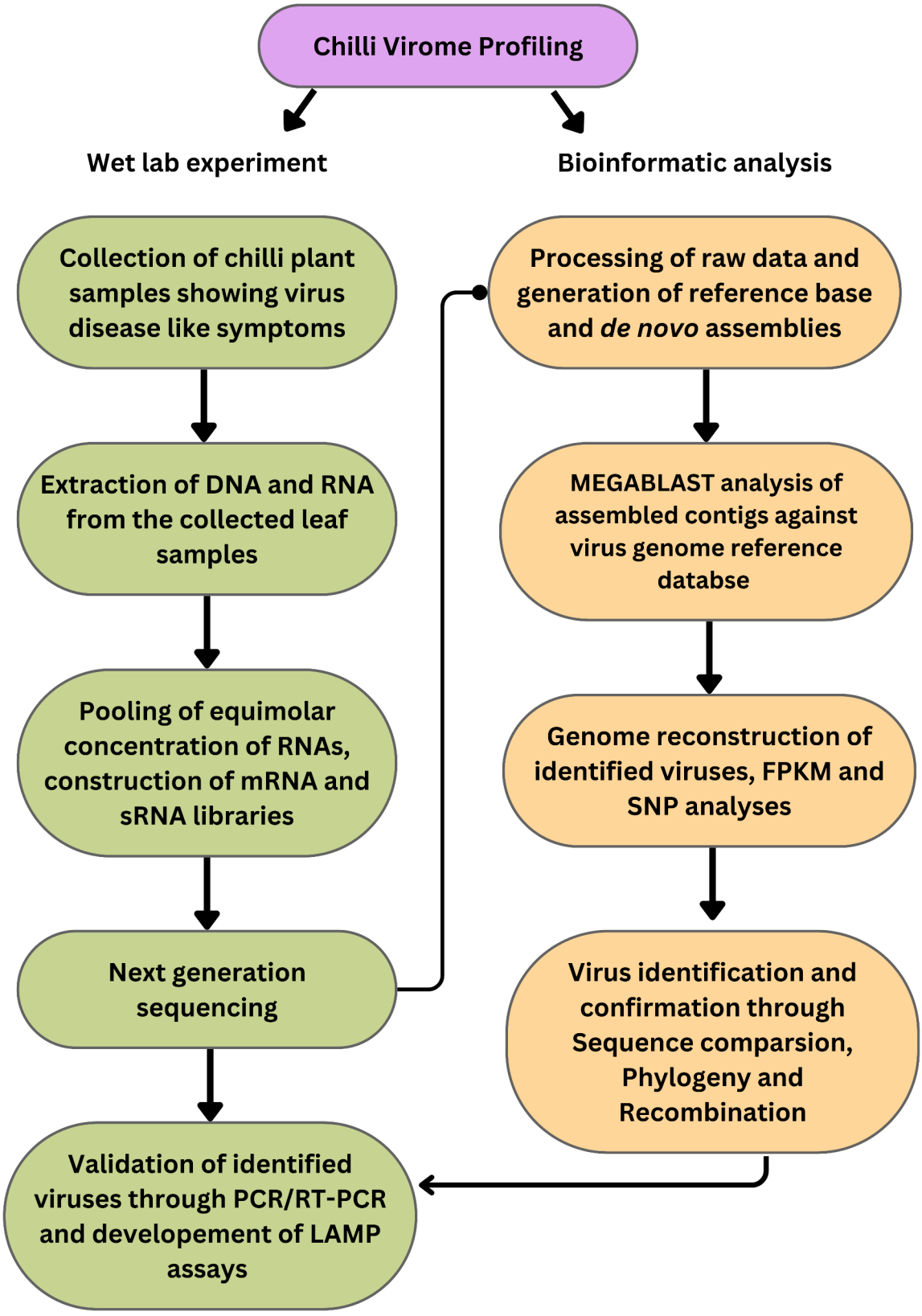

Fig. 2 Virome analyses workflow in chilli. Chilli leaf samples showing varied disease symptoms were collected for virome profiling. Total RNA extracted from each sample was pooled and used for mRNA and sRNA libraries preparation. Libraries of mRNA and sRNA were sequenced through next-generation sequencing (NGS) platforms and obtained raw data was pre-processed. Good quality mRNA and sRNA data was de novo assembled, and obtained assembled contigs were used for identification of viruses

(electrophoresed at 80 V for 1 h), stained with ethidium bromide (10 mg/mL), and visualized using the Gel Doc XR Imaging System (Bio-Rad, USA).

Development of LAMP diagnostic assays for identified viruses

LAMP diagnostic assays were developed for the identified viruses for future use in screening. LAMP primers were designed targeting specific genes, including coat protein (ChiLCV, CMV, PCV-2 and PeVYV), movement protein (GBNV) and poly protein (BPEV) using primer Explorer V5 (https://primerexplorer.jp/e/) (Table 2). The LAMP assays were performed with 25 µL LAMP reaction mixture containing 2 µL of DNA/cDNA (50 ng/µL) template, 1 µL (10 µM) of F3, B3, 2 µL (10 µM) of each FIP, BIP, LF and LB primers, 1.5µL of 10 mM dNTPs, 1.0 µL of 5 M betaine, 6.5 µL of sterile double- distilled water, 2.5 µL of 1X ThermoPol Reaction buffer (20 mM Tris–HCl, 10mM $(NH_4)SO_4$, 10 mM KCl, 2 mM $MgSO_4$, 0.1% Triton X-100, pH 8.8 @ 25 °C), 0.5 µL of 100 mM $MgSO_4$, and 1.0 µL of 8U Bst DNA Polymerase (New England Biologicals, USA). The reaction mixture was incubated in a hot water bath at 60 °C for 1 h followed by 80 °C for 10 min to terminate the reaction. LAMP reaction

Table 2 Details of LAMP and RT-LAMP primers synthesized from genes of identified viruses from the virome analyses of chilli

Virus	Species new name	Primer	Sequence (5′-3′)	Gene used for primer synthesis	Reaction performed
ChiLCV	Begomovirus chillicapsici	ChiLCV-F3	CAAGATTTTGGAGAGGTGTT	Coat protein (CP)	LAMP
		ChiLCV-B3	ATTTACCAGCTTCCTGCT		
		ChiLCV-FIP	CCTCAAAACCTGATAACGATCTCT-CAACATGTTTGACAACGAGC		
		ChiLCV-BIP	CAACGGTTACTGGTGGACAGT-CACAACATAATTATTAACCCTAACG		
		ChiLCV-LF	CTTCACAGTGGCAGTGCTAG		
		ChiLCV-LB	CGTCCAAGGAACAGGCATTAGTT		
CMV	Cucumovirus CMV	CMV-F3	GCAAAGTTCCTGCCTCCTC	Coat protein (CP)	RT-LAMP
		CMV-B3	GGTGCTCAATGTCGACATGA		
		CMV-FIP	ACGCCAGATGCAGCATACTGAT-CCGCCATCTCTGCTATGTTC		
		CMV-BIP	TTCAGCGATGCGCGCTGATA-CGTCTCGAGTGCATCGTC		
		CMV-LF	TACTGGTGAGGCTCCGTCCG		
		CMV-LB	CGCCGTGCTCGTGTATTC		
GBNV	Orthotospovirus arachinecrosis	GBNV-F3	GCAGCTCAATTTAACTAGTGAT	Movement protein (MP)	RT-LAMP
		GBNV-B3	GATATCTGTTTCCCACCAATT		
		GBNV-FIP	TGGTATTTTCAGCAGCTATAGGAGT-GGAGTCAGCTTTGCTTCAG		
		GBNV-BIP	GATGTAGTTCCCATCAACAGAGC-CCTTTGGGTATCATTAATTTGCA		
		GBNV-LF	AGTTCTTGACCCAGGAATACATGA		
		GBNV-LB	TAAAGTCATTCAATCTGCAGCTCTG		
PCV-2	Deltapartitivirus duocapsici	PCV2-F3	TTAGTGACACACCAACTACA	Coat protein (CP)	RT-LAMP
		PCV2-B3	TGAATAAAGCTTGCGTCCT		
		PCV2-FIP	GGCTTTTGCGTTATAGGGGG-GTCTACACAAGACAAGACTGA		
		PCV2-BIP	TCCGAGATTTCAAGAATCGCCG-TTCTATGCTTGGGTGGAC		
		PCV2-LF	CGTAGCTGCCGGTAATTGG		
		PCV2-LB	ACTACCTCGCCGATGATTACGAG		
PeVYV	Polerovirus PeVYV2	PeVYV-F3	GGAACCGAAGAAGCCGAAAT	Coat protein (CP)	RT-LAMP
		PeVYV-B3	CGCTGTGGAAGAGGATTCAC		
		PeVYV-FIP	AGGTGACAGCTCCTGAGGAACT-GTCGAGCAACAGCGAGAC		
		PeVYV-BIP	ATCAGAGAGCATCGCGCTTTCA-ACGAAGCGTATGTTGACCAT		
		PeVYV-LF	GATTGAGTCCTTGTTGAAGACGAAA		
		PeVYV-LB	GGTGGAGTTCTCAAAGCCTACC		
BPEV	Alphaendornavirus capsici	BPEV-F3	TCGTATGGTTACAAAGCCAT	Poly protein	RT-LAMP
		BPEV-B3	ATATCAATCACCTCCAACATG		
		BPEV-FIP	CCTTCAGTTTCGTCTAATAACCACA-GGAATTGCCAAATGTAGAAGC		
		BPEV-BIP	TGACTAAGTCAAGATGTGGAGTCT-CATAATGCATTCAGGCATCA		
		BPEV-LF	ACACGAACACGTCATCGCT		
		BPEV-LB	GTTTGCAACCGCTATTGGTTAGA		

ChiLCV- Chilli leaf curl virus, CMV- Cucumber mosaic virus, GBNV- Groundnut bud necrosis orthotospovirus, PCV-2- Pepper cryptic virus-2, PeVYV- Pepper vein yellows virus and BPEV- Bell pepper alphaendornavirus

products were analyzed on 2% agarose gels (65 V for 3 h) stained with ethidium bromide (10 mg/mL) and viewed on Gel Doc XR imaging system (BioRad, USA). For the visual detection of LAMP reaction products, nucleic acid staining dyes VeriPCR dye (Chromus Biotech PVT Ltd, Bengaluru, India) and hydroxynaphthol blue (HNB) (Sigma Aldrich, USA) were used. 5 µL of 1X VeriPCR dye was added to the PCR reaction mixture after amplification, whereas 0.2 µL of 20 mM HNB was added to the LAMP master mix prior to amplification for visual inspection of the reaction products. Positive result with VeriPCR dye shows green fluorescence and HNB shows a colour change from violet to sky blue.

Results

Symptomatology

Viral disease-like symptoms were observed in chili crops surveyed across several regions in Bengaluru Urban (Mavallipura and Muthagadahalli) and Kolar (Rajakallahalli, Madderi, and Mallandahalli) districts of Karnataka, India. The incidence of viral diseases in the surveyed areas ranged from 26.6 to 47.5%, with the highest disease prevalence recorded in Mavallipura, Bengaluru (47.5%) (Supplementary Table S1). A total of 19 leaf samples were collected from infected chili plants, displaying a variety of viral disease symptoms, including leaf curling, vein banding, mosaic patterns, mottling, shoestring/rat-tail/filiform/leathery leaf morphology, dull coloration, and a bunchy leaf appearance (Fig. 1b).

Identification of viruses

Illumina sequencing generated 15.8 million (5.70 GB) and 31.6 million (4.70 GB) raw sequences for the mRNA and sRNA libraries, respectively. De novo assembly of the mRNA reads using Trinity, sRNA reads with Velvet, and whole transcriptome (WT) reads with SPAdes yielded 38,262, 8,773, and 41,383 total contigs, respectively. MEGABLAST analysis of the assembled contigs against complete viral reference sequences from NCBI GenBank identified both RNA and DNA viruses, as well as their associated satellite viruses (Supplementary Tables S4-S6). MEGABLAST analysis identified six distinct viruses in the pooled 19 leaf samples, including *Begomovirus chillicapsici* (Geminiviridae: Begomovirus), *Cucumovirus CMV* (Bromoviridae: Cucumovirus), *Orthotospovirus arachinecrosis* (Tospoviridae: Orthotospovirus), *Deltapartitivirus duocapsici* (Partitiviridae: Deltapartitivirus), *Polerovirus PeVYV2* (Solemoviridae: Polerovirus), and *Alphaendornavirus capsici* (Endornaviridae: Alphaendornavirus). All six viruses chilli leaf curl virus (ChiLCV), cucumber mosaic virus (CMV), pepper cryptic virus-2 (PCV-2), groundnut bud necrosis orthotospovirus (GBNV), pepper vein yellows virus (PeVYV), and bell pepper alphaendornavirus (BPEV)—were detected in both the mRNA and whole transcriptome libraries, while only five viruses (ChiLCV, CMV, PCV-2, PeVYV, and BPEV) were identified in the sRNA library (Supplementary Table S7A). Additionally, five satellites associated with ChiLCV were detected, including tomato leaf curl Bangladesh betasatellite (ToLCBDB), tomato leaf curl Anand alphasatellite (ToLCAA), croton yellow vein mosaic alphasatellite (CYVMA), ageratum yellow vein Singapore alphasatellite (AYVSA), and tomato leaf curl Virudhunagar alphasatellite strain severe (ToLCVirA) (Supplementary Table S7B). Among the DNA satellites of begomoviruses, ToLCBDB was identified in all three datasets, ToLCAA in the sRNA and WT libraries, CYVMA in the mRNA and sRNA libraries, and AYVSA and ToLCVirA only in the sRNA library (Supplementary Table S7B). Notably, PeVYV and BPEV represent the first reports of these viruses in Indian chilli.

Assembling viral genomes

Virus-associated contigs were extracted from the total contigs of mRNA, sRNA, and whole transcriptome libraries and subsequently assembled into complete or near-complete viral genomes (Table 3; Fig. 3). Mapping these virus-associated contigs to NCBI-designated reference virus genomes resulted in five complete or near-complete viral genome sequences (> 90% coverage), including chilli leaf curl virus (ChiLCV), cucumber mosaic virus (CMV), pepper cryptic virus-2 (PCV-2), pepper vein yellow virus (PeVYV), and bell pepper alphaendornavirus (BPEV) (Table 3). However, for groundnut bud necrosis virus (GBNV), the viral contigs could not be assembled into a complete genome.

Assembly using Trinity for mRNA and SPAdes for whole transcriptome libraries generated longer contigs compared to Velvet's assembly of the sRNA library. Complete or near-complete genomes for ChiLCV, CMV (RNA1, RNA2, and RNA3), PCV-2 (RNA1), BPEV, and PeVYV were assembled using contigs from SPAdes, achieving over 90% genome coverage, with the exception of GBNV L and M segments, which exhibited low genome coverage of 9.1% and 20.2%, respectively (Table 3). Notably, the PeVYV and BPEV genomes were assembled with maximum coverages of 99.5% and 99.9%, respectively. The genomes for PCV-2 RNA2 and GBNV S segment were assembled using Trinity, resulting in genome coverage of 97.7% and 21.8%, respectively (Table 3). All anticipated open reading frames (ORFs) were identified in the assembled complete or near-complete genomes of the identified viruses (Supplementary Table S8).

Table 3 Assembled viruses with their genome lineage and summary of virus-associated contigs in mRNA, sRNA and whole transcriptome (WT) library

Reconstructed virus	Species new name	Segment	Genome	No. of contigs associated with virus			Reference genome	Size (nt)	Genome recovered (nt)	Percentage of genome recovery	FPKM values	Virus copy number (%)
				mRNA	sRNA	WT						
ChiLCV	*Begomovirus chillicapsici*	DNA-A	Circular, ssDNA	4	25	2	NC_004628.1	2754	2510	91.14	166623.14	45.36
CMV	*Cucumovirus CMV*	RNA1	Linear, (+) ssRNA	6	11	18	NC_002034.1	3357	3313	98.69	66337.55	18.06
		RNA2		3	13	7	NC_002035.1	3050	3016	98.89	30678.99	8.35
		RNA3		5	11	24	NC_001440.1	2216	2175	98.15	96986.43	26.4
GBNV	*Orthotospovirus arachinecrosis*	L segment	Linear, (-) ssRNA	9	0	6	NC_003614.1	8911	859	9.64	-	-
		M segment		3	0	5	NC_003620.1	4801	972	20.25	-	-
		S segment		0	0	2	NC_003619.1	3057	667	21.82	-	-
PCV-2	*Deltapartitivirus duocapsici*	RNA1	Linear, dsRNA	1	3	1	NC_034159.1	1609	1585	98.51	1527.73	0.42
		RNA2		1	0	1	NC_034167.1	1525	1490	97.7	763.38	0.21
PeVYV	*Polerovirus PeVYV2*		Linear, (+) ssRNA	6	2	8	NC_055129.1	6028	5998	99.5	1230.51	0.33
BPEV	*Alphaendornavirus capsici*		Linear, dsRNA	1	7	1	NC_039216.1	14,714	14,702	99.92	3234.88	0.88

ChiLCV- Chilli leaf curl virus, CMV- Cucumber mosaic virus, GBNV- Groundnut bud necrosis orthotospovirus, PCV-2- Pepper cryptic virus-2, PeVYV- Pepper vein yellows virus and BPEV- Bell pepper alphaendornavirus

Viral copy number and single nucleotide polymorphism (SNP) analysis

The abundance of viruses was quantified by calculating the number of virus-associated sequences and determining the FPKM values. Among the identified viruses, ChiLCV was the most abundant, comprising 45.36% of the viral copy number, followed by CMV RNA3 at 26.4% of the viral copy number (Table 3). Given the inherent genetic variability of viruses relative to other microorganisms, the single nucleotide polymorphisms (SNPs) within the identified and assembled viral genomes were assessed by mapping the raw sequence reads against complete viral genomes. SNPs were then calculated to analyze genetic variation (Fig. 4). The highest number of SNPs was observed in PeVYV with 196 variants, followed by BPEV with 153 SNPs, CMV RNA1 with 139 SNPs, RNA3 with 60 SNPs, RNA2 with 58 SNPs, PCV-2 RNA1 with 13 SNPs, RNA2 with 9 SNPs, and ChiLCV with 2 SNPs.

Sequence comparison and phylogenetic analyses

The complete or near-complete genomes of ChiLCV, CMV, PCV-2, PeVYV, and BPEV were selected for further analysis. Sequence comparison and phylogenetic analysis were performed using these genomes and their respective reference sequences obtained from NCBI GenBank. The ChiLCV genome (DNA-A: 2510 bp, ON714166) exhibited a 99.28% nucleotide (nt) identity with the ChiLCV isolates from Ahmedabad (JN663846) and Raichur (MK161454), both infecting chilli in India (Supplementary Table S10 and Fig. 5A). For CMV, the three genomic segments (RNA1: 3313 bp, ON722351; RNA2: 3016 bp, ON722352; RNA3: 2175 bp, ON722353) clustered closely with isolates belonging to subgroup IB in the phylogenetic tree and showed the highest nt identity with Indian isolates such as BA (RNA1: MN340990; 98.13%), TNCR2-1 (RNA-2: MT422731; 98.11%), and KP (RNA-3: KM272276; 98.07%) (Supplementary Tables S11-S13 and Fig. 5B). For PCV-2, RNA1 (1609 bp, ON722354) demonstrated 99.68% nt identity and 99.22% amino acid (aa) identity with the Chinese isolate CQ (RNA1, KX905077), while RNA2 (1525 bp, ON645206) shared 99.9% nt identity and 100% aa identity with the South Korean isolate SC10-2 (RNA-2, KY923704) (Supplementary Tables S14-S15 and Fig. 5C). The PeVYV sequence (5998 bp, ON961654) obtained in this study closely clustered with PeVYV isolates from NCBI, exhibiting the highest identity with the Israeli isolate Is (HM439608; 93.88% nt and 89.54–97.78% aa identity) infecting chilli (Supplementary Tables S16-S17 and Fig. 5E). Lastly, the BPEV genome (14702 bp, ON722355) showed 99.97% nt identity with the Panamanian isolate (MZ127290), confirming it as a BPEV isolate infecting chilli (Supplementary Table S18 and Fig. 5D).

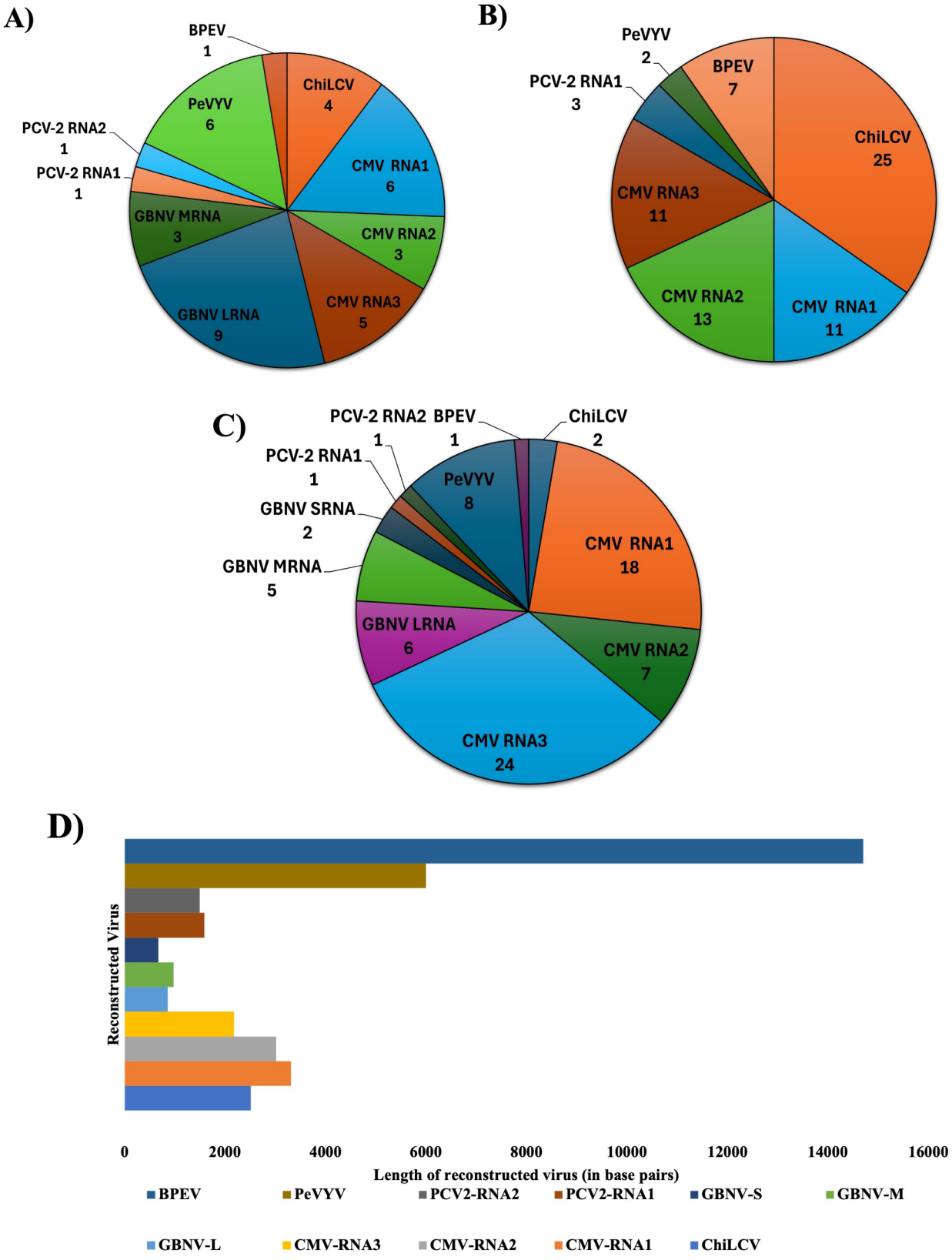

Fig. 3 Pie chart showing the number of contigs associated with each identified virus in **A**) mRNA, **B**) sRNA and **C**) WT library **D**) Graph showing the length of reconstructed viral genomes. Genome organization of reconstructed complete/near complete genomes

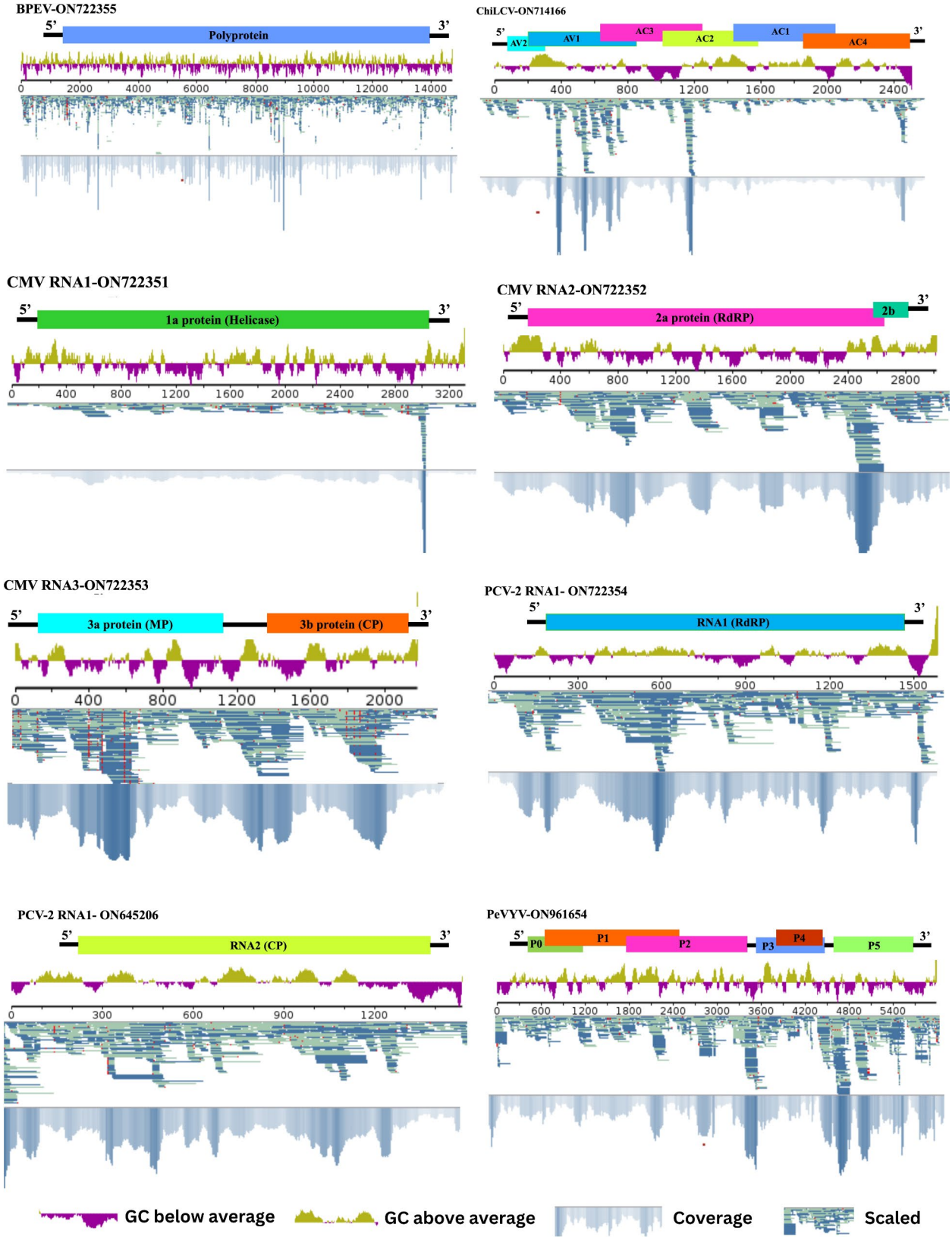

Fig. 4 Genome organization and analysis of single nucleotide variations for identified viruses The SNPs were identified by mapping unmapped reads to individual viral genomes using the BWA aligner and SAMtools, and they were visualized using the tablet program. The red colour markings in the stacked contigs represent SNP regions

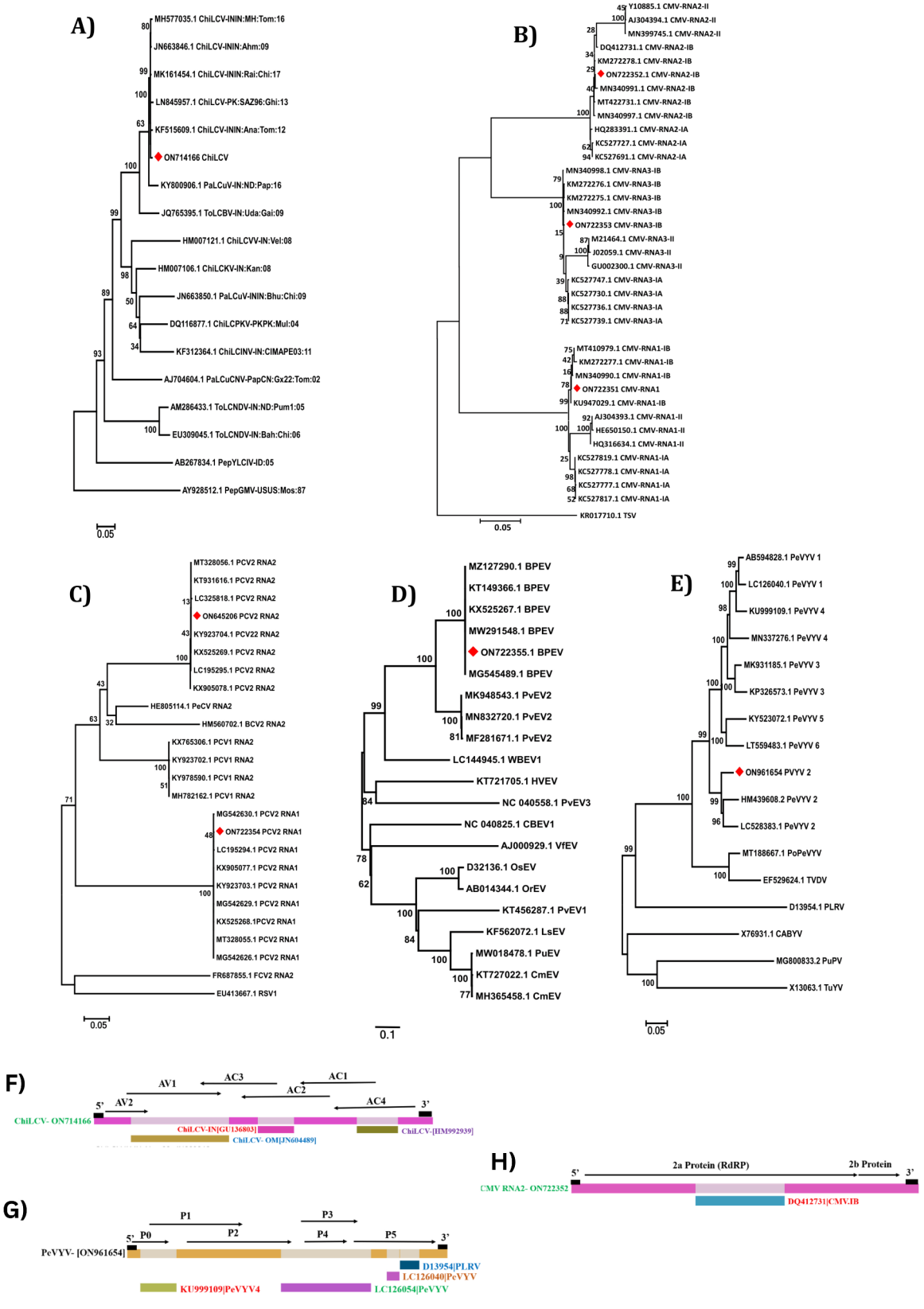

Fig. 5 (See legend on next page.)

(See figure on previous page.)

Fig. 5 Phylogenetic tree from nucleotide sequences of reconstructed complete/near-complete viral genomes **A**) ChiLCV **B**) CMV **C**) PCV2 **D**) BPEV and **E**) PeVY2 with other selected respective reference viruses from NCBI using Neighbourhood Joining (NJ) method and Kimura 2- parameter (K2P) model with 1000 bootstrap replicates in MEGA 11 programme. Recombination breakpoint analysis in reconstructed viral genomes. Recombinants were identified in **F**) ChiLCV, **G**) PeVYV and **H**) CMV RNA2. The recombinants were observed in AV1, AV2 (coat protein) and AC4 regions of ChiLCV, P0, P3 and P5 regions of PeVYV and 2a protein (RdRP) region of CMV RNA2

Detection of recombinants among the identified viruses

Recombination breakpoints were identified in several regions of ChiLCV, CMV, and PeVYV. In ChiLCV, potential breakpoints were detected in the coat protein and AC4 regions of the DNA-A genome (ON714166), while in CMV, recombination breakpoints were observed in the RNA2 segment (ON722352) within the 2a protein region, and in PeVYV, breakpoints were found in the P0, P3, and P5 protein regions (ON961654). These breakpoints were consistently supported by at least three distinct recombination detection algorithms, as shown in Fig. 5F-H. Additional recombination events were noted in CMV RNA3 (ON722353) and BPEV (ON722355) genomes, although these were not corroborated by three or more algorithms. Both intraspecific and interspecific recombination events were observed among ChiLCV, CMV, and PeVYV isolates. Notably, no recombination breakpoints were detected in CMV RNA1 (ON722351), PCV-2 RNA1 (ON722354), or RNA2 (ON645206) genome segments (Supplementary Table S27).

Confirmation of HTS results and development of LAMP assay

The presence of the identified DNA ChiLCV and RNA (CMV, GBNV, PCV-2, PeVYV, and BPEV) viruses in the pooled samples was confirmed through PCR and RT-PCR, respectively. PCR assays successfully amplified the target genes of the identified viruses, producing bands at the expected amplicon lengths, as demonstrated by 1% agarose gel electrophoresis (Fig. 6A-F). Additionally, LAMP (Loop-mediated isothermal amplification) diagnostic assays were successfully developed for both DNA (ChiLCV) and RNA (CMV, GBNV, PCV-2, PeVYV, and BPEV) viruses. LAMP and RT-LAMP assays were performed to detect the presence of DNA and RNA viruses, respectively, in the pooled samples. These assays yielded amplified products in a characteristic ladder-like pattern on 2% agarose gel electrophoresis (Fig. 7A-F). Furthermore, the LAMP assays enabled visual confirmation of the virus's presence by staining with VeriPCR dye, which emitted green fluorescence, and with HNB dye, which caused a colour change from violet to sky blue (Fig. 7A-F).

Discussion

This study aimed to characterize the virome of chilli plants by identifying both known and previously unreported viruses, as well as to develop diagnostic tools. We identified six distinct viruses in 19 pooled leaf samples, including four established viruses (ChiLCV, CMV, GBNV, and PCV-2) and novel viruses (PeVYV and BPEV) previously unrecorded in India. For these viruses, we successfully developed LAMP-based diagnostic assays to facilitate detection.

Chilli plants exhibited a range of disease symptoms, including leaf curling, vein banding, mosaic, mottling, shoestring/rat-tail/filiform/leathery leaves, dull coloration, and a bunchy appearance. To enhance viral sequence detection sensitivity, viral nucleic acids were enriched from total RNA prior to library preparation [50]. Given that total RNA sequencing typically generates a high proportion of host rRNA reads (around 90%), which can obscure viral reads [51], rRNA depletion was performed in this study to improve viral detection. Paired-end raw reads were generated and subjected to de novo assembly using three approaches - mRNAome (Trinity), sRNAome (Velvet), and whole transcriptome (SPAdes) to profile the chilli virome and reduce false negatives comprehensively [52]. Previous studies have demonstrated the efficacy of these diverse tools in optimizing results [53]. BLAST analysis of the assembled reads revealed the presence of six viruses (ChiLCV, CMV, GBNV, PCV-2, PeVYV, and BPEV) and five DNA satellites (ToLCBDB, ToLCAA, CYVMA, AYVSA, and ToLCVirA). Of the six viruses, all were detected in the mRNA and whole transcriptome libraries, while five were identified in the sRNA library, excluding GBNV and PCV-2, likely due to shorter sRNA contigs compared to the mRNA and whole transcriptome libraries [54]. Notably, PeVYV and BPEV are reported here for the first time in chilli plants from India.

The length of viral contigs plays a crucial role in accurate virus identification through high-throughput sequencing (HTS) data [55]. Shorter viral contigs can lead to false identifications during BLAST analysis [56]. To mitigate this, we mapped longer contigs generated by Trinity and SPAdes to NCBI virus genomes, as both assemblers outperformed Velvet, aiding in the reconstruction of complete viral genomes [18]. Mapping these contigs to their respective reference sequences in the NCBI database resulted in the assembly of complete or near-complete genomes for five viruses - ChiLCV, CMV, PCV-2, PeVYV, and BPEV, as well as a partial genome for GBNV. The incomplete genome recovery of certain viruses may be attributed to several factors, including low virus titers, host-mediated viral silencing, poor RNA

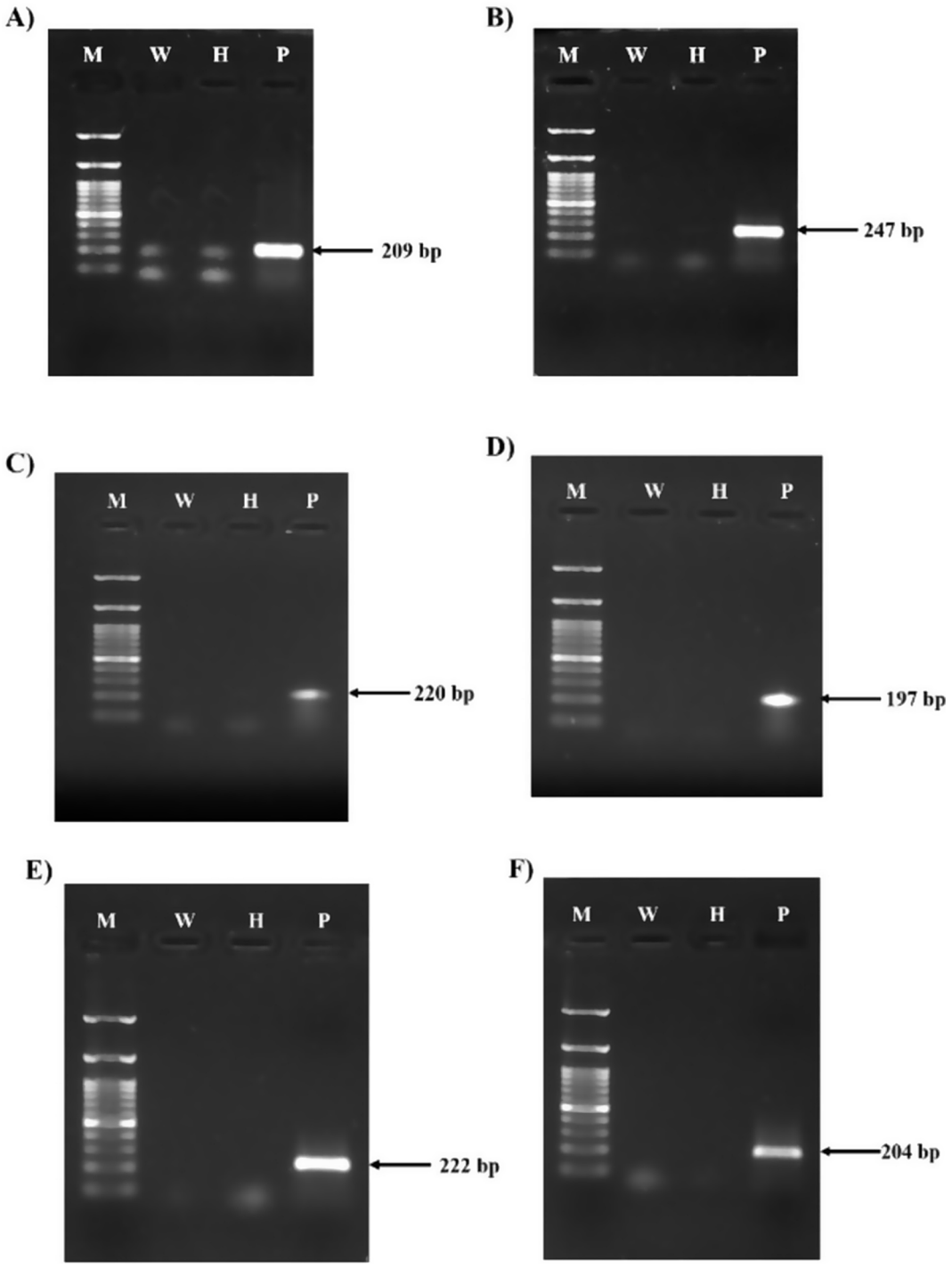

Fig. 6 Validation of identified viruses **A**) ChiLCV with PCR and **B**) CMV, **C**) GBNV, **D**) PCV-2, **E**) PeVYV and **F**) BPEV with RT-PCR assay. PCR/RT-PCR amplification products were checked on 1% agarose gel. Lane M: 100 bp DNA ladder, Lane W: Water control, Lane H: Healthy sample and Lane P: Pooled DNA/RNA sample used for virome analyses. Coat protein gene was amplified in ChiLCV, CMV, PCV-2 and PeVYV, movement protein gene was amplified in GBNV, and polyprotein gene was amplified in BPEV

quality, insufficient sequencing depth, or data analysis limitations [57]. Notably, although ChiLCV is an ssDNA virus, a complete genome was assembled from mRNA transcriptome data, which encompasses most regions of virus-encoded genes. This is consistent with findings by Jo et al. (2017) [31], who successfully identified geminiviruses and satellite DNAs from pepper viromes using mRNA data. Among the identified viruses, ChiLCV exhibited the highest viral copy number and the lowest mutation frequency. This observation aligns with

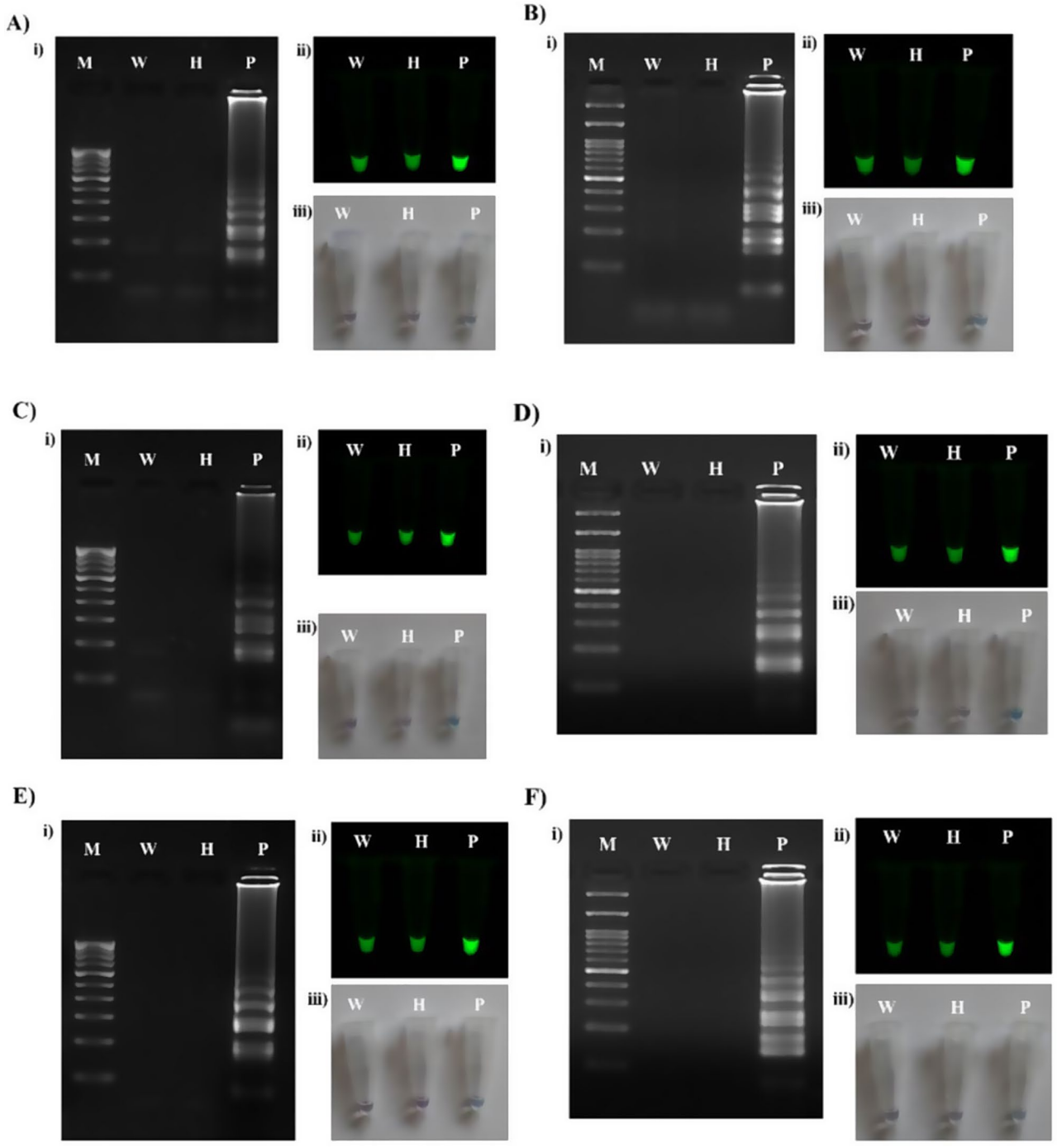

Fig. 7 Development of LAMP diagnostic assays for identified viruses **A**) ChiLCV with LAMP and **B**) CMV, **C**) GBNV, **D**) PCV-2, **E**) PeVYV, **F**) BPEV with RT-LAMP assays. (i) LAMP/RT-LAMP reaction products on 2% agarose gel (ii) Visual inspection with Veri PCR dye. Pooled DNA/RNA sample used for virome analyses showed green fluorescence, whereas in healthy sample and water control no fluorescence was observed. (iii) Visual inspection with hydroxy naphthol blue (HNB). Pooled DNA/RNA sample used for virome analyses showed change in colour from violet to sky blue, whereas in healthy sample and water control no change in violet colour was observed. Lane M: 100 bp DNA ladder, Lane W: Water control, Lane H: Healthy sample and Lane P: Pooled DNA/RNA sample used for virome analyses. Coat protein gene was amplified in ChiLCV, CMV, PCV-2 and PeVYV, movement protein gene was amplified in GBNV, and polyprotein gene was amplified in BPEV

numerous studies that report lower variation in DNA viruses compared to RNA viruses [24, 58].

The complete or near-complete genomes recovered from the data were subjected to sequence comparison, phylogenetic analysis, and recombination detection. The ChiLCV genome (ON714166) was confirmed as a variant of the Ahmedabad and Raichur isolates, in accordance with the current species demarcation criteria for begomoviruses, which stipulate a nucleotide identity threshold of less than 91% for full-length monopartite genomes or full-length DNA-A for bipartite begomoviruses [59]. The CMV genomes (RNA1: ON722351, RNA2: ON722352, RNA3: ON722353) were classified under CMV subgroup IB and identified as variants, with more than 65% sequence identity across the whole genome, consistent with cucumovirus species criteria [60]. The PeVYV genome (ON961654) was identified as a variant of the Israeli PeVYV isolate infecting chilli, with less than a 10% amino acid sequence difference in any gene product [61]. In addition to symptomatic viruses, complete or near-complete genomes of asymptomatic viruses such as PCV-2 and BPEV were also recovered. The PCV-2 genome in this study is a variant of the South Korean and Chinese isolates infecting chilli, meeting the species demarcation criteria of more than 90% amino acid identity in RNA1 (ON722354; RdRP) and over 80% identity in RNA2 (ON645206; CP) for deltapartitiviruses [62]. The BPEV genome (ON722355) exhibited a nucleotide identity of 99.9% with the BPEV isolate from Panama, confirming it as a variant of that isolate and meeting the 75% whole genome sequence identity criterion for alphaendornaviruses [63]. The obtained partial genome of GBNV nucleocapsid (N) gene was compared with other orthotospovirus isolates showed 99.6% nucleotide identity with GBNV isolate from India: Kodigahalli (Supplementary Table S28), confirming it is variant of that isolate meeting the 90% nucleotide similarity in the nucleocapsid gene [64].

Recombination plays a pivotal role in viral emergence, facilitating the exchange or transfer of genomic fragments between closely or distantly related viruses [65]. This process can result in the creation of novel recombinant genomes, which may confer advantages such as enhanced adaptation to new environments, expansion of host range, and/or the development of new symptoms [66]. Our recombination analysis revealed breakpoints in the coat protein and AC4 regions of ChiLCV, the P0, P3, and P5 protein regions of PeVYV, and the 2a protein region of CMV RNA2. In agreement with these findings, Mishra et al. (2020) [67] reported intragenic recombination in the coat protein (V1/AV1) region of ChiLCV infecting *Osteospermum fruticosum* (Cape Daisy). The recombination and mutation-driven re-emergence of begomovirus strains pose significant threats to the global sustainable cultivation of chilli [5]. PeVYV shares substantial nucleotide similarity with tobacco vein-distorting virus (TVDV) in the first half of its genome, suggesting its emergence from a recombination event between TVDV and other poleroviruses [68]. Similarly, Bonnet et al. (2005) [69] highlighted the role of recombination in the evolution of CMV, noting intragenic recombination in all three RNA segments through an analysis of Spanish isolates collected between 1989 and 2002.

The presence of the identified viruses (ChiLCV, CMV, GBNV, PCV-2, PeVYV, and BPEV) in pooled samples was confirmed through PCR and RT-PCR assays. Molecular confirmation of these viruses strengthens virome studies conducted via high-throughput sequencing (HTS), providing further validation of viral detection. In this study, we also developed a loop-mediated isothermal amplification (LAMP) diagnostic assay for these viruses, marking the first such application for future screening. The choice of LAMP was driven by its superior specificity and sensitivity compared to traditional PCR assays [70–72]. LAMP reactions are performed at a constant temperature for approximately one hour, amplifying smaller gene segments [73, 74]. The major advantage of LAMP is the ease of visualizing the amplification products and its quicker turnaround time relative to PCR [75–77]. We successfully amplified the target genes for all identified viruses using both LAMP and RT-LAMP assays.

Conclusions

This study represents the first example of virome profiling in chilli in India, underscoring the advantages of HTS over conventional virus detection methods, particularly during disease outbreaks where the causative agents were previously unknown. Virome analyses in chilli identified six distinct viruses including four previously reported (ChiLCV, CMV, PCV-2 and GBNV) and two newly reported (PeVYV and BPEV) viruses. These findings underscore the critical importance of understanding viral diversity using HTS technologies and highlights the importance of continuous surveillance of viral landscape in agricultural crops. The identification of novel or previously unrecognized viruses offers valuable insights for strategic efforts in chilli breeding and cultivation, ultimately mitigating the risk of future disease outbreaks. Virome profiling in agricultural ecosystems will provide a comprehensive understanding of the presence of both known and novel viruses, facilitating the development of targeted diagnostic assays and reducing the impact of viral diseases on food production.

Abbreviations

ICTV	International Committee on Taxonomy of Viruses
ChiLCV	Chilli leaf curl virus
ChiVMV	Chilli veinal mottle virus
CMV	Cucumber mosaic virus

CaCV	Capsicum chlorosis virus
TMV	Tobacco mosaic virus
PVY	potato virus Y
GBNV	Groundnut bud necrosis virus
TSV	Tobacco streak virus
ELISA	Enzyme-linked immunosorbent assay
PCR	Polymerase chain reaction
HTS	High-throughput sequencing
LAMP	Loop-mediated isothermal amplification
PDI	Percentage of disease incidence
FPKM	Fragments Per Kilobase of transcript per Million
ToLCBDB	Tomato leaf curl Bangladesh betasatellite
ToLCAA	Tomato leaf curl Anand alphasatellite
CYVMA	Croton yellow vein mosaic alphasatellite
AYVSA	Ageratum yellow vein Singapore alphasatellite
ToLCVirA	Tomato leaf curl Virudhunagar alphasatellite strain severe

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12985-025-02713-x.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4

Supplementary Material 5

Acknowledgements
We express our sincere gratitude to Dr Nataraja Karaba, Professor and Head, Department of Crop Physiology UAS, GKVK, Bengaluru, for guiding us and Dr M. K. Prasanna Kumar, Professor, Plant Pathology, UAS, GKVK, Bengaluru, for providing RT-PCR facility as a part of Central Instrumentation Facility. We also thank all the authorities of the University of Agricultural Sciences, GKVK, Bengaluru, for providing research support.

Author contributions
NVR generated the samples for sequencing and done the validation work, NVR, SH and PK performed bioinformatic analyses. MM, HDV, KSS, VV and CRJB performed diversity analysis. MNM edited the manuscript. CNLR conceptualized and obtained funds and provided the overall direction. All the authors are involved in literature mining and manuscript preparation. All authors read and approved the final manuscript.

Funding
We express our gratitude to the University of Agricultural Sciences, G.K.V.K., for funding the publication charges.

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate
This study did not include the use of any animals or otherwise, so did not require ethical approval.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Plant Pathology, College of Agriculture, University of Agricultural Sciences, GKVK, Bengaluru, Karnataka 560065, India
[2]Department of Biotechnology, College of Agriculture, University of Agricultural Sciences, GKVK, Bengaluru, Karnataka 560065, India
[3]Department of Agricultural Statistics, Applied Mathematics and Computer Science, College of Agriculture, University of Agricultural Sciences, GKVK, Bengaluru, Karnataka 560065, India
[4]Division of Plant Protection, ICAR-Indian Institute of Horticultural Research, Hessaraghatta Lake PO, Bangalore, Karnataka 560089, India
[5]Department of Plant Pathology, College of Horticulture, University of Horticultural Sciences, Bagalkot, Bengaluru, Karnataka 560065, India
[6]Agriculture, Health and Environment Department, Natural Resources Institute, University of Greenwich, Medway campus, Chatham, Kent ME4 4TB, UK
[7]Centre for Infectious Diseases, Biological Sciences and Technology Division, CSIR- North East Institute of Science and Technology, Jorhat, Assam 785006, India

Received: 2 February 2025 / Accepted: 25 March 2025
Published online: 24 April 2025

References

1. Khan N, Ahmed MJ, Shah SZA. Comparative analysis of mineral content and proximate composition from Chilli pepper (*Capsicum annuum* L.) germplasm. Pesqui Agropecu Bras. 2019;8.
2. Rao GP, Reddy MG. Overview of yield losses due to plant viruses. In: Awasthi LP, editor. Applied plant virology. Academic; 2020. pp. 531–62.
3. Materska M, Perucka I. Antioxidant activity of the main phenolic compounds isolated from hot pepper fruit (*Capsicum annuum* L). J Agric Food Chem. 2005;53.
4. Olatunji TL, Afolayan AJ. The suitability of Chili pepper (*Capsicum annuum* L.) for alleviating human micronutrient dietary deficiencies: A review. Food Sci Nutr. 2018; 6:2239–51.
5. Shingote PR, Wasule DL, Parma VS, Holkar SK, Karkute SG, Parlawar ND et al. An overview of Chili leaf curl disease: molecular mechanisms, impact, challenges, and disease management strategies in Indian Subcontinent. Front Microbiol. 2022;13.
6. Thakur H, Jindal SK, Sharma A, Dhaliwal MS. Chilli leaf curl virus disease: a serious threat for Chilli cultivation. J Plant Dis Prot. 2018;125:239–49.
7. Kenyon L, Kumar S, Tsai W-S, Hughes J, d'A. Virus diseases of peppers (*Capsicum* spp.) and their control. In: Loebenstein G, Katis N, editors. Advances in Virus Research. Academic press;2014. pp. 297–354.
8. Thomas JE, Gronenborn B, Harding RM, Mandal B, Grigoras I, Randles JW et al. ICTV virus taxonomy profile: *Nanoviridae*. J Gen Virol. 2021;102.
9. Satyanarayana T, Mitchell SE, Reddy DVR, Kresovich S, Jarret R, Naidu RA, et al. The complete nucleotide sequence and genome organization of the M RNA segment of peanut bud necrosis Tospovirus and comparison with other Tospoviruses. J Gen Virol. 1996;77:2347–52.
10. Jain RK, Bag S, Awasthi LP. First report of natural infection of *Capsicum annuum* by tobacco streak virus in India. Plant Pathol. 2005;54:257–257.
11. Krishnareddy M, Rani RU, Kumar KSA, Reddy KM, Pappu HR. Capsicum chlorosis virus (Genus *Tospovirus*) infecting Chili pepper (*Capsicum annuum*) in India. Plant Dis. 2008;92:1469–1469.
12. Senanayake DMJB, 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. J Phytopathol. 2012;160:146–55.
13. Ashwathappa KV, Krishna Reddy M, Venkataravanappa V, Madhavi Reddy K, Hemachandra Reddy P, Lakshminarayana Reddy CN. Genome characterization and host range studies of cucumber mosaic virus belonging to the subgroup IB infecting Chilli in India and screening of Chilli genotypes for identification of resistance. VirusDisease. 2021;32.
14. Dr B, Anjaneya Reddy MV, Praful, Ramachandra RK, Krishna Reddy M, Anjanappa M. Epidemiology of ChiVMV and loss assessment in capsicum (*Capsicum annum* Var. Grossum Sendt). J Hortic Sci. 2022;17:417–23.
15. Boonham N, Kreuze J, Winter S, van der Vlugt R, Bergervoet J, Tomlinson J, et al. Methods in virus diagnostics: from ELISA to next generation sequencing. Virus Res. 2014;186:20–31.
16. Hema M, Konakalla NC. Recent developments in detection and diagnosis of plant viruses. In: Viswanath B, editor. Recent developments in applied microbiology and biochemistry. Elsevier; 2021. pp. 163–80.
17. Pandey V, Srivastava A, Gaur RK. Begomovirus: a curse for the agricultural crops. Archives Phytopathol Plant Prot. 2021;54:949–78.

18. Czotter N, Molnar J, Szabó E, Demian E, Kontra L, Baksa I et al. NGS of virus-derived small RNAs as a diagnostic method used to determine Viromes of Hungarian vineyards. Front Microbiol. 2018;9.
19. Yadav N, Khurana SMP. Plant virus detection and diagnosis: progress and challenges. In: Shukla P, editor. ediotrs. Frontier discoveries and innovations in interdisciplinary microbiology. New Delhi: Springer India; 2016. pp. 97–132.
20. Moreno AB, López-Moya JJ. When viruses play team sports: mixed infections in plants. Phytopathology. 2020;110:29–48.
21. Mehetre GT, Leo VV, Singh G, Sorokan A, Maksimov I, Yadav MK, et al. Current developments and challenges in plant viral diagnostics: a systematic review. Viruses. 2021;13:412.
22. Maree HJ, Fox A, Al Rwahnih M, Boonham N, Candresse T. Application of HTS for routine plant virus diagnostics: state of the Art and challenges. Front Plant Sci. 2018;9:1082.
23. Sommers P, Chatterjee A, Varsani A, Trubl G. Integrating viral metagenomics into an ecological framework. Annu Rev Virol. 2021;8:133–58.
24. Jo Y, Lian S, Chu H, Cho JK, Yoo S-H, Choi H, et al. Peach RNA Viromes in six different Peach cultivars. Sci Rep. 2018;8:1844.
25. Coetzee B, Freeborough M-J, Maree HJ, Celton J-M, Rees DJG, Burger JT. Deep sequencing analysis of viruses infecting grapevines: Virome of a vineyard. Virology. 2010;400:157–63.
26. Jo Y, Choi H, Kyong Cho J, Yoon J-Y, Choi S-K, Kyong Cho W. In Silico approach to reveal viral populations in grapevine cultivar Tannat using transcriptome data. Sci Rep. 2015;5:15841.
27. Kashif M, Pietilä S, Artola K, Jones RAC, Tugume AK, Mäkinen V, et al. Detection of viruses in Sweetpotato from Honduras and Guatemala augmented by deep-Sequencing of small-RNAs. Plant Dis. 2012;96:1430–7.
28. Li R, Gao S, Hernandez AG, Wechter WP, Fei Z, Ling K-S. Deep sequencing of small RNAs in tomato for virus and viroid identification and strain differentiation. PLoS ONE. 2012;7:e37127.
29. Wylie SJ, Li H, Saqib M, Jones MGK. The global trade in fresh produce and the vagility of plant viruses: a case study in Garlic. PLoS ONE. 2014;9:e105044.
30. Jo Y, Choi H, Kim S-M, Kim S-L, Lee BC, Cho WK. Integrated analyses using RNA-Seq data reveal viral genomes, single nucleotide variations, the phylogenetic relationship, and recombination for Apple stem grooving virus. BMC Genomics. 2016;17:579.
31. Jo Y, Choi H, Kim S-M, Kim S-L, Lee BC, Cho WK. The pepper Virome: natural co-infection of diverse viruses and their quasispecies. BMC Genomics. 2017;18:453.
32. Matsumura E, Coletta-Filho H, Nouri S, Falk B, Nerva L, Oliveira T, et al. Deep sequencing analysis of RNAs from citrus plants grown in a citrus sudden death-affected area reveals diverse known and putative novel viruses. Viruses. 2017;9:92.
33. Karadi A, Samineni S, Sajja S, Sharma M, Thudi M, Mallikarjuna BP, et al. Molecular mapping of dry root rot resistance genes in Chickpea (*Cicer arietinum* L). Euphytica. 2021;217:123.
34. Sajeevan RS, Shivanna MB, Nataraja KN. An efficient protocol for total RNA isolation from healthy and stressed tissues of mulberry (*Morus* sp.) and other species. Am J Plant Sci. 2014;05:2057–65.
35. Andrews S. FastQC - A quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Babraham Bioinf. 2010.
36. Krueger F. Trim galore! A wrapper tool around cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files. Babraham Inst. 2015.
37. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 2011;29:644–52.
38. Zerbino DR, Birney E. Velvet: algorithms for de Novo short read assembly using de Bruijn graphs. Genome Res. 2008;18:821–9.
39. Bushmanova E, Antipov D, Lapidus A, Prjibelski AD. RnaSPAdes: a de Novo transcriptome assembler and its application to RNA-Seq data. Gigascience. 2019;8.
40. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics. 2012;28:3150–2.
41. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10:421.
42. Walker PJ, Siddell SG, Lefkowitz EJ, Mushegian AR, Adriaenssens EM, Alfenas-Zerbini P et al. Changes to virus taxonomy and to the international code of virus classification and nomenclature ratified by the International Committee on Taxonomy of Viruses (2021). Arch Virol. 2021;166.
43. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and clustal X version 2.0. Bioinformatics. 2007;23:2947–8.
44. Hall T, Biosciences I, Carlsbad C. BioEdit: an important software for molecular biology. GERF Bull Biosci. 2011;2.
45. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and samtools. Bioinformatics. 2009;25:2078–9.
46. Milne I, Bayer M, Cardle L, Shaw P, Stephen G, Wright F, et al. Tablet–next generation sequence assembly visualization. Bioinformatics. 2010;26:401–2.
47. Muhire BM, Varsani A, Martin DP. SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. PLoS ONE. 2014;9:e108277.
48. Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. Mol Biol Evol. 2021;38:3022–7.
49. Martin DP, Varsani A, Roumagnac P, Botha G, Maslamoney S, Schwab T et al. RDP5: a computer program for analyzing recombination in, and removing signals of recombination from, nucleotide sequence datasets. Virus Evol. 2021;7.
50. Hall RJ, Wang J, Todd AK, Bissielo AB, Yen S, Strydom H, et al. Evaluation of rapid and simple techniques for the enrichment of viruses prior to metagenomic virus discovery. J Virol Methods. 2014;195:194–204.
51. Al Rwahnih M, Daubert S, Golino D, Rowhani A. Deep sequencing analysis of RNAs from a grapevine showing Syrah decline symptoms reveals a multiple virus infection that includes a novel virus. Virology. 2009;387:395–401.
52. Massart S, Chiumenti M, De Jonghe K, Glover R, Haegeman A, Koloniuk I et al. Virus detection by high-throughput sequencing of small RNAs: Large-scale performance testing of sequence analysis strategies. Phytopathology. 2019;109.
53. Hölzer M, Marz M. De Novo transcriptome assembly: A comprehensive cross-species comparison of short-read RNA-Seq assemblers. Gigascience. 2019;8.
54. Maliogka VI, Minafra A, Saldarelli P, Ruiz-García AB, Glasa M, Katis N et al. Recent advances on detection and characterization of fruit tree viruses using high-throughput sequencing technologies. Viruses. 2018;10:436.
55. Kim NY, Lee HJ, Kim HS, Lee SH, Moon JS, Jeong RD. Identification of plant viruses infecting Pear using RNA sequencing. Plant Pathol J. 2021;37.
56. Kutnjak D, Tamisier L, Adams I, Boonham N, Candresse T, Chiumenti M et al. A primer on the analysis of high-throughput sequencing data for detection of plant viruses. Microorganisms. 2021:9:841.
57. Thakur V. Challenges and strategies for next generation sequencing (NGS) data analysis. J Comput Sci Syst Biol. 2010;03.
58. Sidharthan VK, Sevanthi AM, Jaiswal S, Baranwal VK. Robust Virome profiling and whole genome reconstruction of viruses and viroids enabled by use of available mRNA and sRNA-Seq datasets in grapevine (*Vitis vinifera* L). Front Microbiol. 2020;11.
59. Fiallo-Olivé E, Lett J-M, Martin DP, Roumagnac P, Varsani A, Zerbini FM et al. ICTV virus taxonomy profile: *Geminiviridae* 2021. J Gen Virol. 2021;102.
60. Bujarski J, Gallitelli D, García-Arenal F, Pallás V, Palukaitis P, Krishna Reddy M et al. ICTV virus taxonomy profile: *Bromoviridae*. J Gen Virol. 2019;100.
61. Sõmera M, Fargette D, Hébrard E, Sarmiento C. ICTV virus taxonomy profile: *Solemoviridae* 2021. J Gen Virol. 2021;102.
62. Vainio EJ, Chiba S, Ghabrial SA, Maiss E, Roossinck M, Sabanadzovic S et al. ICTV virus taxonomy profile: *Partitiviridae*. J Gen Virol. 2018;99.
63. Valverde RA, Khalifa ME, Okada R, Fukuhara T, Sabanadzovic S. ICTV virus taxonomy profile: *Endornaviridae*. J Gen Virol. 2019;100.
64. Mandal B, Jain RK, Krishnareddy M, Krishna Kumar NK, Ravi KS, Pappu HR. Emerging problems of tospoviruses (*Bunyaviridae*) and their management in the Indian Subcontinent. Plant Dis. 2012;96:468–79.
65. Bujarski JJ. Genetic recombination in plant-infecting messenger-sense RNA viruses: overview and research perspectives. Front Plant Sci. 2013.
66. Martin DP, Lefeuvre P, Varsani A, Hoareau M, Semegni JY, Dijoux B et al. Complex recombination patterns arising during geminivirus coinfections preserve and demarcate biologically important intra-genome interaction networks. PLoS Pathog. 2011;7.
67. Mishra M, Verma RK, Gaur RK. Identification of Chilli leaf curl virus and associated betasatellite infecting *Osteospermum fruticosum* in Rajasthan, India. 3 Biotech. 2020;10.
68. Zhao K, Yin Y, Hua M, Wang S, Mo X, Yuan E et al. Pod pepper vein yellows virus, a new Recombinant Polerovirus infecting *Capsicum frutescens* in Yunnan Province, China. Virol J. 2021;18.
69. Bonnet J, Fraile A, Sacristán S, Malpica JM, García-Arenal F. Role of recombination in the evolution of natural populations of cucumber mosaic virus, a tripartite RNA plant virus. Virology. 2005;332.
70. Saharan P, Khatri P, Dingolia S, Duhan JS, Gahlawat SK. Rapid detection of viruses using loop-mediated isothermal amplification (LAMP): A review. Biotechnology: Prospects Appl. 2013;287–306.

71. Soroka M, Wasowicz B, Rymaszewska A. Loop-mediated isothermal amplification (LAMP): the better sibling of Pcr?? Cells. 2021;10:1931.
72. Wong YP, Othman S, Lau YL, Radu S, Chee HY. Loop-mediated isothermal amplification (LAMP): a versatile technique for detection of micro-organisms. J Appl Microbiol. 2018;124:626–43.
73. Bhat AI, Aman R, Mahfouz M. Onsite detection of plant viruses using isothermal amplification assays. Plant Biotechnol J. 2022;20:1859–73.
74. Bhat AI, Rao GP. Next-Generation sequencing for diagnosis of viruses. Characterization of plant viruses. Springer protocols handbooks. New York, NY: Humana; 2020. pp. 389–95.
75. Goto M, Honda E, Ogura A, Nomoto A, Hanaki K-I. Colorimetric detection of loop-mediated isothermal amplification reaction by using hydroxy naphthol blue. Biotechniques. 2009;46:167–72.
76. Mori Y, Nagamine K, Tomita N, Notomi T. Detection of loop-mediated isothermal amplification reaction by turbidity derived from magnesium pyrophosphate formation. Biochem Biophys Res Commun. 2001;289.
77. Naganur P, Premchand U, Shankarappa KS, Mesta RK, Manjunatha C, Patil CV. Development of a loop-mediated isothermal amplification assay for detection of tomato leaf curl new Delhi virus in ridge gourd [*Luffa acutangula* (L) Roxb]. Int J Curr Microbiol Appl Sci. 2019;8.
78. Kumar RV, Singh AK, Chakraborty S. A new monopartite begomovirus species, Chilli leaf curl Vellanad virus, and associated betasatellites infecting Chilli in the Vellanad region of Kerala, India. New Dis Rep. 2012;25:20–20.
79. Saritha RK, Jain P, Baranwal VK, Jain RK, Srivastava A, Kalia P. First record of pepper cryptic virus 2 in Chilli (*Capsicum annuum*) in India. Virusdisease. 2016;27:327–8.
80. Rialch N, Sharma V, Sharma A, Sharma PN. Characterization and complete nucleotide sequencing of pepper mild mottle virus infecting bell pepper in India. Phytoparasitica. 2015;43:327–37.
81. Sharma SK, Chanu NT, Anand YR, Singh YH, Singh S, Raj C, et al. First report of large cardamom Chirke virus, a macluravirus naturally infecting Chili crop in India. Plant Dis. 2019;103:777–777.
82. Wang J-H, Zhang S-Y, Gong D, Wu Y-P, Zhang Y-L, Yu N-T, et al. First report of Chilli ringspot virus on Chili pepper in China. Plant Dis. 2012;96:462–462.
83. Han Z, Gao M, Yan Z, Hong H, Yin X, Xu S, et al. First report of Paprika mild mottle virus infecting pepper (*Capsicum annuum* L.) in China. Plant Dis. 2024;108:3664.
84. Vargas JA, Hernández E, Barboza N, Mora F, Ramírez P. First report of tomato chlorosis virus infecting sweet pepper in Costa Rica. Plant Dis. 2011;95:1482–1482.
85. González-Alvarez H, Chang-Sidorchuk L, Barboza Vargas N, González Arias G, Martínez-Zubiaur Y. First report of *tomato chlorotic spot virus* infecting pepper and common bean in Cuba. Plant Dis. 2017;101:1064.
86. Svoboda J, Leisova-Svobodova L. First report of *broad bean wilt virus*– 2 in pepper in the Czech Republic. Plant Dis. 2013;97:1261–1261.
87. Martínez RT, Poojari S, Tolin SA, Cayetano X, Naidu RA. First report of *tomato spotted wilt virus* in peppers and tomato in the Dominican Republic. Plant Dis. 2014;98:163–163.
88. Tomassoli L, Manglli A, Ahmad A, Tiberini A, Barba M. First report of pepper vein yellows virus infecting Chilli pepper (*Capsicum* spp.) in Italy. New Dis Rep. 2016;33:22–22.
89. Choi GS, Kim JH, Ryu KH, Choi JK, Chae SY, Kim JS, et al. First report of tobacco mild green mosaic virus infecting pepper in Korea. Plant Pathol J. 2002;18:323–7.
90. Tsai WS, Abdourhamane IK, Kenyon L. First report of *pepper veinal mottle virus* associated with mosaic and mottle diseases of tomato and pepper in Mali. Plant Dis. 2010;94:378–378.
91. Pagán I, Betancourt M, de Miguel J, Piñero D, Fraile A, García-Arenal F. Genomic and biological characterization of Chiltepín yellow mosaic virus, a new tymovirus infecting *Capsicum annuum* Var. Aviculare in Mexico. Arch Virol. 2010;155:675–84.
92. Tahir M, Haider MS, Briddon RW. Chili leaf curl betasatellite is associated with a distinct Recombinant begomovirus, pepper leaf curl Lahore virus, in capsicum in Pakistan. Virus Res. 2010;149:109–14.
93. Bandjo-Oreshkovikj K, Rusevski R, Kuzmanovska B, Jankulovska M, Popovski Z. Molecular detection and identification of alfalfa mosaic virus (AMV) on pepper cultivated in open fields in R. Macedonia Genetika. 2017;49:1047–57.
94. Kamran A, Lotos L, Amer MA, Al-Saleh MA, Alshahwan IM, Shakeel MT, et al. Characterization of pepper leafroll chlorosis virus, a new Polerovirus causing yellowing disease of bell pepper in Saudi Arabia. Plant Dis. 2018;102:318–26.
95. Tomašechová J, Hančinský R, Predajňa L, Kraic J, Mihálik D, Šoltys K, et al. High-throughput sequencing reveals bell pepper endornavirus infection in pepper (*Capsicum annum*) in Slovakia and enables its further molecular characterization. Plants. 2019;9:41.
96. Rubio L, Luis-Arteaga M, Cambra M, Serra J, Moreno P, Guerri J. First report of broad bean wilt virus 1 in Spain. Plant Dis. 2002;86:698–698.
97. Janssen D, Saez E, Segundo E, Martín G, Gil F, Cuadrado IM. *Capsicum annuum* – a new host of parietaria mottle virus in Spain. Plant Pathol. 2005;54:567–567.
98. Fondong VN. Geminivirus protein structure and function. Mol Plant Pathol. 2013;14:635–49.
99. Sabanadzovic S, Valverde RA. Properties and detection of two cryptoviruses from pepper (*Capsicum annuum*). Virus Genes. 2011;43:307–12.
100. Wintermantel WM, Hladky LL. Complete genome sequence and biological characterization of Moroccan pepper virus (MPV) and reclassification of lettuce necrotic stunt virus as MPV. Phytopathology. 2013;103:501–8.

Publisher's note